

# Mecánica Analítica: Notas de Clase

Rodolfo Alexander Diaz Sanchez  
Universidad Nacional de Colombia  
Departamento de Física  
Bogotá, Colombia

21 de enero de 2012

# Índice general

<b>Introducción</b>	<b>XIV</b>
<b>1. Elementos básicos de Mecánica Newtoniana</b>	<b>1</b>
1.1. Cinemática . . . . .	1
1.2. Dinámica: Leyes de Newton . . . . .	2
1.3. Trabajo y energía . . . . .	3
1.4. Torque y momento angular . . . . .	4
1.5. Dinámica de un sistema de partículas . . . . .	5
1.5.1. Definición de centro de masa . . . . .	5
1.5.2. Sistemas de partículas no aislados . . . . .	5
1.5.3. Momento angular y torque de un sistema de partículas . . . . .	7
1.5.4. Trabajo y energía de un sistema de partículas . . . . .	10
1.5.5. Conservación de la energía de un sistema de partículas . . . . .	11
1.5.6. Transformación de la energía cinética al sistema-C a partir del sistema de laboratorio . . . . .	13
1.6. Ejercicios . . . . .	13
<b>2. Principio de D'Alembert y ecuaciones de Lagrange</b>	<b>14</b>
2.1. Ligaduras . . . . .	14
2.2. Principio de D'Alembert . . . . .	16
2.3. Coordenadas generalizadas y ecuaciones de Lagrange . . . . .	17
2.3.1. Energías cinética y potencial en coordenadas generalizadas . . . . .	21
2.3.2. Una simetría gauge o de calibración para el Lagrangiano . . . . .	21
2.4. Ecuaciones de Lagrange para potenciales generalizados . . . . .	22
2.5. Ecuaciones de Lagrange con fuerzas disipativas . . . . .	24
2.6. Algunas características de las cantidades generalizadas . . . . .	25
2.7. Relación entre sistemas coordinados y sistemas de referencia . . . . .	26
2.8. Ejemplos de uso de la formulación Lagrangiana . . . . .	27
2.8.1. Partícula en el espacio . . . . .	27
2.8.2. Máquina de Atwood . . . . .	28
2.8.3. Cuenta sobre un alambre . . . . .	29
2.8.4. Gauge electromagnético en la formulación Lagrangiana . . . . .	30
2.8.5. Un sistema ligado de dos masas . . . . .	30
2.8.6. Aro sobre plano inclinado deslizante . . . . .	32
2.8.7. Un potencial generalizado para una fuerza central . . . . .	33
2.8.8. Partícula inmersa en un fluido . . . . .	35
2.9. Ventajas del formalismo Lagrangiano . . . . .	36
2.10. Ejercicios . . . . .	36

<b>3. Cálculo variacional y multiplicadores de Lagrange</b>	<b>38</b>
3.1. Algunos problemas prácticos de naturaleza variacional . . . . .	38
3.1.1. Minimización del tiempo de caída de una partícula . . . . .	38
3.1.2. Minimización de una superficie de revolución . . . . .	39
3.2. Aspectos fundamentales del cálculo de variaciones . . . . .	40
3.2.1. Cálculo variacional en una dimensión . . . . .	40
3.2.2. Cálculo de variaciones multidimensional . . . . .	43
3.3. Solución de los problemas de aplicación planteados . . . . .	43
3.3.1. Minimización de la longitud de arco . . . . .	44
3.3.2. Minimización del tiempo de caída de una partícula: la braquistócrona . . . . .	44
3.3.3. Minimización de una superficie de revolución . . . . .	45
3.4. Ligaduras y multiplicadores de Lagrange (opcional) . . . . .	46
3.4.1. Generalización a un conjunto arbitrario de variables y de ligaduras . . . . .	48
3.5. Problemas variacionales con ligaduras (opcional) . . . . .	49
3.6. Ejercicios . . . . .	52
<b>4. Principio variacional de Hamilton y Ecs. de Lagrange</b>	<b>53</b>
4.1. Aplicación del cálculo de variaciones al principio de Hamilton . . . . .	54
4.2. Extensión del principio de Hamilton a algunos sistemas no holónomos . . . . .	54
4.2.1. Significado físico de los multiplicadores de Lagrange: fuerzas de ligadura . . . . .	56
4.2.2. Formalismo de los multiplicadores para ligaduras holónomas . . . . .	57
4.3. Relación entre el principio diferencial de D'Alembert y el Principio variacional de Hamilton . . . . .	58
4.4. Aplicación del principio de Hamilton con coordenadas dependientes . . . . .	59
4.4.1. Bloque deslizante sobre una semiesfera . . . . .	59
4.4.2. Aro sobre plano inclinado con condición de rodadura . . . . .	61
4.4.3. Esfera en un hueco cilíndrico . . . . .	62
4.5. Características básicas de una formulación variacional . . . . .	64
4.6. Principio variacional para Lagrangianos que contienen a $\ddot{q}$ (opcional) . . . . .	64
4.7. Ejercicios . . . . .	66
<b>5. Simetrías y cantidades conservadas (Lagrange)</b>	<b>69</b>
5.1. Teoremas de conservación y propiedades de simetría . . . . .	69
5.1.1. Momento lineal y coordenadas globales de traslación . . . . .	71
5.1.2. Momento angular y coordenadas globales de rotación . . . . .	73
5.1.3. Consideraciones generales sobre simetrías asociadas a coordenadas cíclicas y cantidades conservadas . . . . .	75
5.2. Función energía y conservación de la energía . . . . .	75
5.2.1. Relación entre energía y función energía . . . . .	76
5.2.2. Función energía con fuerzas disipativas . . . . .	77
5.3. Teorema de Noether para sistemas discretos (opcional) . . . . .	78
5.3.1. Comentarios sobre el teorema de Noether . . . . .	82
5.4. Ejemplos de aplicación del teorema de Noether . . . . .	83
5.4.1. Invarianza ante traslación temporal y conservación de la energía . . . . .	83
5.4.2. Invarianza ante traslación espacial y conservación del momento lineal . . . . .	84
5.4.3. Invarianza ante rotaciones espaciales y la conservación del momento angular . . . . .	86
5.4.4. Transformaciones de Galileo . . . . .	88
5.5. Ejercicios . . . . .	90

<b>6. Ecuaciones de Movimiento de Hamilton</b>	<b>92</b>
6.1. Consideraciones generales . . . . .	92
6.2. Transformaciones de Legendre . . . . .	93
6.3. Generación del Hamiltoniano y Ecuaciones de Hamilton . . . . .	94
6.4. Algoritmo matricial para la obtención del Hamiltoniano . . . . .	95
6.4.1. Hamiltoniano para un cuerpo sometido a una fuerza central en coordenadas esféricas . . . . .	96
6.4.2. Hamiltoniano de una carga no relativista inmersa en un campo electromagnético . . . . .	98
6.5. Forma Simpléctica de las Ecuaciones de Hamilton . . . . .	99
6.6. Coordenadas cíclicas y teoremas de conservación . . . . .	100
6.7. El Hamiltoniano en diferentes sistemas coordenados . . . . .	101
6.7.1. Hamiltoniano de un sistema masa resorte en diferentes sistemas coordenados . . . . .	102
6.8. Problemas de aplicación de las ecuaciones de Hamilton . . . . .	103
6.8.1. Partícula sobre superficie cilíndrica . . . . .	103
6.8.2. Ejemplo de aplicación del algoritmo matricial . . . . .	106
6.8.3. Péndulo sujeto a una recorrido parabólico . . . . .	107
6.9. Procedimiento de Routh . . . . .	110
6.9.1. Partícula sometida a un potencial central atractivo por el método de Routh . . . . .	112
6.10. Ecs. de Hamilton a partir de un principio variacional . . . . .	113
6.11. El principio de mínima acción (opcional) . . . . .	115
6.11.1. Algunas aplicaciones del principio de mínima acción . . . . .	118
6.12. Ejercicios . . . . .	119
<b>7. Transformaciones canónicas</b>	<b>120</b>
7.1. Transf. Canónicas y el principio de Hamilton modificado . . . . .	121
7.2. Funciones generadoras de una transformación canónica . . . . .	122
7.3. Ejemplos de transformaciones canónicas . . . . .	125
7.4. Transformaciones canónicas para el oscilador armónico . . . . .	127
7.5. Transf. Canónicas con la forma simpléctica de las Ecs. de Hamilton . . . . .	129
7.5.1. Ejemplos de transformaciones canónicas con método matricial . . . . .	131
7.6. Método simpléctico para T.C's dependientes del tiempo . . . . .	132
7.7. Ejemplos de transformaciones canónicas . . . . .	134
7.7.1. Transformación canónica por conjugación compleja . . . . .	134
7.7.2. Transformación canónica de rotación . . . . .	135
7.7.3. Un sistema con dos grados de libertad . . . . .	137
7.8. Ejercicios . . . . .	138
<b>8. Corchetes de Poisson y otros invariantes canónicos</b>	<b>140</b>
8.1. Corchetes de Poisson . . . . .	140
8.2. Propiedades de los corchetes de Poisson . . . . .	142
8.3. Corchetes de Lagrange . . . . .	143
8.4. Otros invariantes canónicos . . . . .	144
8.5. Condi. simpléctica y funciones generatrices (opcional) . . . . .	144
8.6. Ecuaciones de movimiento con corchetes de Poisson . . . . .	146
8.7. Constantes de movimiento con corchetes de Poisson . . . . .	147
8.8. Constantes de mov. evaluadas por corchetes de Poisson . . . . .	148
8.8.1. Sistema con dos grados de libertad . . . . .	148
8.8.2. Otro sistema con dos grados de libertad . . . . .	149
8.8.3. Constante de movimiento del oscilador armónico . . . . .	151
8.9. Transf. Canónicas infinitesimales y corchetes de Poisson . . . . .	152
8.10. Cambios activo y pasivo de una función del sistema bajo una T.C. . . . .	154
8.11. Cambio del Hamiltoniano bajo una transformación canónica . . . . .	155

8.12. Cantidades conservadas e invarianzas del Hamiltoniano . . . . .	156
8.12.1. El momento lineal total como generador de TCI's que generan traslaciones . . . . .	158
8.12.2. El momento angular total como generador de TCI's que generan rotaciones . . . . .	158
8.13. Construcción de TC's finitas a partir de TCI's . . . . .	160
8.13.1. Aplicación del operador evolución temporal en el movimiento uniformemente acelerado . . . . .	162
8.13.2. Aplicación del operador evolución temporal en el movimiento armónico simple . . . . .	162
8.13.3. Aplicación del operador evolución paramétrica para la generación de rotaciones . . . . .	163
8.14. Propiedades de los corchetes de Poisson de los momentos angulares . . . . .	164
8.14.1. Ejemplos de aplicación . . . . .	166
8.15. Ejercicios . . . . .	167
<b>9. Teoría de Hamilton-Jacobi y variables acción-ángulo</b> . . . . .	<b>168</b>
9.1. Ec. de Hamilton-Jacobi para la función principal de Hamilton . . . . .	168
9.1.1. Interpretación física de las soluciones de Hamilton-Jacobi . . . . .	170
9.2. Solución del oscilador armónico por el método de Hamilton-Jacobi . . . . .	171
9.2.1. Oscilador armónico unidimensional con el método de Hamilton Jacobi . . . . .	171
9.2.2. Oscilador armónico bidimensional anisotrópico . . . . .	173
9.2.3. Oscilador armónico bidimensional isotrópico . . . . .	174
9.3. Ec. de H-J para la función característica de Hamilton . . . . .	175
9.4. Paralelismo entre los dos formalismos de H-J . . . . .	177
9.5. Separación de variables en la ecuación de Hamilton-Jacobi . . . . .	179
9.5.1. Coordenadas ignorables y separabilidad . . . . .	180
9.5.2. Condiciones más generales para la separabilidad . . . . .	181
9.6. Fuerzas centrales en el formalismo de Hamilton-Jacobi . . . . .	182
9.6.1. Problema bidimensional . . . . .	183
9.6.2. La dinámica de las fuerzas centrales como problema tridimensional . . . . .	184
9.7. Otros problemas de aplicación con el formalismo de H-J . . . . .	186
9.7.1. Partícula sometida a potencial armónico y campo magnético . . . . .	186
9.7.2. Partícula bajo potencial conservativo en coordenadas elipsoidales . . . . .	188
9.8. Variables acción-ángulo para sistemas con un grado de libertad . . . . .	191
9.8.1. Formulación general de la variables acción-ángulo en una dimensión . . . . .	193
9.9. Problemas de variables acción-ángulo con un grado de libertad . . . . .	195
9.9.1. El oscilador armónico unidimensional en variables acción-ángulo . . . . .	195
9.9.2. Partícula en movimiento periódico en una dimensión bajo un potencial $V(x) = F x $ . . . . .	196
9.10. Variables acción-ángulo para sistemas completamente separables <sup>1</sup> . . . . .	197
9.10.1. Movimientos periódicos múltiples de libración . . . . .	199
9.10.2. Movimientos cuasi periódicos múltiples de rotación . . . . .	200
9.10.3. Movimientos periódicos simples y múltiples tipo libración . . . . .	201
9.10.4. Variables acción-ángulo para sistemas degenerados . . . . .	203
9.11. Comentarios finales sobre las variables acción-ángulo . . . . .	206
9.12. El problema de Kepler en variables acción-ángulo . . . . .	206
9.12.1. Variables acción-ángulo teniendo en cuenta la degeneración . . . . .	211
9.13. Ejercicios . . . . .	212
<b>10. Fuerzas centrales</b> . . . . .	<b>214</b>
10.1. Reducción al problema de dos partículas desacopladas . . . . .	214
10.2. Ecuaciones de movimiento y primeras integrales . . . . .	216
10.3. Problema unidimensional asociado y clasificación de órbitas . . . . .	220
10.4. Análisis de curvas de potencial efectivo . . . . .	221
10.4.1. Potencial efectivo para interacción kepleriana . . . . .	221

---

<sup>1</sup>En lo que sigue del capítulo no adoptaremos la convención de suma de índices repetidos a menos que se indique lo contrario.

10.4.2. Potencial efectivo equivalente para dos cuerpos no interactuantes . . . . .	223
10.4.3. Potencial atractivo proporcional al inverso del cubo de la distancia . . . . .	226
10.4.4. Potencial efectivo para fuerza restauradora lineal . . . . .	227
10.4.5. Consideraciones generales sobre curvas de potencial efectivo . . . . .	227
10.5. El teorema del virial . . . . .	228
10.5.1. Otras aplicaciones del teorema del virial . . . . .	230
10.6. Ecuación de la órbita y potenciales integrables . . . . .	230
10.7. Condición para órbitas circulares estables e inestables . . . . .	233
10.8. Órbitas circulares perturbadas a primer orden . . . . .	235
10.9. Órbitas circulares perturbadas y condic. para órbitas cerradas . . . . .	237
10.10. Órbitas circulares perturbadas con variables ac.-ang. (Opcional) . . . . .	238
10.11. El problema de Kepler: Ley del inverso al cuadrado (atractiva) . . . . .	239
10.12. Solución para la órbita en el problema de Kepler . . . . .	240
10.12.1. Clasificación de las órbitas según los valores de $E$ y $l$ . . . . .	242
10.12.2. Condición de circularidad . . . . .	243
10.12.3. Órbitas elípticas . . . . .	243
10.13. Movimiento en el tiempo en el problema de Kepler . . . . .	244
10.13.1. Dependencia temporal en el caso parabólico . . . . .	245
10.13.2. Dependencia temporal para el movimiento elíptico . . . . .	245
10.13.3. Tercera ley de Kepler . . . . .	247
10.14. Vector de Laplace-Runge-Lenz . . . . .	248
10.15. Parametrización de las órbitas keplerianas en el espacio . . . . .	250
10.16. Problema de Kepler en variables acción-ángulo revisado (opcional) . . . . .	252
10.16.1. Determinación del Hamiltoniano en términos de variables de acción . . . . .	253
10.16.2. Relación entre variables acción-ángulo y variables orbitales . . . . .	254
10.17. Ejercicios . . . . .	258
<b>11. Colisiones y dispersión</b> . . . . .	<b>260</b>
11.1. Colisiones y dispersiones generales . . . . .	260
11.1.1. Caso especial 1: reacción de captura . . . . .	263
11.1.2. Caso especial 2, blanco en reposo . . . . .	263
11.2. Dispersión en un campo de fuerzas centrales . . . . .	265
11.2.1. Dispersión de Rutherford . . . . .	269
11.2.2. Características generales de la sección eficaz . . . . .	270
11.3. Dispersión vista por el laboratorio y el CM (blanco en reposo) . . . . .	274
11.3.1. Relación entre el ángulo de dispersión medido por el laboratorio y el medido por el centro de masa . . . . .	274
11.3.2. Caracterización del factor $\rho$ de la colisión . . . . .	276
11.3.3. Sección eficaz en términos de los dos ángulos de dispersión . . . . .	277
11.4. Ejercicios . . . . .	279
<b>12. Interludio: Matrices, vectores y tensores</b> . . . . .	<b>282</b>
12.1. Propiedades básicas de las matrices . . . . .	282
12.1.1. Determinantes y trazas de matrices . . . . .	285
12.1.2. Matrices rectangulares . . . . .	286
12.2. Interpretación activa y pasiva de las transf. lineales . . . . .	287
12.2.1. Transformaciones de similaridad . . . . .	288
12.3. Problema de valores propios . . . . .	289
12.3.1. El problema de la degeneración de valores propios . . . . .	291
12.4. Propiedades básicas de las matrices ortogonales . . . . .	293
12.4.1. Matrices ortogonales y norma de vectores complejos (opcional) . . . . .	294

12.4.2. Transformaciones ortogonales propias e impropias . . . . .	295
12.5. Vector asociado a una matriz antisimétrica real $3 \times 3$ . . . . .	297
12.6. Propiedades de paridad de vectores y escalares . . . . .	298
12.7. Transformaciones ortogonales propias . . . . .	300
12.7.1. Matrices ortogonales reales propias en tres dimensiones . . . . .	301
12.8. Matriz adjunta . . . . .	302
12.9. Matrices unitarias y cambios de base (opcional) . . . . .	303
12.10. Matrices con espectro completo . . . . .	306
12.11. Matrices Hermíticas y simétricas reales . . . . .	307
12.11.1. Matrices reales simétricas . . . . .	309
12.11.2. Problema de valores propios de matrices hermíticas en tres dimensiones . . . . .	309
12.11.3. Matrices simétricas reales en $\mathbb{R}^3$ . . . . .	310
12.12. Matrices normales (opcional) . . . . .	312
12.13. Matrices positivas y definidas positivas . . . . .	314
12.13.1. Matrices simétricas reales que son positivas . . . . .	315
12.14. Problema de valores propios modificado para matrices positivas . . . . .	316
12.14.1. Diagonalización simultánea de dos formas cuadráticas . . . . .	319
12.15. Interpretación geométrica de la diagonalización de dos matrices . . . . .	320
12.15.1. Argumentación por geometría analítica . . . . .	320
12.15.2. Argumentación por geometría de Riemann . . . . .	324
12.16. Tensores cartesianos . . . . .	328
12.17. Los grupos $O(3)$ y $SO(3)$ y la definición de tensores cartesianos . . . . .	328
12.18. Tensores de $SO(3)$ . . . . .	329
12.19. Tensores de $O(3)$ y el concepto de pseudotensor . . . . .	331
12.19.1. Los tensores de Kronecker $\delta_{ij}$ y de Levi-Civita $\varepsilon_{ijk}$ . . . . .	332
12.20. Diadas y afinores . . . . .	333
12.21. Contracción de tensores, reglas del cociente . . . . .	334
12.22. Ejercicios . . . . .	335
<b>13. Cinemática del cuerpo rígido . . . . .</b>	<b>338</b>
13.1. Coordenadas independientes de un cuerpo rígido . . . . .	338
13.2. Asignación de los grados de libertad de un cuerpo rígido . . . . .	340
13.3. Transformaciones ortogonales . . . . .	342
13.4. Ángulos de Euler . . . . .	343
13.5. Parámetros de Cayley-Klein (Opcional) . . . . .	346
13.6. Teorema de Euler para el movimiento del cuerpo rígido . . . . .	347
13.7. Rotaciones finitas . . . . .	351
13.7.1. Forma matricial de la fórmula de rotación . . . . .	352
13.7.2. Relación entre la parametrización eje-ángulo $(\mathbf{n}, \Phi)$ y los ángulos de Euler . . . . .	353
13.8. Rotaciones infinitesimales . . . . .	354
13.9. Rotaciones finitas e infinitesimales en convención quiral derecha . . . . .	357
13.9.1. Construcción de rotaciones finitas por integración de rotaciones infinitesimales . . . . .	359
13.10. Velocidad angular en términos de los ángulos de Euler . . . . .	360
13.11. Razón de cambio de un vector visto por sistemas rotantes . . . . .	361
13.11.1. Razón de cambio por argumentos vectoriales . . . . .	361
13.11.2. Razón de cambio por argumentos algebraicos . . . . .	362
13.11.3. Segunda derivada en el sistema rotante . . . . .	363
13.12. Sistemas no inerciales rotantes . . . . .	365
13.12.1. La tierra como sistema rotante . . . . .	366
13.12.2. Superficie de un líquido rotante . . . . .	367
13.12.3. Cuenta sobre un alambre (revisado) . . . . .	369

13.12.4. Deflexión de una masa en caída libre . . . . .	369
13.13. Ejercicios . . . . .	371
<b>14. Ecuaciones de movimiento del cuerpo rígido</b>	<b>372</b>
14.1. Momento ang. y energía cinét. alrededor de un punto . . . . .	372
14.1.1. Momento angular, velocidad angular y tensor de inercia . . . . .	374
14.2. Tensor de inercia y momento de inercia . . . . .	376
14.2.1. Teorema de los ejes paralelos y de los ejes perpendiculares . . . . .	378
14.3. Compendio de propiedades del tensor de inercia . . . . .	380
14.4. Ecuaciones de Euler para el movimiento de un cuerpo rígido . . . . .	385
14.4.1. Ecuaciones de Euler con el formalismo de Lagrange . . . . .	387
14.5. Precesión libre . . . . .	388
14.5.1. Construcción de Poincot para la precesión libre . . . . .	388
14.5.2. Elipsoide de Binet y el momento angular . . . . .	390
14.5.3. Elipsoide de Binet, rotación estacionaria y condiciones para la rotación estable . . . . .	391
14.5.4. Solución algebraica para la precesión libre con simetría axial . . . . .	393
14.5.5. Estabilidad de sólidos irregulares con precesión libre por métodos algebraicos . . . . .	394
14.6. La peonza simétrica pesada con un punto fijo . . . . .	395
14.6.1. Planteamiento del Lagrangiano . . . . .	395
14.6.2. Reducción del problema a cuadraturas . . . . .	397
14.6.3. Análisis cualitativo del movimiento . . . . .	399
14.6.4. Análisis cuantitativo aproximado de la peonza rápida . . . . .	403
14.6.5. Peonza con precesión regular . . . . .	407
14.6.6. Peonza inicialmente vertical . . . . .	411
14.6.7. Efectos de fricción, torques adicionales y aplicaciones de la peonza simétrica pesada . . . . .	412
14.7. Ejercicios . . . . .	414
<b>15. Oscilaciones</b>	<b>416</b>
15.1. Pequeñas oscilaciones y equilibrio estable . . . . .	416
15.2. Solución de las Ecs. de mov. como problema de valores propios . . . . .	419
15.2.1. Un ejemplo con dos grados de libertad . . . . .	421
15.3. Problema de valores propios con degeneración . . . . .	424
15.3.1. Un ejemplo bidimensional con degeneración . . . . .	425
15.4. Frecuencias de vibración libre y coordenadas normales . . . . .	425
15.5. Vibraciones libres de una molécula triatómica . . . . .	429
15.5.1. Modos normales de frecuencia cero: traslaciones y rotaciones rígidas . . . . .	431
15.5.2. Vectores propios de la ecuación secular . . . . .	432
15.5.3. Modos normales y modos reales de la molécula triatómica . . . . .	432
15.5.4. Análisis cualitativo de vibraciones transversales y longitudinales . . . . .	435
15.6. Modos normales puros y soluciones físicas asociadas . . . . .	436
15.7. Vibraciones forzadas y amortiguadas . . . . .	437
15.7.1. Vibraciones forzadas . . . . .	437
15.7.2. Vibraciones amortiguadas . . . . .	439
15.7.3. Vibraciones amortiguadas forzadas . . . . .	442
15.8. Ejemplos de oscilaciones anarmónicas (opcional) . . . . .	444
15.8.1. Ecuaciones del péndulo amortiguado forzado de amplias oscilaciones . . . . .	444
15.8.2. Soluciones del péndulo amortiguado forzado de amplias oscilaciones . . . . .	445
15.9. Ejercicios . . . . .	450



<b>16. Relatividad especial</b>	<b>451</b>
16.1. Propiedades de las transformaciones de Lorentz puras . . . . .	451
16.1.1. Transformaciones de Lorentz puras y matrices ortogonales complejas . . . . .	453
16.2. Transformaciones de Lorentz restringidas . . . . .	455
16.2.1. Transformaciones de Lorentz restringidas: Boosts y rotaciones . . . . .	456
16.2.2. Composición de boosts . . . . .	457
16.2.3. Precesión de Thomas . . . . .	460
16.3. Transf. de Lorentz en espacios de Riemann . . . . .	461
16.4. El concepto de formulación covariante en Física . . . . .	464
16.5. Formulaciones covariantes en el espacio de Minkowski . . . . .	466
16.5.1. Cuadrivectores notables . . . . .	468
16.6. Fuerza y momento en relatividad especial . . . . .	470
16.7. Energía y relación momento-energía en relatividad especial . . . . .	472
16.8. Formulación Lagrangiana de la mecánica relativista . . . . .	475
16.9. Formulación no manifiestamente covariante . . . . .	475
16.9.1. Movimiento relativista bajo una fuerza constante (hiperbólico) . . . . .	477
16.9.2. Oscilador armónico unidimensional relativista . . . . .	479
16.9.3. Movimiento de partícula cargada en un campo magnético . . . . .	481
16.10. Formulaciones lagrangianas covariantes . . . . .	481
16.11. Ejercicios . . . . .	485
<b>17. Teoría canónica de perturbaciones</b>	<b>486</b>
17.1. Variación de ctes para perturbaciones dependientes del tiempo . . . . .	487
17.2. Perturb. dependiente del tiempo en términos de los parámetros de mov. . . . .	488
17.3. Variación periódica y variación secular de una perturbación . . . . .	489
17.4. Ejemplos en teoría de perturb. dependiente del tiempo . . . . .	490
17.4.1. Oscilador armónico como perturbación de la partícula libre . . . . .	490
17.4.2. Péndulo plano con amplitud finita . . . . .	493
17.4.3. Perturbaciones en el problema de Kepler acotado . . . . .	497
17.5. Perturbaciones indep. del tiempo: primer orden con un grado de libertad . . . . .	499
17.5.1. Péndulo plano con oscilación finita usando método de perturbación independiente del tiempo . . . . .	503
17.6. Perturb. indep. del tiempo: orden superior y varios grados de libertad . . . . .	504
17.6.1. Oscilador anarmónico unidimensional . . . . .	508
17.7. Perturb. indep. del tiempo con degeneración . . . . .	509
17.8. Aspectos cualitativos de la teoría clásica de perturbaciones . . . . .	511
17.9. Invariantes adiabáticos . . . . .	512
17.9.1. Invarianza adiabática del oscilador armónico . . . . .	514
17.9.2. Variación asintótica de $J$ para el oscilador armónico (opcional) . . . . .	517
17.9.3. Un invariante exacto del oscilador armónico con frecuencia dependiente del tiempo (opcional) . . . . .	517
17.9.4. Invariantes adiabáticos de partículas cargadas en campos electromagnéticos . . . . .	519
17.10. Ejercicios . . . . .	522
<b>18. Formulación Lagrangiana y Hamiltoniana para campos</b>	<b>523</b>
18.1. Ecuación de continuidad . . . . .	523
18.2. Transición de un sistema discreto a un sistema continuo . . . . .	527
18.3. Formulación lagrang. para una dimensión y una variable de campo . . . . .	530
18.4. Formulación lagrang. para 3 dim. y varias variables de campo . . . . .	531
18.5. El tensor esfuerzo energía y teoremas de conservación asociados . . . . .	534
18.5.1. Interpretación física de $T_0^\mu$ : densidad de energía y vector de Poynting . . . . .	535

18.5.2.	Interpretación física de $T_i^j$ : densidad de momento lineal y tensor de esfuerzos . . . . .	536
18.5.3.	Energía y momento total del sistema . . . . .	540
18.5.4.	Densidad de momento angular y momento angular total . . . . .	541
18.6.	Formulación Hamiltoniana para medios continuos . . . . .	542
18.6.1.	Propiedades básicas de $\mathcal{H}$ . . . . .	544
18.6.2.	Densidades generalizadas y corchetes de Poisson . . . . .	545
18.6.3.	Formulación por corchetes de Poisson utilizando descomposición de Fourier . . . . .	546
18.7.	Ejemplos de teorías de campos . . . . .	549
18.7.1.	Un modelo juguete . . . . .	549
18.8.	Teoría de campos relativista . . . . .	551
18.9.	Algunas teorías de campos relativistas . . . . .	555
18.9.1.	Campo escalar complejo . . . . .	556
18.9.2.	Ecuación de seno Gordon . . . . .	559

# Prefacio

Estas notas de clase tienen como objetivo ser una guía para un curso de mecánica analítica, en donde los principios de la mecánica clásica se examinan a la luz de las formulaciones Lagrangiana y Hamiltoniana y las variantes que de ellas se derivan. Estas formulaciones ubican a la energía en el papel fundamental que las fuerzas tienen en la formulación Newtoniana. Quizás el aspecto más atractivo de éstas formulaciones consiste en su poder para enlazar las simetrías de un sistema con las constantes de movimiento, y en la riqueza de estrategias para extraer información del sistema sin resolver explícitamente las ecuaciones de movimiento. Estos aspectos se enfatizan fuertemente a lo largo del texto, aunque en algunos casos se estudia la solución completa de acuerdo con la conveniencia y la simplicidad de estas soluciones. Se ha pretendido enfatizar en aspectos que en opinión del autor, han presentado fuerte dificultad en el desarrollo de las clases. A manera de ejemplo, la discusión del álgebra matricial que precede al estudio de la cinemática del cuerpo rígido es considerablemente extensa y detallada haciendo énfasis tanto en lo geométrico como en lo algebraico. Debe notarse sin embargo que el contenido de esta sección de álgebra matricial va más allá de las necesidades del curso presente, lo cual tiene como fin preparar al estudiante para trabajar no solo en los espacios euclidianos  $\mathbb{R}^n$  sino también en los espacios unitarios  $\mathbb{C}^n$  que juegan un papel fundamental en mecánica cuántica. Los capítulos que se incluyen son en opinión del autor de gran importancia para la formación general del Físico y constituyen el punto de partida de muchas ramas de la Física.

Es muy claro a lo largo de la lectura de las notas, que éstas últimas se han generado con una influencia considerable del clásico texto de Herbert Goldstein, especialmente de la segunda y tercera edición. No obstante, existen cambios de enfoque y/o presentación de numerosas unidades temáticas, debidos a la influencia de otros autores (tales como José, Saletan, Cromer, Whittaker, Marion etc.), así como de algunos abordajes propios del autor. A manera de ejemplo, los ángulos de Euler se han introducido de manera que no solo quede claro el algoritmo de rotación, sino la necesidad de dicho algoritmo. Se ha realizado un considerable esfuerzo por presentar de manera clara la filosofía e implementación del método de Hamilton-Jacobi. La mayor parte de herramientas matemáticas necesarias se han aislado en capítulos independientes a fin de dar más flexibilidad al texto y con el fin de que el lector las capture en su esencia y no las asocie a problemas muy específicos de la Física, lo cual dificulta en general la aplicación de estas herramientas en otros escenarios de la Física diferentes a los aprendidos. Algunas secciones se han indicado como opcionales, a fin de facilitar al lector una primera lectura, y al mismo tiempo, darle al texto la riqueza necesaria para ir más allá de lo estrictamente básico, sugiriendo caminos que incentiven la curiosidad del lector.

Para una adecuada comprensión de estas notas, el lector debe tener conocimientos a nivel introductorio sobre mecánica newtoniana, así como de álgebra lineal y cálculo diferencial e integral. En algunos pasajes aislados se asume un conocimiento básico de electricidad, magnetismo y ondas.

El capítulo 1 es un repaso de la mecánica clásica en la llamada “formulación Newtoniana” en donde la fuerza es la cantidad dinámica central. El capítulo 2 nos presenta el principio de D’Alembert y las ecuaciones de Lagrange, enfatizando que estas formulaciones apuntan a resolver dos problemas importantes en la dinámica de sistemas clásicos: (a) excluir a las ligaduras de la formulación debido a la dificultad que usualmente se presenta para obtenerlas y (b) trabajar solo con los grados de libertad independientes, evitando las coordenadas redundantes. El capítulo 3 es un suplemento matemático sobre el cálculo variacional. A pesar de que este

suplemento será aplicado mayormente en el llamado principio variacional de Hamilton, la exposición muestra la posibilidad de aplicar esa herramienta matemática en otros ámbitos de la Física. El capítulo 4 trata sobre la formulación integral del formalismo de Lagrange, cuya formulación diferencial fué presentada en el capítulo 2. En el capítulo 5, se presenta una de las ventajas notables de la formulación de Lagrange, a saber la explotación sistemática de las simetrías del sistema y su relación con las cantidades conservadas.

En el capítulo 6 se presenta la formulación de Hamilton de la mecánica clásica en la cual se sustituyen las coordenadas y velocidades generalizadas por las coordenadas generalizadas y momentos conjugados, se discutirá la ventaja de trabajar con este nuevo sistema de variables (denominadas variables canónicas), así como las ventajas de pasar de un sistema de  $n$  ecuaciones diferenciales de segundo orden, a un sistema de  $2n$  ecuaciones diferenciales de primer orden. Puesto que la función Hamiltoniana se puede expresar en cualquier conjunto de variables canónicas, es necesario estudiar la relación que hay entre los diversos conjuntos de variables canónicas, así como las transformaciones que nos llevan de un conjunto de variables canónicas a otro, este será el tema del capítulo 7. Una forma alternativa de presentar las ecuaciones de Hamilton, es a través de diversos invariantes canónicos, entre los cuales se destacan los llamados “corchetes de Poisson”, con los cuales las simetrías y constantes de movimiento se pueden identificar con gran facilidad, este será el tema del capítulo 8. Por otra parte, una elección adecuada de variables canónicas puede conducirnos a trivializar las ecuaciones de movimiento (ecuaciones de Hamilton), en cuyo caso la tarea principal consiste en encontrar la transformación del sistema canónico original al sistema canónico que trivializa dichas ecuaciones, esta es la esencia de la teoría de Hamilton-Jacobi, presentada en el capítulo 9.

En este punto culmina la descripción de estas formulaciones alternativas de la mecánica clásica. En los capítulos posteriores, se describen problemas clásicos específicos que se abordarán con una formulación Lagrangiana y/o Hamiltoniana. En el capítulo 10 se aborda el problema de las fuerzas centrales, y el tipo de órbitas que dichas fuerzas generan. Como caso particular importante, se estudia en detalle el problema de Kepler. En el capítulo 11, se estudia la teoría de colisiones y dispersión, con especial énfasis en la teoría de la dispersión por potenciales centrales. Otro interludio matemático es presentado en el capítulo 12 concerniente a matrices, vectores y tensores cartesianos, en este capítulo se desarrollan diversas herramientas matemáticas necesarias para capítulos subsecuentes. En particular para el capítulo 13 de cinemática del cuerpo rígido, en el cual se analiza el movimiento general de un cuerpo rígido, utilizando la parametrización por ángulos de Euler y la parametrización eje-ángulo. En tal capítulo se analiza además la razón de cambio de un vector visto por un sistema rotante, lo cual nos permitirá estudiar con naturalidad a los sistemas no inerciales rotantes. En el capítulo 14 se estudia la dinámica del cuerpo rígido, privilegiando el estudio de la precesión libre y de la peonza simétrica pesada con punto fijo. El capítulo 15 sobre oscilaciones, estudia el problema de la obtención de los modos normales de oscilación (desacople de las ecuaciones diferenciales), de un conjunto de osciladores acoplados, como un problema de valores propios modificado que involucra matrices positivas. Este abordaje del problema posee la ventaja (con respecto a la forma tradicional utilizada en los cursos de oscilaciones y ondas) que es extendible a coordenadas generalizadas, incluso si éstas no están asociadas a un sistema ortogonal de vectores unitarios.

En el capítulo 16 se introduce la formulación Lagrangiana y Hamiltoniana de la relatividad especial. Se ha utilizado el espacio con eje temporal imaginario  $ict$ , con métrica trivial. A pesar de que la mayor parte de textos de relatividad modernos privilegian la métrica  $g_{\mu\nu}$  con eje real, el uso de un eje temporal imaginario nos permitirá aprovechar la teoría de matrices ortogonales desarrollada en los capítulos 12 y 13. Por otra parte, dado que no se abordará la relatividad general, ambas escogencias poseen aproximadamente las mismas ventajas. En todo caso, la “traducción” entre las dos formulaciones se presenta en la sección 16.3.

En el capítulo 17 se introduce la teoría canónica de perturbaciones dependiente e independiente del tiempo, basada en el formalismo de Hamilton-Jacobi. Se analizará además el fenómeno de la invarianza adiabática, para lo cual las variables acción-ángulo serán particularmente ventajosas.

Finalmente, el capítulo 18 estudia la formulación Lagrangiana y Hamiltoniana para la mecánica de medios continuos (teoría clásica de campos), tanto en el régimen no-relativista como en el relativista, estableciendo la ecuación de continuidad para el flujo de cualquier medio continuo, y la conservación de la carga generalizada. Este formalismo es aplicable a todo tipo de medios continuos tales como campos electromagnéticos, campos de presiones, de temperatura, fluidos etc. No obstante, en el capítulo 18 nos restringimos a describir modelos

suficientemente simples para ilustrar los principios fundamentales.

El material cubierto en estas notas está pensado para dos cursos cada uno de 16 semanas con una intensidad de 4 horas semanales. Complementado quizás, con una introducción a la teoría del caos clásico. Por supuesto, los capítulos 3 y 12 relacionados con herramientas matemáticas, podrán tomarse a diferentes ritmos o grados de detalle, dependiendo del nivel de preparación de los lectores en estos temas. Varias distribuciones en la presentación de los temas son posibles. Si se toma el orden de capítulos en éstas notas, un primer curso puede ser hasta el capítulo 11, siendo los capítulos remanentes para un segundo curso. Sin embargo, una vez estudiado el capítulo 6 es posible saltar directamente al capítulo 10 y continuar la secuencia de capítulos obviando algunas secciones (por ejemplo, la sección 10.16 en la que se trata el problema de Kepler con variables acción-ángulo), con excepción del capítulo 17, el cual depende fuertemente de la formulación de Hamilton-Jacobi.

Vale la pena resaltar que el material presente son notas de clase y no un libro de texto. Por esta razón aparecen algunos desarrollos en excesivo detalle, ya que fueron producto del ejercicio directo de preparación para la clase. Espero que tales desarrollos no se conviertan en un distractor para el lector, quien puede obviar estos detalles de ser necesario. No obstante, considero que las notas en su presente forma están autocontenidas para ser usadas en una clase, o para autoaprendizaje.

Quiero expresar mis agradecimientos a los estudiantes del Departamento de Física de la Universidad Nacional de Colombia, sede Bogotá, por sus contínuas contribuciones y observaciones sobre el texto y el curso en general. Al profesor Eduardo Brieva, quien fuera mi instructor de mecánica analítica en mis años de estudiante, y a quien debo mi comprensión de buena parte del material aquí presentado. A los profesores John Morales y William Herrera por las discusiones sobre varios temas que contribuyeron a madurar el presente texto. A toda mi familia por su constante apoyo cuando las vicisitudes parecían oscurecer el camino. A mis hijos Iris Soraya y David Leandro, por ser siempre una fuerza motora de mi existencia.

Rodolfo Alexander Diaz Sanchez  
Facultad de Ciencias. Departamento de Física  
Universidad Nacional de Colombia  
Bogotá, Enero de 2012.

# Introducción

En las formulaciones Lagrangiana y Hamiltoniana de la Mecánica Clásica (así como en sus formulaciones derivadas), existen varias estrategias que aportan un considerable valor agregado con respecto a la formulación Newtoniana. En las presentes notas se ha procurado enfatizar reiteradamente en aquellos puntos que en opinión del autor, constituyen los valores agregados más fuertes.

Discutiremos brevemente dos aspectos que constituyen la motivación de una formulación Lagrangiana: (a) La eliminación de las fuerzas de ligadura, de las ecuaciones de movimiento y (b) el uso del mínimo número posible de coordenadas. Para comprender la motivación del inciso (a) bastará que el lector examine con cuidado un problema como el de una partícula que desliza sobre una trayectoria hiperbólica, e intente encontrar el valor de la fuerza normal (fuerza de ligadura) que mantiene a la partícula sobre la trayectoria en cuestión. Para el inciso (b) basta con decir que cuando un conjunto de  $N$  partículas están ligadas (por ejemplo si las distancias entre ellas son constantes), el número de coordenadas independientes es menor que  $3N$ , pero en la formulación Newtoniana tendremos que plantear las ecuaciones para las  $3N$  coordenadas, obteniendo así información redundante, para posteriormente incorporar la ligadura. Hay entonces un considerable ahorro al elaborar una formulación en donde de entrada se trabaja solo sobre coordenadas independientes. Esta misma filosofía se conserva en la formulación Hamiltoniana.

Otra ventaja de estas formulaciones consiste en que permitirá un uso más sistemático de las simetrías del sistema para extraer información total o parcial de éste.

De otra parte, aunque a través del texto se estudian problemas físicos específicos, es también común abordar temas introduciendo un “pensamiento físico abstracto”, en el sentido de que ciertos aspectos estructurales nos darán información parcial del sistema, independiente de los detalles de éste. A manera de ejemplo: para muchos sistemas físicos se puede construir una cantidad denominada Lagrangiano, y que depende de un conjunto de coordenadas generalizadas  $q_i$ , velocidades generalizadas  $\dot{q}_i$  y el tiempo

$$L = L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n; t)$$

supongamos que un Lagrangiano es tal que aparece la velocidad generalizada  $\dot{q}_k$  pero no aparece su coordenada generalizada asociada  $q_k$ , cuando esto ocurre existe una cantidad que es constante de movimiento, denominada momento conjugado a  $q_k$

$$p_k \equiv \frac{\partial L}{\partial \dot{q}_k} = cte$$

esta característica solo depende de un aspecto estructural del Lagrangiano, no de los detalles del sistema, ni siquiera importa si el sistema es mecánico, eléctrico o de otra naturaleza.

Otro aspecto que nos introduce en el pensamiento físico abstracto es la introducción constante de **cantidades generalizadas**. Las coordenadas generalizadas son simplemente las variables mínimas independientes de un sistema y no tienen que ser necesariamente variables de posición. Así mismo,  $\dot{q}$  no es necesariamente una velocidad lineal. Una densidad generalizada  $\rho(\mathbf{x}, t)$  es cantidad de “carga generalizada” por unidad de volumen, donde la carga generalizada es cualquier cantidad física escalar tal como la carga eléctrica, la masa, la energía, la probabilidad etc. A esta cantidad escalar se le puede asociar una propiedad de transporte a través de una **densidad de corriente generalizada**, no importa si se transporta energía, masa, carga eléctrica, probabilidad etc. La dinámica de estas densidades y densidades de corriente generalizada serán válidas para éstos y muchos otros escenarios al tiempo. En particular, la formulación de la ecuación de continuidad adquirirá un poder extraordinario con esta forma de **pensamiento generalizado**.

Acorde con lo anterior, se ha procurado mantener un balance entre el “pensamiento físico específico” y el “pensamiento físico abstracto”, competencias ambas indispensables en la formación del físico y otros profesionales afines.

Por otra parte, es muy común enfocar un curso de mecánica analítica como un “puente” necesario para abordar los cursos de mecánica cuántica. En opinión del autor, ésta no debe constituir la única motivación para dictar un curso de esta naturaleza. Ciertamente los postulados de la mecánica cuántica requieren del conocimiento de la formulación Hamiltoniana y también se puede abordar con el formalismo Lagrangiano. Sin embargo, la mecánica clásica posee numerosos problemas abiertos puros y aplicados (caos, mecánica de fluidos, teoría de perturbaciones clásica etc.), que constituyen también un campo de acción plausible para el físico, y para los cuales el lenguaje que se aborda es usualmente el descrito en estos cursos. Basta con observar que históricamente, los formalismos Lagrangiano y Hamiltoniano precedieron en varias décadas al nacimiento de la mecánica cuántica.





# Elementos básicos de Mecánica Newtoniana

## 1.1. Cinemática

La cinemática trata de la descripción del movimiento de los cuerpos sin referencia a las causas de dicho movimiento. El tratamiento será breve sin una discusión detallada de los conceptos. Para detalles, ver por ejemplo las referencias. [2, 3]. Asumiremos que tenemos una idea intuitivamente clara de los conceptos de espacio, tiempo y masa.

El primer concepto que se construye es el de vector posición. Una partícula puntual ocupa un punto específico en el espacio, si elegimos un sistema de referencia, podemos trazar un vector desde el origen de dicho sistema hasta el punto donde se ubica la partícula, y lo denominamos **vector posición**. La posición entendida como un punto geométrico en el espacio, no es un vector como tal (no tiene dirección, magnitud, ni sentido), lo cual se refleja en el hecho de que el vector posición depende del origen elegido para el sistema coordenado. Cuando una partícula se desplaza desde un punto descrito por el vector posición  $\mathbf{r}_0$  hasta otro descrito por  $\mathbf{r}_f$ , podemos describir el movimiento de esta partícula a través del **vector desplazamiento**  $\Delta\mathbf{r}$ , como un vector que va desde  $\mathbf{r}_0$  hacia  $\mathbf{r}_f$ . Este vector indica la dirección del desplazamiento y la distancia recorrida (magnitud del vector).

$$\Delta\mathbf{r} \equiv \mathbf{r}_f - \mathbf{r}_0$$

vale la pena mencionar que  $\Delta\mathbf{r}$  sí es un vector como tal, lo cual se refleja en el hecho de que  $\Delta\mathbf{r}$  es independiente del origen elegido.

Ahora definimos el vector velocidad, como el cambio de posición (o desplazamiento) por unidad de tiempo

$$\mathbf{v} = \frac{\mathbf{r}_f - \mathbf{r}_0}{t_f - t_0} \equiv \frac{\Delta\mathbf{r}}{\Delta t}$$

si queremos conocer el valor de la velocidad del móvil en forma mas detallada, partimos el intervalo anterior en intervalos mas finos, y definimos una velocidad para cada intervalo

$$\mathbf{v} = \frac{\mathbf{r}(t_i + \Delta t_i) - \mathbf{r}(t_i)}{\Delta t_i} \equiv \frac{\Delta\mathbf{r}_i}{\Delta t_i}$$

la velocidad instantánea se define como un paso al límite

$$\lim_{\Delta t_i \rightarrow 0} \frac{\Delta\mathbf{r}_i}{\Delta t_i} = \mathbf{v}_{inst}$$

a partir de la definición de velocidad se obtiene

$$\mathbf{v}_i \Delta t_i = \Delta\mathbf{x}_i$$

si sumamos sobre todos los intervalos

$$\sum_{i=1}^N \mathbf{v}_i \Delta t_i = \sum_{i=1}^N \Delta \mathbf{x}_i$$

tomando el límite cuando  $N \rightarrow \infty$  y haciendo el paso al continuo

$$\int_{t_0}^{t_f} \mathbf{v} dt = \int_{\mathbf{r}_0}^{\mathbf{r}_f} d\mathbf{r} \Rightarrow \mathbf{r}_f - \mathbf{r}_0 = \int_{t_0}^{t_f} \mathbf{v} dt$$

también es útil definir la razón de cambio de la velocidad, a través del vector aceleración en la forma

$$\bar{\mathbf{a}} = \frac{\Delta \mathbf{v}}{\Delta t}; \quad \mathbf{a}_{inst} = \frac{d\mathbf{v}}{dt}$$

donde  $\bar{\mathbf{a}}$  denota la aceleración promedio, en tanto que  $\mathbf{a}_{inst}$  es la aceleración instantánea, usualmente esta última se denota simplemente como  $\mathbf{a}$ . Un argumento similar al anterior nos lleva a la ecuación

$$\int_{t_0}^{t_f} \mathbf{a} dt = \int_{\mathbf{v}_0}^{\mathbf{v}_f} d\mathbf{v} \Rightarrow \mathbf{v}_f - \mathbf{v}_0 = \int_{t_0}^{t_f} \mathbf{a} dt$$

estas ecuaciones y algunas combinaciones especiales de ellas nos proveen el marco para la descripción del movimiento de los cuerpos.

## 1.2. Dinámica: Leyes de Newton

La dinámica de las partículas puntuales está dictaminada por las leyes de Newton, haremos una descripción muy breve sin ninguna pretensión de discusión. Para una discusión detallada, ver por ejemplo la referencia [3].

**Primera Ley:** Existen un conjunto de sistemas de referencia llamados inerciales tales que toda partícula aislada, posee velocidad constante con respecto a dichos sistemas. El reposo es naturalmente un caso particular de velocidad constante nula.

**Segunda Ley:** Nos establece la forma de cuantificar la interacción de una partícula con el resto del universo, se enfatiza que esta ley solo es válida para partículas puntuales que por definición tienen masa constante.

$$\mathbf{F} = m\mathbf{a}$$

esta ley también contiene el principio de superposición de las fuerzas, según el cual la fuerza neta o resultante sobre una partícula es la suma vectorial de cada fuerza aplicada como si cada una de ellas actuara sola. Esto significa que no hay efectos de interferencia entre las distintas fuerzas aplicadas sobre la partícula.

**Tercera Ley:** Cuando una partícula  $A$  hace una fuerza  $\mathbf{F}_{AB}$  sobre una partícula  $B$  entonces la fuerza sobre  $A$  debida a  $B$  (denotada como  $\mathbf{F}_{BA}$ ) está relacionada con  $\mathbf{F}_{AB}$  en la forma

$$\mathbf{F}_{AB} = -\mathbf{F}_{BA}$$

esta ley tiene implícita la propagación instantánea de señales por lo cual su validez es muy limitada. En su forma fuerte, la fuerza es de naturaleza central, sin embargo existen fuerzas que solo cumplen esta ley en su forma débil, es decir los pares de fuerzas son opuestos pero no van a lo largo de la línea que une a las partículas. Finalmente, en otros casos la ley no se cumple en ninguna de sus versiones, lo cual ocurre cuando el tiempo de propagación de la interacción es significativo.

Por otro lado, las leyes anteriores se pueden sustituir por sus equivalentes en términos del concepto de momento lineal definido como el producto de la masa por la velocidad  $\mathbf{p} \equiv m\mathbf{v}$ . La primera ley nos dice que en los sistemas inerciales el momento lineal de una partícula aislada es constante, la segunda ley se escribiría de la forma  $\mathbf{F} = d\mathbf{p}/dt$  y la tercera ley sería sustituida por el principio de conservación del momento para un sistema aislado de partículas. Estas leyes tienen un rango de validez más amplio que la formulación original, aunque hay que usar un concepto extendido de momento lineal.

A pesar de que las leyes de Newton me dan en principio una descripción completa de la evolución de los sistemas, tienen el limitante de que requieren el conocimiento de las fuerzas en función del tiempo, en la práctica es más usual que se conozca la fuerza en función de la posición, lo cual nos lleva al concepto de trabajo.

### 1.3. Trabajo y energía

Si arrastramos un cuerpo con una fuerza aplicada  $\mathbf{F}$  y queremos construir un concepto físico que dependa del desplazamiento del cuerpo, vemos que solo la componente de la fuerza paralela al desplazamiento contribuye a éste. Tal hecho nos induce a construir el concepto de trabajo instantáneo de la forma

$$dW = \mathbf{F} \cdot d\mathbf{r}$$

la fuerza es 100% efectiva cuando es paralela al desplazamiento, 0% efectiva cuando es perpendicular, y su contribución es negativa cuando la proyección de la fuerza sobre el desplazamiento tiene sentido opuesto a tal desplazamiento. En una trayectoria arbitraria el trabajo que realiza una fuerza  $\mathbf{F}$  sobre una partícula viene dada por una integral de línea con límites en los extremos  $A$  y  $B$  de la trayectoria

$$W = \int_A^B \mathbf{F} \cdot d\mathbf{r}$$

nótese que  $\mathbf{F}$  es una de las fuerzas aplicadas sobre la partícula, y no necesariamente corresponde a la fuerza resultante. Sin embargo, cuando la fuerza en cuestión es **la resultante** sobre la partícula, la segunda ley de Newton conduce automáticamente al **teorema fundamental del trabajo y la energía**

$$\int_A^B \mathbf{F} \cdot d\mathbf{r} = \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 \quad (1.1)$$

la cual nos indica que sin importar la trayectoria seguida por la partícula, el trabajo realizado por la fuerza resultante sobre ésta equivale al cambio en la cantidad  $(1/2)mv^2$  que denominamos la energía cinética de la partícula. Es indispensable tener claro que el teorema fundamental del trabajo y la energía solo es aplicable a la fuerza resultante sobre la partícula y no a una de las fuerzas aplicadas. A priori se podría pensar que esta formulación es estéril cuando la queremos aplicar a una fuerza sobre una partícula, dado que el cálculo del trabajo requiere conocer la trayectoria de ésta, lo cual presupone que de alguna forma el problema ya está resuelto. Sin embargo, hay tres razones por las cuales la formulación es útil a pesar de lo anterior

1. Con frecuencia, existen fuerzas de ligadura que obligan a la partícula a seguir una trayectoria dada (e.g. péndulo, montaña rusa), de modo que conocemos la trayectoria aunque no conozcamos el valor de la fuerza de ligadura, ni otras variables dinámicas del sistema (velocidad o aceleración en función de la posición o del tiempo).
2. Existen fuerzas para las cuales la evaluación de la integral de línea no requiere del conocimiento de la trayectoria sino solo de los puntos inicial y final. Esto nos lleva al concepto de **Fuerza conservativa**
3. En el caso de la fuerza resultante el teorema fundamental del trabajo y la energía nos permite encontrar el trabajo que dicha fuerza hace sobre la partícula, conociendo únicamente las velocidades en los extremos de la trayectoria así como la masa de la partícula.

**Definition 1** *Una fuerza conservativa es aquella para la cual el trabajo asociado no depende de la trayectoria seguida por la partícula sino solo de la posición final e inicial*

$$\int_A^B \mathbf{F} \cdot d\mathbf{r} = U(\mathbf{r}_A) - U(\mathbf{r}_B) \quad (1.2)$$

la función escalar  $U(\mathbf{r})$  se conoce como energía potencial. Por otro lado, si la fuerza resultante es conservativa, podemos combinar la definición de conservatividad con el teorema fundamental del trabajo y la energía y se obtiene

$$\begin{aligned}\frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 &= U(\mathbf{r}_A) - U(\mathbf{r}_B) \Rightarrow \\ \frac{1}{2}mv_A^2 + U(\mathbf{r}_A) &= \frac{1}{2}mv_B^2 + U(\mathbf{r}_B)\end{aligned}$$

esto conduce al **teorema de conservación de la energía mecánica**. Es necesario enfatizar que la conservatividad requiere que la energía potencial definida en (1.2) dependa únicamente de la posición. Si la energía potencial es función explícita del tiempo, entonces la suma de  $E_k + U$  todavía me define la energía total del sistema, pero esta cantidad ya no se conserva en general. En otros casos la energía potencial puede depender de la velocidad, aceleración etc. Finalmente, en algunos casos no existe ninguna función escalar que pueda dar cuenta del trabajo realizado. En ninguno de estos casos se conserva la energía.

Retomando la definición (1.2), vemos que a la energía potencial se le puede agregar una constante arbitraria sin alterar el contenido Físico de ésta, ya que lo que es relevante físicamente es el cambio en la energía potencial y no su valor en sí. Es fácil demostrar que para que una fuerza sea conservativa, cada una de estas afirmaciones es condición necesaria y suficiente

$$\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r}) \quad ; \quad \nabla \times \mathbf{F}(\mathbf{r}) = 0 \quad ; \quad \int_A^B \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = U(\mathbf{r}_A) - U(\mathbf{r}_B) \quad ; \quad \oint \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = 0$$

donde todas estas expresiones deben cumplirse para todo  $\mathbf{r} \in \mathbb{R}^3$  o para toda trayectoria en  $\mathbb{R}^3$ . Las dos primeras son condiciones en todo el espacio y las dos siguientes para toda trayectoria (general y cerrada respectivamente). En todas estas ecuaciones, se debe enfatizar que no debe haber dependencia temporal explícita. Las fuerzas conservativas mas importantes son las fuerzas constantes y las fuerzas centrales. Dentro de las no conservativas el rozamiento es la mas destacable.

En el tratamiento de fuerzas centrales existe una cantidad que se conserva y que resulta muy útil en el tratamiento de este tipo de fuerzas: el momento angular

## 1.4. Torque y momento angular

Si para una fuerza central elegimos el origen en el punto de convergencia de la fuerza, tenemos claramente que la cantidad  $\mathbf{r} \times \mathbf{F}$  es nula. Llamemos a esta cantidad el torque de la partícula relativo a el origen  $O$ , ya que es con respecto a este origen que se construye el vector posición  $\mathbf{r}$ . Esta cantidad denotada por  $\vec{\tau}$  se puede escribir como la derivada temporal total de otra cantidad: el llamado momento angular y definido por  $\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$  veamos

$$\begin{aligned}\frac{d}{dt}(\mathbf{r} \times \mathbf{p}) &= \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} = m\mathbf{v} \times \mathbf{v} + \mathbf{r} \times \mathbf{F} \Rightarrow \\ \frac{d\mathbf{L}}{dt} &= \vec{\tau}\end{aligned}\tag{1.3}$$

para una fuerza central con origen en el punto de convergencia, el torque es cero y el momento angular es una constante de movimiento. Aunque los conceptos de torque y momento angular de una partícula surgen de manera natural en el caso de fuerzas centrales, son extensibles a cualquier tipo de fuerza y la relación (1.3) es válida en general. En particular si no hay torque sobre la partícula el momento angular se conserva, de la misma forma que el momento lineal se conserva ante la ausencia o anulación de las fuerzas.

Finalmente, es necesario insistir en la fuerte dependencia que el torque y el momento angular tienen con respecto al origen coordenado elegido, lo cual se manifiesta a través de su dependencia del vector posición  $\mathbf{r}$ .

De esta forma hay tres cantidades cuya conservación será mas adelante extensible a sistemas de partículas, la energía, el momento lineal y el momento angular

## 1.5. Dinámica de un sistema de partículas

### 1.5.1. Definición de centro de masa

Sea un sistema de partículas con masas  $m_1, m_2, \dots, m_n$  y con velocidades  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ . Teniendo en cuenta la definición de momento lineal para una partícula, es natural definir el momento lineal total del sistema de partículas como la suma vectorial de los momentos de sus partículas, es decir:

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 + \dots = m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 + \dots + m_n \mathbf{v}_n = \sum_{i=1}^n m_i \mathbf{v}_i$$

definamos la velocidad del centro de masa  $\mathbf{v}_{CM}$  del sistema de manera que:

$$\mathbf{P} \equiv M \mathbf{v}_{CM} \quad (1.4)$$

donde  $M$  es la masa total. Es decir, de manera que el momento del sistema sea el mismo que le correspondería al caso en que toda la masa estuviera concentrada en el llamado centro de masa, cuya velocidad es  $\mathbf{v}_{CM}$ . Por lo tanto:

$$\mathbf{v}_{CM} = \frac{\mathbf{P}}{M} = \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 + \dots + m_n \mathbf{v}_n}{m_1 + m_2 + \dots + m_n} = \frac{\sum_{i=1}^n m_i \mathbf{v}_i}{M} \quad (1.5)$$

si suponemos que las masas son independientes de la velocidad,  $\mathbf{v}_{CM}$  corresponde a la velocidad asociada a la posición:

$$\mathbf{r}_{CM} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2 + \dots + m_n \mathbf{r}_n}{m_1 + m_2 + \dots + m_n} = \frac{\sum_{i=1}^n m_i \mathbf{r}_i}{M} \quad (1.6)$$

la cual define la posición del centro de masa del sistema relativa a algún observador  $O$ , que mide los vectores de posición y velocidad  $\mathbf{r}_i, \mathbf{v}_i$ . Ahora bien, es uno de los principios más fundamentales de la naturaleza el llamado **principio de conservación del momento** el cual establece que si un sistema de partículas está aislado, su momento total es constante, y como suponemos que la masa no depende de la velocidad, tenemos de la ecuación (1.4) que  $\mathbf{v}_{CM} = cte$  de modo que: **El centro de masa de un sistema de partículas aislado se mueve con velocidad constante con respecto a un sistema inercial.**

Un sistema de referencia muy particular es aquél que no rota con respecto a un sistema inercial y cuyo origen coincide con el propio centro de masa del sistema de partículas. Si colocamos nuestro sistema de referencia sobre el centro de masa, obviamente  $\mathbf{v}_{CM} = 0$  por tanto, el momento total  $\mathbf{P}$  del sistema de partículas es cero. Por brevedad designaremos de ahora en adelante al sistema de referencia del centro de masa como sistema- $C$  de referencia. De modo que podemos escribir:

$$\mathbf{P}_{CM} = \sum_{i=1}^n \mathbf{p}_i = 0 \quad (\text{en el sistema } - C \text{ de referencia}) \quad (1.7)$$

Este sistema de referencia es muy importante dado que muchos fenómenos físicos pueden ser descritos más fácilmente en el sistema de referencia centro de masa que en el sistema del laboratorio.

### 1.5.2. Sistemas de partículas no aislados

Supongamos que tenemos un sistema  $S$  de  $n$  partículas, que interactúa con las  $m$  partículas de otro sistema  $S'$ . Ahora supondremos que los sistemas  $S$  y  $S'$  juntos forman un sistema aislado. Como el sistema  $S + S'$  es aislado, su momento se conserva, luego:

$$\mathbf{P} = \sum_{i=1}^n \mathbf{p}_i + \sum_{j=1}^m \mathbf{p}'_j = \mathbf{P}_S + \mathbf{P}_{S'} = cte \quad (1.8)$$

esto significa que cualquier cambio en el momento de  $S$  debe venir acompañado de un cambio en el momento de  $S'$ , a fin de mantener la suma constante.

$$\Delta \mathbf{P}_S = -\Delta \mathbf{P}_{S'} \quad (1.9)$$

Luego, la interacción entre los sistemas  $S$  y  $S'$  se puede describir como un intercambio de momento. Al tomar la derivada temporal de (1.8) obtenemos:

$$\begin{aligned} \frac{d}{dt} (\mathbf{P}_S + \mathbf{P}_{S'}) &= 0 \\ \frac{d\mathbf{P}_S}{dt} &= -\frac{d\mathbf{P}_{S'}}{dt} \end{aligned} \quad (1.10)$$

haciendo una extrapolación, es natural llamar a la derivada temporal del momento total del sistema, como la fuerza externa ejercida sobre  $S$  (en analogía al caso de una partícula), es decir:

$$\frac{d\mathbf{P}_S}{dt} = \mathbf{F}_{ext} \quad \text{ó} \quad \frac{d(\sum_{i=1}^n \mathbf{P}_i)}{dt} = \mathbf{F}_{ext} \quad (1.11)$$

la denominación de externa se debe al hecho de que es producida por su interacción con  $S'$ . Las fuerzas internas que existen en  $S$  debidas a la interacción entre sus partículas no producen ningún cambio en el momento total en virtud del principio de conservación del momento, pues si quitamos las interacciones externas, se tendría que el momento total  $\mathbf{P}_S$  del sistema permanecería constante y por tanto  $\frac{d\mathbf{P}_S}{dt} = 0$  es decir  $\mathbf{F}_{ext} = 0$ . De allí se concluye que las fuerzas internas de  $S$  no contribuyen a la cantidad  $d\mathbf{P}_S/dt$ . Utilizando las ecuaciones (1.10) y (1.11) tenemos que

$$\mathbf{F}_{ext} = -\mathbf{F}'_{ext} \quad (1.12)$$

donde  $\mathbf{F}'_{ext}$  es la fuerza externa sobre  $S'$  la cual es ejercida por las partículas del sistema  $S$ . Esta es la extensión de la ley de acción y reacción para las interacciones entre  $S$  y  $S'$ .

Por otro lado como  $\mathbf{P}_S = M\mathbf{v}_{CM}$  (donde  $CM$  define el centro de masa del sistema  $S$  y no del sistema compuesto  $S + S'$ ) tenemos:

$$M \frac{d\mathbf{v}_{CM}}{dt} = \mathbf{F}_{ext} \quad (1.13)$$

con lo cual podemos definir de manera natural la aceleración del centro de masa:

$$\mathbf{a}_{CM} = \frac{d\mathbf{v}_{CM}}{dt} \Rightarrow \quad (1.14)$$

$$M\mathbf{a}_{CM} = \mathbf{F}_{ext} \quad (1.15)$$

Nuevamente análogo al caso de una partícula. Reuniendo las ecuaciones (1.4) (1.14) y (1.15) se puede concluir que:

*La dinámica del centro de masa de un sistema de partículas es equivalente al de una partícula de masa igual a la masa total del sistema concentrada en dicho centro de masa, y sobre la cual se aplicara una fuerza equivalente a la suma vectorial de las fuerzas externas sobre el sistema.*

Adicionalmente, la interacción entre los sistemas  $S$  y  $S'$  puede ser descrita formalmente de manera análoga al caso de dos partículas según se ve en las Ecs. (1.8), (1.9), (1.10), (1.11) y (1.12). En este punto queda por tanto, mas que justificada la introducción del concepto de centro de masa.

No obstante, vale la pena aclarar que la formulación anterior no resuelve el problema dinámico completo para el sistema  $S$ . Para resolver formalmente el problema dinámico de cada partícula, vamos a relacionar  $\mathbf{F}_{ext}$  con las fuerzas que actúan sobre cada partícula. De aquí en adelante nos olvidaremos del sistema  $S'$  cuya influencia sobre  $S$  estará representada por la fuerzas externas sobre  $S$ . Sea  $\mathbf{p}_i$  el momento lineal de la partícula  $i$  de masa  $m_i$ ; sea  $\mathbf{F}_i^{(e)}$  la fuerza externa resultante sobre dicha partícula, y  $\mathbf{F}_{ij}$  la fuerza interna que

la partícula  $j$  también del sistema  $S$ , ejerce sobre la partícula  $i$ . La fuerza total ejercida sobre la partícula  $i$  es por tanto igual a la fuerza externa  $\mathbf{F}_i^{(e)}$  más la suma de las fuerzas internas de todas las partículas  $j$  del sistema  $S$  (excepto la propia partícula  $i$ ) aplicando la segunda ley de Newton a esta partícula se tiene:

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i^{(e)} + \sum_{j \neq i}^n \mathbf{F}_{ij} \quad (1.16)$$

asumiremos de aquí en adelante que  $\mathbf{F}_{ii} = 0$ . Resolviendo esta ecuación para cada partícula  $i$  del sistema, se obtiene la solución dinámica completa de éste. Obsérvese que aquí sí aparecen explícitamente las fuerzas internas. Ahora sumemos todas las ecuaciones de todas las partículas de la siguiente manera:

$$\begin{aligned} \sum_{i=1}^n \frac{d\mathbf{p}_i}{dt} &= \sum_{i=1}^n \mathbf{F}_i^{(e)} + \sum_{i=1}^n \sum_{j \neq i}^n \mathbf{F}_{ij} \Rightarrow \\ \frac{d(\sum_{i=1}^n \mathbf{p}_i)}{dt} &= \sum_{i=1}^n \mathbf{F}_i^{(e)} + \sum_{i=1}^n \sum_{j \neq i}^n \mathbf{F}_{ij} \end{aligned} \quad (1.17)$$

pero  $\sum_i \sum_j \mathbf{F}_{ij}$  se puede escribir como:

$$\sum_{i=1}^n \sum_{j \neq i}^n \mathbf{F}_{ij} = \frac{1}{2} \sum_{i=1}^n \sum_{j > i}^n (\mathbf{F}_{ij} + \mathbf{F}_{ji}) \quad (1.18)$$

sustituyendo (1.18) en (1.17) y usando el principio de acción y reacción, la ecuación (1.17) nos queda:

$$\frac{d(\sum_{i=1}^n \mathbf{p}_i)}{dt} = \sum_{i=1}^n \mathbf{F}_i^{(e)}$$

comparando con la ecuación (1.11) se concluye que la fuerza externa sobre un sistema de partículas es la suma de las fuerzas externas sobre cada una de las partículas del sistema.

### 1.5.3. Momento angular y torque de un sistema de partículas

Para un sistema  $S$  de varias partículas definimos el torque y el momento angular de cada partícula  $i$  de  $S$ :

$$\boldsymbol{\tau}_i \equiv \mathbf{r}_i \times \mathbf{F}_i \quad ; \quad \mathbf{L}_i \equiv \mathbf{r}_i \times \mathbf{p}_i$$

y de manera completamente análoga al caso de una partícula se puede probar que

$$\frac{d\mathbf{L}_i}{dt} = \boldsymbol{\tau}_i$$

supongamos además que la partícula está sujeta a las fuerzas internas que sobre ella realizan las otras partículas del sistema  $S$  más una fuerza externa resultante. Luego, la fuerza resultante sobre la partícula  $i$  es  $\mathbf{F}_i^{(e)} + \sum_j \mathbf{F}_{ij}$  y su torque resultante es:

$$\frac{d\mathbf{L}_i}{dt} = \boldsymbol{\tau}_i = \mathbf{r}_i \times \left( \mathbf{F}_i^{(e)} + \sum_{j \neq i}^n \mathbf{F}_{ij} \right)$$

en analogía con la definición de momento lineal total, definimos el torque total  $\boldsymbol{\tau}$  como la suma vectorial de los torques individuales de las partículas, similarmente definimos el momento angular total  $\mathbf{L}$ . Escribimos entonces:

$$\boldsymbol{\tau} = \sum_{i=1}^n \boldsymbol{\tau}_i = \sum_{i=1}^n \frac{d\mathbf{L}_i}{dt} = \frac{d(\sum_{i=1}^n \mathbf{L}_i)}{dt} = \frac{d\mathbf{L}}{dt}$$

por otro lado:

$$\begin{aligned} \boldsymbol{\tau} &= \sum_{i=1}^n \left[ \mathbf{r}_i \times \left( \mathbf{F}_i^{(e)} + \sum_{j \neq i}^n \mathbf{F}_{ij} \right) \right] \\ &= \sum_{i=1}^n \left( \mathbf{r}_i \times \mathbf{F}_i^{(e)} \right) + \sum_{i=1}^n \left( \mathbf{r}_i \times \sum_{j \neq i}^n \mathbf{F}_{ij} \right) \end{aligned}$$

por otra parte, utilizando la ley de acción y reacción, puede demostrarse por inducción que:

$$\sum_{i=1}^n \left( \mathbf{r}_i \times \sum_{j \neq i}^n \mathbf{F}_{ij} \right) = \sum_{i=1}^{n-1} \sum_{j>i}^n [(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij}] \quad (1.19)$$

por tanto, el torque total queda:

$$\boldsymbol{\tau} = \sum_{i=1}^n \left( \mathbf{r}_i \times \mathbf{F}_i^{(e)} \right) + \sum_{i=1}^{n-1} \sum_{j>i}^n [(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij}]$$

Si suponemos especialmente que las fuerzas internas  $\mathbf{F}_{ij}$  actúan a lo largo de los vectores relativos  $\mathbf{r}_{ij}$  se tiene que el segundo término de la derecha se anula quedando

$$\boldsymbol{\tau} = \sum_{i=1}^n \left( \mathbf{r}_i \times \mathbf{F}_i^{(e)} \right)$$

donde  $\mathbf{r}_i \times \mathbf{F}_i^{(e)}$  representa el torque debido a las fuerzas externas  $\mathbf{F}_i^{(e)}$  sobre cada partícula  $i$  entonces:

$$\boldsymbol{\tau} = \sum_{i=1}^n \boldsymbol{\tau}_{i,ext} = \boldsymbol{\tau}_{ext}$$

de modo que

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau}_{ext} \quad (1.20)$$

que constituye la relación fundamental de la mecánica rotacional, obsérvese la analogía entre las Ecs. (1.20) y (1.11). Si  $\boldsymbol{\tau}_{ext} = 0$ , se encuentra que:

$$\frac{d\mathbf{L}}{dt} = 0 \Rightarrow \mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2 + \dots + \mathbf{L}_n = cte \quad (1.21)$$

la ecuación (1.21) constituye la ley de conservación del momento angular. La cual nos indica que si en un sistema las fuerzas externas sobre él son tales que su torque total es cero, su momento angular permanece constante. En particular, el momento angular se conservará cuando el sistema esté aislado.

Vale anotar que la ley de conservación del momento angular ha mostrado ser universalmente válida a pesar de nuestra suposición inicial de que las fuerzas vayan a lo largo de las líneas que unen a las partículas (propiedad que no cumplen por ejemplo las fuerzas magnéticas). De manera que aún en los casos en que nuestra suposición inicial no es válida, la ley de conservación del momento angular se ha cumplido en todos los procesos observados hasta ahora, aunque con un concepto extendido de momento angular.



### Relación entre el momento angular de un sistema de partículas relativo al sistema-C y el relativo al laboratorio

Las posiciones de las partículas relativas al laboratorio (denotadas por  $\mathbf{r}_i$ ) y las asociadas al centro de masa (denotadas por  $\mathbf{r}'_i$ ) se pueden relacionar fácilmente, y se puede ver que:

$$\mathbf{r}_i = \mathbf{r}'_i + \mathbf{r}_{CM} \quad (1.22)$$

$$\mathbf{v}_i = \mathbf{v}'_i + \mathbf{v}_{CM} \quad (1.23)$$

siendo  $\mathbf{r}_{CM}$  y  $\mathbf{v}_{CM}$  la posición y velocidad del centro de masa medidas por el laboratorio. Multiplicando (1.23) por  $m_i$  se obtiene

$$\mathbf{p}_i = \mathbf{p}'_i + m_i \mathbf{v}_{CM} \quad (1.24)$$

utilizando estas relaciones, obtenemos:

$$\begin{aligned} \mathbf{L} &= \sum_{i=1}^n (\mathbf{r}_i \times \mathbf{p}_i) = \sum_{i=1}^n [(\mathbf{r}'_i + \mathbf{r}_{CM}) \times (\mathbf{p}'_i + m_i \mathbf{v}_{CM})] \\ \mathbf{L} &= \sum_{i=1}^n (\mathbf{r}'_i \times \mathbf{p}'_i) + \mathbf{r}_{CM} \times \sum_{i=1}^n \mathbf{p}'_i + \left( \sum_{i=1}^n m_i (\mathbf{r}'_i + \mathbf{r}_{CM}) \right) \times \mathbf{v}_{CM} \\ \mathbf{L} &= \sum_{i=1}^n (\mathbf{r}'_i \times \mathbf{p}'_i) + \mathbf{r}_{CM} \times \sum_{i=1}^n \mathbf{p}'_i + \left( \sum_{i=1}^n m_i \mathbf{r}_i \right) \times \mathbf{v}_{CM} \end{aligned}$$

pero  $\sum_i \mathbf{p}'_i = 0$  de acuerdo con (1.7) y utilizando (1.6) se tiene que  $(\sum_i m_i \mathbf{r}_i) = M \mathbf{r}_{CM}$  de modo que:

$$\begin{aligned} \mathbf{L} &= \sum_{i=1}^n (\mathbf{r}'_i \times \mathbf{p}'_i) + M \mathbf{r}_{CM} \times \mathbf{v}_{CM} \\ \mathbf{L} &= \mathbf{L}_{CM} + \mathbf{r}_{CM} \times \mathbf{P} \end{aligned} \quad (1.25)$$

el término  $\mathbf{L}_{CM}$  es el momento angular relativo al centro de masa, o momento angular interno; esto último debido a que el momento angular así medido es una propiedad del sistema e independiente del observador. El segundo término a la derecha suele denominarse momento angular externo relativo al sistema-L (laboratorio), y equivale al momento angular (medido desde el sistema laboratorio) correspondiente a una partícula de masa  $M$  colocada en la posición del centro de masa y con la velocidad del CM. Por ejemplo, cuando un lanzador arroja una pelota rotando, el momento angular debido a la rotación está dado por  $\mathbf{L}_{CM}$ , mientras que el momento angular debido a la traslación de la pelota está dado por  $M \mathbf{r}_{CM} \times \mathbf{v}_{CM}$  con  $M$  la masa de la bola. Para la Tierra  $\mathbf{L}_{CM}$  es debido a su rotación y  $M \mathbf{r}_{CM} \times \mathbf{v}_{CM}$  es debido a su traslación alrededor del Sol.

### Relación entre el torque externo alrededor del centro de masa y el torque alrededor del laboratorio

Con un argumento similar al anterior se puede calcular la relación entre el torque externo medido por el laboratorio y el medido por el centro de masa

$$\boldsymbol{\tau}_{ext} = \boldsymbol{\tau}_{CM} + \mathbf{r}_{CM} \times \mathbf{F}_{ext} \quad (1.26)$$

$\mathbf{r}_{CM} \times \mathbf{F}_{ext}$  es el torque de traslación y es equivalente al torque medido en el sistema-L para una partícula ubicada en el centro de masa sometida a la fuerza externa resultante de todo el sistema  $\mathbf{F}_{ext}$ .

Pero teniendo en cuenta el resultado (1.25) y derivando temporalmente se obtiene:

$$\frac{d\mathbf{L}}{dt} = \frac{d\mathbf{L}_{CM}}{dt} + \mathbf{r}_{CM} \times M \frac{d\mathbf{v}_{CM}}{dt}$$

y tomando las Ecs. (1.13) y (1.20):

$$\boldsymbol{\tau}_{ext} = \frac{d\mathbf{L}_{CM}}{dt} + \mathbf{r}_{CM} \times \mathbf{F}_{ext} \quad (1.27)$$

Y comparando (1.26) y (1.27) se obtiene:

$$\frac{d\mathbf{L}_{CM}}{dt} = \boldsymbol{\tau}_{CM} \quad (1.28)$$

esta relación es funcionalmente idéntica a (1.20) pero con la diferencia de que (1.20) es válida sólo cuando el torque y el momento angular se miden a partir de un punto fijo (usualmente el origen) en un sistema inercial de referencia, en tanto que la relación (1.28) es válida incluso si el sistema de referencia-C no es inercial<sup>1</sup>. Esta ecuación resultará muy útil para estudiar el movimiento del cuerpo rígido.

#### 1.5.4. Trabajo y energía de un sistema de partículas

La fuerza resultante sobre la partícula  $i$  es:

$$\mathbf{F}_i = \mathbf{F}_i^{(e)} + \sum_{j \neq i}^n \mathbf{F}_{ij} = m_i \mathbf{a}_i$$

donde  $\mathbf{F}_i^{(e)}$  es la resultante de las fuerzas externas. El diferencial de trabajo es:

$$dW_i = m_i \mathbf{a}_i \cdot d\mathbf{r}_i = \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i + \sum_{j \neq i}^n \mathbf{F}_{ij} \cdot d\mathbf{r}_i$$

el diferencial de trabajo total realizado sobre el sistema, es entonces la suma de los diferenciales de trabajo de partícula individual:

$$dW = \sum_{i=1}^n m_i \mathbf{a}_i \cdot d\mathbf{r}_i = \sum_{i=1}^n \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i + \sum_{i=1}^n \sum_{j \neq i}^n \mathbf{F}_{ij} \cdot d\mathbf{r}_i$$

aplicando una identidad análoga a (1.19)

$$\sum_{i=1}^n \sum_{j \neq i}^n \mathbf{F}_{ij} \cdot d\mathbf{r}_i = \sum_{i=1}^{n-1} \sum_{j>i}^n [(\mathbf{dr}_i - \mathbf{dr}_j) \cdot \mathbf{F}_{ij}] = \sum_{i=1}^{n-1} \sum_{j>i}^n (\mathbf{dr}_{ij} \cdot \mathbf{F}_{ij})$$

por otro lado  $\mathbf{a}_i \cdot d\mathbf{r}_i = \frac{d\mathbf{v}_i}{dt} \cdot d\mathbf{r}_i = d\mathbf{v}_i \cdot \frac{d\mathbf{r}_i}{dt} = d\mathbf{v}_i \cdot \mathbf{v}_i$  con lo cual queda:

$$dW = \sum_{i=1}^n m_i d\mathbf{v}_i \cdot \mathbf{v}_i = \sum_{i=1}^n \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i + \sum_{i=1}^{n-1} \sum_{j>i}^n (\mathbf{dr}_{ij} \cdot \mathbf{F}_{ij})$$

integrando a ambos lados la ecuación queda:

$$\int dW = \sum_{i=1}^n \int_{v_{0i}}^{v_i} m_i \mathbf{v}_i \cdot d\mathbf{v}_i = \sum_{i=1}^n \int_{A_i}^{B_i} \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i + \sum_{i=1}^{n-1} \sum_{j>i}^n \int_{A_{ij}}^{B_{ij}} (\mathbf{dr}_{ij} \cdot \mathbf{F}_{ij})$$

<sup>1</sup>Sin embargo, la relación (1.28) no es válida si el sistema-C está rotando con respecto al sistema inercial [4].

obsérvese que estrictamente aquí se realizan  $n$  integrales dado que existen  $n$  trayectorias seguidas por cada una de las  $n$  partículas. El primer término de la derecha es el trabajo realizado por las fuerzas externas a  $S$ . El segundo término corresponde al trabajo hecho por las fuerzas internas.

$$\int dW = \sum_{i=1}^n \left( \frac{1}{2} m_i \mathbf{v}_i^2 - \frac{1}{2} m_i \mathbf{v}_{0i}^2 \right) = W_{ext} + W_{int}$$

$$W = \sum_{i=1}^n \frac{1}{2} m_i v_i^2 - \sum_{i=1}^n \frac{1}{2} m_i v_{0i}^2 = W_{ext} + W_{int}$$

pero  $\sum_{i=1}^n \frac{1}{2} m_i v_i^2$  es la energía cinética del sistema en un instante dado. Por tanto el trabajo total realizado por las fuerzas externas e internas cuando el sistema de partículas se desplaza desde la configuración de posiciones  $(\mathbf{x}_{01}, \dots, \mathbf{x}_{0n})$  hasta la configuración de posiciones  $(\mathbf{x}_{f1}, \dots, \mathbf{x}_{fn})$  a través de las trayectorias  $(\mathbf{x}_1(t), \dots, \mathbf{x}_n(t))$  es igual al cambio de energía cinética entre estas dos configuraciones de posición, sin importar la naturaleza de las trayectorias. Hemos recuperado por tanto el teorema fundamental del trabajo y la energía para el caso de un sistema de partículas, el cual enunciaremos de la manera siguiente:

**El trabajo total efectuado por las fuerzas externas e internas sobre un sistema de partículas es igual al cambio producido en su energía cinética.**

Por tanto escribiremos:

$$E_k - E_{k,0} = W_{ext} + W_{int} \quad (1.29)$$

### 1.5.5. Conservación de la energía de un sistema de partículas

Supongamos ahora que las fuerzas internas son conservativas, en tal caso se tiene que:

$$\int_{A_{ij}}^{B_{ij}} (d\mathbf{r}_{ij} \cdot \mathbf{F}_{ij}) = E_{p0,int}(\mathbf{r}_{ij}) - E_{p,int}(\mathbf{r}_{ij})$$

pero como se trata de interacción de pares de partículas, la energía potencial depende únicamente del vector relativo entre los pares de partículas  $i, j$ . Si suponemos especialmente que la fuerza  $\mathbf{F}_{ij}$  va a lo largo del vector  $\mathbf{r}_{ij}$  que une a las dos partículas, se tiene que la energía potencial dependerá sólo de la distancia entre las dos partículas:

$$\int_{A_{ij}}^{B_{ij}} (d\mathbf{r}_{ij} \cdot \mathbf{F}_{ij}) = E_{p0,int}^{ij}(r_{ij}) - E_{p,int}^{ij}(r_{ij}) \quad , \quad r_{ij} \equiv \|\mathbf{r}_{ij}\|$$

en este caso, la energía potencial está definida por pares de partículas, a esto se refiere el supraíndice  $i, j$  en la ecuación anterior. Por tanto, las fuerzas internas producen un trabajo equivalente a:

$$\sum_{i=1}^{n-1} \sum_{j>i}^n \left[ E_{p0,int}^{ij}(r_{ij}) - E_{p,int}^{ij}(r_{ij}) \right]$$

y como sabemos que la suma de fuerzas conservativas es también conservativa se tiene que:

$$W_{int} = \sum_{i=1}^{n-1} \sum_{j>i}^n E_{p0,int}^{ij}(r_{ij}) - \sum_{i=1}^{n-1} \sum_{j>i}^n E_{p,int}^{ij}(r_{ij}) = E_{p0,int} - E_{p,int}$$

que al sustituirlo en (1.29) nos da:

$$E_k - E_{k,0} = W_{ext} + E_{p0,int} - E_{p,int} \Rightarrow$$

$$[E_k + E_{p,int}] - [E_{p0,int} + E_{k,0}] = W_{ext}$$

definamos ahora la energía propia del sistema como:

$$U \equiv E_k + E_{p,int} \quad (1.30)$$

el nombre de energía propia indica que esta es una propiedad del sistema (y no de agentes externos). Podemos reescribir la ecuación anterior como:

$$U - U_0 = W_{ext} \quad (1.31)$$

es decir, el cambio en la energía propia del sistema es igual al trabajo realizado por las fuerzas externas sobre el sistema de partículas.

Como la energía potencial  $E_{p,int}(r_{ij})$  solo depende de la distancia entre las dos partículas  $i, j$  se tiene entonces que  $E_{p,int}^{ij} = E_{p,int}^{ji}$  por lo tanto:

$$\sum_{i=1}^{n-1} \sum_{j>i}^n E_{p,int}^{ij}(r_{ij}) = \frac{1}{2} \sum_{i=1}^n \sum_{j \neq i}^n E_{p,int}^{ij}(r_{ij})$$

regresando a la Ec. (1.31) se tiene que si  $W_{ext} = 0$ , la energía propia es igual al principio y al final de cualquier proceso, de modo que podemos enunciar el siguiente principio de conservación:

**Si las fuerzas externas no realizan ningún trabajo, se tiene que la energía propia del sistema se conserva (siendo las fuerzas internas conservativas).**

Finalmente, supongamos que también las fuerzas externas son conservativas de manera que:

$$W_{ext} = E_{p,ext0} - E_{p,ext}$$

la Ec. (1.31) queda:

$$\begin{aligned} U - U_0 &= E_{p,ext0} - E_{p,ext} \Rightarrow \\ (U + E_{p,ext}) - (U_0 + E_{p,ext0}) &= 0 \end{aligned}$$

es decir, la cantidad  $U + E_{p,ext}$  se conserva para cualquier proceso realizado por las fuerzas externas e internas. Como en el cálculo de  $U + E_{p,ext}$  intervienen todas las fuerzas sobre  $S$ , podemos denominar a ésta cantidad como la energía total del sistema

$$E = U + E_{p,ext} = E_k + E_{p,int} + E_{p,ext}$$

es decir la energía total del sistema es igual a la energía cinética mas la energía potencial de las fuerzas internas y externas, y se tiene que:

$$E - E_0 = 0$$

es decir la energía total se conserva para cualquier proceso realizado por las fuerzas internas y externas siempre y cuando ambos tipos de fuerzas sean conservativos. Hemos deducido entonces el teorema de conservación de la energía para un sistema de partículas en donde las fuerzas involucradas son todas conservativas<sup>2</sup>. En el caso de una sola partícula, la energía total se reduce a la expresión correcta ya que cuando solo hay una partícula  $E_{p,int} = 0$  (puesto que la energía potencial interna se debe a la interacción por pares, de modo que debe haber al menos dos partículas) y  $E_{p,ext}$  es la energía potencial correspondiente a la fuerza resultante sobre la partícula.

---

<sup>2</sup>En su forma más general, la conservación de la energía adquiere el carácter de principio y su validez es universal, mas allá de las suposiciones realizadas para demostrar el teorema.

### 1.5.6. Transformación de la energía cinética al sistema-C a partir del sistema de laboratorio

La energía cinética depende del observador. No obstante, dicha cantidad medida a partir del centro de masa del sistema sería una cantidad independiente del observador. Puesto que la energía potencial interna no depende del sistema de referencia, ya que sólo es función de las coordenadas relativas entre las partículas, podemos definir una energía del sistema que sea independiente del observador a través de la relación:

$$U_{int} \equiv E_{k,CM} + E_{p,int}$$

que denominamos energía interna del sistema de partículas. Vemos pues que la energía interna es la energía propia medida por el sistema de referencia del centro de masa. En general, por la energía del sistema  $S$  se entiende que se refiere a su energía interna.

Pero como en la realidad las cantidades se miden desde el laboratorio, es necesario conocer la forma en que la energía cinética se transforma cuando pasamos del sistema del laboratorio al sistema centro de masa. El procedimiento es similar al del caso del momento angular y se deja como ejercicio al lector demostrar que:

$$\begin{aligned} E_k &= E_{k,CM} + \frac{1}{2}Mv_{CM}^2 \\ \sum_{i=1}^n \frac{1}{2}m_i v_i^2 &= \sum_{i=1}^n \frac{1}{2}m_i v_i'^2 + \frac{1}{2}Mv_{CM}^2 \end{aligned} \quad (1.32)$$

el primer término de la derecha es el debido al movimiento de las partículas del sistema con respecto al centro de masa, el segundo término es el debido al movimiento del sistema como un todo (movimiento de su centro de masa con respecto al laboratorio). La energía propia de  $S$  vendría dada entonces por:

$$U = U_{int} + \frac{1}{2}Mv_{CM}^2$$

esta relación muestra que la energía interna es el menor valor que puede tomar la energía propia, y este valor lo toma cuando se mide en el sistema  $-C$  de referencia.

## 1.6. Ejercicios

1. Demuestre por argumentos puramente cinemáticos, que la velocidad inicial  $\mathbf{v}_0$  y la velocidad final  $\mathbf{v}_f$  de una partícula, están relacionadas con su aceleración  $\mathbf{a}$ , por medio de la ecuación

$$\mathbf{v}_f^2 - \mathbf{v}_0^2 = \int_{\mathbf{r}_0}^{\mathbf{r}_f} \mathbf{a} \cdot d\mathbf{r}$$

donde  $\mathbf{r}_0$  y  $\mathbf{r}_f$  son la posición inicial y final de la partícula respectivamente. Nótese que al multiplicar esta ecuación por la masa, se obtiene el teorema fundamental del trabajo y la energía, Ec. (1.1).

2. La velocidad de escape de una partícula en la tierra, es la velocidad inicial mínima que requiere una partícula que está en la superficie de la tierra, para poder escapar del campo gravitatorio terrestre. Ignorando la resistencia del aire, calcule esta velocidad de escape por argumentos de energía.
3. Demuestre la Ec. (1.19) por inducción matemática.
4. Demuestre que la relación entre el torque externo medido por el laboratorio y el medido por el centro de masa viene dado por la Ec. (1.26). (Para comentarios adicionales sobre algunas sutilezas de esta relación, ver la Ref. [4]).
5. Demuestre que la energía cinética de un sistema de partículas vista por el laboratorio es la energía cinética vista por el centro de masa, mas la energía cinética que tendría una partícula con la masa total del sistema y la velocidad del centro de masa (ver Ec. 1.32).

# Principio de D'Alembert y ecuaciones de Lagrange

Las ecuaciones de Lagrange son un formalismo equivalente a la formulación Newtoniana de la mecánica pero que tiene una serie de ventajas tanto operativas como formales que veremos en el transcurso de nuestros desarrollos. Es usual en Física realizar una formulación diferencial y una formulación integral equivalente de cierto formalismo. Por ejemplo, las ecuaciones de Maxwell se pueden escribir en forma integral en la cual aparecen integrales de línea, de superficie y de volumen, o alternativamente se pueden escribir en forma diferencial por medio de divergencias, rotacionales y derivadas parciales del tiempo. Similarmente, la segunda ley de Newton tiene una versión diferencial

$$\mathbf{F} = \frac{d\mathbf{P}}{dt}$$

y una forma integral equivalente que se conoce como teorema del impulso **I**:

$$\mathbf{F} dt = d\mathbf{P} \Rightarrow \int_0^T \mathbf{F} dt = \mathbf{P} - \mathbf{P}_0 \equiv \mathbf{I}$$

la experiencia muestra que las formulaciones integrales y diferenciales tienen cada una sus ventajas y desventajas y en general resultan un buen complemento para obtener un buen panorama de la teoría subyacente. En esta misma tónica queremos obtener el nuevo formalismo a partir de una formulación diferencial y una formulación integral. La formulación diferencial se cimienta en el llamado **principio de D'Alembert** en tanto que la formulación integral se basa en el **principio variacional de Hamilton**.

Nos ocuparemos primero de la versión diferencial, pero antes de discutir el principio de D'Alembert que la genera, debemos estudiar el papel de las ligaduras en la Física.

## 2.1. Ligaduras

Aunque las ecuaciones de movimiento (1.16) nos dictaminan formalmente toda la evolución del sistema, desde el punto de vista operativo es en general difícil conocer todas las fuerzas aplicadas sobre cada partícula. Con frecuencia, los sistemas están sometidos a ligaduras que obligan a la partícula o sistema de partículas a moverse en ciertas trayectorias o a restringir su movimiento a ciertas regiones específicas. Tal es el caso de sistemas como la montaña rusa o el péndulo en donde la trayectoria del móvil está determinada por la normal y la tensión de la cuerda respectivamente, obsérvese que en estos casos particulares no es fácil encontrar a priori el valor de las fuerzas de ligadura. Otro caso común es el de un gas en un contenedor, en tal caso la ligadura se manifiesta como la exclusión de la región exterior al contenedor como posible región de movimiento para las partículas.

Hay varias formas de clasificar las ligaduras. Una de las clasificaciones más útiles consiste en la caracterización de las ligaduras como holónomas y no holónomas. Las ligaduras holónomas son aquellas que se pueden escribir como ecuaciones que conectan las coordenadas de las partículas y tal vez el tiempo, de la forma

$$f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = 0 \tag{2.1}$$

un ejemplo sencillo lo constituye el cuerpo rígido para el cual las ligaduras se expresan de la forma

$$(\mathbf{r}_i - \mathbf{r}_j)^2 - c_{ij}^2 = 0$$

siendo  $\mathbf{r}_i$ ,  $\mathbf{r}_j$  posiciones de un par de partículas del cuerpo rígido y siendo  $c_{ij}$  sus distancias (constantes). La montaña rusa y el péndulo son también ejemplos de ligaduras holónomas, ya que el móvil está obligado a seguir una trayectoria específica. La ecuación de la trayectoria actúa entonces como la ligadura.

Toda ligadura que no cumpla la condición (2.1), se denomina no holónoma. En el ejemplo del contenedor de gas, asumiendo por simplicidad un contenedor esférico de radio  $a$ , tendríamos una ligadura de la forma

$$r^2 - a^2 \leq 0$$

es decir una desigualdad. Naturalmente hay infinidad de ligaduras no holónomas ya que solo se requiere que no se cumpla una ecuación de la forma (2.1).

Otra clasificación importante es en ligaduras reónomas (que contienen el tiempo como variable explícita) y esclerónomas (el tiempo no aparece explícitamente en la ligadura). Un bloque que se desliza por un plano inclinado, donde este último está fijo en el espacio, es una ligadura esclerónoma. Por otro lado, si el plano inclinado se mueve de una manera prescrita<sup>1</sup>, la ligadura es reónoma. Nótese sin embargo, que si el movimiento del plano inclinado es solo debido a la fuerza de reacción del bloque, la dependencia temporal entra a través de las coordenadas que describen a la curva que hace el bloque; en tal caso, la ligadura como tal es esclerónoma.

Como ya se mencionó, en general es difícil hallar las fuerzas de ligadura. Adicionalmente, si tenemos  $N$  partículas, las  $3N$  coordenadas necesarias para determinar las posiciones de las  $N$  partículas no son en general independientes, de modo que los grados de libertad son menores a  $3N$ . En el caso más general puede ser muy difícil saber cuántas coordenadas independientes hay, pero en el caso de las ligaduras holónomas, este conteo es muy sencillo ya que si tenemos  $k$  ecuaciones holónomas, el número de grados de libertad será  $3N - k$ . Por tanto, cuando las ligaduras son holónomas, se pueden encontrar un conjunto de **coordenadas generalizadas** que consisten en un conjunto de  $3N - k$  grados de libertad independientes que denotaremos por  $q_1, \dots, q_{3N-k}$  las cuales se pueden escribir en términos de las antiguas coordenadas  $\mathbf{r}_1, \dots, \mathbf{r}_N$  con ecuaciones de la forma

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \dots, q_{3N-k}, t) \quad ; \quad i = 1, \dots, N \quad (2.2)$$

que contienen implícitamente las ligaduras. Si a este sistema le añadimos las ecuaciones de ligadura

$$f_j(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0 \quad ; \quad j = 1, \dots, k \quad (2.3)$$

el conjunto de transformaciones debe ser invertible de modo que cada  $q_i$  se puede escribir en términos de las antiguas coordenadas  $\mathbf{r}_i$  y el tiempo. Es importante notar que en el caso más general, las coordenadas generalizadas no se pueden agrupar en triplas que formen un vector euclidiano.

Un ejemplo que ilustra todos estos detalles es el péndulo doble que se muestra en la figura 2.1. Por simplicidad asumamos que las lentejas se mueven en un plano. Las coordenadas cartesianas de las dos lentejas nos dan seis escalares que especifican las posiciones de ambas. Sin embargo, es claro que no todas estas componentes son independientes. Los ángulos  $\theta_1$  y  $\theta_2$  representan un conjunto de grados de libertad (coordenadas independientes) que determinan completamente la posición de las dos lentejas bajo las ligaduras ya mencionadas. Existen en efecto 4 ecuaciones de ligadura, dos de ellas asociadas al hecho de que las lentejas están en un plano (e.g.  $z_1 = 0$  y  $z_2 = 0$ ) y otras dos que nos dicen que las longitudes de las cuerdas son constantes<sup>2</sup>. Para este sistema, el par  $(\theta_1, \theta_2)$  no forma un vector euclidiano en el plano de movimiento, también se vé otro aspecto interesante: las coordenadas generalizadas no necesariamente tienen dimensiones de longitud (en este caso son adimensionales), en general las coordenadas generalizadas pueden tener cualquier tipo de dimensión.

Dado que cuando las ligaduras son holónomas las coordenadas dependientes se pueden eliminar, estos problemas son accesibles para solucionarse al menos formalmente. Las ligaduras no holónomas deben tratarse

<sup>1</sup>Usualmente debido a alguna fuerza externa que introduce la dependencia temporal en la ecuación de ligadura.

<sup>2</sup>En este caso, dos ligaduras están asociadas a condiciones iniciales (la ausencia de una componente  $z$  de las velocidades iniciales de las lentejas), y otras dos están asociadas a fuerzas de ligadura (las tensiones de las cuerdas).

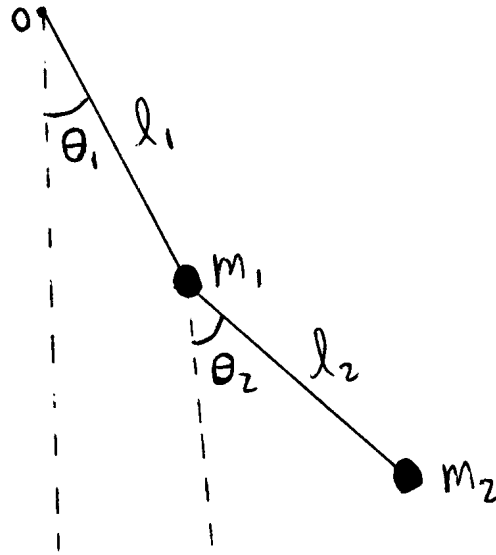


Figura 2.1: Ilustración de un péndulo doble. Un conjunto conveniente de coordenadas son los ángulos  $\theta_1$  y  $\theta_2$  medidos con respecto a la vertical, siempre que el movimiento se restrinja a un plano.

cada una por aparte y no hay una estrategia general de solución. Es importante enfatizar que para que las ligaduras sean holónomas es necesario que las funciones  $f$  no contengan como argumento, derivadas u operadores diferenciales de la posición<sup>3</sup>.

De momento nos limitaremos a ligaduras holónomas para las cuales es más sencillo el conteo de grados de libertad. Recordemos que una segunda dificultad cuando nos encontramos con ligaduras es el hecho de que las fuerzas de ligadura son desconocidas y en general difíciles de hallar. Para obviar este problema sería deseable obtener un formalismo en el cual las fuerzas de ligadura no estén incluidas. Esta es la principal motivación para introducir el llamado **principio de D'Alembert**.

## 2.2. Principio de D'Alembert

Introduciremos primero el concepto de desplazamiento virtual, concebido como el cambio de configuración de un sistema como resultado de un cambio infinitesimal de las coordenadas, denotado como  $\delta \mathbf{r}_i$ , y consistente con las fuerzas y ligaduras impuestas sobre el sistema en un instante dado  $t$ . A diferencia del desplazamiento real, el cual ocurre en un intervalo de tiempo  $dt$ , el tiempo se considera fijo en los desplazamientos virtuales, de modo que se ignoran los posibles cambios en las fuerzas y en las ligaduras que pueden provenir de la evolución temporal. Comencemos por escribir la segunda ley de Newton, y el trabajo virtual asociado a las fuerzas sobre una partícula  $i$

$$\mathbf{F}_i - \dot{\mathbf{p}}_i = 0 \Rightarrow \sum_{i=1}^N (\mathbf{F}_i - \dot{\mathbf{p}}_i) \cdot \delta \mathbf{r}_i = 0$$

<sup>3</sup>Cuando hay operadores diferenciales es en algunos casos posible tener ligaduras integrables que permitan reducir las variables dependientes con facilidad. Incluso cuando las ligaduras no son integrables, el método de multiplicadores de Lagrange puede hacer esta función, pero volveremos sobre eso más adelante.



como la idea es eliminar a las fuerzas de ligadura de la formulación, haremos la separación de  $\mathbf{F}_i$  entre las fuerzas aplicadas  $\mathbf{F}_i^{(a)}$  y las fuerzas de ligadura  $\mathbf{f}_i$

$$\begin{aligned}\mathbf{F}_i &= \mathbf{F}_i^{(a)} + \mathbf{f}_i \\ \Rightarrow \sum_{i=1}^N \left( \mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i \right) \cdot \delta \mathbf{r}_i + \sum_{i=1}^N \mathbf{f}_i \cdot \delta \mathbf{r}_i &= 0\end{aligned}$$

a continuación nos restringiremos a sistemas en los cuales los trabajos virtuales netos asociados a las fuerzas de ligadura se anulan. Es fácil ver que esta condición se cumple en gran número de casos, por ejemplo la tensión de la cuerda no realiza trabajo sobre la lenteja de un péndulo (ni virtual ni real). En una montaña rusa sin rozamiento, la normal no hace trabajo sobre el móvil. En un cuerpo rígido las fuerzas de ligadura son fuerzas internas que obligan a las partículas a conservar sus distancias relativas, lo cual conduce a que no hayan cambios en la energía interna, es decir que no hay trabajo realizado por estas fuerzas internas de ligadura. Cuando intervienen fuerzas de fricción por deslizamiento no se cumple esta condición de tal manera que debemos excluir esta situación de nuestra formulación actual<sup>4</sup>. Sin embargo, el rozamiento estático de rodadura no viola esta condición ya que no realiza trabajo real ni virtual. Vale la pena anotar que si una partícula está restringida a una superficie o curva que a su vez se desplaza en el tiempo, la fuerza de ligadura es instantáneamente perpendicular a la superficie o curva de modo que el trabajo virtual es cero, aunque el trabajo real en un intervalo  $dt$  no es necesariamente cero (ver Refs. [5, 6] y problema 2.2). Esta última observación justifica la introducción de los desplazamientos virtuales, ya que eliminan muchas fuerzas de ligadura que no se eliminan con los desplazamientos reales.

Asumiendo entonces que las fuerzas de ligadura no producen trabajos virtuales, se tiene que

$$\sum_{i=1}^N \left( \mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i \right) \cdot \delta \mathbf{r}_i = 0 \quad (2.4)$$

a la expresión (2.4), se le conoce como **principio de D'Alembert**. Hemos logrado nuestro objetivo en el sentido de excluir las fuerzas de ligadura de la formulación. De aquí en adelante omitiremos el supraíndice <sup>(a)</sup> sobreentendiendo que las fuerzas involucradas excluyen a las ligaduras. No obstante, los coeficientes de los desplazamientos virtuales  $\delta \mathbf{r}_i$  no son necesariamente cero, ya que los desplazamientos virtuales  $\delta \mathbf{r}_i$  al ser compatibles con las ligaduras no son en general independientes, están conectados por las ecuaciones de ligadura. El siguiente paso es entonces encontrar un conjunto de **coordenadas generalizadas**  $q_j$  que sean independientes y solo tengan en cuenta los verdaderos grados de libertad del sistema. En tal caso los desplazamientos virtuales  $\delta q_j$  serán independientes y podremos aseverar que los coeficientes asociados a estos desplazamientos deben ser nulos. La existencia de estas coordenadas independientes solo se puede garantizar cuando las ligaduras son holónomas.

## 2.3. Coordenadas generalizadas y ecuaciones de Lagrange

Supondremos entonces que las ligaduras son holónomas. En consecuencia, es posible encontrar un conjunto  $q_1, q_2, \dots, q_n$  de coordenadas generalizadas independientes que junto con el tiempo, me caracterizan completamente la configuración y dinámica del sistema físico, y que están conectadas con las coordenadas originales en la forma

$$\mathbf{r}_i = \mathbf{r}_i(q_1, q_2, \dots, q_n, t) \quad ; \quad i = 1, \dots, N \quad (2.5)$$

usando la regla de la cadena en la Ec. (2.5), podemos escribir la velocidad de la partícula  $i$ -ésima como

$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt} = \sum_{k=1}^n \frac{\partial \mathbf{r}_i}{\partial q_k} \dot{q}_k + \frac{\partial \mathbf{r}_i}{\partial t} \quad (2.6)$$

---

<sup>4</sup>Estrictamente las fuerzas de fricción no son fuerzas de ligadura. Sin embargo, en el caso de la fricción deslizante, la magnitud de dicha fuerza depende de la magnitud de la fuerza de ligadura (la normal), con lo cual, la fuerza aplicada (fricción) introduce una dependencia de la fuerza de ligadura, que no puede ser desacoplada. En el caso de la fricción estática, este problema solo aparece si dicha fuerza adquiere su valor máximo.

donde hemos hecho la suposición de que  $\partial_i q_k = dq_k/dt \equiv \dot{q}_k$  más adelante discutiremos esta suposición. Teniendo en cuenta de nuevo la Ec. (2.5), los desplazamientos virtuales originales se conectan con los desplazamientos virtuales en coordenadas generalizadas a través de la relación

$$\delta \mathbf{r}_i = \sum_{j=1}^n \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \quad (2.7)$$

en donde no se incluye la variación temporal por la definición de desplazamiento virtual. Como ya vimos antes, cuando la ligadura cambia con el tiempo, es posible que solo el trabajo virtual se anule, pero no el trabajo real. El trabajo virtual en términos de las coordenadas generalizadas queda entonces

$$\sum_{i=1}^N \mathbf{F}_i \cdot \delta \mathbf{r}_i = \sum_{i=1}^N \sum_{j=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j = \sum_{j=1}^n Q_j \delta q_j \quad (2.8)$$

$$Q_j \equiv \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (2.9)$$

en analogía con la expresión original del trabajo virtual, al término  $Q_j$  se le llama la **fuerza generalizada**. Nótese que  $Q_j$  no necesariamente tiene dimensiones de fuerza al igual que las  $q_j$  no tienen necesariamente unidades de longitud, pero el producto  $Q_j \delta q_j$  debe tener unidades de trabajo.

Nos ocuparemos ahora del segundo término en (2.4)

$$\sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i = \sum_{i=1}^N m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i$$

utilizando (2.7) resulta

$$\sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i = \sum_{i=1}^N \sum_{j=1}^n m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j$$

y usando la identidad

$$m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \frac{d}{dt} \left( m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right)$$

se obtiene

$$\sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i = \sum_{i=1}^N \sum_{j=1}^n \left[ \frac{d}{dt} \left( m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \right] \delta q_j \quad (2.10)$$

examinemos el último término de esta ecuación

$$\frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) = \sum_{k=1}^n \left[ \frac{\partial}{\partial q_k} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \right] \dot{q}_k + \frac{\partial}{\partial t} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right)$$

intercambiando el orden de las derivadas parciales y usando (2.6), resulta

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) &= \sum_{k=1}^n \left[ \frac{\partial}{\partial q_j} \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) \right] \dot{q}_k + \frac{\partial}{\partial q_j} \left( \frac{\partial \mathbf{r}_i}{\partial t} \right) = \frac{\partial}{\partial q_j} \left[ \sum_{k=1}^n \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) \dot{q}_k + \left( \frac{\partial \mathbf{r}_i}{\partial t} \right) \right] \\ \frac{d}{dt} \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) &= \frac{\partial \mathbf{v}_i}{\partial q_j} \end{aligned} \quad (2.11)$$

en lo anterior se asumió que  $\partial \dot{q}_k / \partial q_j = 0$ , suposición que discutiremos más adelante. Ahora usando la Ec. (2.6), tenemos

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \sum_{k=1}^n \left[ \dot{q}_k \frac{\partial}{\partial \dot{q}_j} \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) + \frac{\partial \dot{q}_k}{\partial \dot{q}_j} \frac{\partial \mathbf{r}_i}{\partial q_k} \right] + \frac{\partial}{\partial \dot{q}_j} \left( \frac{\partial \mathbf{r}_i}{\partial t} \right) = \sum_{k=1}^n \left[ \dot{q}_k \frac{\partial}{\partial q_k} \left( \frac{\partial \mathbf{r}_i}{\partial \dot{q}_j} \right) + \delta_{kj} \frac{\partial \mathbf{r}_i}{\partial q_k} \right] + \frac{\partial}{\partial t} \left( \frac{\partial \mathbf{r}_i}{\partial \dot{q}_j} \right)$$

y suponiendo que  $\partial \mathbf{r}_i / \partial \dot{q}_j = 0$ , se obtiene

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \equiv \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (2.12)$$

sustituyendo (2.11) y (2.12) en (2.10) se obtiene

$$\begin{aligned} \sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i &= \sum_{i=1}^N \sum_{j=1}^n \left[ \frac{d}{dt} \left( m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \right) - m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_j} \right] \delta q_j \\ \sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \delta \mathbf{r}_i &= \sum_{j=1}^n \left\{ \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{q}_j} \left( \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right) \right] - \frac{\partial}{\partial q_j} \left( \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right) \right\} \delta q_j \end{aligned} \quad (2.13)$$

y reemplazando (2.8) y (2.13) en (2.4) resulta

$$- \sum_{j=1}^n \left\{ \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{q}_j} \left( \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right) \right] - \frac{\partial}{\partial q_j} \left( \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right) - Q_j \right\} \delta q_j = 0$$

denotaremos la energía cinética del sistema con  $T$  de modo que el principio de D'Alembert queda

$$\sum_{j=1}^n \left\{ \left[ \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] - Q_j \right\} \delta q_j = 0 \quad (2.14)$$

a lo largo de estos desarrollos hemos realizado las siguientes suposiciones

$$\frac{\partial q_k}{\partial t} = \frac{dq_k}{dt} \quad ; \quad \frac{\partial \dot{q}_k}{\partial q_j} = 0 \quad ; \quad \frac{\partial \mathbf{r}_i}{\partial \dot{q}_j} = 0 \quad (2.15)$$

para comprenderlas, observemos que cuando las ligaduras son holónomas, la tercera de estas condiciones es una consecuencia directa del carácter holónomo de las ligaduras como se vé en las Ecs. (2.5). Adicionalmente, dichas ecuaciones nos muestran que siempre es posible obtener un conjunto de coordenadas independientes que nos dan cuenta de los verdaderos grados de libertad del sistema. En tal caso cada una de las coordenadas generalizadas puede depender a lo más del tiempo, de modo que

$$\begin{aligned} q_i &= q_i(t) \Rightarrow \frac{\partial q_i}{\partial t} = \frac{dq_i}{dt} \\ \dot{q}_i &= \dot{q}_i(t) \Rightarrow \frac{\partial \dot{q}_i}{\partial q_j} = 0 \end{aligned} \quad (2.16)$$

vemos entonces que las dos primeras suposiciones (2.15) provienen de las condiciones (2.16) que resultan cuando las coordenadas son independientes, en tanto que la tercera de las suposiciones (2.15) proviene directamente del carácter holónomo de las ligaduras.

Por otra parte, el hecho de que las  $q_k$  solo sean funciones del tiempo, implica además que podemos realizar un desplazamiento virtual para una sola coordenada  $q_k$  sin violar las ligaduras<sup>5</sup>. Al ser independientes todos los desplazamientos virtuales, podemos hacer que todos los  $\delta q$ 's sean nulos excepto un  $\delta q_j$  específico, por tanto su coeficiente asociado en la ecuación (2.14) debe ser nulo. Procediendo de la misma forma con cada coordenada, concluimos que

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j \quad ; \quad j = 1, \dots, n \quad (2.17)$$

hay un número  $n$  de estas ecuaciones, donde  $n$  es el número de grados de libertad (y de coordenadas generalizadas).

<sup>5</sup>Nótese la importancia de que el desplazamiento sea virtual, pues si el desplazamiento es real, la dependencia temporal de las coordenadas no permite que en general se mueva una sola coordenada en el sistema.

Ahora veamos el caso en el cual cada una de las fuerzas  $\mathbf{F}_i$  son derivables de una función potencial escalar<sup>6</sup>  $V$

$$\mathbf{F}_i = -\nabla_i V = -\left(\frac{\partial V}{\partial x_i}, \frac{\partial V}{\partial y_i}, \frac{\partial V}{\partial z_i}\right) \quad ; \quad i = 1, \dots, N \quad (2.18)$$

en este caso las fuerzas generalizadas definidas en (2.9), se escriben como

$$\begin{aligned} Q_j &= \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum_{i=1}^N \nabla_i V \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \\ Q_j &= -\sum_{i=1}^N \left[ \frac{\partial V}{\partial x_i} \frac{\partial x_i}{\partial q_j} + \frac{\partial V}{\partial y_i} \frac{\partial y_i}{\partial q_j} + \frac{\partial V}{\partial z_i} \frac{\partial z_i}{\partial q_j} \right] \end{aligned} \quad (2.19)$$

por otro lado, imponemos como condición adicional, que dicho potencial sea función solamente de las posiciones y el tiempo

$$V = V(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = V(x_1, \dots, x_N, y_1, \dots, y_N, z_1, \dots, z_N; t) \quad (2.20)$$

cuando los argumentos del potencial son los dados por la Ec. (2.20), el término a la derecha de la Ec. (2.19) coincide con la derivada parcial de  $V$  con respecto a  $q_j$  con lo cual se obtiene<sup>7</sup>

$$Q_j = -\frac{\partial V}{\partial q_j} \quad (2.21)$$

reemplazando la expresión (2.21) en (2.17) y teniendo en cuenta que el potencial  $V$  no depende de las velocidades generalizadas  $\dot{q}_j$ <sup>8</sup> podemos escribir

$$\frac{d}{dt} \left[ \frac{\partial (T - V)}{\partial \dot{q}_j} \right] - \frac{\partial (T - V)}{\partial q_j} = 0 \quad ; \quad j = 1, \dots, n$$

definimos entonces una nueva función

$$L \equiv T - V \quad (2.22)$$

conocida como el Lagrangiano  $L$  asociado al sistema. Con lo cual obtenemos finalmente

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}_j} \right] - \frac{\partial L}{\partial q_j} = 0 \quad ; \quad j = 1, \dots, n \quad (2.23)$$

las  $n$  ecuaciones así obtenidas se conocen como **ecuaciones de Lagrange**. Obsérvese que estas ecuaciones no requieren que las fuerzas sean conservativas, ya que el potencial puede ser función explícita del tiempo.

En la formulación Lagrangiana ya vemos algunas ventajas operativas como son: (1) No aparecen las fuerzas de ligadura, y las fuerzas aplicadas usualmente son parámetros de entrada. Es decir, normalmente conocemos la forma funcional de las fuerzas aplicadas. (2) Al no haber coordenadas dependientes, el número de ecuaciones es el menor posible. (3) Las ecuaciones son escalares ya que están basadas en la energía. (4) Las ecuaciones de Lagrange son idénticas en forma en cualquier sistema de coordenadas generalizadas<sup>9</sup>.

<sup>6</sup>Debemos tener presente que  $\mathbf{F}_i$  incluye las fuerzas internas y externas sobre la partícula  $i$ . Adicionalmente, el potencial  $V$  está asociado a todo el sistema de partículas y no a una sola partícula, razón por la cual las fuerzas sobre cada partícula se pueden extraer de él como se observa en la Ec. (2.18).

<sup>7</sup>Si el potencial dependiera por ejemplo de los  $\dot{x}_i$ , entonces  $\partial V / \partial q_j$  también tendría términos de la forma  $\sum_i \frac{\partial V}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial q_j}$  de manera que la igualdad (2.21) ya no se cumple.

<sup>8</sup>Cuando las ligaduras son holónomas, la transformación entre coordenadas me garantiza que si el potencial no depende de las velocidades en las coordenadas originales, tampoco depende de las velocidades generalizadas en el nuevo sistema coordenado.

<sup>9</sup>Por ejemplo, en coordenadas cartesianas las ecuaciones de Newton tienen la forma  $F = m\ddot{x}$ . Sin embargo, al cambiar a coordenadas polares  $F_\theta \neq m\ddot{\theta}$ , y la forma de las ecuaciones cambia con el cambio en el sistema coordenado.

### 2.3.1. Energías cinética y potencial en coordenadas generalizadas

En virtud de que las ecuaciones de Lagrange implican derivadas en las coordenadas generalizadas, es necesario conocer el valor de la energía cinética y potencial en términos de dichas coordenadas. Utilizando (2.6), la energía cinética en términos de coordenadas generalizadas queda

$$T = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 = \sum_{i=1}^N \frac{1}{2} m_i \left( \sum_{j=1}^n \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t} \right)^2$$

al expandir el binomio obtenemos

$$\begin{aligned} T &= \sum_{i=1}^N \frac{1}{2} m_i \left[ \left( \sum_{j=1}^n \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j \right) \cdot \left( \sum_{k=1}^n \frac{\partial \mathbf{r}_i}{\partial q_k} \dot{q}_k \right) + 2 \frac{\partial \mathbf{r}_i}{\partial t} \cdot \left( \sum_{j=1}^n \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j \right) + \left( \frac{\partial \mathbf{r}_i}{\partial t} \right)^2 \right] \\ T &= \sum_{i=1}^N \frac{1}{2} m_i \left( \frac{\partial \mathbf{r}_i}{\partial t} \right)^2 + \sum_{j=1}^n \left[ \sum_{i=1}^N m_i \frac{\partial \mathbf{r}_i}{\partial t} \cdot \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \right] \dot{q}_j + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \left[ \sum_{i=1}^N m_i \left( \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \cdot \left( \frac{\partial \mathbf{r}_i}{\partial q_k} \right) \right] \dot{q}_j \dot{q}_k \end{aligned}$$

vemos que la energía cinética contiene un término independiente de las velocidades generalizadas, así como otro lineal y otro cuadrático en dichas velocidades. Es entonces conveniente escribir la energía cinética en la forma

$$\begin{aligned} T &= T_0 + T_1 + T_2 = M_0 + \sum_{j=1}^n M_j \dot{q}_j + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n M_{jk} \dot{q}_j \dot{q}_k \quad (2.24) \\ M_0 &\equiv \sum_{i=1}^N \frac{1}{2} m_i \left( \frac{\partial \mathbf{r}_i}{\partial t} \right)^2 \quad ; \quad M_j \equiv \sum_{i=1}^N m_i \frac{\partial \mathbf{r}_i}{\partial t} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad ; \quad M_{jk} \equiv \sum_{i=1}^N m_i \frac{\partial \mathbf{r}_i}{\partial q_j} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \end{aligned}$$

si las ecuaciones de transformación (2.5) no dependen explícitamente del tiempo (ligaduras holónomas y esclerónomas), solo el término cuadrático sobrevive.

En cuanto al potencial, su forma explícita en coordenadas generalizadas depende de cada sistema en particular. En realidad, la forma del potencial es la que usualmente sugiere las coordenadas generalizadas a usar.

### 2.3.2. Una simetría gauge o de calibración para el Lagrangiano

Hemos visto que cuando las fuerzas sobre el sistema (excepto quizás las de ligadura) se pueden escribir como potenciales que dependen de la posición y el tiempo, el Lagrangiano contiene toda la información física del sistema. No obstante, el Lagrangiano no se construyó como un observable, sino como una función generadora de las ecuaciones dinámicas del sistema. De esto se concluye que si existe otra función generadora  $L'$  que me lleve a las mismas ecuaciones de movimiento (para un conjunto dado de coordenadas generalizadas) esta función contiene la misma información física del Lagrangiano, por tanto  $L'$  bien puede considerarse como otro Lagrangiano igualmente válido para el mismo sistema Físico y el mismo sistema coordinado. En realidad, la función Lagrangiana definida por (2.22) no es la única que conduce a las ecuaciones de movimiento (2.23). Efectivamente se puede demostrar que si  $\Omega(q_1, \dots, q_n, t)$  es una función diferenciable en todos sus argumentos, la nueva función Lagrangiana definida por

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d\Omega(q, t)}{dt} \quad (2.25)$$

conduce a las mismas ecuaciones de movimiento. En algunos casos particulares otras redefiniciones son posibles. Las ecuaciones de Lagrange poseen entonces una simetría **gauge** o de **calibración** para el Lagrangiano. Para verificar que el Lagrangiano  $L'$  conduce a las mismas ecuaciones de movimiento que  $L$ , es suficiente demostrar

que las ecuaciones de Lagrange para un Lagrangiano de la forma  $\bar{L} = d[\Omega(\mathbf{q}, t)]/dt$ , son triviales, es decir dan cero en ambos miembros.

$$\begin{aligned}
& \frac{d}{dt} \left( \frac{\partial \bar{L}}{\partial \dot{q}_i} \right) - \frac{\partial \bar{L}}{\partial q_i} = \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} \left( \frac{d\Omega(\mathbf{q}, t)}{dt} \right) - \frac{\partial}{\partial q_i} \left( \frac{d\Omega(\mathbf{q}, t)}{dt} \right) \\
&= \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} \left( \sum_{j=1}^n \frac{\partial \Omega}{\partial q_j} \dot{q}_j + \frac{\partial \Omega}{\partial t} \right) - \frac{\partial}{\partial q_i} \left( \sum_{j=1}^n \frac{\partial \Omega}{\partial q_j} \dot{q}_j + \frac{\partial \Omega}{\partial t} \right) \\
&= \frac{d}{dt} \left( \sum_{j=1}^n \frac{\partial \Omega}{\partial q_j} \delta_{ij} \right) - \left( \sum_{j=1}^n \frac{\partial^2 \Omega}{\partial q_i \partial q_j} \dot{q}_j + \frac{\partial^2 \Omega}{\partial q_i \partial t} \right) \\
&= \frac{d}{dt} \left( \frac{\partial \Omega}{\partial q_i} \right) - \left\{ \sum_{j=1}^n \left[ \frac{\partial}{\partial q_j} \left( \frac{\partial \Omega}{\partial q_i} \right) \right] \dot{q}_j + \frac{\partial}{\partial t} \left( \frac{\partial \Omega}{\partial q_i} \right) \right\} \\
&= \sum_{j=1}^n \left[ \frac{\partial}{\partial q_j} \left( \frac{\partial \Omega}{\partial q_i} \right) \right] \dot{q}_j + \frac{\partial}{\partial t} \left( \frac{\partial \Omega}{\partial q_i} \right) - \left\{ \sum_{j=1}^n \left[ \frac{\partial}{\partial q_j} \left( \frac{\partial \Omega}{\partial q_i} \right) \right] \dot{q}_j + \frac{\partial}{\partial t} \left( \frac{\partial \Omega}{\partial q_i} \right) \right\} \\
&= 0
\end{aligned}$$

donde hemos usado el hecho de que  $\Omega$  no depende de  $\dot{q}_i$ . Adicionalmente, hemos tenido en cuenta que las derivadas parciales se pueden intercambiar, siempre y cuando éstas últimas sean continuas en todo el intervalo espacio temporal en cuestión<sup>10</sup>.

## 2.4. Ecuaciones de Lagrange para potenciales generalizados

Las ecuaciones (2.17) se pueden escribir sintéticamente en términos de un operador diferencial lineal

$$\hat{O}_j T = Q_j \quad ; \quad \hat{O}_j \equiv \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{q}_j} \right] - \frac{\partial}{\partial q_j} \quad (2.26)$$

claramente, de esta ecuación podemos llegar a la ecuación de Lagrange (2.23) siempre y cuando la fuerza generalizada  $Q_j$  se pueda escribir como

$$Q_j = \hat{O}_j U \quad ; \quad \hat{O}_j \equiv \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{q}_j} \right] - \frac{\partial}{\partial q_j} \quad (2.27)$$

siendo  $U$  una función escalar de  $q_k, \dot{q}_k, t$ . Cuando se cumple la condición (2.27) la podemos sustituir en (2.26) y obtener

$$\begin{aligned}
\hat{O}_j T &= \hat{O}_j U \Rightarrow \hat{O}_j (T - U) = 0 \\
&\Rightarrow \hat{O}_j L = 0 \quad ; \quad L \equiv T - U
\end{aligned} \quad (2.28)$$

Dada la definición de  $\hat{O}_j$  Ec. (2.26) se vé que la Ec. (2.28) es idéntica en forma a la Ec. (2.23). Al escalar  $U$  se le denomina **potencial generalizado**. Este potencial puede depender de las velocidades generalizadas  $\dot{q}_k$  y el caso  $\mathbf{F}_i = -\nabla_i U(q_k, t)$  está incluido como caso particular.

El caso especial mas importante en la Física de estos potenciales generalizados lo constituye el potencial generalizado asociado a la fuerza de Lorentz. Sea una carga puntual  $q$  de masa  $m$  y que viaja con velocidad  $\mathbf{v}$ . Asumamos que esta carga se propaga en el vacío en presencia de un campo eléctrico  $\mathbf{E}$  y un campo magnético  $\mathbf{B}$ , el valor de la fuerza instantánea que experimenta dicha carga es

$$\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} \quad (2.29)$$

<sup>10</sup>De este procedimiento se vé que si  $\Omega$  dependiera de los  $\dot{q}_i$ , los Lagrangianos relacionados en (2.25), no serían equivalentes.

donde  $\mathbf{E}(\mathbf{r}, t)$  y  $\mathbf{B}(\mathbf{r}, t)$  son funciones continuas en sus argumentos. Estos campos se pueden generar de un potencial escalar  $\phi(\mathbf{r}, t)$  y uno vectorial  $\mathbf{A}(\mathbf{r}, t)$

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \quad ; \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (2.30)$$

verifiquemos que el potencial

$$U = q\phi - q\mathbf{A} \cdot \mathbf{v} \quad (2.31)$$

es un potencial generalizado adecuado para reproducir la fuerza generalizada de Lorentz. En nuestro caso, dado que el problema no tiene ligaduras, lo más cómodo es usar las coordenadas cartesianas como coordenadas generalizadas, de modo que las coordenadas originales y las generalizadas coinciden. Las fuerzas generalizadas son simplemente las componentes de la fuerza original (recordemos que las fuerzas generalizadas son cantidades escalares). Calculemos la cantidad  $\widehat{O}_j U$ , en lo que sigue asumimos convención de suma sobre índices repetidos y tendremos en cuenta que  $x_i$  y  $\dot{x}_i = v_i$  son independientes entre sí. Finalmente, asumiremos que  $\mathbf{A}$  y  $\phi$  solo son funciones de  $x_i, t$  pero no son funciones de  $\dot{x}_i$ .

$$U = q\phi - q\mathbf{A} \cdot \mathbf{v} = q\phi - qA_i v_i$$

$$\begin{aligned} \frac{\partial U}{\partial x_j} &= q \frac{\partial \phi}{\partial x_j} - q \frac{\partial A_i}{\partial x_j} v_i - q A_i \frac{\partial v_i}{\partial x_j} = q \frac{\partial \phi}{\partial x_j} - q \frac{\partial A_i}{\partial x_j} v_i \\ \frac{\partial U}{\partial \dot{x}_j} &= \frac{\partial U}{\partial v_j} = q \frac{\partial \phi}{\partial v_j} - q \frac{\partial A_i}{\partial v_j} v_i - q A_i \cdot \frac{\partial v_i}{\partial v_j} = -q A_j \\ \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{x}_j} \right) &= -q \frac{\partial A_j}{\partial x_i} \dot{x}_i - q \frac{\partial A_j}{\partial t} = -q \frac{\partial A_j}{\partial x_i} v_i - q \frac{\partial A_j}{\partial t} \end{aligned}$$

usando las Ecs. (2.26, 2.27), tenemos

$$\begin{aligned} Q_j &= \widehat{O}_j U = -\frac{\partial U}{\partial x_j} + \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{x}_j} \right) \\ &= -q \frac{\partial \phi}{\partial x_j} + q \frac{\partial A_i}{\partial x_j} v_i - q \frac{\partial A_j}{\partial x_i} v_i - q \frac{\partial A_j}{\partial t} \\ &= -q \left( \frac{\partial \phi}{\partial x_j} + \frac{\partial A_j}{\partial t} \right) + q \left( \frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} \right) v_i \end{aligned}$$

definiendo el tensor de Levi-Civita  $\varepsilon_{ijk}$  podemos escribir lo anterior en notación mas compacta

$$\begin{aligned} Q_j &= -q(\partial_j \phi + \partial_t A_j) + q(\partial_j A_i - \partial_i A_j) v_i = -q(\partial_j \phi + \partial_t A_j) + q\varepsilon_{jin} \varepsilon_{nrs} v_i \partial_r A_s \\ Q_j &= -q(\nabla \phi + \partial_t \mathbf{A})_j + q\varepsilon_{jin} v_i (\nabla \times \mathbf{A})_n = -q(\nabla \phi + \partial_t \mathbf{A})_j + q[\mathbf{v} \times (\nabla \times \mathbf{A})]_j \end{aligned}$$

teniendo en cuenta (2.30) la fuerza generalizada  $Q_j$  queda

$$Q_j = q\mathbf{E}_j + q[\mathbf{v} \times \mathbf{B}]_j$$

de modo que componente a componente, este potencial genera correctamente la fuerza de Lorentz Ec. (2.29). El Lagrangiano se escribe

$$L \equiv T - U = \frac{1}{2} m \dot{x}_i \dot{x}_i - q\phi + qA_i \dot{x}_i \quad (2.32)$$

tomemos la ecuación de Lagrange asociada a la coordenada generalizada  $x_k$  con  $k = 1, 2, 3$

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{x}_k} \right] - \frac{\partial L}{\partial x_k} = 0 \Rightarrow \frac{d}{dt} [m\dot{x}_k + qA_k] + q\partial_k \phi - q(\partial_k A_n) \dot{x}_n = 0$$

la ecuación de movimiento se puede escribir como

$$m\ddot{x}_k = q(v_n \partial_k A_n) - q \left( \partial_k \phi + \frac{dA_k}{dt} \right) \quad (2.33)$$

dado que  $\mathbf{A}$  solo depende de la posición y el tiempo se tiene que

$$\frac{dA_k}{dt} = \frac{\partial A_k}{\partial t} + \frac{\partial A_k}{\partial x_n} \frac{\partial x_n}{\partial t} = \frac{\partial A_k}{\partial t} + \frac{\partial A_k}{\partial x_n} \frac{dx_n}{dt} = \partial_t A_k + v_n \partial_n A_k$$

de modo que la derivada total de  $A_k$  se puede expresar en términos de la denominada “derivada convectiva”

$$\frac{dA_k}{dt} = \partial_t A_k + \mathbf{v} \cdot \nabla A_k = \partial_t A_k + v_n \partial_n A_k$$

reemplazando en (2.33)

$$\begin{aligned} m\ddot{x}_k &= qv_n \partial_k A_n - q(\partial_t A_k + v_n \partial_n A_k) - q\partial_k \phi \\ m\ddot{x}_k &= qv_n (\partial_k A_n - \partial_n A_k) - q(\partial_t A_k + \partial_k \phi) \\ m\ddot{x}_k &= q[\mathbf{v} \times (\nabla \times \mathbf{A})]_k - q \left[ \left( \frac{\partial \mathbf{A}}{\partial t} \right)_k + (\nabla \phi)_k \right] \end{aligned}$$

usando (2.30) la ecuación queda

$$m\ddot{x}_k = q[\mathbf{v} \times \mathbf{B}]_k + q\mathbf{E}_k$$

y se observa que las ecuaciones de Lagrange reproducen la expresión para la fuerza de Lorentz Ec. (2.29).

## 2.5. Ecuaciones de Lagrange con fuerzas disipativas

En general las fuerzas generalizadas se pueden descomponer en dos términos de la siguiente forma

$$Q_j = Q_j^U + Q_j' \quad ; \quad Q_j^U = \widehat{O}_j U$$

de tal manera que  $Q_j^U$  proviene de algún potencial generalizado en tanto que  $Q_j'$  es un término que no se puede escribir en términos de ningún potencial generalizado. Si introducimos esta expresión en (2.26)

$$\begin{aligned} \widehat{O}_j T &= \widehat{O}_j U + Q_j' \Rightarrow \\ \widehat{O}_j (T - U) &= Q_j' \end{aligned} \quad (2.34)$$

podemos entonces definir un Lagrangiano  $L = T - U$  y una fuerza generalizada  $Q_j'$  que no proviene de ningún potencial, de tal manera que usando (2.26), la Ec. (2.34) queda

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}_j} \right] - \frac{\partial L}{\partial q_j} = Q_j' \quad (2.35)$$

esta ecuación de Lagrange generalizada es útil por ejemplo para el caso de fuerzas de fricción. Con frecuencia, la fricción viscosa se modela como una interacción proporcional a la velocidad, de modo que su componente  $x$  se escribe como

$$F_{fx} = -k_x v_x \quad (2.36)$$

y similarmente para las otras componentes. Si el fluido es anisotrópico, las constantes  $k_x, k_y, k_z$  serán en general diferentes. Supondremos sin embargo, que el medio es homogéneo y que por tanto  $k_x, k_y$  y  $k_z$  no dependen de la posición ni del tiempo. Las fuerzas de fricción de este tipo se pueden derivar de una expresión de la forma

$$F = \frac{1}{2} \sum_{i=1}^N (k_x v_{ix}^2 + k_y v_{iy}^2 + k_z v_{iz}^2) \quad (2.37)$$



donde la suma es sobre las partículas del sistema. Si el fluido en donde están inmersas las partículas es homogéneo, entonces  $F$  será función de las velocidades únicamente. A esta expresión se le conoce como **función de disipación de Rayleigh**. A partir de (2.37), la componente  $x$  de la fuerza de fricción sobre la  $i$ -ésima partícula se escribe

$$F_{f_{ix}} = -\frac{\partial F}{\partial v_{ix}}$$

y análogamente para las otras componentes. Simbólicamente se puede escribir

$$\mathbf{F}_{if} = -\nabla_{v_i} F = -\nabla_{\dot{x}_i} F$$

para ver el significado físico de la función de disipación, calculemos el trabajo realizado por el fluido sobre el sistema

$$\begin{aligned} dW_f &= \sum_{i=1}^N \mathbf{F}_{if} \cdot d\mathbf{r}_i = \sum_{i=1}^N \mathbf{F}_{if} \cdot \mathbf{v}_i dt = -\sum_{i=1}^N (k_x v_{ix}, k_y v_{iy}, k_z v_{iz}) \cdot (v_{ix}, v_{iy}, v_{iz}) dt \\ dW_f &= -\sum_{i=1}^N (k_x v_{ix}^2 + k_y v_{iy}^2 + k_z v_{iz}^2) dt = -2F dt \end{aligned}$$

de lo cual se vé que  $2F$  corresponde a la rata de disipación de energía debida a la fricción. La componente de la fuerza generalizada resultante de la fuerza de fricción está dada por

$$Q_j = \sum_{i=1}^N \mathbf{F}_{fi} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum_{i=1}^N \nabla_{v_i} F \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

usando (2.12), y teniendo en cuenta que la función de disipación solo depende de los  $\dot{\mathbf{r}}_i$  y por tanto de los  $\dot{q}_i$ , obtenemos

$$Q_j = -\sum_{i=1}^N \nabla_{\dot{x}_i} F \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = -\frac{\partial F}{\partial \dot{q}_j}$$

Las ecuaciones de Lagrange con un término de disipación quedan

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} + \frac{\partial F}{\partial \dot{q}_j} = 0 \quad (2.38)$$

Vemos que cuando existen fuerzas generalizadas que no provienen de un potencial generalizado (denotadas por  $Q'_j$ ), el Lagrangiano no contiene toda la información física del sistema, ya que solo la parte de  $Q_j$  que sí proviene de potenciales generalizados es absorbida en él. Este hecho resulta claro de la Ec. (2.35), y para el caso específico de fuerzas disipativas esto se vé de la Ec. (2.38) en donde la función de disipación de Rayleigh contiene información física que no posee el Lagrangiano.

## 2.6. Algunos aspectos peculiares de las coordenadas y las fuerzas generalizadas

El concepto de fuerza generalizada, definida a través de las coordenadas generalizadas  $q_j$  resultó ser

$$Q_j \equiv \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (2.39)$$

observemos en primer lugar que las coordenadas  $q_j$  no están necesariamente asociadas a una partícula. Por ejemplo, si en el péndulo doble de la Fig. 2.1 tengo una lenteja adicional en la mitad de una de las varillas, la coordenada generalizada digamos  $\theta_1$  me describe la posición de dos de las lentejas. Tampoco es necesario que la coordenada generalizada determine la posición de una o varias de las partículas en forma directa, por ejemplo

una o más coordenadas del centro de masa del sistema puede ser útil como coordenada generalizada, y es bien sabido que no necesariamente debe haber presencia de masa en este punto. Es quizás más apropiado decir que cada coordenada  $q_j$  está asociada al sistema. De esto se desprende que una fuerza generalizada tampoco está necesariamente asociada a una partícula o incluso a un subsistema, solo podemos decir en general que está asociada a una coordenada generalizada  $q_j$ . Recordemos además que  $q_j, Q_j$  pueden tener en principio cualquier dimensión siempre y cuando  $Q_j \delta q_j$  tenga dimensiones de energía. En la Ec. (2.39),  $\mathbf{F}_i$  corresponde a la resultante de las fuerzas internas y externas sobre la partícula  $i$ , pero hay una suma sobre todas las partículas (las posiciones y fuerzas iniciales  $\mathbf{r}_i, \mathbf{F}_i$  sí están asociadas a partículas), esto enfatiza el hecho de que  $Q_j$  no está necesariamente asociado a una partícula.

Con esta misma filosofía, debemos comprender que un desplazamiento compatible con las ligaduras (real o virtual) de una sola coordenada independiente, no necesariamente implica el desplazamiento de una sola partícula del sistema. Tomemos de nuevo el péndulo doble de la Fig. 2.1, los ángulos  $\theta_1$  y  $\theta_2$  que cada lenteja hace con la vertical son un conjunto posible de coordenadas generalizadas independientes. Un desplazamiento virtual  $\delta\theta_2$  compatible con las ligaduras y que mantenga fijo a  $\theta_1$  es claro que implica solo el movimiento de la lenteja  $m_2$  permaneciendo  $m_1$  en su lugar. Sin embargo, un desplazamiento virtual  $\delta\theta_1$  de la lenteja  $m_1$  compatible con las ligaduras y que deje fija la coordenada  $\theta_2$  claramente requiere el movimiento de la otra lenteja, pues mover la lenteja 1 sin mover la lenteja 2 viola las ligaduras y modifica el valor de  $\theta_2$ . Es claro sin embargo, que al menos virtualmente existe un movimiento simultáneo de las dos lentejas que es compatible con las ligaduras y que varía la coordenada  $\theta_1$  y deja fija la coordenada  $\theta_2$ . Tal movimiento consiste en desplazar  $m_1$  de modo que  $l_1$  permanezca constante y que la lenteja  $m_2$  se mueva de tal forma que el vector relativo  $\mathbf{r}_{12}$  que une a las dos masas, ejecute una translación paralela. Nuevamente no hay una asociación directa entre coordenadas generalizadas y partículas del sistema.

## 2.7. Relación entre sistemas coordenados y sistemas de referencia

Un conjunto de tres ejes coordenados linealmente independientes que convergen en un punto (origen) forman un sistema de referencia respecto al cual se pueden medir cantidades físicas tales como velocidades, desplazamientos, aceleraciones, fuerzas, torques etc. Existen infinitos sistemas de referencia inerciales y no inerciales. Una vez fijado el sistema de referencia (usualmente inercial), existen infinitos sistemas de coordenadas que se pueden construir. En particular, para describir un sistema físico particular es posible usar diversos sistemas de coordenadas generalizadas **para un mismo sistema de referencia**. Por ejemplo, las transformaciones (2.5) describen un cambio de sistema coordenado pero bajo el mismo sistema de referencia. En la mayor parte del tratamiento de este texto se trabajará con transformaciones de coordenadas sin cambio en el sistema de referencia, a menos que se indique lo contrario.

Notemos en particular que aunque el Lagrangiano puede tener un comportamiento funcional diferente cuando se cambia de sistema coordenado (sin cambio en el sistema de referencia), su valor numérico debe permanecer intacto ya que la energía cinética y potencial no sufren ningún cambio (en su valor numérico). En contraste, si se cambia de sistema de referencia claramente pueden cambiar la energía cinética y la potencial, e incluso es posible que haya que agregar fuerzas ficticias, por tanto un cambio en el sistema de referencia puede alterar tanto el valor numérico como el comportamiento funcional del Lagrangiano. Es necesario entonces diferenciar muy bien entre un cambio de sistema coordenado y un cambio de sistema de referencia.

## 2.8. Ejemplos de uso de la formulación Lagrangiana

### 2.8.1. Partícula en el espacio

Una partícula de masa  $m$ , se mueve bajo la acción de una fuerza  $\mathbf{F}$ . Dado que no hay ligaduras, usaremos coordenadas cartesianas a manera de coordenadas generalizadas. La energía cinética se escribe

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad ; \quad \frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0$$

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x} \quad ; \quad \frac{\partial T}{\partial \dot{y}} = m\dot{y} \quad ; \quad \frac{\partial T}{\partial \dot{z}} = m\dot{z}$$

las fuerzas generalizadas Ec. (2.39) nos dan

$$Q_j \equiv \sum_{i=1}^1 \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial q_j} \quad ; \quad q_j = x, y, z$$

$$Q_x = (F_x, F_y, F_z) \cdot \frac{\partial}{\partial x}(x, y, z) = F_x$$

similarmente  $Q_y = F_y$  y  $Q_z = F_z$ . Con lo cual las ecuaciones de movimiento (2.17) quedan

$$\frac{d}{dt}(m\dot{x}) = F_x \quad ; \quad \frac{d}{dt}(m\dot{y}) = F_y \quad ; \quad \frac{d}{dt}(m\dot{z}) = F_z$$

con lo cual llegamos a las ecuaciones de movimiento de Newton.

Veamos las ecuaciones de la misma partícula pero ahora en dos dimensiones y en coordenadas polares, la transformación a estas coordenadas generalizadas es

$$x = r \cos \theta \quad ; \quad y = r \sin \theta$$

las velocidades son

$$\dot{x} = \dot{r} \cos \theta - r\dot{\theta} \sin \theta \quad ; \quad \dot{y} = \dot{r} \sin \theta + r\dot{\theta} \cos \theta$$

para calcular la energía cinética calculamos  $\mathbf{v}^2$

$$\mathbf{v}^2 = (\dot{x}^2 + \dot{y}^2) = (\dot{r} \cos \theta - r\dot{\theta} \sin \theta)^2 + (\dot{r} \sin \theta + r\dot{\theta} \cos \theta)^2$$

$$\mathbf{v}^2 = \dot{r}^2 (\cos^2 \theta + \sin^2 \theta) + r^2 \dot{\theta}^2 (\sin^2 \theta + \cos^2 \theta) - 2r\dot{r}\dot{\theta} \cos \theta \sin \theta + 2r\dot{r}\dot{\theta} \sin \theta \cos \theta$$

$$\mathbf{v}^2 = \dot{r}^2 + r^2 \dot{\theta}^2$$

por tanto<sup>11</sup>

$$T = \frac{1}{2}m(\dot{r}^2 + r^2 \dot{\theta}^2) \quad (2.40)$$

ahora calculemos las fuerzas generalizadas, Ecs. (2.39)

$$Q_r = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial r} = \mathbf{F} \cdot \frac{\partial (r \mathbf{u}_r)}{\partial r} = \mathbf{F} \cdot \mathbf{u}_r + \mathbf{F} \cdot r \frac{\partial \mathbf{u}_r}{\partial r} = \mathbf{F} \cdot \mathbf{u}_r = F_r \quad (2.41)$$

$$Q_\theta = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial \theta} = \mathbf{F} \cdot \frac{\partial (r \mathbf{u}_r)}{\partial \theta} = \mathbf{F} \cdot r \frac{\partial \mathbf{u}_r}{\partial \theta} = \mathbf{F} \cdot r \mathbf{u}_\theta = rF_\theta \quad (2.42)$$

obsérvese que no todas las fuerzas (y coordenadas) generalizadas tienen las mismas dimensiones.  $Q_r$  tiene dimensiones de fuerza en tanto que  $Q_\theta$  tiene dimensiones de torque (o trabajo), las coordenadas generalizadas

<sup>11</sup>El lector puede también calcular la energía cinética a partir de las Ecs. (2.24).

respectivas tienen dimensiones de longitud y son adimensionales respectivamente, de modo que los productos  $Q_j \delta q_j$  siempre tienen dimensiones de trabajo. Escribamos las dos ecuaciones de Lagrange (2.17)<sup>12</sup>

$$\frac{\partial T}{\partial r} = mr\dot{\theta}^2 ; \quad \frac{\partial T}{\partial \dot{r}} = m\dot{r} ; \quad \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{r}} \right) = m\ddot{r} \Rightarrow$$

$$m\ddot{r} - mr\dot{\theta}^2 = F_r$$

el segundo término es el de aceleración centrípeta. Veamos la ecuación asociada a  $\theta$

$$\frac{\partial T}{\partial \theta} = 0 ; \quad \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta} ; \quad \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\theta}} \right) = mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta}$$

la ecuación resulta

$$mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} = rF_\theta$$

se puede demostrar que el término de la izquierda es la derivada temporal del momento angular, en tanto que el término de la derecha es el torque aplicado. Hemos llegado entonces a la expresión en componentes de la Ec. (1.3).

### 2.8.2. Máquina de Atwood

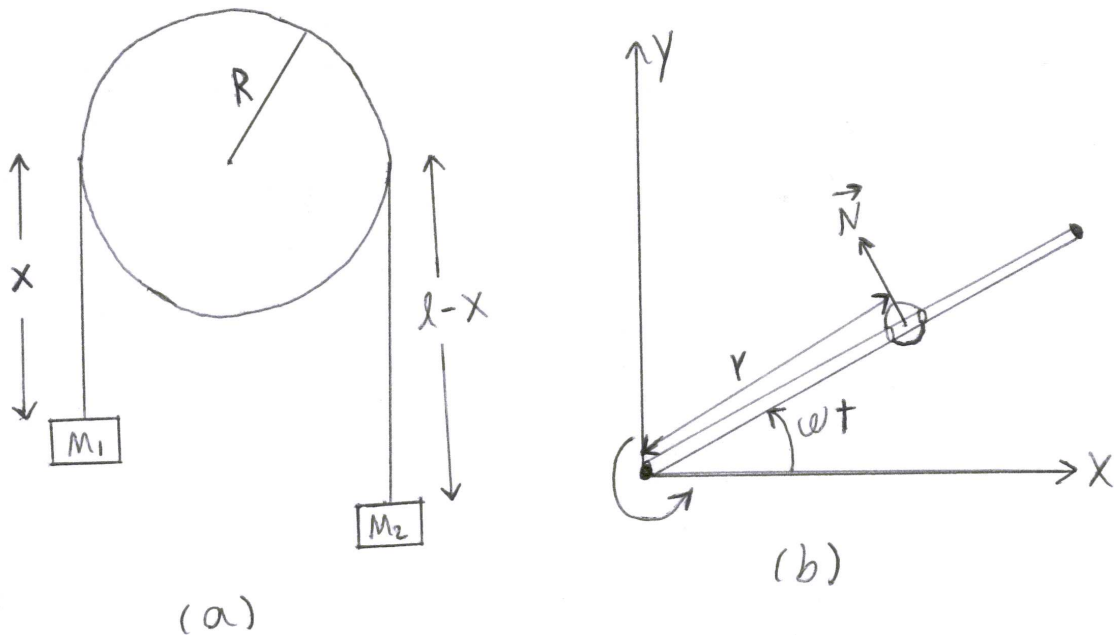


Figura 2.2: (a) Máquina de Atwood. La longitud  $L = l_1 + l_2 + \pi R$  de la cuerda es constante, y por tanto también lo es  $l = l_1 + l_2$ . (b) Vista aérea de una cuenta sobre un alambre que rota con velocidad angular constante  $\omega$ , en ausencia de un campo gravitacional.

La máquina de Atwood mostrada en la Fig. 2.2a, consiste en una polea fija, por la cual pasa una cuerda que se ata a dos masas  $m_1$  y  $m_2$ . Asumiremos que la cuerda no desliza sobre la polea y que la masa de la polea y de la cuerda son despreciables. La ligadura (longitud constante de la cuerda) es holónoma y esclerónoma. Solo hay una coordenada independiente  $x$ , la otra es fijada por la ligadura de longitud constante de la cuerda. Usando el origen en el centro de la polea, la energía potencial es

$$V = -M_1gx - M_2g(l - x) ; \quad T = \frac{1}{2}(M_1 + M_2)\dot{x}^2$$

<sup>12</sup>En este problema no podemos usar las ecuaciones de Lagrange dadas por (2.22, 2.23), puesto que no conocemos la forma específica de la fuerza y por tanto, no sabemos si ésta posee un potencial asociado.

el Lagrangiano queda

$$L \equiv T - V = \frac{1}{2} (M_1 + M_2) \dot{x}^2 + M_1 g x + M_2 g (l - x)$$

solo hay una coordenada generalizada  $x$ , y por tanto una sola ecuación de movimiento del tipo (2.23)

$$\begin{aligned} \frac{\partial L}{\partial x} &= (M_1 - M_2) g ; \quad \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = (M_1 + M_2) \ddot{x} \Rightarrow \\ (M_1 + M_2) \ddot{x} &= (M_1 - M_2) g \end{aligned}$$

de aquí sale el valor de la aceleración el cual coincide con el ya obtenido por métodos tradicionales. Obsérvese que la fuerza de ligadura de tensión no aparece en el formalismo y no se puede obtener directamente de él.

### 2.8.3. Cuenta sobre un alambre

Una cuenta se desliza sin rozamiento por un alambre que rota uniformemente alrededor de un eje fijo perpendicular al alambre (ver Fig. 2.2b). Este es un ejemplo de ligadura holónoma reónoma, ya que la posición de la cuenta depende explícitamente del tiempo<sup>13</sup>

$$x = r \cos \omega t ; \quad y = r \sin \omega t \quad (2.43)$$

donde  $\omega$  es la velocidad angular de rotación del alambre y  $r$  la distancia entre la cuenta y el eje de rotación. Aquí estamos tomando un sistema coordenado polar con  $\theta = \omega t$  de modo que  $\dot{\theta} = \omega$ . El potencial es nulo ya que asumimos que el sistema no está en un campo gravitacional, de modo que el Lagrangiano es  $T$ . La energía cinética se puede tomar de (2.40)<sup>14</sup>

$$T = \frac{1}{2} m (\dot{r}^2 + r^2 \omega^2) \quad (2.44)$$

vemos que en este problema hay una sola coordenada generalizada  $r$ , puesto que el tiempo es un parámetro y no se considera una coordenada. La ecuación de movimiento es

$$m\ddot{r} - mr\omega^2 = 0 \quad \Rightarrow \quad \ddot{r} - \omega^2 r = 0 \quad (2.45)$$

la solución general es de la forma

$$r(t) = Ae^{\omega t} + Be^{-\omega t} ; \quad \dot{r}(t) = \omega Ae^{\omega t} - \omega Be^{-\omega t} \quad (2.46)$$

si la cuenta está inicialmente en reposo sobre el alambre entonces  $\dot{r}(0) = 0$  y por tanto  $A = B$ . La solución queda entonces

$$r = A (e^{\omega t} - e^{-\omega t}) = C \sinh \omega t \quad (2.47)$$

y la cuenta se mueve exponencialmente hacia afuera para tiempos suficientemente largos, la constante  $C$  se determina con la posición inicial. De nuevo el método no nos da directamente la fuerza de ligadura que actúa sobre la cuenta. Dicha ligadura (fuerza normal  $N$ ) se puede hallar con la expresión del momento angular

$$\begin{aligned} L &= |\mathbf{r} \times \mathbf{p}| = r m v_\theta = r m r \omega = m r^2 \omega = m \omega C^2 \sinh^2 \omega t \\ \frac{dL}{dt} &= \tau = 2m\omega^2 C^2 \sinh \omega t \cosh \omega t \end{aligned}$$

que produce la fuerza normal  $N = \tau/r$ , de modo que la fuerza de ligadura es

$$\begin{aligned} N &= \frac{\tau}{r} = \frac{2m\omega^2 C^2 \sinh \omega t \cosh \omega t}{C \sinh \omega t} \\ N &= 2m\omega^2 C \cosh \omega t \end{aligned} \quad (2.48)$$

la cual actúa en dirección perpendicular al alambre y al eje de rotación.

<sup>13</sup>Como es usual, esta dependencia explícita con el tiempo se debe a algún agente externo, que mantiene el alambre a velocidad angular constante.

<sup>14</sup>La energía cinética (2.44), contiene un término independiente de  $\dot{r}$ . Esto se debe a que las ecuaciones de transformación (2.43) de coordenadas cartesianas a coordenadas generalizadas dependen explícitamente del tiempo (ver discusión en la sección 2.3.1).

### 2.8.4. Gauge electromagnético en la formulación Lagrangiana

Los potenciales vectorial  $\mathbf{A}(\mathbf{r}, t)$  y escalar  $\phi(\mathbf{r}, t)$  en electrodinámica, contienen toda la información física de los campos eléctricos y magnéticos. Por otro lado, es bien sabido que el campo electromagnético es invariante ante la transformación gauge

$$\mathbf{A}(\mathbf{r}, t) \rightarrow \mathbf{A}(\mathbf{r}, t) + \nabla\Psi(\mathbf{r}, t) \quad ; \quad \phi(\mathbf{r}, t) \rightarrow \phi(\mathbf{r}, t) - \frac{\partial\Psi(\mathbf{r}, t)}{\partial t} \quad (2.49)$$

donde  $\Psi(\mathbf{r}, t)$  es una función diferenciable en todos sus argumentos, pero por lo demás arbitraria. Por otro lado, ya hemos escrito el Lagrangiano asociado a una partícula cargada  $q$ , inmersa en un campo electromagnético caracterizado por los potenciales  $\mathbf{A}(\mathbf{r}, t)$  y  $\phi(\mathbf{r}, t)$ , Ec. (2.32) Pág. 23

$$L = T - q\phi(\mathbf{r}, t) + q\mathbf{A}(\mathbf{r}, t) \cdot \mathbf{v}(t) \quad (2.50)$$

Queremos ver el efecto que las transformaciones (2.49) tienen sobre este Lagrangiano y sobre las ecuaciones de Lagrange. Con las transformaciones (2.49), el Lagrangiano (2.50) se convierte en

$$\begin{aligned} L' &= T - q\phi'(\mathbf{r}, t) + q\mathbf{A}'(\mathbf{r}, t) \cdot \mathbf{v}(t) = T - q \left[ \phi(\mathbf{r}, t) - \frac{\partial\Psi(\mathbf{r}, t)}{\partial t} \right] + q [\mathbf{A}(\mathbf{r}, t) + \nabla\Psi(\mathbf{r}, t)] \cdot \mathbf{v}(t) \\ L' &= (T - q\phi + q\mathbf{A} \cdot \mathbf{v}) + q \frac{\partial\Psi}{\partial t} + q \mathbf{v} \cdot \nabla\Psi \\ L' &= L + q \left[ \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \Psi(\mathbf{r}, t) \end{aligned}$$

ahora bien, el término en paréntesis cuadrados corresponde a la llamada “derivada convectiva” que es igual a la derivada total en el tiempo teniendo en cuenta que el campo  $\Psi$  **depende solo de las coordenadas y el tiempo**, y no por ejemplo de  $\dot{\mathbf{r}}$  u otras derivadas de orden superior (esto se puede verificar fácilmente con la regla de la cadena). Podemos escribir entonces

$$L' = L + \frac{d}{dt} [q\Psi(\mathbf{r}, t)]$$

pero esta transformación no afecta a las ecuaciones de movimiento como se discutió en la sección 2.3.2, mostrando que las transformaciones (2.49) dejan invariante la Física como era de esperarse.

### 2.8.5. Un sistema ligado de dos masas

Sean dos masas idénticas  $m$  unidas por una barra rígida de peso despreciable y longitud  $l$ . El centro de masa del sistema está restringido a moverse sobre un círculo de radio  $a$  como lo indica la Fig. 2.3a. El campo gravitacional es perpendicular al plano del círculo. Calcularemos el Lagrangiano en un sistema adecuado de coordenadas generalizadas.

Podemos suponer sin pérdida de generalidad que el círculo yace sobre el plano  $XY$ , y el eje  $Z$  va en sentido contrario al campo gravitacional. De la Fig. 2.3a, es claro que la rapidez del centro de masa es  $v_C = |a\dot{\psi}|$  y por tanto la energía cinética del centro de masa es

$$T_{CM} = \frac{1}{2} (m + m) v_C^2 = ma^2 \dot{\psi}^2$$

La energía cinética del sistema vista por el laboratorio, es la suma de la energía cinética del centro de masa mas la energía cinética con respecto al centro de masa, como se aprecia en la Ec. (1.32) Pág 13. Para encontrar la energía cinética del sistema con respecto al CM, recurrimos a la Fig. 2.3b, en donde describimos la posición de una de las partículas en un sistema coordenado con origen en el CM, tal que los ejes  $X_C Y_C Z_C$  son instantáneamente paralelos a los ejes  $XYZ$ . En el CM, los momentos de ambas partículas son opuestos y como las masas son iguales, las velocidades son iguales y opuestas, la energía cinética vista por el CM es

$$T' = \frac{1}{2} m v'^2 + \frac{1}{2} m v'^2 = m v'^2$$

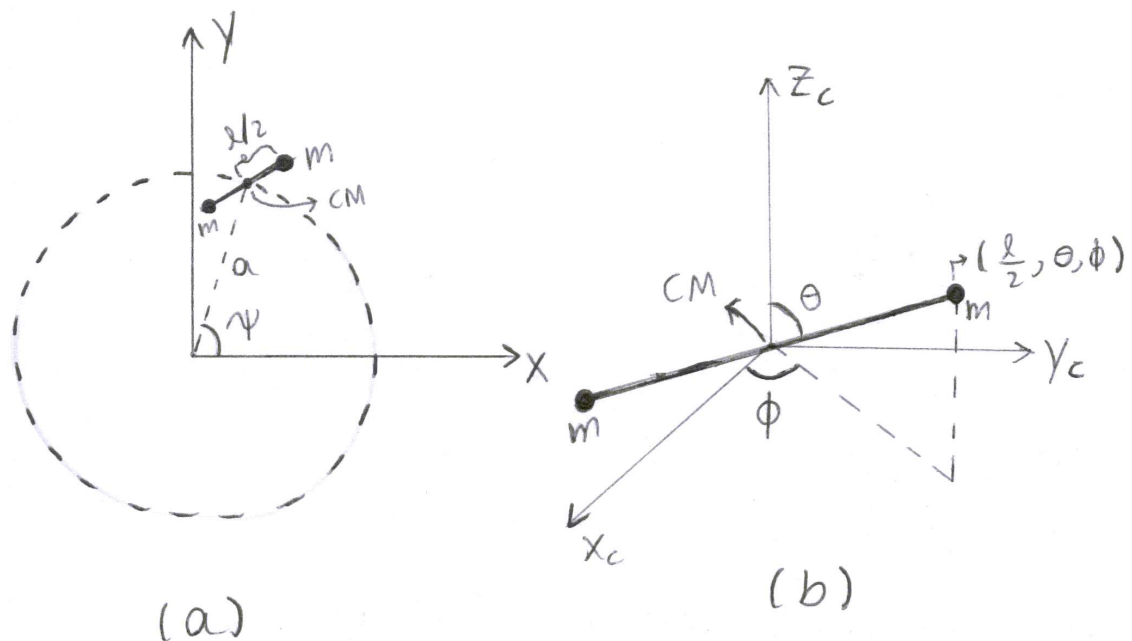


Figura 2.3: (a) Sistema de dos masas unidas por una barra sin masa. El círculo esta sobre el plano  $XY$  y por tanto, también el centro de masa. Aunque cada masa puede tener una componente  $Z$ . (b) Sistema de coordenadas esféricas para determinar la posición de una de las masas con respecto al centro de masa.

donde  $v'$  es la rapidez de cada partícula vista por el CM. Construyendo coordenadas esféricas  $r, \theta, \phi$  sobre este nuevo sistema de referencia, es claro que  $r = l/2$  de modo que para la partícula en cuestión, la posición está dada por

$$x = \frac{l}{2} \sin \theta \cos \phi ; \quad y = \frac{l}{2} \sin \theta \sin \phi ; \quad z = \frac{l}{2} \cos \theta$$

de modo que

$$\dot{x} = \frac{l}{2} (\dot{\theta} \cos \phi \cos \theta - \dot{\phi} \sin \theta \sin \phi) ; \quad \dot{y} = \frac{l}{2} (\dot{\theta} \sin \phi \cos \theta + \dot{\phi} \sin \theta \cos \phi) ; \quad \dot{z} = -\frac{l}{2} \dot{\theta} \sin \theta$$

la rapidez al cuadrado de cada partícula vista por el centro de masa es

$$v'^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$$

calculando cada término al cuadrado tenemos

$$\begin{aligned} \dot{x}^2 &= \frac{l^2}{4} \left[ \dot{\theta}^2 \cos^2 \phi \cos^2 \theta - 2\dot{\phi}\dot{\theta} \cos \phi \cos \theta \sin \theta \sin \phi + \dot{\phi}^2 \sin^2 \theta \sin^2 \phi \right] \\ \dot{y}^2 &= \frac{l^2}{4} \left[ \dot{\theta}^2 \sin^2 \phi \cos^2 \theta + 2\dot{\phi}\dot{\theta} \sin \phi \cos \theta \sin \theta \cos \phi + \dot{\phi}^2 \sin^2 \theta \cos^2 \phi \right] \\ \dot{z}^2 &= \frac{l^2}{4} \dot{\theta}^2 \sin^2 \theta \end{aligned}$$

adicionándolos se obtiene

$$v'^2 = \frac{l^2}{4} [\dot{\theta}^2 \cos^2 \theta (\cos^2 \phi + \sin^2 \phi) + \dot{\phi}^2 \sin^2 \theta (\sin^2 \phi + \cos^2 \phi) + \dot{\theta}^2 \sin^2 \theta]$$

$$v'^2 = \frac{l^2}{4} [\dot{\theta}^2 (\cos^2 \theta + \sin^2 \theta) + \dot{\phi}^2 \sin^2 \theta]$$

$$v'^2 = \frac{l^2}{4} (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta)$$

y la energía cinética total queda

$$T = T_{CM} + T' = ma^2 \dot{\psi}^2 + mv'^2$$

$$T = ma^2 \dot{\psi}^2 + m \frac{l^2}{4} (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta)$$

teniendo en cuenta que la componente  $Z$  de cada partícula es igual en  $XYZ$  que en  $X_C Y_C Z_C$  y que en el sistema de referencia del CM las posiciones de las dos partículas son opuestas, tenemos  $V = mgz - mgz = 0$ . De modo que el Lagrangiano coincide con la energía cinética.

Puede verse que  $\psi, \theta, \phi$  son coordenadas generalizadas independientes ya que cada una se puede mover en un desplazamiento virtual, sin modificar las otras coordenadas y sin violar las ligaduras. Podemos verlo también por un conteo de grados de libertad: los seis grados de libertad originales para dos partículas, se pueden traducir en tres grados de libertad de la posición del centro de masa, y tres más del vector posición relativo entre las dos partículas. La ligadura de mover el CM en un círculo, nos lleva a un solo grado de libertad para fijar el CM (el ángulo  $\psi$ ), los tres grados de libertad del vector relativo se reducen a dos por la ligadura de distancia constante que impone la barra.

### 2.8.6. Aro sobre plano inclinado deslizante

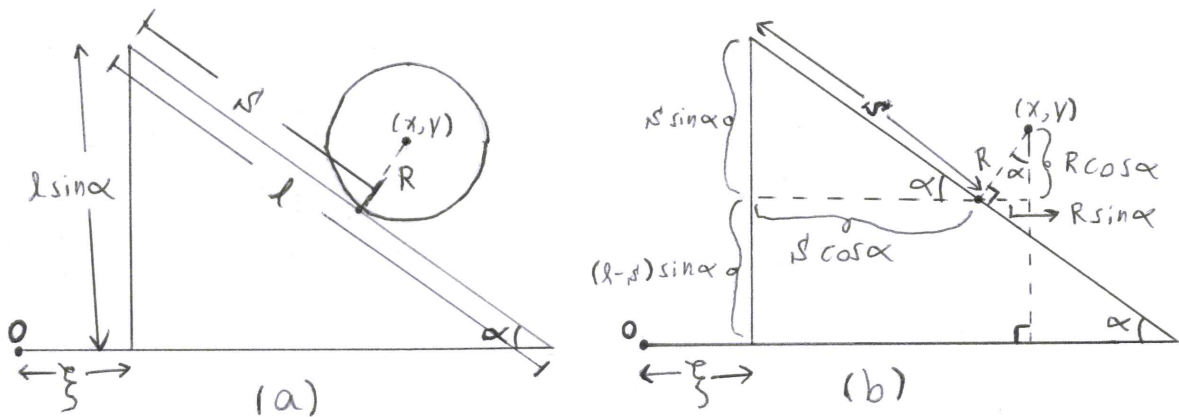


Figura 2.4: (a) Aro que rueda sin deslizar sobre un plano inclinado o cuña que desliza sobre el suelo sin rozamiento. (b) Ilustración de las coordenadas generalizadas  $(\xi, S)$  y de la geometría básica del problema.

Un aro de masa  $m$  y radio  $R$  rueda sin deslizar sobre un plano inclinado o cuña de masa  $M$  que hace un ángulo  $\alpha$  con la horizontal como se vé en la Fig. 2.4. Encuentre el Lagrangiano y las ecuaciones de Lagrange si el plano inclinado puede deslizar sin fricción a lo largo del suelo.

Tomaremos como coordenadas generalizadas una coordenada  $\xi$  que indica la posición del plano inclinado (desde el origen hasta el vértice en ángulo recto del plano inclinado), y una coordenada  $S$  medida desde el vértice superior de la cuña, hasta el punto de contacto del aro con la cuña. La longitud total de la cuña se



denota por  $l$ , y las coordenadas del centro del aro con respecto al origen serán  $x, y$ . La Fig. 2.4 nos muestra que estas coordenadas vienen dadas por

$$x = \xi + S \cos \alpha + R \sin \alpha \quad ; \quad y = R \cos \alpha + (l - S) \sin \alpha$$

usando estas coordenadas, teniendo en cuenta que el momento de inercia de un aro es  $I = mR^2$  y que  $S = R\phi$  por la condición de rodadura, la energía cinética del aro quedará en la forma

$$T_{hoop} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I\dot{\phi}^2 = \frac{1}{2}m \left[ (\dot{\xi} + \dot{S} \cos \alpha)^2 + (-\dot{S} \sin \alpha)^2 \right] + \frac{1}{2}(mR^2) \frac{\dot{S}^2}{R^2}$$

de modo que

$$\begin{aligned} T_{hoop} &= \frac{1}{2}m \left[ \dot{\xi}^2 + \dot{S}^2 \cos^2 \alpha + 2\dot{\xi}\dot{S} \cos \alpha + \dot{S}^2 \sin^2 \alpha \right] + \frac{1}{2}m\dot{S}^2 \\ &= \frac{1}{2}m \left[ 2\dot{S}^2 + \dot{\xi}^2 + 2\dot{\xi}\dot{S} \cos \alpha \right] \end{aligned}$$

pero la energía cinética total debe incluir el movimiento translacional de la cuña a lo largo de  $x$

$$T_{plane} = \frac{1}{2}M\dot{\xi}^2$$

la energía total es entonces

$$T = T_{hoop} + T_{plane} = m\dot{S}^2 + \frac{1}{2}(m + M)\dot{\xi}^2 + m\dot{\xi}\dot{S} \cos \alpha$$

y la energía potencial nos da

$$V = mgy = mg[R \cos \alpha + (l - S) \sin \alpha]$$

no es necesario incluir la energía potencial asociada a la cuña ya que ésta no cambia. El Lagrangiano queda entonces

$$L = m\dot{S}^2 + \frac{1}{2}(m + M)\dot{\xi}^2 + m\dot{\xi}\dot{S} \cos \alpha - mg[R \cos \alpha + (l - S) \sin \alpha] \quad (2.51)$$

De modo que

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{S}} \right) &= \frac{d}{dt} (2m\dot{S} + m\dot{\xi} \cos \alpha) = 2m\ddot{S} + m\ddot{\xi} \cos \alpha \quad ; \quad \frac{\partial L}{\partial S} = mg \sin \alpha \\ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\xi}} \right) &= \frac{d}{dt} [(m + M)\dot{\xi} + m\dot{S} \cos \alpha] = (m + M)\ddot{\xi} + m\ddot{S} \cos \alpha \quad ; \quad \frac{\partial L}{\partial \xi} = 0 \end{aligned}$$

con lo cual se obtienen las ecuaciones de Lagrange para  $S, \xi$

$$2m\ddot{S} + m\ddot{\xi} \cos \alpha - mg \sin \alpha = 0 \quad (2.52)$$

$$(m + M)\ddot{\xi} + m\ddot{S} \cos \alpha = 0 \quad (2.53)$$

Es importante tener en cuenta que la energía cinética translacional del aro no se puede escribir en la forma  $\frac{1}{2}m\dot{S}^2$ , ya que  $S$  está medida con respecto al vértice superior de la cuña, que no define un sistema inercial. Esto debido al movimiento (en general acelerado) de la cuña con respecto al origen fijo en el suelo (que sí se supone inercial).

### 2.8.7. Un potencial generalizado para una fuerza central

Supongamos que una partícula se mueve en un plano bajo la influencia de una fuerza central dada por

$$\mathbf{F} = \frac{1}{r^2} \left( 1 - \frac{\dot{r}^2 - 2\ddot{r}r}{c^2} \right) \mathbf{u}_r \quad (2.54)$$

donde  $r$  es la distancia de la partícula al centro de fuerzas y  $c$  la velocidad de la luz en el vacío. Esta expresión representa la fuerza entre dos cargas en la electrodinámica de Weber. Queremos plantear un Lagrangiano asociado a este sistema.

La partícula se mueve en un plano, pero no tiene ligaduras dentro del plano. Lo natural es entonces utilizar coordenadas generalizadas polares  $r$  y  $\theta$ . Ya calculamos las expresiones para las fuerzas generalizadas de una partícula no ligada en coordenadas polares Ecs. (2.41, 2.42) Pág. 27. Puesto que esta fuerza es central tenemos que  $Q_\theta = rF_\theta = 0$  y  $Q_r = F_r = F$ .

Se puede verificar que el potencial dado por

$$U(r, \dot{r}) = \frac{1}{r} + \frac{\dot{r}^2}{c^2 r} \quad (2.55)$$

es un potencial generalizado válido para reproducir el valor de la fuerza generalizada  $Q_r = F$ , donde  $F$  es dado en (2.54). Para verlo partimos de las expresiones (2.26, 2.27) para el potencial y la fuerza generalizados, de modo que la fuerza generalizada asociada al potencial en (2.55) es

$$\begin{aligned} Q_r &= F_r = -\frac{\partial U}{\partial r} + \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{r}} \right) \\ \frac{\partial U}{\partial r} &= -\frac{1}{r^2} - \frac{\dot{r}^2}{c^2 r^2} ; \quad \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{r}} \right) = \frac{d}{dt} \left( \frac{2\dot{r}}{c^2 r} \right) = \frac{2\dot{r}}{c^2} \left( -\frac{\dot{r}}{r^2} \right) + \frac{2\ddot{r}}{c^2 r} = -\frac{2\dot{r}^2}{c^2 r^2} + \frac{2\ddot{r}}{c^2 r} \end{aligned} \quad (2.56)$$

con lo cual

$$-\frac{\partial U}{\partial r} + \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{r}} \right) = \frac{1}{r^2} + \frac{\dot{r}^2}{c^2 r^2} - \frac{2\dot{r}^2}{c^2 r^2} + \frac{2\ddot{r}}{c^2 r} = \frac{1}{r^2} - \frac{\dot{r}^2}{c^2 r^2} + \frac{2\ddot{r}}{c^2 r} = \frac{1}{r^2} \left( 1 - \frac{\dot{r}^2 - 2\ddot{r}r}{c^2} \right)$$

que reproduce la expresión (2.54) de la fuerza. Para  $q_j \equiv \theta$ , también se cumple la condición (2.27) obteniéndose cero en ambos miembros. Pues no debe perderse de vista que  $U$  debe reproducir todas las fuerzas generalizadas asociadas al problema.

Hay dos puntos que vale la pena enfatizar: (a) Los potenciales generalizados deben reproducir la fuerza generalizada y no la fuerza real (aunque en este caso ambas coinciden en magnitud). (b) El potencial generalizado no tiene porqué ser único, ya que solo se busca una solución a la ecuación diferencial (2.27), sin condiciones iniciales ni de frontera. La no unicidad del potencial generalizado es de esperarse, puesto que incluso el potencial "tradicional" no es único. Por supuesto, el potencial generalizado tiene un gauge más complejo, ya que si redefinimos el potencial generalizado en la forma

$$U'(q, \dot{q}, t) = U(q, \dot{q}, t) + W(q, \dot{q}, t) \quad (2.57)$$

en donde  $W(q, \dot{q}, t)$  satisface la ecuación

$$\frac{d}{dt} \left( \frac{\partial W(q, \dot{q}, t)}{\partial \dot{q}_j} \right) - \frac{\partial W(q, \dot{q}, t)}{\partial q_j} = 0 \quad \forall q_j \text{ del sistema} \quad (2.58)$$

es claro que  $U'(q, \dot{q}, t)$  reproduce las mismas fuerzas generalizadas que  $U(q, \dot{q}, t)$ . En particular  $W = cte$  es un gauge posible.

Usando la energía cinética (2.40) para una partícula en coordenadas polares y el potencial generalizado (2.55) el Lagrangiano queda

$$L = T - U = \frac{1}{2}m \left( \dot{r}^2 + r^2\dot{\theta}^2 \right) - \frac{1}{r} \left( 1 + \frac{\dot{r}^2}{c^2} \right)$$

¿puede el lector encontrar una forma sistemática de encontrar un potencial generalizado, por medio de la ecuación (2.56) y la expresión (2.54) para la fuerza?.

### 2.8.8. Partícula inmersa en un fluido

Sea una partícula de masa  $m$  que cae verticalmente bajo la influencia de la gravedad y de una fuerza viscosa de la forma  $\mathbf{F} = -k\mathbf{v}$ , debida a su inmersión en un fluido (el aire), donde  $\mathbf{v}$  es la velocidad instantánea de la partícula. Plantearemos las ecuaciones de movimiento para esta partícula a partir de su Lagrangiano y de la función de disipación de Rayleigh. El problema es unidimensional no ligado, de modo que el Lagrangiano es simplemente

$$L = \frac{1}{2}m\dot{z}^2 - mgz$$

Pero debido a la presencia de una fuerza de fricción, el Lagrangiano no contiene toda la información Física del sistema. Es necesario incorporar la función de disipación de Rayleigh. La ecuación de movimiento viene dada por (2.38)

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{z}} \right) - \frac{\partial L}{\partial z} + \frac{\partial F}{\partial \dot{z}} &= 0 \\ m\ddot{z} + mg + k_z \dot{z} &= 0 \end{aligned} \quad (2.59)$$

donde hemos usado la función de disipación  $F$  de la Ec. (2.37), para una partícula. Puesto que la fuerza viscosa aumenta con la rapidez, llega un momento en el cual la fuerza viscosa se anula con el peso (al menos si la caída dura el suficiente tiempo). Esta condición de cancelación nos llevará entonces a la velocidad terminal de la partícula

$$mg + k_z \dot{z}_{term} = 0 \Rightarrow \dot{z}_{term} = -\frac{mg}{k_z} \quad (2.60)$$

cuando la partícula alcanza esta velocidad, ésta ya no experimenta una aceleración y continúa con esta velocidad hasta que haga contacto con el piso. Esto se aprecia combinando las Ecs. (2.59, 2.60) para obtener  $\ddot{z}_{term} = 0$ . La Ec. (2.59) se puede resolver para todo tiempo haciendo el cambio de variable

$$Z = \dot{z} + \frac{mg}{k_z} ; \quad \dot{Z} = \ddot{z}$$

con lo cual la Ec. (2.59) queda

$$\ddot{z} + g + \frac{k_z}{m}\dot{z} = 0 \Rightarrow \dot{Z} + \frac{k_z}{m}Z = 0 \Rightarrow \frac{\dot{Z}}{Z} = -\frac{k_z}{m}$$

cuya solución es

$$\begin{aligned} \ln Z &= -\frac{k_z}{m}t + B \Rightarrow Z = Ce^{-\frac{k_z}{m}t} \Rightarrow \\ \dot{z} + \frac{mg}{k_z} &= Ce^{-\frac{k_z}{m}t} \end{aligned} \quad (2.61)$$

tomando la condición inicial  $\dot{z}(t=0) = v_0$ , en esta ecuación se tiene que

$$v_0 + \frac{mg}{k_z} = C$$

con lo cual la ecuación (2.61) queda

$$\dot{z} = \left( \frac{mg}{k_z} + v_0 \right) e^{-\frac{k_z}{m}t} - \frac{mg}{k_z} \quad (2.62)$$

la velocidad terminal se puede obtener haciendo  $t \rightarrow \infty$ , con lo cual se reproduce adecuadamente el valor dado en la Ec. (2.60). Nótese que por los dos caminos seguidos, es claro que la velocidad terminal es independiente del valor (y del signo) de  $v_0$ . En particular, si  $|v_0| > |\dot{z}_{term}|$  y la partícula va inicialmente hacia abajo ( $v_0 < 0$ ), la fuerza viscosa disminuye la rapidez de la partícula hasta que alcanza el valor de la velocidad terminal, esto debido a que inicialmente la fuerza viscosa supera en magnitud al peso. El lector puede obtener fácilmente  $z(t)$  integrando (2.62).

## 2.9. Ventajas del formalismo Lagrangiano

El Lagrangiano es invariante ante un cambio en las coordenadas generalizadas y es mucho más conveniente para definir la mecánica o en general la Física de sistemas continuos. Por otro lado, diferentes sistemas físicos (incluso no necesariamente mecánicos) pueden exhibir Lagrangianos semejantes<sup>15</sup>. Tomemos el ejemplo del siguiente Lagrangiano

$$L = \frac{1}{2} \sum_j L_j \dot{q}_j^2 + \frac{1}{2} \sum_{j \neq k} M_{jk} \dot{q}_j \dot{q}_k - \sum_j \frac{q_j^2}{2C_j} + \sum_j V_j(t) q_j \quad (2.63)$$

con la siguiente función de disipación

$$F = \frac{1}{2} \sum_j R_j \dot{q}_j^2 \quad (2.64)$$

Se puede demostrar que el Lagrangiano (2.63) y la función de disipación (2.64) junto con las ecuaciones de Lagrange con término disipativo Ecs. (2.38), reproducen las ecuaciones diferenciales que describen a un conjunto de circuitos RLC con fuentes, y acoplados a través de las inductancias mutuas  $M_{jk}$ . Los valores  $L_j$ ,  $C_j$ ,  $V_j$  denotan inductancias, capacitancias y voltajes de fuentes respectivamente.  $q_j$  denota las cargas eléctricas que en este caso han sido tomadas como las variables dinámicas (y por tanto, como las coordenadas generalizadas), este no es un ejemplo mecánico en el sentido de que las cargas no denotan de ningún modo posiciones de las partículas o subsistemas. No obstante, el sistema tiene un análogo mecánico: la inductancia es una medida de la resistencia al cambio en el flujo de carga, de modo que es un término de inercia, caso similar el de las inductancias mutuas<sup>16</sup>. El capacitor actúa como una fuente de energía potencial del tipo oscilador armónico simple<sup>17</sup>, la resistencia proporciona un término disipativo tipo Stokes (proporcional a la velocidad generalizada), y finalmente la fuerza electromotriz equivale a un forzamiento externo que se traduce en una energía potencial de la forma  $q_j V_j$  (siendo  $V_j$  el voltaje en un instante).

Finalmente, cabe resaltar que en el caso de ligaduras holónomas, en el cual es posible encontrar un conjunto mínimo de coordenadas independientes, el sistema de ecuaciones Lagrangianas se reduce con respecto al formalismo Newtoniano. Esto se debe a que cada ecuación de movimiento está asociada a una coordenada y no a una partícula, y claramente el número de coordenadas independientes es menor o igual que  $3N$ , siendo  $N$  el número de partículas.

## 2.10. Ejercicios

1. Tres masas se acoplan por medio de cuerdas en serie, para formar un péndulo triple similar al de la Fig. 2.1. (a) Encuentre un conjunto apropiado de coordenadas generalizadas independientes. (b) Describa un desplazamiento virtual para cada coordenada, de modo que las otras se mantengan fijas. (c) Plantee el Lagrangiano asociado. Puede asumir que todas las masas se mueven en un plano.
2. Sea un sistema de referencia inercial  $S$  y un plano inclinado que está en reposo con respecto a  $S$ . Sobre el plano inclinado desliza un bloque de masa  $m$ . Tomemos ahora otro sistema de referencia inercial  $S'$  que se mueve a una velocidad constante  $\mathbf{v} = -v\mathbf{u}_x$  con respecto a  $S$ . (a) Demuestre que la fuerza normal del plano inclinado sobre el bloque realiza trabajo (real) sobre el bloque, visto por el sistema de referencia  $S'$ . (b) Demuestre que el trabajo virtual de la normal sobre el bloque es cero, tanto en  $S$  como en  $S'$ . (Para más comentarios ver Refs. [5, 6]).

<sup>15</sup> Hay que aclarar sin embargo que esta característica no solo la presenta la formulación Lagrangiana, en realidad las ecuaciones diferenciales en la mecánica Newtoniana ya presentan estas analogías.

<sup>16</sup> La suma de los términos con autoinductancias e inductancias mutuas dan una “energía cinética” que solo depende cuadráticamente de las coordenadas generalizadas. Lo cual equivale en el caso mecánico a escenarios con ligaduras esclerónomas (ver sección 2.3.1, Ec. 2.24).

<sup>17</sup> Efectivamente, se requiere una cierta energía para cargar el condensador, esta energía queda almacenada y se puede intercambiar en otra forma de energía.

3. Demuestre que las ecuaciones de Lagrange del tipo dado en (2.17) se pueden escribir equivalentemente en la forma

$$\frac{\partial \dot{T}}{\partial \dot{q}_j} - 2 \frac{\partial T}{\partial q_j} = Q_j$$

usualmente conocidas como ecuaciones de Nielsen.

4. Encuentre el Lagrangiano y las ecuaciones de Lagrange de un péndulo esférico. Esto es, una masa puntual en un campo gravitacional, que se mueve de modo que su distancia a un punto fijo permanece constante (debido por ejemplo, a una varilla rígida sin masa). Plantee un Lagrangiano para pequeñas oscilaciones en las coordenadas generalizadas apropiadas.
5. Obtenga el Lagrangiano y las ecuaciones de movimiento del péndulo doble ilustrado en la Fig. 2.1 Pág. 16, asumiendo que el movimiento se realiza en un plano. Plantee el Lagrangiano en una aproximación de pequeñas oscilaciones para ambas coordenadas generalizadas  $\theta_1$  y  $\theta_2$ .
6. Supongamos una masa  $m_1$  que se mueve con movimiento circular uniforme, debido a que está atada a una cuerda de longitud constante. Por otro lado, una partícula sometida a una fuerza central atractiva de la forma  $\mathbf{F} = -kr^{-2}\mathbf{u}_r$ , puede realizar un movimiento circular uniforme con las condiciones iniciales apropiadas (ver capítulo 10). No obstante, la fuerza que mantiene la distancia constante en el primer problema (tensión de la cuerda) se considera una fuerza de ligadura. En contraste, aunque la fuerza  $\mathbf{F} = -kr^{-2}\mathbf{u}_r$  mantenga constante la distancia de la partícula al centro de fuerzas, ésta NO se trata como fuerza de ligadura. ¿En que consiste la diferencia?
7. Encuentre un potencial generalizado  $U'$  que difiera de manera no trivial del potencial  $U$  en la Ec. (2.55) y que genere las mismas fuerzas generalizadas.

## Suplemento matemático: cálculo de variaciones y multiplicadores de Lagrange

Existen muchos problemas de minimización o maximización que requieren de extender las herramientas tradicionales del cálculo. Es frecuente encontrar problemas en los cuales una trayectoria completa se mapea en un número real y queremos encontrar la trayectoria que minimiza o maximiza dicho número. En el presente capítulo estudiaremos el cálculo variacional, como herramienta para resolver esta clase de situaciones. Por otra parte, cuando un proceso de optimización está sujeto a ligaduras, el método de multiplicadores de Lagrange es una herramienta sistemática para encontrar los extremos de una función sin violar las ligaduras impuestas. Veremos más adelante que la combinación de las dos herramientas nos generará una formulación lagrangiana con la capacidad de incluir y resolver las fuerzas de ligadura.

### 3.1. Algunos problemas prácticos de naturaleza variacional

Vamos a enunciar algunos problemas de índole práctica que no se pueden resolver con los métodos tradicionales del cálculo. Aunque el principal objetivo del desarrollo del cálculo variacional en este texto será la discusión del principio variacional de Hamilton enunciado en el capítulo 4, esta herramienta matemática tiene muchas otras aplicaciones de modo que lo desarrollaremos en forma independiente al tema central.

#### 3.1.1. Minimización del tiempo de caída de una partícula

Consideremos una partícula que se mueve en un campo gravitacional constante y que partiendo del reposo viaja desde un punto  $(x_1, y_1)$  hasta otro punto más bajo  $(x_2, y_2)$  con la condición de que los puntos no se encuentran sobre la misma vertical. Encuentre la curva que permita que la partícula haga este recorrido en el menor tiempo posible.

**Planteamiento:** Tomemos a  $X$  como el eje vertical hacia abajo y a  $Y$  como el eje horizontal a la derecha. Por simplicidad hacemos coincidir al punto  $(x_1, y_1)$  con el origen. Tomando el cero de potencial en el origen, y teniendo en cuenta que la partícula parte del reposo, se tiene que la energía mecánica de la partícula viene dada por

$$T + V = \frac{1}{2}mv^2 - mgx = 0$$

de modo que la rapidez está dada por  $v = \sqrt{2gx}$ , el tiempo de recorrido es

$$t = \int dt = \int_{(x_1, y_1)}^{(x_2, y_2)} \frac{ds}{v} = \int_{(x_1, y_1)}^{(x_2, y_2)} \frac{\sqrt{(dx)^2 + (dy)^2}}{\sqrt{2gx}} = \int_{(x_1, y_1)}^{(x_2, y_2)} \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2gx}} dx$$

donde  $ds$  denota una longitud de arco infinitesimal de la trayectoria, la expresión queda finalmente

$$t = \int_{x_1=0}^{x_2} \sqrt{\frac{1+y^2}{2gx}} dx \quad ; \quad \dot{y}(x) \equiv \frac{dy}{dx} \quad (3.1)$$

por tanto será necesario encontrar el valor de  $y(x)$  que hace que la integral (3.1) nos dé el menor valor posible. En cálculo ordinario los problemas de minimización usualmente consisten en encontrar el valor de un punto (por ejemplo sobre la recta real) de modo que una cierta cantidad (función) nos de un valor mínimo al evaluarla en ese punto. En contraste, aquí requerimos encontrar una trayectoria completa que haga que cierta cantidad sea mínima, puesto que lo que tenemos en este problema es una trayectoria completa que se mapea en un número (el tiempo).

### 3.1.2. Minimización de una superficie de revolución

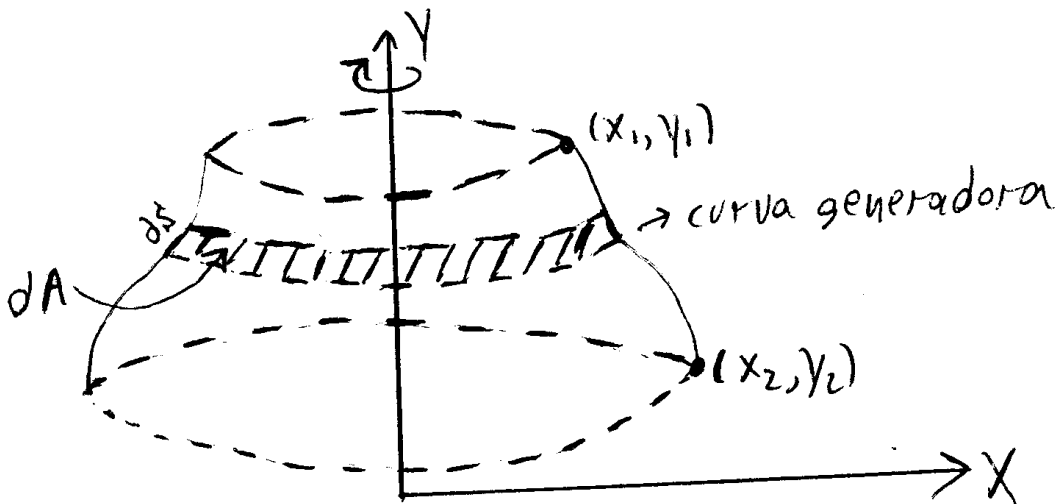


Figura 3.1: Sólido de revolución generado alrededor de  $Y$ .  $(x_1, y_1)$  y  $(x_2, y_2)$  son puntos fijos y se busca la curva generadora que pase por estos puntos y que minimice la superficie lateral del sólido.

Consideremos la superficie lateral generada por una curva que une dos puntos fijos  $(x_1, y_1)$  y  $(x_2, y_2)$  cuando se revoluciona alrededor de un eje coplanar con los puntos y la curva (ver Fig 3.1). La idea es encontrar la ecuación de la curva en cuestión a fin de que el área lateral generada por este sólido de revolución sea mínima.

**Planteamiento:** Asumiremos que la curva con extremos fijos se revoluciona alrededor del eje  $Y$ . Para calcular el área lateral total de revolución, calculamos primero el diferencial de área  $dA$  sobre una pequeña tira como se muestra en la figura 3.1. Dicho diferencial viene dado por

$$dA = 2\pi x ds = 2\pi x \sqrt{(dx)^2 + (dy)^2} = 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

nuevamente  $ds$  es la longitud infinitesimal de arco asociada a la curva. La expresión final para el área queda

$$A = 2\pi \int_{x_1}^{x_2} x \sqrt{1 + \dot{y}^2} dx \quad ; \quad \dot{y} \equiv \frac{dy}{dx} \quad (3.2)$$

de nuevo la minimización del área consiste en encontrar una trayectoria completa que minimice tal cantidad.

### 3.2. Aspectos fundamentales del cálculo de variaciones

Un rasgo general de los ejemplos hasta aquí propuestos es que es necesario resolver el problema de como hallar la trayectoria que hace que el valor de una cierta cantidad (tiempo de caída, superficie etc.) sea estacionario (mínimo, máximo, punto de inflexión etc.). Este es uno de los problemas fundamentales del cálculo variacional. Si denotamos por  $J$  a la cantidad genérica (tiempo de caída, superficie, etc.) que queremos minimizar podemos ver que las Ecs. (3.1, 3.2) tienen la siguiente estructura genérica

$$J = \int_{x_1}^{x_2} f(y, \dot{y}, x) dx \quad ; \quad y = y(x), \quad \dot{y} \equiv \frac{dy}{dx} \quad (3.3)$$

a la cantidad  $y$  la llamaremos coordenada generalizada y la variable independiente  $x$  será considerada un parámetro que modula la forma de la trayectoria  $y(x)$  y de su derivada  $\dot{y}(x)$ . Para propósitos futuros es útil considerar el caso en el cual tenemos más de una coordenada generalizada, de modo que la relación (3.3) se extiende en la forma

$$J = \int_{x_1}^{x_2} f(y_1(x), \dots, y_n(x), \dot{y}_1(x), \dots, \dot{y}_n(x), x) dx \quad (3.4)$$

donde las  $y_i$  son coordenadas generalizadas independientes entre sí y solo pueden depender del parámetro  $x$ . Definiremos ahora un **espacio de configuraciones** que se construye asignando un eje a cada coordenada generalizada independiente  $y_i$  de modo que tenemos un espacio cartesiano  $n$  dimensional, es necesario enfatizar que **NO** se asigna un eje asociado al parámetro  $x$ . Para un valor fijo de  $x$  una configuración completa de coordenadas  $y_i$  corresponde a una  $n$ -upla y por tanto a un punto en el espacio de configuraciones. Cuando se mueve el parámetro  $x$  cambia la posición del punto en el espacio de configuraciones trazando una trayectoria en dicho espacio. Si queremos minimizar o maximizar  $J$  debemos encontrar el conjunto de trayectorias unidimensionales  $y_i(x)$  que hagan que la integral (3.4) sea mínima o máxima. Nótese que en el lenguaje del espacio de configuraciones, el problema de minimizar o maximizar  $J$  se convierte en el problema de encontrar la trayectoria en el espacio de configuraciones que minimice tal cantidad. Esta visión del problema resulta útil en la medida que nos permite visualizar una única trayectoria de minimización o maximización en lugar de muchas trayectorias independientes.

Replanteando el problema de la minimización de  $J$ , tal minimización significa que entre todas las trayectorias posibles que el sistema puede seguir en el espacio de configuraciones al barrer el parámetro entre los extremos  $x_1$  y  $x_2$ , el sistema trazará la trayectoria que haga que el valor de la integral (3.4) sea estacionario (estrictamente no sabemos aún si es mínimo máximo o punto de inflexión). Para una integral de línea, el término estacionario significa que el valor de la integral cuando se toma un cierto camino, tiene el mismo valor dentro de infinitesimales de primer orden que cuando se toman caminos vecinos a éste (caminos que difieren del original por desplazamientos generalizados infinitesimales). Recordemos que las trayectorias de las que hablamos aquí, son hipertrayectorias en el espacio  $n$  dimensional de configuraciones. Para integrales de línea este es el equivalente en funciones ordinarias a puntos con derivada cero. En este caso en lugar de un punto tenemos una trayectoria entera.

#### 3.2.1. Cálculo variacional en una dimensión

Trabajaremos el problema inicialmente en una dimensión. Sea una función  $f(y, \dot{y}, x)$  definida sobre un camino  $y = y(x)$  donde  $x$  actúa como un parámetro, y  $y$  es una coordenada generalizada.  $\dot{y}$  denota la derivada respecto al parámetro  $x$ . Queremos encontrar un camino particular  $y(x)$  de tal modo que la integral de línea  $J$  de la función  $f$  entre  $x_1$  y  $x_2$

$$J = \int_{x_1}^{x_2} f(y, \dot{y}, x) dx \quad ; \quad y = y(x), \quad \dot{y} \equiv \frac{dy}{dx} \quad (3.5)$$

adquiera un valor estacionario relativo a caminos vecinos que difieren infinitesimalmente del original. En nuestro caso, consideraremos solo los caminos que cumplen la condición  $y(x_1) = y_1, y(x_2) = y_2$ , ya que esta condición siempre se impone en el principio de Hamilton. Esto se denomina condición de extremos fijos. Si hacemos



una gráfica de  $y$  vs  $x$ , para varias trayectorias posibles, encontramos que todas ellas deben converger en los extremos de la gráfica, a los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$ . Nótese que la gráfica  $y$  vs  $x$  no representa el espacio de configuraciones, ya que este último no contiene ningún eje que represente al parámetro. En este caso, el espacio de configuraciones es de una sola dimensión, todos los caminos posibles son líneas rectas que conectan a  $y_1$  con  $y_2$ . En consecuencia, lo que distingue a los diferentes caminos es su dependencia con el parámetro  $x$ . Vale decir que la coordenada  $y$  es generalizada y por tanto no tiene necesariamente dimensiones de longitud.

Ahora parametrizaremos el problema de tal manera que se pueda hacer uso del cálculo ordinario. Cuando partimos del camino correcto, la variación de  $J$  debe ser cero con respecto al cambio a una familia de caminos vecinos que rotularemos con alguna variable infinitesimal  $\alpha$ . Denotaremos a las trayectorias vecinas como  $y(x, \alpha)$ , donde  $y(x, 0)$  es el camino correcto. Por ejemplo, si seleccionamos una función arbitraria  $\eta(x)$  pero que cumpla con la condición de que se anula en los extremos i.e. en  $x = x_1$  y  $x = x_2$  una posible familia de caminos vecinos estaría definida por

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x) \quad (3.6)$$

claramente, esta familia de funciones converge en sus extremos gracias a la condición impuesta a  $\eta(x)$  de nulidad en los extremos. Asumiremos de aquí en adelante que el camino correcto  $y(x, 0)$  y la función  $\eta(x)$  son de clase  $C^2$  (contínuas y no singulares hasta la segunda derivada en todos sus argumentos) en el intervalo  $[x_1, x_2]$ . Con cualquier familia de curvas en la vecindad de  $y(x)$ , el variacional  $J$  es función de  $\alpha$

$$J(\alpha) = \int_{x_1}^{x_2} f(y(x, \alpha), \dot{y}(x, \alpha), x) dx \quad (3.7)$$

dado que hemos definido que la curva correcta se encuentre cuando  $\alpha = 0$ , la condición de punto estacionario para  $J$  se puede ahora expresar en términos de cantidades del cálculo ordinario en la forma

$$\left(\frac{dJ}{d\alpha}\right)_{\alpha=0} = 0 \quad (3.8)$$

derivando la Ec. (3.7) bajo el signo integral de la forma usual, tenemos

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial y} \frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}} \frac{\partial \dot{y}}{\partial \alpha} \right) dx \quad (3.9)$$

naturalmente se asume que  $\partial x / \partial \alpha = 0$  ya que  $\alpha$  es un parámetro que me caracteriza cada curva en tanto que  $x$  se mueve a lo largo de cada curva (con  $\alpha$  fijo) claramente los dos son independientes<sup>1</sup>. Veamos la segunda integral en el miembro derecho de la Ec. (3.9)

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \frac{\partial \dot{y}}{\partial \alpha} dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \left[ \frac{\partial}{\partial \alpha} \left( \frac{\partial y}{\partial x} \right) \right] dx = \int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \left[ \frac{\partial}{\partial x} \left( \frac{\partial y}{\partial \alpha} \right) \right] dx \quad (3.10)$$

para integrar por partes elegiremos

$$u = \frac{\partial f}{\partial \dot{y}} \quad ; \quad dv = \left[ \frac{\partial}{\partial x} \left( \frac{\partial y}{\partial \alpha} \right) \right] dx$$

con lo cual  $du = d\left(\frac{\partial f}{\partial \dot{y}}\right) = \left[\frac{d}{dx}\left(\frac{\partial f}{\partial \dot{y}}\right)\right] dx$  puesto que en la integral solo se hace variación en el parámetro  $x$ . Por otro lado,  $v = \left(\frac{\partial y}{\partial \alpha}\right)$ , y al integrar por partes queda

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \left[ \frac{\partial}{\partial x} \left( \frac{\partial y}{\partial \alpha} \right) \right] dx = \frac{\partial f}{\partial \dot{y}} \frac{\partial y}{\partial \alpha} \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] \frac{\partial y}{\partial \alpha} dx \quad (3.11)$$

donde la continuidad hasta segundas derivadas de  $y(x, \alpha)$  asegura que el intercambio de las derivadas parciales es posible. Teniendo en cuenta que todas las curvas deben pasar a través de los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$ , se

<sup>1</sup>En la Ec. (3.6) el término  $\alpha\eta(x)$  que determina la desviación del camino variado con respecto al camino real, tiene toda su dependencia de la variable  $x$  en la función  $\eta(x)$ .

llega a que la derivada parcial de  $y$  con respecto a  $\alpha$  deben ser cero en  $x_1$  y en  $x_2$ . En consecuencia, el primer término de la derecha en (3.11) se anula y por tanto, la integral (3.10) queda

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \frac{\partial \dot{y}}{\partial \alpha} dx = - \int_{x_1}^{x_2} \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] \frac{\partial y}{\partial \alpha} dx \quad (3.12)$$

reemplazando (3.12) en (3.9) se obtiene

$$\frac{dJ}{d\alpha} = \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] \right\} \frac{\partial y}{\partial \alpha} dx \quad (3.13)$$

Combinando (3.8) con (3.13) se vé que la condición para que  $J$  adquiera un valor estacionario es

$$\left( \frac{dJ}{d\alpha} \right)_{\alpha=0} = \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] \right\} \left( \frac{\partial y}{\partial \alpha} \right)_{\alpha=0} dx = 0 \quad (3.14)$$

por otro lado, la expresión  $(\partial y / \partial \alpha)_{\alpha=0}$  es una función arbitraria de  $x$  excepto por exigencias de continuidad hasta la segunda derivada y la condición de extremos fijos. Por ejemplo, para el conjunto particular de familias definidas por (3.6) el factor  $(\partial y / \partial \alpha)_{\alpha=0}$  viene dado por la función arbitraria  $\eta(x)$ . Aplicaremos ahora a la Ec. (3.14) el lema fundamental del cálculo de variaciones, según el cual si se cumple

$$\int_{x_1}^{x_2} M(x) \eta(x) dx = 0 \quad (3.15)$$

para todas las funciones arbitrarias  $\eta(x)$  continuas hasta la segunda derivada, entonces  $M(x)$  debe ser idénticamente cero en el intervalo  $(x_1, x_2)$ . La demostración formal se puede encontrar en los textos de cálculo variacional pero se puede dar una visión eurística de este lema: Imaginemos que construimos una función  $\eta(x)$  que es positiva en una vecindad de un cierto punto y cero en las otras regiones, en consecuencia la integral (3.15) solo será válida si  $M$  es cero en este punto arbitrariamente escogido. Por tanto,  $M$  debe ser cero en todo el intervalo de integración. Usando este lema para la expresión (3.14) se obtiene

$$\frac{\partial f}{\partial y} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] = 0 \quad (3.16)$$

esta ecuación diferencial nos da entonces como solución la trayectoria  $y = y(x)$  que deja estacionaria la cantidad  $J$ . Ahora bien, la desviación infinitesimal de un camino dado con respecto al camino correcto  $y(x, 0)$  en el punto  $x$  viene dado por

$$\left( \frac{\partial y(x, \alpha)}{\partial \alpha} \right)_{\alpha=0} d\alpha \equiv \delta y(x) \quad (3.17)$$

como en esta variación el parámetro  $x$  es fijo, el desplazamiento anterior corresponde a los desplazamientos virtuales de las coordenadas generalizadas discutidos en la formulación de Lagrange (en la formulación de Lagrange, el parámetro es el tiempo y permanece fijo cuando se hace un desplazamiento virtual de la coordenada generalizada) por eso la notación  $\delta y$ . Por otro lado, la variación de  $J$  con respecto a su valor cuando se toma el camino correcto es

$$\left( \frac{dJ}{d\alpha} \right)_{\alpha=0} d\alpha \equiv \delta J \quad (3.18)$$

la estacionaridad del variacional  $J$  cuando se evalúa sobre el camino correcto se manifiesta entonces como  $\delta J = 0$ . Multiplicando la Ec. (3.14) por  $d\alpha$  y usando (3.17) y (3.18) resulta

$$\delta J = \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \right] \right\} \delta y dx = 0 \quad (3.19)$$

requiriendo que  $y(x)$  satisfaga la ecuación (3.16). La variación de una integral de línea usualmente se denota como  $\delta$  para diferenciarla de la variación de una función ordinaria. La Ec. (3.19) no contiene información nueva con respecto a (3.14) pero resulta más adecuada para una extensión del formalismo cuando tenemos  $n$  coordenadas generalizadas como veremos a continuación.

### 3.2.2. Cálculo de variaciones multidimensional

Veamos la extensión al caso de  $n$  coordenadas generalizadas. El variacional lo escribimos como

$$J = \int_1^2 f(y_1(x), \dots, y_n(x), \dot{y}_1(x), \dots, \dot{y}_n(x), x) dx \quad (3.20)$$

donde todas las variables  $y_i, \dot{y}_i$  son independientes entre sí, y solo pueden depender del parámetro  $x$ . En este caso tenemos una hipertrayectoria correcta y una familia posible de hipertrayectorias vecinas se puede escribir como

$$y_i(x, \alpha) = y_i(x, 0) + \alpha \eta_i(x) \quad ; \quad i = 1, \dots, n \quad (3.21)$$

donde  $y_i(x, 0)$  con  $i = 1, \dots, n$ , denotan la solución (trayectoria correcta en el espacio de configuraciones). Las  $\eta_i(x)$  son independientes entre sí, pero deben anularse en los extremos y ser continuas hasta la segunda derivada, por lo demás son completamente arbitrarias. Un cálculo semejante al anterior nos conduce a

$$\frac{\partial J}{\partial \alpha} d\alpha = \int_1^2 \sum_{i=1}^n \left( \frac{\partial f}{\partial y_i} \frac{\partial y_i}{\partial \alpha} d\alpha + \frac{\partial f}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial \alpha} d\alpha \right) dx$$

de nuevo el segundo término se puede integrar por partes

$$\int_1^2 \frac{\partial f}{\partial \dot{y}_i} \frac{\partial^2 y_i}{\partial \alpha \partial x} dx = \frac{\partial f}{\partial \dot{y}_i} \frac{\partial y_i}{\partial \alpha} \Big|_1^2 - \int_1^2 \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}_i} \right) \right] \frac{\partial y_i}{\partial \alpha} dx \quad (3.22)$$

y una vez más el primer término a la derecha se anula cuando imponemos la condición de extremos fijos. La variación  $\delta J$  con respecto a la trayectoria correcta del sistema, queda entonces

$$\left( \frac{\partial J}{\partial \alpha} \right)_{\alpha=0} d\alpha = \delta J = \sum_{i=1}^n \int_1^2 \left( \frac{\partial f}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}_i} \right) \right] \right) \delta y_i dx \quad (3.23)$$

donde análogamente al caso unidimensional

$$\delta y_i = \left( \frac{\partial y_i}{\partial \alpha} \right)_{\alpha=0} d\alpha \quad (3.24)$$

dado que las coordenadas generalizadas son independientes, sus variaciones  $\delta y_i$  también lo son<sup>2</sup>. Por ejemplo, para la familia de curvas vecinas definidas por la Ec. 3.21, equivale a que los  $\eta_i(x)$  sean independientes. En virtud de la independencia de los  $\delta y_i$ , si la suma sobre  $i$  en la Ec. (3.23) es nula, debe ser nulo cada sumando. Apelando entonces al lema fundamental del cálculo variacional, se llega a que la condición de estacionaridad  $\delta J = 0$  nos lleva a

$$\frac{\partial f}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}_i} \right) \right] = 0 \quad , \quad i = 1, \dots, n \quad (3.25)$$

las soluciones  $y_i(x)$  forman entonces la trayectoria en el espacio de configuraciones que hace de  $J$  un valor estacionario. Estas ecuaciones se conocen como las ecuaciones de Euler-Lagrange. Se pueden estudiar generalizaciones a variacionales en donde por ejemplo la función  $f$  dependa de derivadas superiores de  $y$ , existan varios parámetros  $x_j$  o se relaje la condición de extremos fijos. Para el análisis del principio de Hamilton que se verá en el capítulo 4, las condiciones aquí trabajadas son lo suficientemente generales.

## 3.3. Solución de los problemas de aplicación planteados

Las Ecs. (3.16, 3.25) son generales y aplicables a muchos sistemas incluso no mecánicos. Estamos entonces preparados para resolver los problemas propuestos al principio del capítulo y otros de características similares.

<sup>2</sup>Puesto que las coordenadas generalizadas  $y_i$  son independientes, cada una se puede mover sin mover las otras y sin violar posibles ligaduras. Y puesto que en el proceso  $x$  es fijo, los  $\delta y_i$  corresponden a los desplazamientos virtuales que vimos en el principio de D'Alembert.

### 3.3.1. Minimización de la longitud de arco

Antes de resolver los problemas planteados, veamos un problema de gran simplicidad y valor pedagógico. Dados dos puntos fijos en un plano, encontrar la trayectoria que minimiza la longitud de arco. La longitud de arco diferencial para una curva bidimensional se obtiene fácilmente

$$(ds)^2 = (dx)^2 + (dy)^2 \Rightarrow ds = \sqrt{1 + \dot{y}^2} dx \quad ; \quad \dot{y} \equiv \frac{dy}{dx}$$

la longitud de arco total se escribe entonces como

$$S = \int_{x_1}^{x_2} \sqrt{1 + \dot{y}^2} dx \quad (3.26)$$

comparando (3.5) con (3.26) es claro que

$$f(y, \dot{y}, x) = \sqrt{1 + \dot{y}^2}$$

de las expresiones

$$\frac{\partial f}{\partial y} = 0 \quad ; \quad \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) = \frac{d}{dx} \left( \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} \right)$$

usando la Ec. (3.16) resulta

$$\frac{d}{dx} \left( \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} \right) = 0$$

el término entre paréntesis debe ser constante

$$\left( \frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} \right) = c$$

ahora bien, para que se cumpla esta relación es necesario que a su vez  $\dot{y}$  sea una constante relacionada con  $c$  (aunque esta relación no es relevante), de lo cual se obtiene la ecuación de la trayectoria

$$\dot{y} = a \Rightarrow y = ax + b$$

que corresponde a una línea recta, las constantes de integración  $a$  y  $b$  se determinan con los puntos fijos de la trayectoria. Estrictamente, solo se ha demostrado que la trayectoria da un valor estacionario para la longitud de arco pero sabemos por intuición que corresponde a un mínimo. En realidad se puede encontrar el tipo de extremo a través de la segunda derivada funcional.

### 3.3.2. Minimización del tiempo de caída de una partícula: la braquistócrona

En la sección 3.1.1 planteábamos el problema de una partícula que se mueve en un campo gravitacional constante desde un punto fijo inicial a otro final y se pedía encontrar la curva que minimice el tiempo de caída. El planteamiento del problema nos llevó a la Ec. (3.1)

$$t = \int_0^{x_2} \sqrt{\frac{1 + \dot{y}^2}{2gx}} dx$$

dado que la constante  $\sqrt{2g}$  no afecta el problema, identificamos la función  $f$  como

$$f = \sqrt{\frac{1 + \dot{y}^2}{x}} \quad (3.27)$$

como  $\frac{\partial f}{\partial y} = 0$  la ecuación de Euler Lagrange queda

$$\frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) = 0 \quad \Rightarrow \quad \frac{\partial f}{\partial \dot{y}} = cte \equiv \frac{1}{\sqrt{2a}}$$

la parametrización de la constante de integración se hace por comodidad. Realizando  $\partial f / \partial \dot{y}$  en (3.27) resulta

$$\begin{aligned} \frac{\dot{y}}{\sqrt{x(1+\dot{y}^2)}} &= \frac{1}{\sqrt{2a}} \quad \Rightarrow \quad \frac{\dot{y}^2}{x(1+\dot{y}^2)} = \frac{1}{2a} \quad \Rightarrow \quad 2a\dot{y}^2 = x(1+\dot{y}^2) \\ \dot{y}^2 &= \frac{x}{(2a-x)} = \frac{x^2}{(2ax-x^2)} \end{aligned}$$

que se puede escribir en la forma

$$y = \int \frac{x \, dx}{\sqrt{2ax-x^2}}$$

haciendo el cambio de variable

$$x = a(1 - \cos \theta) \quad ; \quad dx = a \sin \theta \, d\theta \quad (3.28)$$

la integral queda de la forma

$$\begin{aligned} y &= \int \frac{a(1 - \cos \theta)(a \sin \theta \, d\theta)}{\sqrt{2a[a(1 - \cos \theta)] - [a(1 - \cos \theta)]^2}} = \int \frac{a(1 - \cos \theta) \sin \theta \, d\theta}{\sqrt{(2 - 2\cos \theta) - (1 - 2\cos \theta + \cos^2 \theta)}} \\ y &= \int \frac{\sin \theta}{\sqrt{1 - \cos^2 \theta}} a(1 - \cos \theta) \, d\theta = \int a(1 - \cos \theta) \, d\theta \end{aligned}$$

cuya solución es

$$y = a(\theta - \sin \theta) + b \quad (3.29)$$

Con  $\theta = 0$ , las Ecs. (3.28, 3.29) nos dan  $x = 0$  y  $y = b$ . Si el punto inicial coincide con el origen entonces  $y = 0$  cuando  $x = 0$ . Por tanto  $b = 0$ , y la constante  $a$  se ajusta para que la curva pase por el punto  $(x_2, y_2)$ . Las ecuaciones (3.28, 3.29) para  $x$  e  $y$  quedan entonces

$$x = a(1 - \cos \theta) \quad ; \quad y = a(\theta - \sin \theta)$$

que coinciden con las ecuaciones paramétricas del cicloide, con parámetro  $\theta$ . Este perfil específico del cicloide se denomina una braquistócrona.

### 3.3.3. Minimización de una superficie de revolución

En la sección 3.1.2 tomamos una superficie generada por una trayectoria que une dos puntos fijos y que se revoluciona alrededor de un eje coplanar con los puntos y la curva. La idea era encontrar la ecuación de la trayectoria en cuestión que minimizara el área generada. El planteamiento nos llevó a la Ec. (3.2)

$$A = 2\pi \int_{x_1}^{x_2} x \sqrt{1 + \dot{y}^2} \, dx$$

La función  $f$  vendrá dada por

$$f = x \sqrt{1 + \dot{y}^2}$$

y empleamos las ecuaciones de Euler Lagrange (3.16) para lo cual calculamos

$$\frac{\partial f}{\partial y} = 0 \quad ; \quad \frac{\partial f}{\partial \dot{y}} = \frac{x\dot{y}}{\sqrt{1 + \dot{y}^2}}$$

con lo cual la Ec. (3.16) nos da

$$\begin{aligned} \frac{d}{dx} \left[ \frac{xy}{\sqrt{(1+y^2)}} \right] &= 0 \Rightarrow \frac{xy}{\sqrt{(1+y^2)}} = cte \equiv a \Rightarrow x^2 y^2 = a^2 (1+y^2) \\ &\Rightarrow y^2 (x^2 - a^2) = a^2 \end{aligned}$$

despejando  $y$

$$y = \frac{a}{\sqrt{(x^2 - a^2)}} ; y = \int \frac{a dx}{\sqrt{(x^2 - a^2)}}$$

esta integral nos da

$$y = a \cosh^{-1} \left( \frac{x}{a} \right) + b$$

donde  $a$  y  $b$  son constantes de integración que se determinan requiriendo que la curva pase por los puntos  $(x_1, y_1)$  y  $(x_2, y_2)$ . Esta ecuación se puede invertir para escribir

$$x = a \cosh \left( \frac{y - b}{a} \right)$$

que se reconoce como la ecuación de la catenaria, que es la curva que se forma cuando una cuerda flexible cuelga libremente entre dos puntos fijos en un campo gravitacional uniforme.

Nótese que si definimos  $x' = x/a$  y  $y' = y/a$  obtenemos

$$x' = \cosh (y' - b/a)$$

obtenemos una ecuación con un solo parámetro  $b/a$ . Es decir, un reescalamiento de la curva nos deja un solo parámetro relevante.

Para algunos pares de puntos fijos se puede encontrar un conjunto único de constantes de integración  $a$  y  $b$ . Para otros puntos es posible encontrar dos catenarias a través de los puntos fijos en tanto que en otros casos no hay solución para los valores de  $a$  y  $b$ . Adicionalmente, solo hemos demostrado que la curva deja a la superficie en un valor estacionario, de modo que la catenaria no siempre representa un valor mínimo, en algunos casos representa un punto de inflexión. Para algunas combinaciones de pares de puntos fijos el mínimo absoluto en la superficie de revolución se genera a partir de una curva compuesta de segmentos de línea recta. Esta clase de solución no se puede encontrar con el formalismo presentado aquí, ya que en nuestra formulación las  $f(\alpha, x)$  y las  $\eta(x)$  se consideran continuas hasta la segunda derivada, en tanto que la curva descrita arriba no es derivable en los puntos de “quiebre” entre segmentos.

Las anteriores consideraciones muestran las restricciones que se derivan de nuestra formulación sobre la condición variacional estacionaria.

### 3.4. Ligaduras y multiplicadores de Lagrange (opcional)

A continuación estudiaremos un método sistemático para tratar problemas en los que se pretende encontrar los extremos de una función en donde sus argumentos están sujetos a ligaduras. La técnica de los multiplicadores de Lagrange que ilustraremos aquí, no está necesariamente asociada a problemas variacionales, pero se puede combinar con estos últimos, especialmente cuando tenemos ligaduras integrales.

Supongamos que queremos encontrar el extremo de una función de tres variables independientes  $x_1, x_2, x_3$ ; en tal caso la condición de extremo (o punto de silla) es

$$df = 0 \tag{3.30}$$

es decir

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \tag{3.31}$$

dado que las variables son independientes, podemos variar por ejemplo  $x_1$  sin variar  $x_2$  y  $x_3$  de modo que  $dx_2 = dx_3 = 0$  con lo cual  $\partial_{x_1} f = 0$ . Procediendo idénticamente con las otras variables vemos que la condición necesaria y suficiente para tener un extremo es

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0 \quad (3.32)$$

sin embargo ocurre con frecuencia que las variables que se utilizan a priori para describir al sistema no son independientes. En muchos casos la ligadura se puede escribir en forma de una ecuación de la forma

$$\varphi(x_1, x_2, x_3) = 0 \quad (3.33)$$

en este caso la Ec. (3.32) no necesariamente se cumple ya que las variables no se pueden mover independientemente. Una forma de tratar el problema es despejar una variable de modo que nos quedamos con dos variables independientes y hacemos el tratamiento normal con dos variables independientes. Sin embargo, este despeje puede ser muy difícil en la práctica, por lo cual exploraremos un método alternativo conocido como el método de los multiplicadores indeterminados de Lagrange.

De la Ec. (3.33), es claro que  $d\varphi = 0$ , de modo que se cumple una relación similar a (3.31) que escribiremos en la forma

$$-\frac{\partial \varphi}{\partial x_3} dx_3 = \frac{\partial \varphi}{\partial x_1} dx_1 + \frac{\partial \varphi}{\partial x_2} dx_2$$

con lo cual tenemos

$$\begin{aligned} df &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial_{x_3} f}{\partial_{x_3} \varphi} \frac{\partial \varphi}{\partial x_3} dx_3 \\ &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 - \frac{\partial_{x_3} f}{\partial_{x_3} \varphi} \left( \frac{\partial \varphi}{\partial x_1} dx_1 + \frac{\partial \varphi}{\partial x_2} dx_2 \right) \\ df &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \lambda \left( \frac{\partial \varphi}{\partial x_1} dx_1 + \frac{\partial \varphi}{\partial x_2} dx_2 \right) ; \lambda \equiv -\frac{\partial_{x_3} f}{\partial_{x_3} \varphi} \end{aligned} \quad (3.34)$$

donde hemos supuesto que  $\partial_{x_3} \varphi \neq 0$ . Nótese que con este procedimiento, hemos logrado que la expresión de  $df$  quede en términos solo de los diferenciales independientes  $dx_1$  y  $dx_2$ . Escribamos ahora

$$df + \lambda d\varphi = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 + \lambda \left( \frac{\partial \varphi}{\partial x_1} dx_1 + \frac{\partial \varphi}{\partial x_2} dx_2 + \frac{\partial \varphi}{\partial x_3} dx_3 \right)$$

puesto que  $df = d\varphi = 0$  tenemos que  $df + \lambda d\varphi = 0$ . Teniendo en cuenta esto y redistribuyendo términos, se tiene

$$\left( \frac{\partial f}{\partial x_1} + \lambda \frac{\partial \varphi}{\partial x_1} \right) dx_1 + \left( \frac{\partial f}{\partial x_2} + \lambda \frac{\partial \varphi}{\partial x_2} \right) dx_2 + \left( \frac{\partial f}{\partial x_3} + \lambda \frac{\partial \varphi}{\partial x_3} \right) dx_3 = 0 \quad (3.35)$$

ahora tenemos en cuenta que dos de las variables son independientes, digamos  $x_1, x_2$ . Nótese que el multiplicador  $\lambda$  en la Eq. (3.34) ha sido definido de modo que

$$\frac{\partial f}{\partial x_3} + \lambda \frac{\partial \varphi}{\partial x_3} = 0 \quad (3.36)$$

siempre que  $\partial_{x_3} \varphi \neq 0$ . La Eq. (3.35) queda de la forma

$$\left( \frac{\partial f}{\partial x_1} + \lambda \frac{\partial \varphi}{\partial x_1} \right) dx_1 + \left( \frac{\partial f}{\partial x_2} + \lambda \frac{\partial \varphi}{\partial x_2} \right) dx_2 = 0$$

y dado que  $x_1$  y  $x_2$  se pueden variar independientemente, cada término entre paréntesis debe anularse

$$\left( \frac{\partial f}{\partial x_1} + \lambda \frac{\partial \varphi}{\partial x_1} \right) = 0 ; \left( \frac{\partial f}{\partial x_2} + \lambda \frac{\partial \varphi}{\partial x_2} \right) = 0 \quad (3.37)$$

cuando se cumplen las Ecs. (3.36, 3.37) se tiene que  $df = 0$  y  $f$  es un extremo (o punto de silla). Si añadimos la ecuación de ligadura (3.33) tendremos cuatro ecuaciones con cuatro incógnitas  $x_1, x_2, x_3, \lambda$ . Nótese sin embargo que en general nuestro interés es calcular los  $x_i$ , de modo que  $\lambda$  no necesita ser hallado. Por esta razón, a  $\lambda$  se le denomina **multiplicador indeterminado de Lagrange**. Es claro que el método falla cuando todos los coeficientes de  $\lambda$  se anulan en el extremo i.e. si

$$\left. \frac{\partial \varphi}{\partial x_i} \right|_{x_i=x_{i,0}} = 0 \quad \text{para } i = 1, 2, 3;$$

donde  $(x_{1,0}, x_{2,0}, x_{3,0})$  es el punto donde se ubica el extremo de  $f$ . En este caso resulta imposible despejar  $\lambda$ .

En lo anterior hemos identificado a  $f$  como la función que toma un valor extremo y  $\varphi$  la ecuación que expresa la ligadura. Sin embargo, la forma de las Ecs. (3.36, 3.37) nos muestran que podemos intercambiar los papeles de estas funciones.

### 3.4.1. Generalización a un conjunto arbitrario de variables y de ligaduras

Supongamos que tenemos una función con  $n$  variables  $f = f(x_1, \dots, x_n)$  y que dichas variables están sujetas a  $m$  ligaduras de la forma

$$\varphi_k(x_1, \dots, x_n) = 0 \quad ; \quad k = 1, \dots, m \quad (3.38)$$

con  $m \leq n$ . Multipliquemos cada ecuación de ligadura por un factor  $\lambda_k$

$$\lambda_k \varphi_k(x_1, \dots, x_n) = 0 \quad ; \quad k = 1, \dots, m \quad (\text{no suma})$$

es claro que

$$\begin{aligned} df + \sum_{k=1}^m \lambda_k d\varphi_k &= \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i + \sum_{k=1}^m \lambda_k \left[ \sum_{i=1}^n \frac{\partial \varphi_k}{\partial x_i} dx_i \right] \\ df + \sum_{k=1}^m \lambda_k d\varphi_k &= \sum_{i=1}^n \left\{ \frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} \right\} dx_i \end{aligned}$$

y dado que  $df = d\varphi_k = 0$  tenemos que  $df + \sum_{k=1}^m \lambda_k \varphi_k = 0$  de modo que

$$\sum_{i=1}^n \left\{ \frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} \right\} dx_i = 0 \quad (3.39)$$

como los  $x_i$  no son independientes, los términos entre paréntesis no necesariamente son nulos. No obstante, podemos ahora elegir las  $m - n$  primeras variables como independientes, en tanto que las  $m$  últimas están determinadas por las ligaduras (3.38). Ahora aprovechamos el carácter indeterminado de los  $\lambda_k$  para eliminar las coordenadas dependientes, lo cual haremos exigiendo que los  $\lambda_k$  tomen valores tales que se satisfagan las ecuaciones

$$\frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} = 0 \quad ; \quad i = n - m + 1, \dots, n \quad (3.40)$$

para las últimas  $m$  variables. Al reemplazar (3.40) en (3.39) se obtiene

$$\sum_{i=1}^{n-m} \left\{ \frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} \right\} dx_i = 0$$

y como las coordenadas involucradas en esta ecuación son las independientes, podemos afirmar que cada término entre paréntesis es cero de modo que

$$\frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} = 0 \quad ; \quad i = 1, \dots, m - n \quad (3.41)$$



y uniendo las Ecs. (3.40, 3.41) se obtiene simplemente

$$\frac{\partial f}{\partial x_i} + \sum_{k=1}^m \lambda_k \frac{\partial \varphi_k}{\partial x_i} = 0 \quad ; \quad i = 1, \dots, n \quad (3.42)$$

siendo  $m$  el número de ligaduras y  $n$  el número de variables  $x_i$ . Hay un multiplicador  $\lambda_k$  por cada ligadura  $\varphi_k$ . Las Ecs. (3.42) junto con las ligaduras (3.38) nos dan un conjunto de  $n + m$  ecuaciones con  $n + m$  incógnitas (las  $n$  variables y los  $m$  multiplicadores  $\lambda_k$ ).

### Minimización de la superficie de un cilindro con volumen constante

Encontraremos el radio  $R$  y la altura  $H$  de un cilindro circular recto, que minimice su área superficial total, manteniendo constante el volumen  $V_0$  del cilindro.

La función a minimizar es el área superficial total del cilindro que depende de las variables  $R$  y  $H$

$$f(R, H) = 2\pi R^2 + 2\pi RH \quad (3.43)$$

la ecuación de ligadura (3.38) corresponde en este caso a la restricción de mantener el volumen constante, y se puede escribir como

$$\varphi(R, H) = \pi R^2 H - V_0 = 0 \quad (3.44)$$

puesto que solo hay una ecuación de ligadura, hay un solo multiplicador indeterminado  $\lambda$ . Reemplazando (3.43, 3.44) en (3.42) las ecuaciones para  $f(R, H)$  con multiplicadores para las variables  $R$  y  $H$  son

$$\begin{aligned} \frac{\partial f(R, H)}{\partial H} + \lambda \frac{\partial \varphi}{\partial H} &= 0 \quad ; \quad \frac{\partial f(R, H)}{\partial R} + \lambda \frac{\partial \varphi}{\partial R} = 0 \\ 2\pi R + \lambda \pi R^2 &= 0 \quad ; \quad \pi R + 2\pi \lambda RH = 0 \\ 2 + \lambda R &= 0 \quad ; \quad 1 + 2\lambda H = 0 \end{aligned}$$

en la última línea hemos tomado la solución no trivial  $R \neq 0$ , con lo cual se obtiene

$$\lambda = -\frac{2}{R} \quad ; \quad R = 4H$$

reemplazando esto último en la ecuación de ligadura (3.44)

$$\begin{aligned} \pi (4H)^2 H &= V_0 \Rightarrow \\ H &= \sqrt[3]{\frac{V_0}{16\pi}} \quad ; \quad R = \sqrt{\frac{4V_0}{\pi}} \end{aligned}$$

### 3.5. Problemas variacionales con ligaduras (opcional)

Tomemos de nuevo el problema variacional en el cual buscamos un camino que haga que la integral (3.20)

$$J = \int_1^2 f(y_1(x), \dots, y_n(x), \dot{y}_1(x), \dots, \dot{y}_n(x), x) dx \quad (3.45)$$

sea estacionaria. Siendo las  $y_i$  coordenadas generalizadas que solo pueden depender del parámetro  $x$ . La codición de estacionaridad es

$$\delta J = \delta \int_1^2 f(y_1(x), \dots, y_n(x), \dot{y}_1(x), \dots, \dot{y}_n(x), x) dx = 0 \quad (3.46)$$

asumamos sin embargo, que el problema debe respetar ciertas restricciones que se manifiestan en ecuaciones de ligadura de la forma

$$\varphi_k(y_1, y_2, \dots, y_n; x) = 0 \quad ; \quad k = 1, 2, \dots, m \quad (3.47)$$

multiplicando cada una de las ecuaciones (3.47) por un multiplicador indeterminado que en general depende del parámetro, i.e.  $\lambda_k = \lambda_k(x)$ , sumando sobre  $k$  e integrando en el mismo rango que en la ecuación 3.45, se obtiene

$$\int_1^2 \sum_{k=1}^m \lambda_k(x) \varphi_k(y_1, y_2, \dots, y_n; x) dx = 0$$

es claro por tanto que

$$\delta \int_1^2 \sum_{k=1}^m \lambda_k(x) \varphi_k(y_1, y_2, \dots, y_n; x) dx = 0 \quad (3.48)$$

en algunas ocasiones, las ligaduras de por sí adquieren forma integral i.e.

$$\int_1^2 \varphi_k(y_1, y_2, \dots, y_n; x) dx = cte \quad ; \quad k = 1, 2, \dots, m$$

en este caso podemos multiplicar cada una de estas ecuaciones por un  $\lambda_k$  (que en este caso no dependería del parámetro  $x$ ), sumar sobre  $k$  y aplicar la variación quedando

$$\delta \int_1^2 \sum_{k=1}^m \lambda_k \varphi_k(y_1, y_2, \dots, y_n; x) dx = 0 \quad (3.49)$$

nótese que las Ecs. (3.48, 3.49) son esencialmente idénticas excepto por la dependencia de  $\lambda_k$  del parámetro  $x$  en la Ec. (3.48). Sumando cualquiera de estas dos ecuaciones con la Ec. (3.46) resulta

$$\delta \int_1^2 \left[ f(y_i, \dot{y}_i, x) + \sum_{k=1}^m \lambda_k \varphi_k(y_i; x) \right] dx = 0$$

redefiniendo

$$g(y_i, \dot{y}_i, x) \equiv f(y_i, \dot{y}_i, x) + \sum_{k=1}^m \lambda_k \varphi_k(y_i; x)$$

tenemos que

$$\delta \int_1^2 g(y_i, \dot{y}_i, x) dx = 0$$

y con el mismo procedimiento que nos llevó de la Ec. (3.20) a la Ec. (3.23), obtenemos

$$\delta \int_1^2 g(y_i, \dot{y}_i, x) dx = \sum_{i=1}^n \int_1^2 \left( \frac{\partial g}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial g}{\partial \dot{y}_i} \right) \right] \right) \delta y_i dx = 0 \quad (3.50)$$

en esta caso no podemos decir que cada integrando es cero ya que los desplazamientos virtuales  $\delta y_i$  no son independientes. De nuevo, usamos el carácter indeterminado de los multiplicadores de Lagrange para exigir que los términos asociados a las coordenadas dependientes se anulen

$$\frac{\partial g}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial g}{\partial \dot{y}_i} \right) \right] = 0 \quad , \quad i = n - m + 1, \dots, n \quad (3.51)$$

con lo cual la Ec. (3.50) queda

$$\sum_{i=1}^{n-m} \int_1^2 \left( \frac{\partial g}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial g}{\partial \dot{y}_i} \right) \right] \right) \delta y_i dx = 0$$

y dado que las coordenadas que quedan son independientes, podemos afirmar que cada integrando es cero, y junto a las Ecs. (3.51) esto nos lleva a

$$\frac{\partial g}{\partial y_i} - \left[ \frac{d}{dx} \left( \frac{\partial g}{\partial \dot{y}_i} \right) \right] = 0, \quad i = 1, \dots, n \quad (3.52)$$

$$g(y_i, x) \equiv f(y_i, \dot{y}_i, x) + \sum_{k=1}^m \lambda_k \varphi_k(y_i; x) \quad (3.53)$$

los multiplicadores de Lagrange tienen aplicaciones en múltiples escenarios de la Física más allá de los que utilizaremos en el presente texto (ver por ejemplo la Ref. [7]).

### Minimización del área lateral de un sólido de revolución

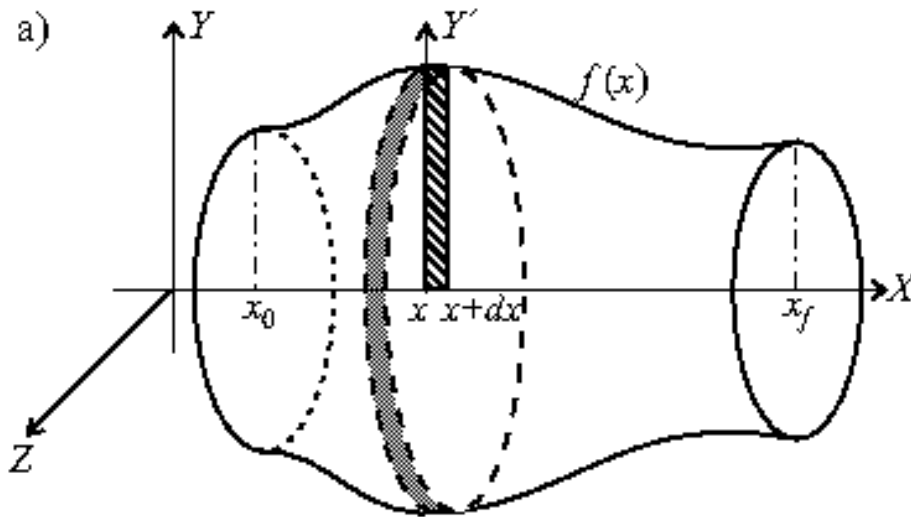


Figura 3.2: Sólido de revolución generado por la función  $f(x)$  definida en el intervalo  $[x_0, x_f]$ , cuando se rota alrededor del eje  $X$ .

Sea una función  $f(x)$  definida en un intervalo  $[x_0, x_f]$  de longitud  $L$ , de tal manera que genera un sólido de revolución alrededor del eje  $X$  como el de la Fig. 3.2. Queremos encontrar una función  $f(x)$  de tal manera que el sólido generado posea la mínima área lateral, pero de tal forma que el volumen  $V_0$  permanezca constante.

Un diferencial de área lateral, es el área lateral del cilindro de radio  $f(x)$  y altura  $dx$  como se vé en la Fig. 3.2, i.e.  $dA = 2\pi f(x) dx$ . El área lateral es entonces

$$2\pi \int_{x_0}^{x_f} f(x) dx = A \quad (3.54)$$

Un diferencial de volumen del sólido, es el volumen del cilindro de radio  $f(x)$  y altura  $dx$  de modo que  $dV = \pi f(x)^2 dx$ . La ligadura de volumen constante queda entonces en la forma

$$\pi \int_{x_0}^{x_f} f(x)^2 dx = V_0 \quad (3.55)$$

en este caso la función  $f(x)$  que se pretende optimizar es la coordenada generalizada y  $x$  es el parámetro del cual depende. Tenemos entonces las asignaciones

$$y(x) \rightarrow f(x) \quad , \quad \dot{y} \rightarrow \frac{df(x)}{dx} \equiv \dot{f}$$

Para evitar confusiones, conviene hacer un cambio de notación en la Ecs. (3.52, 3.53)

$$\frac{\partial G}{\partial f} - \left[ \frac{d}{dx} \left( \frac{\partial G}{\partial \dot{f}} \right) \right] = 0 \quad ; \quad G(f, \dot{f}, x) \equiv F(f, \dot{f}, x) + \sum_{k=1}^m \lambda_k \varphi_k(y_i; x) \quad (3.56)$$

dado que la cantidad a optimizar es el área lateral, y el volumen actúa como la ligadura, podemos escribir

$$J \equiv A = \int_{x_0}^{x_f} F(f, \dot{f}, x) \, dx \quad ; \quad \int_{x_0}^{x_f} \varphi(f, \dot{f}, x) \, dx = V_0 \quad (3.57)$$

comparando las Ecs. (3.57) con las Ecs. (3.54, 3.55), tenemos que

$$F(f, \dot{f}, x) = F(f) = 2\pi f \quad ; \quad \varphi(f, \dot{f}, x) = \varphi(f) = \pi f^2 \quad (3.58)$$

es decir  $F$  y  $\varphi$  solo dependen de la coordenada generalizada y no dependen explícitamente del parámetro  $x$  ni de  $\dot{f}$ . Sustituyendo (3.58) en la segunda de las Ecs. (3.56), tenemos entonces

$$G(f, \dot{f}, x) = G(f) = 2\pi f + \lambda \pi f^2$$

que al reemplazar en la primera de las Ecs. (3.56) nos da

$$2\pi + 2\pi\lambda f = 0 \quad ; \quad f = -\frac{1}{\lambda}$$

dado que  $\lambda$  es constante en el caso de ligaduras integrales, tenemos que  $f(x) = -\lambda^{-1} = cte$  y el sólido de revolución es un cilindro cuyo radio es  $|\lambda^{-1}|$ . Nótese que efectivamente  $f(x)$  y  $-f(x)$  generan el mismo sólido de revolución. Definiremos entonces  $f(x) \equiv R > 0$ . El radio del cilindro se puede determinar de la ligadura (3.55)

$$\pi \int_{x_0}^{x_f} R^2 \, dx = \pi R^2 L = V_0 \Rightarrow f(x) \equiv R = \sqrt{\frac{V_0}{\pi L}}$$

### 3.6. Ejercicios

1. Encuentre el radio  $R$  y la altura  $H$  de un cilindro circular recto, que minimice su área lateral, manteniendo constante el volumen  $V_0$  del cilindro.
2. Sea una elipse descrita por la ecuación

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1.$$

Encuentre el rectángulo inscrito de mayor área. Muestre que el cociente entre el área de este rectángulo y el área de la elipse es  $2/\pi$ .

3. Sea  $f(x)$  el generador de un sólido de revolución alrededor del eje  $X$  como en la Fig. 3.2. Encuentre la función  $f(x)$  que genere un sólido de revolución con el mínimo volumen, manteniendo fijo el valor del área lateral.
4. Para un sólido de revolución como el de la Fig. 3.2, con densidad  $\rho$  constante, los momentos de inercia con respecto a los ejes  $X$  e  $Y$  vienen dados por (ver problema 1 pág 414, y Ref. [8])

$$I_X = \frac{\pi\rho}{2} \int_{x_0}^{x_f} f(x)^4 \, dx \quad ; \quad I_Y = \frac{I_X}{2} + \pi\rho \int_{x_0}^{x_f} x^2 f(x)^2 \, dx$$

utilizando ligadura de masa constante para el sólido encuentre (a) La función  $f(x)$  que minimiza  $I_X$ , (b) la función  $f(x)$  que minimiza a  $I_Y$ . (c) ¿A que sólidos de revolución corresponden estas funciones?.

5. Encuentre la curva  $f(x)$  de longitud  $L$ , limitada por abajo por el eje  $X$ , que pasa por los puntos  $(-a, 0)$  y  $(a, 0)$ , y que encierra la mayor área. Este tipo de problemas se conocen como problemas isoperimétricos o problemas de Dido.

# Principio variacional de Hamilton y ecuaciones de Lagrange

En el capítulo 2, las ecuaciones de Lagrange se derivaron de un principio diferencial (principio de D'Alembert), puesto que se basaba en desplazamientos virtuales infinitesimales de cada coordenada generalizada del sistema, a partir de cierta configuración instantánea. En el presente capítulo, derivaremos las ecuaciones de Lagrange de un principio integral en el cual a partir de cierta configuración instantánea, se considerarán desplazamientos finitos de cada coordenada generalizada en cierto intervalo de tiempo, y se realizarán variaciones virtuales infinitesimales de estos movimientos.

Dado que las coordenadas generalizadas dan en general todas las variables independientes del sistema para determinar la configuración de éste, es útil describir el movimiento del sistema a través de dichas variables usando el tiempo como parámetro. Si tenemos  $n$  coordenadas generalizadas  $q_i$  es útil construir un hiperespacio cartesiano  $n$  dimensional donde cada eje coordenado representa a una de las  $q_i$ . Una configuración instantánea del sistema estará representada por un punto en este hiperespacio, que llamaremos **espacio de configuraciones**. Ahora bien, la evolución del sistema (el paso del tiempo) se representa a través del movimiento de este punto en el espacio de configuraciones, dibujando una curva en dicho espacio. En este espíritu cuando hablemos del **movimiento del sistema** entre los tiempos  $t_1$  y  $t_2$  estaremos hablando del movimiento del punto en el espacio de configuraciones, y la **trayectoria de movimiento del sistema** será la curva trazada por este punto. Por supuesto el tiempo actúa como parámetro en esta hipercurva y a cada punto de la trayectoria puede estar asociado más de un valor del tiempo<sup>1</sup>. El espacio de configuraciones no tiene necesariamente ninguna conexión con el espacio físico tridimensional, sus coordenadas no necesariamente definen longitudes y las  $n$ -uplas, en este espacio no necesariamente definen vectores euclidianos. De la misma forma la hipercurva que define la trayectoria de movimiento del sistema no necesariamente está asociada a la trayectoria de una partícula real, tengamos en cuenta que un punto en el hiperespacio de configuraciones y su movimiento, describen la configuración y movimiento de todo el sistema.

De las condiciones para derivar las ecuaciones de Lagrange, surge de manera natural la siguiente definición: **Un sistema monogénico**, es aquél para el cual todas las fuerzas (excepto tal vez las de ligadura) son derivables de un potencial escalar generalizado que es función de las coordenadas generalizadas, las velocidades generalizadas y el tiempo, en el sentido de las Ecs. (2.26, 2.27). Nótese que esta definición se traduce en el hecho de que para esta clase de sistemas, el Lagrangiano contiene toda la información Física del sistema. Por ejemplo, un sistema con fuerzas disipativas no es monogénico ya que la fuerza de rozamiento viscosa descrita por la ecuación (2.36) no se puede escribir en términos de un potencial generalizado, esto se traduce en el hecho de que el Lagrangiano no contiene toda la dinámica del sistema y dicha información debe ser completada con la función de disipación de Rayleigh como se puede ver en la Ec. (2.38).

Volviendo a los sistemas monogénicos, cuando el potencial generalizado solo es función de las coordena-

---

<sup>1</sup>Decimos que el tiempo es un parámetro porque no construimos un eje asociado al tiempo en el espacio de configuraciones. El movimiento del parámetro tiempo regula el barrido del punto sobre la curva en el espacio de configuraciones.

das generalizadas, tenemos un sistema conservativo. A continuación enunciaremos un principio asociado al movimiento del sistema, cuando éste es monogénico

**Principio de Hamilton:** El movimiento en el espacio de configuraciones de un sistema monogénico desde un punto  $P_1$  en el tiempo  $t_1$  hasta un punto  $P_2$  en el tiempo  $t_2$ , es tal que la integral de línea

$$I \equiv \int_{t_1}^{t_2} L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) dt \quad (4.1)$$

tiene un valor estacionario para la trayectoria real del movimiento.

A la integral anterior se le llama la **acción** y el Lagrangiano se define en la forma usual  $L \equiv T - U$ . Esto significa que entre todas las trayectorias posibles que el sistema puede seguir en el espacio de configuraciones entre los tiempos  $t_1$  y  $t_2$ , el sistema viajará a través de la trayectoria que haga que el valor de la integral (4.1) sea estacionario. Para una integral de línea, el término estacionario significa que el valor de la integral cuando se toma un cierto camino, tiene el mismo valor dentro de infinitesimales de primer orden que cuando se toman caminos vecinos a éste (caminos que difieren del original por desplazamientos generalizados infinitesimales). Insistimos en que todas las trayectorias de las que hablamos aquí, son realmente hipertrayectorias en el espacio de configuraciones y no están asociadas en general a ninguna trayectoria real. Para integrales de línea este es el equivalente en funciones ordinarias a puntos con derivada cero. En este caso en lugar de un punto tenemos una trayectoria entera.

De acuerdo con el principio de Hamilton, la trayectoria del sistema es tal que la variación de la acción para valores fijos de  $t_1$  y  $t_2$  es nula

$$\delta I = \delta \int_{t_1}^{t_2} L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) dt = 0 \quad (4.2)$$

cuando las ligaduras son holónomas el principio de Hamilton es condición necesaria y suficiente para la validez de las ecuaciones de Lagrange. Una ventaja de esta formulación es que la acción es invariante ante una transformación de un sistema de coordenadas generalizadas a otro. Adicionalmente, veremos que el formalismo variacional permitirá incluir las ligaduras en el formalismo y tratar algunos problemas con ligaduras no holónomas. Finalmente, una formulación variacional es más adecuada en el tratamiento de campos y sistemas continuos.

## 4.1. Aplicación del cálculo de variaciones al principio de Hamilton

En el principio de Hamilton el sistema debe ser monogénico de modo que el Lagrangiano no depende de derivadas temporales de  $q$  más altas que la primera i.e.  $L = L(q_i, \dot{q}_i, t)$ . La condición de extremos fijos está implícita en su enunciado, y finalmente el único parámetro es el tiempo. De lo anterior es claro que el problema de la estacionaridad de la acción es un problema variacional como el tratado en la sección 3.2 del capítulo 3. Comparando las ecuaciones (3.20) y (4.1), vemos que podemos utilizar las ecuaciones (3.25) con las siguientes asignaciones

$$x \rightarrow t, \quad y_i \rightarrow q_i, \quad f(y_i, \dot{y}_i, x) \rightarrow L(q_i, \dot{q}_i, t)$$

con lo cual las Ecs. (3.25) conducen a las ecuaciones de Lagrange

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad ; \quad i = 1, \dots, n \quad (4.3)$$

al derivar estas ecuaciones debemos suponer que las coordenadas generalizadas son todas independientes, lo cual siempre es posible cuando las ligaduras son holónomas.

## 4.2. Extensión del principio de Hamilton a algunos sistemas no holónomos

Si recordamos el procedimiento hecho en las secciones (3.2.1) y (3.2.2) observaremos que en la deducción de las ecuaciones de Lagrange a partir del principio de Hamilton, el hecho de que la ligadura sea holónoma

solo se usa en el último paso cuando se considera que todas las coordenadas generalizadas (y sus desplazamientos virtuales) son independientes entre sí. Sin embargo, cuando las ligaduras son no holónomas, las coordenadas generalizadas ya no serán independientes y no es posible reducirlas por medio de ecuaciones de la forma  $f_\alpha(q_1, \dots, q_n, t) = 0$ . De modo que en general tendremos que trabajar con un sistema coordinado no independiente.

Desde el punto de vista del principio variacional, esto influye en la forma en que se construyen los caminos variados. Un desplazamiento  $\delta y$  (o  $\delta q$ ), nos lleva de un punto en el camino real a otro punto sobre un camino variado, cuando las coordenadas son independientes es el camino variado final el que importa y no la forma como se construye. Pero cuando tenemos coordenadas no independientes relacionadas entre sí por ligaduras, la forma en que se construye el camino influye ya que los desplazamientos virtuales en general no respetan las ligaduras. Por tanto resulta importante si el camino variado fué construído con desplazamientos que respetan o no las ligaduras.

Un conjunto de ligaduras no holónomas para el cual es susceptible el tratamiento variacional, es el conjunto de ligaduras de la forma

$$\sum_{k=1}^n a_{lk} dq_k + a_{lt} dt = 0 \quad ; \quad l = 1, \dots, m \quad (4.4)$$

es decir una relación lineal entre los diferenciales de las coordenadas y el tiempo. El índice  $l$  determina el número de ecuaciones  $l = 1, \dots, m$ . Los coeficientes  $a_{lk}$ ,  $a_{lt}$  pueden ser funciones de las coordenadas y el tiempo. Estas ligaduras son en general no integrables a menos que se cumplan las relaciones

$$a_{lk} = \frac{\partial f}{\partial q_k} \quad ; \quad a_{lt} = \frac{\partial f}{\partial t} \quad (4.5)$$

para alguna función  $f = f(q_i, t)$ . En tal caso la ligadura es realmente holónoma. Tomaremos el caso general en donde las relaciones (4.5) no necesariamente se cumplen. En principio se podría pensar en construir un camino variado a través de desplazamientos virtuales infinitesimales del camino real que sean compatibles con la ligadura (4.4). Sin embargo se ha demostrado que no se puede construir tal camino variado, a menos que las ligaduras sean integrables en cuyo caso dichas ligaduras son realmente holónomas. Construiremos de todas formas un principio variacional en el cual los desplazamientos virtuales generarán los caminos variados aunque no sea en forma compatible con las ligaduras. Como los desplazamientos virtuales son fijos en el tiempo, las ligaduras para los desplazamientos virtuales se escriben

$$\sum_{k=1}^n a_{lk} \delta q_k = 0 \quad ; \quad l = 1, \dots, m \quad (4.6)$$

el camino variado en general no satisface las ecuaciones (4.4). La idea es ahora reducir los desplazamientos virtuales a los independientes. Para eliminar los desplazamientos virtuales sobrantes se utiliza el método de los **multiplicadores indeterminados de Lagrange**. Para ello multiplicamos la ecuación (4.6) por una cantidad indeterminada  $\lambda_l$  que puede ser función de las coordenadas y del tiempo

$$\lambda_l \sum_{k=1}^n a_{lk} \delta q_k = 0 \quad (4.7)$$

por supuesto, hay  $m$  cantidades  $\lambda_l$  una para cada ecuación de ligadura. Si asumimos que el principio de Hamilton es también válido para sistemas no holónomos<sup>2</sup>, y usamos las Ecs. (3.23), entonces el principio de Hamilton conduce a

$$- \int_1^2 dt \sum_{k=1}^n \left[ \frac{\partial L}{\partial q_k} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k = 0 \quad (4.8)$$

---

<sup>2</sup>La naturaleza de la ligadura no aparece en el principio de Hamilton. Solo se exige que el sistema sea monogénico.

con el fin de poder introducir los  $\lambda_l$  dentro del principio de Hamilton, sumamos la Ec. (4.7) sobre los valores de  $l$  e integramos en el tiempo entre  $t_1$  y  $t_2$

$$\int_1^2 dt \left[ \sum_{k=1}^n \sum_{l=1}^m \lambda_l a_{lk} \delta q_k \right] = 0 \quad (4.9)$$

ahora sumamos (4.8) y (4.9) obteniéndose

$$\int_1^2 dt \sum_{k=1}^n \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} + \sum_{l=1}^m \lambda_l a_{lk} \right] \delta q_k = 0 \quad (4.10)$$

las  $\delta q_k$  siguen siendo dependientes, ya que están ligadas por las  $m$  ecuaciones (4.6). De modo que las  $n - m$  primeras de ellas se pueden elegir en forma independiente, en tanto que las  $m$  últimas están determinadas por las Ecs. (4.6). Para eliminar los desplazamientos sobrantes, aprovechamos la arbitrariedad en las cantidades  $\lambda_l$ . Exigiremos entonces que estas cantidades satisfagan las ecuaciones

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} + \sum_{l=1}^m \lambda_l a_{lk} = 0 \quad ; \quad k = n - m + 1, \dots, n \quad (4.11)$$

es decir, para las últimas  $m$  coordenadas generalizadas. Reemplazando (4.11) en (4.10) se obtiene

$$\int_1^2 dt \sum_{k=1}^{n-m} \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} + \sum_{l=1}^m \lambda_l a_{lk} \right] \delta q_k = 0$$

pero ahora las  $\delta q_k$  que intervienen son las independientes, de modo que se puede afirmar que

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} + \sum_{l=1}^m \lambda_l a_{lk} = 0 \quad ; \quad k = 1, \dots, n - m. \quad (4.12)$$

las Ecs. (4.11, 4.12) nos dan el sistema completo de ecuaciones de Lagrange para sistemas no holónomos

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} + \sum_{l=1}^m \lambda_l a_{lk} = 0 \quad ; \quad k = 1, \dots, n. \quad (4.13)$$

este sistema tiene  $n$  ecuaciones con  $n + m$  incógnitas ( $n$  coordenadas y  $m$  multiplicadores). Las  $m$  ecuaciones faltantes serían las ecuaciones de ligadura (reales) Ecs. (4.4), que enlazan las  $q_k$ . No obstante, es más conveniente para la solución del sistema tener estas ligaduras en forma de ecuaciones diferenciales de primer orden

$$\sum_{k=1}^n a_{lk} \dot{q}_k + a_{lt} = 0 \quad ; \quad l = 1, \dots, m \quad (4.14)$$

de modo que las ecuaciones (4.13, 4.14) definen el sistema completo de  $n + m$  ecuaciones e incógnitas.

#### 4.2.1. Significado físico de los multiplicadores de Lagrange: fuerzas de ligadura

Ahora nos preguntamos por el significado de los multiplicadores de Lagrange. Para ello supongamos que quitamos las ligaduras del sistema y las sustituimos por fuerzas aplicadas que mantienen intacta la dinámica de éste. Las ecuaciones de movimiento quedan entonces de la forma

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = Q_k^L \quad ; \quad k = 1, \dots, n \quad (4.15)$$

donde  $Q_k^L$  son las fuerzas generalizadas asociadas a las fuerzas aplicadas que emulan a las de ligadura. Naturalmente, estas fuerzas aplicadas deben ser iguales a las de ligadura para mantener intacta la dinámica del



sistema<sup>3</sup>. Por otro lado, la invarianza de la dinámica también requiere que las ecuaciones (4.13) y (4.15) sean idénticas. Esto conduce a

$$Q_k^L = - \sum_{l=1}^m \lambda_l a_{lk} \quad ; \quad k = 1, \dots, n \quad (4.16)$$

esto significa que esta formulación tiene una información adicional que no habíamos obtenido con las anteriores formulaciones: las fuerzas de ligadura. Para ver con mayor claridad la relación entre los  $\lambda_l$  y las fuerzas reales de ligadura, aplicamos la definición de fuerza generalizada a las Ecs. (4.16)

$$\sum_{i=1}^N \mathbf{F}_i^L \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = - \sum_{l=1}^m \lambda_l a_{lk} \quad ; \quad k = 1, \dots, n \quad (4.17)$$

donde  $\mathbf{F}_i^L$  representa las fuerzas reales de ligadura. Vemos que en la relación interviene una suma sobre todas las fuerzas reales de ligadura y otra suma sobre todos los multiplicadores, la relación entre multiplicadores y fuerzas reales de ligadura es entonces bastante indirecta.

#### 4.2.2. Formalismo de los multiplicadores para ligaduras holónomas

Por supuesto las ligaduras no holónomas descritas por (4.4) no son las más generales pero cubren una gran cantidad de casos que se presentan en los sistemas reales. Por otro lado, las ligaduras (4.4) incluyen a las ligaduras holónomas cuando las ecuaciones se vuelven integrables. Efectivamente, una ligadura de la forma

$$f_l(q_1, \dots, q_n, t) = 0 \quad ; \quad l = 1, \dots, m \quad (4.18)$$

es equivalente a la ecuación diferencial

$$\left( \sum_k \frac{\partial f_l}{\partial q_k} dq_k \right) + \frac{\partial f_l}{\partial t} dt = 0 \quad ; \quad l = 1, \dots, m \quad (4.19)$$

que es igual a (4.4) con las asignaciones<sup>4</sup>

$$a_{lk} = \frac{\partial f_l}{\partial q_k} \quad ; \quad a_{lt} = \frac{\partial f_l}{\partial t} \quad (4.20)$$

reemplazando estas relaciones en (4.13) se obtiene

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = - \sum_{l=1}^m \lambda_l \frac{\partial f_l}{\partial q_k} \quad ; \quad k = 1, \dots, n. \quad (4.21)$$

de modo que la Ec. (4.21) corresponde a las ecuaciones de Lagrange con multiplicadores indeterminados en el caso de ligaduras holónomas. Las fuerzas generalizadas provenientes de ligaduras holónomas se escriben entonces

$$Q_k^L \equiv - \sum_{l=1}^m \lambda_l \frac{\partial f_l}{\partial q_k} \quad (4.22)$$

estas mismas ecuaciones se pueden obtener partiendo directamente de la forma de la ligadura holónoma  $f_l(q, t) = 0$  e introduciendo un término de la forma  $\sum_l \lambda_l f_l = 0$  en el principio de Hamilton, con un procedimiento similar al realizado para ligaduras no holónomas.

<sup>3</sup>Para llegar a la Ec. (4.15), podemos seguir el procedimiento de la sección 2.2, pero de modo que en la Ec. (2.4) no se excluyan las fuerzas de ligadura de la formulación. Seguimos entonces el procedimiento que nos lleva de (2.4) hasta (2.17), de tal modo que  $Q_j$  en ésta última ecuación de puede descomponer como  $Q_j = Q_j^{(a)} + Q_j^L$ , donde  $Q_j^{(a)}$  proviene de las fuerzas aplicadas originales, y  $Q_j^L$  está asociada a las nuevas fuerzas aplicadas que emulan a las de ligadura. Si  $Q_j^{(a)} = -\partial_{q_j} V$  ó  $Q_j^{(a)} = \hat{O}_j U$  como en las Ecs. (2.21, 2.27), podemos construir un lagrangiano  $L = T - U$  donde el potencial incluye solo a las fuerzas aplicadas originales, con lo cual se llega a la Ec. (4.15).

<sup>4</sup>La ecuación (4.19) es equivalente a (4.18), siempre y cuando al integrar (4.19) exijamos que la constante de integración sea nula.

La introducción de multiplicadores de Lagrange en ligaduras holónomas se justifica en uno de estos casos (1) Comenzamos con coordenadas no independientes y queremos eliminar las dependientes<sup>5</sup> (2) queremos hallar las fuerzas de ligadura, es necesario tener en cuenta que los multiplicadores de Lagrange solo extraen el valor de la magnitud de las fuerzas de ligadura, la dirección se debe determinar por consideraciones físicas. Esto se debe a la arbitrariedad para definir el signo del multiplicador, en realidad se puede hacer el cambio  $\lambda_l \rightarrow -\lambda_l$  (siempre y cuando sea para todos los  $\lambda_l$ ), y la formulación es igualmente consistente.

### 4.3. Relación entre el principio diferencial de D'Alembert y el Principio variacional de Hamilton

Aunque no es evidente, la versión del principio de Hamilton tanto para sistemas holónomos como no holónomos, también requiere que el trabajo virtual de las fuerzas de ligadura sea nulo. Para ver esto, escribamos el principio de Hamilton en la forma

$$\delta \int_{t_1}^{t_2} L dt = \delta \int_{t_1}^{t_2} T dt - \delta \int_{t_1}^{t_2} U dt = 0$$

siguiendo los procedimientos variacionales ya descritos teniendo en cuenta que el potencial generalizado es función de las coordenadas generalizadas, las velocidades generalizadas y el tiempo, se tiene

$$\delta \int_{t_1}^{t_2} T dt = \int_{t_1}^{t_2} \sum_{k=1}^n \left[ \frac{\partial U}{\partial q_k} - \frac{d}{dt} \left( \frac{\partial U}{\partial \dot{q}_k} \right) \right] \delta q_k dt$$

ahora teniendo en cuenta (2.26) y (2.27) resulta

$$\delta \int_{t_1}^{t_2} T dt = - \int_{t_1}^{t_2} \sum_{k=1}^n Q_k \delta q_k dt \quad (4.23)$$

en esta forma el principio de Hamilton dice que la diferencia de la integral temporal de la energía cinética entre dos caminos vecinos es igual a menos la integral temporal del trabajo realizado en los desplazamientos virtuales entre los caminos. El trabajo calculado proviene solo de las fuerzas que derivan del potencial generalizado (ya que implícitamente el potencial del Lagrangiano solo contiene fuerzas aplicadas y no de ligadura). Por tanto, se requiere que las ligaduras no realicen trabajos virtuales a fin de que las fuerzas generalizadas asociadas a dichas ligaduras no entren en el miembro derecho de la Ec. (4.23). En consecuencia, si queremos mantener el principio de Hamilton tanto para el caso holónimo como el no holónimo, es necesario que las fuerzas adicionales de ligadura no holónomas no trabajen en desplazamientos virtuales  $\delta q_k$ .

Otra manera de ver que las fuerzas de ligadura no entran en el principio variacional de Hamilton, se obtiene partiendo de las ecuaciones (4.6) de ligadura para desplazamiento virtual

$$\sum_{k=1}^n a_{lk} \delta q_k = 0 \quad , \quad l = 1, \dots, m \quad (4.24)$$

multiplicando estas ecuaciones por  $-\lambda_l$  y sumando sobre  $l$ , obtenemos

$$- \sum_{l=1}^m \lambda_l \sum_{k=1}^n a_{lk} \delta q_k = 0 \quad \Rightarrow \quad \sum_{k=1}^n \left[ - \sum_{l=1}^m \lambda_l a_{lk} \right] \delta q_k = 0 \quad (4.25)$$

y sustituyendo (4.16) en (4.25) se obtiene

$$\sum_k Q_k^L \delta q_k = 0 \quad (4.26)$$

<sup>5</sup>Cuando las ligaduras son holónomas, siempre es formalmente posible encontrar las coordenadas mínimas independientes. Sin embargo, en algunos casos específicos la tarea puede ser muy difícil. Los multiplicadores de Lagrange son un método sistemático alternativo para llegar a las coordenadas independientes mínimas.

que nos conduce a la nulidad de los trabajos virtuales de las fuerzas de ligadura.

Por supuesto esta demostración se restringe a ligaduras holónomas o no holónomas que cumplan la condición (4.24). En nuestro caso, la ligadura no-holónoma que más usaremos será la correspondiente a la condición de rodadura, que claramente cumple con la condición (4.24)<sup>6</sup> y no realiza trabajo virtual.

## 4.4. Ejemplos de aplicación del principio de Hamilton a sistemas con coordenadas dependientes

### 4.4.1. Bloque deslizante sobre una semiesfera

Comencemos con un ejemplo de ligadura holónoma. Sea una semiesfera de radio  $a$ , cuya base está montada en la tierra. Una masa  $M$  está en el tope de la semiesfera y no sufre fricción. La masa desliza debido a un ligero desplazamiento. Sea  $z$  el eje vertical y el plano  $xz$  será el plano de movimiento. Si medimos el ángulo  $\theta$  a partir del eje  $z$ , el Lagrangiano y las ligaduras los podemos escribir en coordenadas cartesianas o polares. Para tener en cuenta las ligaduras es necesario introducir coordenadas generalizadas que no son independientes y que estén relacionadas con las ligaduras, usemos entonces las coordenadas  $r$  y  $\theta$  de las cuales solo una de ellas es realmente independiente, la ligadura de que  $r$  es constante se debe a la fuerza normal. El Lagrangiano y las ligaduras en coordenadas cartesianas es

$$\begin{aligned} L &= \frac{1}{2}M(\dot{x}^2 + \dot{z}^2) - Mgz \quad ; \quad F_1(x, y) = a - \sqrt{x^2 + z^2} = 0 \\ x &= r \sin \theta \quad ; \quad \dot{x} = \dot{r} \sin \theta + r\dot{\theta} \cos \theta \quad ; \quad z = r \cos \theta \quad ; \quad \dot{z} = \dot{r} \cos \theta - r\dot{\theta} \sin \theta \end{aligned}$$

en tanto que el Lagrangiano y las ligaduras en coordenadas polares resulta

$$\begin{aligned} L &= \frac{1}{2}M(\dot{r}^2 + r^2\dot{\theta}^2) - Mgr \cos \theta \\ f_1(r, \theta) &= a - r = 0 \end{aligned} \quad (4.27)$$

la ligadura nos muestra que la coordenada  $r$  no es independiente. Así que podemos reducir a  $\theta$  el conjunto de coordenadas generalizadas independientes. No obstante, escribiremos las ecuaciones de Lagrange usando las dos coordenadas  $r, \theta$ . Calcularemos además las fuerzas generalizadas asociadas a las ligaduras con base en la Ec. (4.22)

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{r}} \right) &= M\ddot{r} \quad ; \quad \frac{\partial L}{\partial r} = Mr\dot{\theta}^2 - Mgr \cos \theta \quad ; \quad Q_r^L = -\lambda_1 \frac{\partial f_1}{\partial r} = \lambda_1 \\ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}} \right) &= \frac{d}{dt} (Mr^2\dot{\theta}) = Mr^2\ddot{\theta} + 2Mr\dot{r}\dot{\theta} \quad ; \quad \frac{\partial L}{\partial \theta} = Mgr \sin \theta \quad ; \quad Q_\theta^L = -\lambda_1 \frac{\partial f_1}{\partial \theta} = 0 \end{aligned}$$

aplicando las Ecs. (4.21, 4.22) se tiene

$$M\ddot{r} - Mr\dot{\theta}^2 + Mgr \cos \theta = \lambda_1 \quad (4.28)$$

$$Mr^2\ddot{\theta} + 2Mr\dot{r}\dot{\theta} - Mgr \sin \theta = 0 \quad (4.29)$$

las ecuaciones de Lagrange (4.28, 4.29) junto con la ecuación de ligadura (4.27) forman un sistema de 3 ecuaciones y tres incógnitas ( $r, \theta, \lambda_1$ ), cuya solución es única bajo las condiciones iniciales adecuadas. Hay un solo multiplicador  $\lambda_1$  ya que solo tenemos una ligadura. Aplicando la ligadura llegamos a  $r = a$  y  $\dot{r} = \ddot{r} = 0$ , y sustituyendolo en la ecuaciones (4.28, 4.29) obtenemos

$$\begin{aligned} -Ma\dot{\theta}^2 + Mgr \cos \theta &= \lambda_1 \quad ; \quad Ma^2\ddot{\theta} - Mga \sin \theta = 0 \\ \dot{\theta}^2 - \frac{g}{a} \cos \theta &= -\frac{\lambda_1}{Ma} \quad ; \quad \ddot{\theta} = \frac{g}{a} \sin \theta \end{aligned} \quad (4.30)$$

<sup>6</sup>Ver ejercicios 1, 2, de la Pág. 66.

este es un problema de dos ecuaciones con dos incógnitas que tiene solución única bajo las condiciones iniciales adecuadas. Una solución parcial se puede obtener con el ansatz<sup>7</sup>

$$\dot{\theta}^2 = A - B \cos \theta \quad (4.31)$$

Derivando (4.31) se obtiene

$$2\dot{\theta}\ddot{\theta} = B\dot{\theta} \sin \theta \Rightarrow \ddot{\theta} = \frac{B}{2} \sin \theta$$

comparando esta expresión con la segunda de las Ecs. (4.30) resulta  $B = 2g/a$  con lo cual el ansatz (4.31) nos da

$$\dot{\theta}^2 = A - \frac{2g}{a} \cos \theta$$

Usando condiciones iniciales de la forma  $\theta = \dot{\theta} = 0$  en  $t = 0$ <sup>8</sup> nos da  $A = 2g/a$  de modo que<sup>9</sup>

$$\dot{\theta}^2 = -\frac{2g}{a} \cos \theta + \frac{2g}{a} \quad (4.32)$$

ahora reemplazando (4.32) en la primera de las Ecs. (4.30) se tiene que

$$\begin{aligned} -\frac{2g}{a} \cos \theta + \frac{2g}{a} - \frac{g}{a} \cos \theta &= -\frac{\lambda_1}{Ma} \\ \frac{g}{a} (2 - 3 \cos \theta) &= -\frac{\lambda_1}{Ma} \end{aligned}$$

con lo cual queda finalmente

$$\lambda_1 = Mg (3 \cos \theta - 2)$$

usando los métodos tradicionales de solución se puede ver que el multiplicador corresponde numéricamente al valor de la fuerza de ligadura normal. Esto también se puede ver haciendo uso de las Ecs. (4.17, 4.20), teniendo en cuenta que hay una sola fuerza de ligadura y un solo multiplicador

$$\mathbf{N} \cdot \frac{\partial \mathbf{r}}{\partial q_k} = -\lambda_1 \frac{\partial f_1}{\partial q_k}$$

aplicando esta ecuación para  $q_k = r$  y usando las Ecs. (4.27) resulta

$$\begin{aligned} \mathbf{N} \cdot \frac{\partial (r\mathbf{u}_r)}{\partial r} &= -\lambda_1 \frac{\partial f_1}{\partial r} \Rightarrow (N\mathbf{u}_r) \cdot \mathbf{u}_r = \lambda_1 \\ \lambda_1 &= N \end{aligned}$$

donde hemos usado el hecho de que la fuerza normal va a lo largo de  $\mathbf{u}_r$ . Se puede ver que estrictamente el multiplicador solo nos puede proporcionar la magnitud de la normal ya que la fuerza de ligadura (vectorial) aparece en un producto punto que nos hace perder información sobre la dirección, a esto se le suma la ambigüedad del signo del multiplicador.

Este problema ilustra muchas características de la técnica de multiplicadores de Lagrange

<sup>7</sup>Este ansatz está inspirado en el hecho de que  $\ddot{\theta} = (g/a) \sin \theta$  se asemeja a la ecuación de un péndulo simple (con  $\ddot{\theta} \rightarrow \ddot{\theta}$  y  $\theta \rightarrow -\theta$ ) de oscilaciones no necesariamente pequeñas, en tal problema  $\dot{\theta}^2$  se puede obtener por conservación de la energía mecánica de lo cual se obtendría una expresión similar a nuestro ansatz Ec. (4.31).

<sup>8</sup>Naturalmente,  $\dot{\theta}$  debe ser ligeramente distinto de cero para que exista movimiento. Sin embargo, usamos estas condiciones teniendo en cuenta que en ausencia de rozamiento, la velocidad angular se puede aproximar arbitrariamente a cero.

<sup>9</sup>La Ec. (4.32) también se puede obtener así:  $\ddot{\theta} = \frac{d\dot{\theta}}{dt} = \frac{d\dot{\theta}}{d\theta} \frac{d\theta}{dt} = \dot{\theta} \frac{d\dot{\theta}}{d\theta}$ . De la segunda de las Ecs. (4.30) y usando la expresión anterior tenemos que

$$\int_0^\theta \ddot{\theta}' d\theta' = \frac{g}{a} \int_0^\theta \sin \theta' d\theta' \Rightarrow \int_0^\theta \dot{\theta}' d\dot{\theta}' = -\frac{g}{a} \cos \theta' \Big|_0^\theta$$

con lo cual se llega a la Ec. (4.32).

1. Las ecuaciones de Lagrange solo se pueden solucionar teniendo en cuenta las ecuaciones de ligadura, para formar el sistema de  $n + m$  ecuaciones e incógnitas.
2. Obsérvese que el método para calcular las fuerzas generalizadas de ligadura conduce automáticamente a encontrar las coordenadas espúreas. En este caso existe un multiplicador que elimina la coordenada sobrante  $r$ .
3. La fuerza de ligadura pudo ser hallada y está asociada a  $\lambda_1$ . Debe tenerse en cuenta sin embargo, que no hay una relación uno a uno entre multiplicadores y fuerzas generalizadas de ligadura, en realidad cada fuerza generalizada de ligadura es una combinación lineal de todos los  $\lambda_i$  como lo muestran las Ecs. (4.16, 4.22). La relación entre fuerzas reales de ligadura y multiplicadores es más indirecta todavía.

#### 4.4.2. Aro sobre plano inclinado con condición de rodadura

Veamos un ejemplo con la condición de rodadura, como es un aro de radio “ $a$ ” que rueda sin deslizar sobre un plano inclinado que está fijo en el suelo. La ligadura implícita en la rodadura es realmente holónoma<sup>10</sup> pero la técnica de cálculo será idéntica al caso de ligaduras no holónomas de la forma (4.4). La otra ligadura de que el aro está sobre el plano inclinado está implícita en la escogencia de las coordenadas generalizadas

Definimos el eje  $x$  paralelo a la superficie del plano inclinado hacia abajo y un ángulo  $\theta$  que barre el radio vector del aro. Como coordenadas generalizadas se escoge  $x, \theta$ . La condición de rodadura provee una ligadura entre ambas

$$dx - a d\theta = 0 \quad (4.33)$$

la energía cinética tiene una parte traslacional y una rotacional

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}I\dot{\theta}^2 = \frac{1}{2}M(\dot{x}^2 + K^2\dot{\theta}^2)$$

escribirlo en términos del radio de giro  $K$ , permite resolver de una vez el mismo problema con figuras tales como el aro, la esfera, el cilindro, el disco e incluso figuras como el elipsoide de revolución (ver ejercicio 3, Pág. 67). De momento haremos  $K = a$  (aro con densidad constante). La energía potencial es

$$V = Mg(l - x) \sin \phi_0$$

siendo  $l$  la longitud del plano inclinado y  $\phi_0$  su inclinación. El Lagrangiano queda

$$L = \frac{1}{2}M(\dot{x}^2 + a^2\dot{\theta}^2) - Mg(l - x) \sin \phi_0 \quad (4.34)$$

con lo cual tenemos

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) &= M\ddot{x} \quad ; \quad \frac{\partial L}{\partial x} = Mg \sin \phi_0 \\ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}} \right) &= Ma^2\ddot{\theta} \quad ; \quad \frac{\partial L}{\partial \theta} = 0 \end{aligned}$$

dado que hay una ecuación de ligadura, se requiere solo un multiplicador de Lagrange. Comparando (4.33) con (4.4) los coeficientes que acompañan a la ligadura son

$$a_{1\theta} = -a \quad ; \quad a_{1x} = 1 \quad ; \quad a_{1t} = 0 \quad (4.35)$$

las fuerzas generalizadas de ligadura se obtienen de (4.16) y (4.35)

$$Q_x^L = -\lambda_1 a_{1x} = -\lambda_1; \quad Q_\theta^L = -\lambda_1 a_{1\theta} = a\lambda_1$$

<sup>10</sup>Esto se debe a que esta es una condición de rodadura en donde el aro se restringe a moverse en un plano. Cuando esta restricción deja de ser cierta, la ligadura es realmente no-holónoma, como se aprecia en los ejercicios 1, 2, Pág. 66.

de modo que las ecuaciones de Lagrange junto con las ecuaciones de ligadura quedan

$$M\ddot{x} - Mg \sin \phi_0 = -\lambda_1 \quad ; \quad Ma^2\ddot{\theta} = a\lambda_1 \quad ; \quad a\dot{\theta} = \dot{x} \quad (4.36)$$

constituye un conjunto de tres ecuaciones para  $x, \theta, \lambda_1$ . Diferenciando la ligadura respecto al tiempo se obtiene  $a\ddot{\theta} = \ddot{x}$ , sustituyendo en la segunda de las Ecs. (4.36) se tiene que  $M\ddot{x} = \lambda_1$  al sustituir esto en la primera de las Ecs. (4.36) queda

$$\ddot{x} = \frac{g \sin \phi_0}{2}$$

con lo cual

$$\ddot{\theta} = \frac{g \sin \phi_0}{2a} \quad ; \quad \lambda_1 = \frac{Mg \sin \phi_0}{2} \quad (4.37)$$

puesto que hay un solo multiplicador, podemos obtener la magnitud de la fuerza de ligadura usando una sola fuerza generalizada de ligadura, digamos  $Q_x^L$ . La fuerza (real) de ligadura está dada por  $\mathbf{F}^L = \mathbf{N} + \mathbf{F}_r$  siendo  $\mathbf{N}$  la normal y  $\mathbf{F}_r$  la fuerza de rozamiento estático. Utilizando las ecuaciones (4.16, 4.17) tenemos que

$$Q_x^L = \mathbf{F}^L \cdot \frac{\partial \mathbf{r}}{\partial x} = (\mathbf{N} + \mathbf{F}_r) \cdot \frac{\partial \mathbf{r}}{\partial x} = -\lambda_1 a_{1x}$$

es fácil ver que  $\mathbf{r} = x\mathbf{u}_x$ , por tanto  $\partial_x \mathbf{r} = \mathbf{u}_x$ , utilizando esto y la Ec. (4.35) resulta<sup>11</sup>

$$(\mathbf{N} + \mathbf{F}_r) \cdot \mathbf{u}_x = -\lambda_1$$

ahora teniendo en cuenta que  $\mathbf{N} = N\mathbf{u}_y$  y que  $\mathbf{F}_r = -\|\mathbf{F}_r\| \mathbf{u}_x$  se tiene que

$$-\|\mathbf{F}_r\| = -\lambda_1 \Rightarrow F_r = \frac{Mg \sin \phi_0}{2}$$

donde hemos usado la Ec. (4.37). Nótese que hemos obtenido información sobre la fuerza de rozamiento pero no sobre la normal. Esto era de esperarse, ya que la fuerza de rozamiento estático es la que genera la condición de rodadura, que fue la ecuación que usamos como ligadura.

#### 4.4.3. Esfera en un hueco cilíndrico

Una esfera de radio  $\rho$  está restringida a rodar sin deslizar en la mitad inferior de la superficie interna de un hueco cilíndrico de radio interior  $R$  (ver Fig. 4.1). Determine el Lagrangiano, la ecuación de ligadura y las ecuaciones de Lagrange con multiplicadores indeterminados.

Utilizaremos las coordenadas  $\theta$  y  $\phi$  indicadas en la figura 4.1. Estas coordenadas no son independientes ya que la condición de rodadura establece una relación entre ellas. El centro de masa de la esfera se mueve en un arco circular con radio  $R - \rho$  y velocidad angular  $\dot{\theta}$ , de modo que la velocidad al cuadrado del centro de masa de la esfera es  $v^2 = [(R - \rho)\dot{\theta}]^2$ . Las energías cinética y potencial en estas coordenadas se escriben en la forma

$$T = \frac{1}{2}m [(R - \rho)\dot{\theta}]^2 + \frac{1}{2}I\dot{\phi}^2 \quad ; \quad V = mg[R - (R - \rho) \cos \theta]$$

siendo  $m$  la masa de la esfera y tomamos  $V = 0$  en el punto más bajo del cilindro hueco<sup>12</sup>.  $I$  es el momento de inercia de la esfera con respecto a su diámetro, cuyo valor es  $(2/5)m\rho^2$ , el Lagrangiano queda entonces en la forma

$$L = T - V = \frac{1}{2}m(R - \rho)^2\dot{\theta}^2 + \frac{1}{5}m\rho^2\dot{\phi}^2 - mg[R - (R - \rho) \cos \theta]$$

<sup>11</sup>A priori podemos pensar que la posición debe ser  $\mathbf{r} = x\mathbf{u}_x + a\mathbf{u}_y$  que es la posición del centro de masa. Sin embargo, la fuerza de rozamiento se aplica sobre el punto de contacto sobre la superficie de modo que  $\mathbf{r} = x\mathbf{u}_x$  es el punto de aplicación de dicha fuerza (y también de la normal). En todo caso, esta diferencia no altera las derivadas.

<sup>12</sup>Podríamos tomar el cero de potencial en el punto más bajo del centro de masa de la esfera, pero esto supone un corrimiento del potencial en una constante, lo cual es físicamente irrelevante.

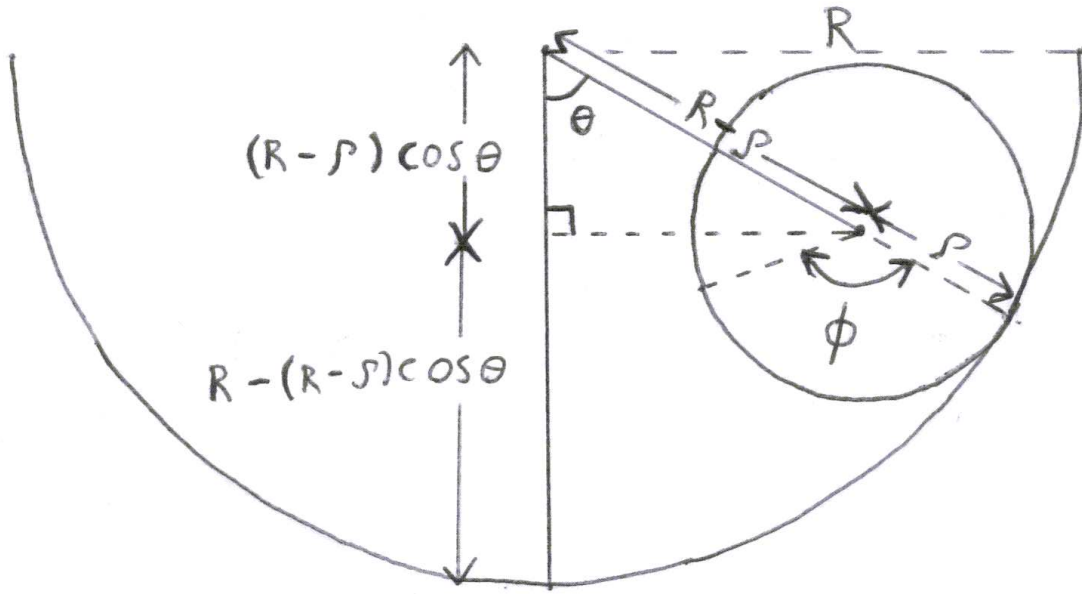


Figura 4.1: Esfera de radio  $\rho$ , que rueda sin deslizar en la mitad inferior de una superficie cilíndrica de radio  $R$ .

la condición de rodadura nos da la ligadura

$$f(\theta, \phi) = R\theta - \rho\phi = 0 \quad (4.38)$$

existe entonces un solo multiplicador de Lagrange y las ecuaciones de Lagrange con multiplicadores quedan

$$\frac{\partial L}{\partial \theta} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}} \right) = \lambda \frac{\partial f}{\partial \theta} ; \quad \frac{\partial L}{\partial \phi} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\phi}} \right) = \lambda \frac{\partial f}{\partial \phi}$$

que explícitamente nos da

$$-mg(R - \rho) \sin \theta - m(R - \rho)^2 \ddot{\theta} = \lambda R \quad (4.39)$$

$$-\frac{2}{5}m\rho^2 \ddot{\phi} = -\lambda\rho \quad (4.40)$$

de (4.40) se obtiene

$$\lambda = \frac{2}{5}m\rho \ddot{\phi} \quad (4.41)$$

y usando (4.38), encontramos

$$\lambda = \frac{2}{5}m\rho \left( \frac{R}{\rho} \ddot{\theta} \right) = \frac{2}{5}mR\ddot{\theta} \quad (4.42)$$

sustituyendo (4.42) en (4.39) se encuentra la ecuación de movimiento con respecto a  $\theta$ .

$$\begin{aligned} -mg(R - \rho) \sin \theta - m(R - \rho)^2 \ddot{\theta} &= \left( \frac{2}{5}mR\ddot{\theta} \right) R \\ -g(R - \rho) \sin \theta &= \left[ \frac{2}{5}R^2 + (R - \rho)^2 \right] \ddot{\theta} \\ \ddot{\theta} &= -\frac{g(R - \rho) \sin \theta}{\left[ \frac{2}{5}R^2 + (R - \rho)^2 \right]} \\ \ddot{\theta} &= -\omega^2 \sin \theta \end{aligned}$$

$\omega$  es la frecuencia de pequeñas oscilaciones dada por

$$\omega = \sqrt{\frac{g(R-\rho)}{\left[\frac{2}{5}R^2 + (R-\rho)^2\right]}}$$

la ecuación para  $\theta$  es la de un péndulo simple de amplias oscilaciones. Una vez resuelto  $\theta(t)$ , la ecuación de  $\phi$  se obtiene directamente de (4.38), así mismo la solución para el multiplicador se obtiene de (4.42). ¿Cuál es la asociación entre fuerzas de ligadura y el multiplicador en este problema?.

#### 4.5. Características básicas de una formulación variacional

Una ventaja notable de la formulación variacional es su independencia con respecto a un cambio en las coordenadas generalizadas<sup>13</sup>, ya que el Lagrangiano es independiente de esta escogencia. Por otro lado a partir del principio de Hamilton se puede ver fácilmente porque la adición de una derivada total de la forma  $dF(q, t)/dt$  en el Lagrangiano no afecta la dinámica, cuando este término se integra en el tiempo, el resultado solo depende de los puntos inicial y final los cuales son fijos. Esto equivale a añadir una constante a la acción, que claramente no afecta la condición de estacionaridad.

Muchas teorías de la Física mas allá de la mecánica subyacen en un principio variacional. Con frecuencia ocurre que la necesidad de modificar el contenido Físico de una teoría sugiere la forma de hacer el mismo cambio en otras teorías. Tal es el caso de la cuantización de partículas que indicó como construir posteriormente el proceso de cuantización de los campos, basados en un principio variacional y una formulación Lagrangiana.

#### 4.6. Principio variacional para Lagrangianos que contienen a $\ddot{q}$ (opcional)

Ciertos problemas de la mecánica clásica están descritos por Lagrangianos que contienen derivadas temporales de  $q_i$  más altas que la primera, este es el caso en ciertas aplicaciones de la teoría del caos. Es usual denominar como mecánica generalizada a este tipo de problemas. En particular cuando el Lagrangiano es de la forma  $L(q, \dot{q}, \ddot{q}, t)$  se suele hablar de “mecánica de la sacudida”. Asumiremos que para un Lagrangiano de la forma  $L(q, \dot{q}, \ddot{q}, t)$ , el principio de Hamilton se cumple con variación cero en los extremos tanto en las  $q_i$  como en las  $\dot{q}_i$ , en cuyo caso hablaremos de un principio variacional de Hamilton extendido.

Veremos entonces como son las ecuaciones de movimiento para un Lagrangiano de la forma

$$L = L(q_i, \dot{q}_i, \ddot{q}_i, t) \quad (4.43)$$

bajo el postulado de que se cumple el principio variacional de Hamilton extendido. La acción queda entonces de la forma

$$I = \int_1^2 L(q_i, \dot{q}_i, \ddot{q}_i, t) dt$$

y se tiene que

$$\frac{\partial I}{\partial \alpha} d\alpha = \int_1^2 \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i} \frac{\partial q_i}{\partial \alpha} d\alpha + \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \alpha} d\alpha + \frac{\partial L}{\partial \ddot{q}_i} \frac{\partial \ddot{q}_i}{\partial \alpha} d\alpha \right) dt \quad (4.44)$$

Asumiremos suma sobre índices repetidos cuando los índices aparezcan. De momento omitiremos los índices para simplificar los cálculos. En analogía con la Ec. (3.17, 3.18) tenemos

$$\frac{\partial I}{\partial \alpha} d\alpha = \delta I \quad ; \quad \delta q = \frac{\partial q}{\partial \alpha} d\alpha$$

con lo cual la Ec. (4.44) queda

<sup>13</sup>Debe aclararse sin embargo, que el espacio de configuraciones y la trayectoria en dicho espacio, sí dependen de las coordenadas generalizadas empleadas.



$$\delta I = \int_1^2 \left( \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial \alpha} d\alpha + \frac{\partial L}{\partial \ddot{q}} \frac{\partial \ddot{q}}{\partial \alpha} d\alpha \right) dt \quad (4.45)$$

Integrando por partes el término de la mitad nos da en analogía con la Ec. (3.11)

$$\int_1^2 \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial \alpha} dt = \int_1^2 \frac{\partial L}{\partial \dot{q}} \frac{\partial^2 q}{\partial \alpha \partial t} dt = \left. \frac{\partial L}{\partial \dot{q}} \frac{\partial q}{\partial \alpha} \right|_1^2 - \int_1^2 \frac{\partial q}{\partial \alpha} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) dt \quad (4.46)$$

el primer término a la derecha es cero debido a la condición de extremo fijo en las  $q_i$ . sustituyendo (4.46) en la expresión (4.45) resulta

$$\delta I = \int_1^2 \left[ \frac{\partial L}{\partial q} \delta q - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) \delta q + \frac{\partial L}{\partial \ddot{q}} \frac{\partial \ddot{q}}{\partial \alpha} d\alpha \right] dt \quad (4.47)$$

donde hemos usado de nuevo la definición  $\delta q = (\partial q / \partial \alpha) d\alpha$ . El último integrando requiere dos integraciones por partes

$$\int_1^2 \frac{\partial L}{\partial \ddot{q}} \frac{\partial \ddot{q}}{\partial \alpha} dt = \int_1^2 \frac{\partial L}{\partial \ddot{q}} \frac{\partial^2 \dot{q}}{\partial \alpha \partial t} dt = \left. \frac{\partial L}{\partial \ddot{q}} \frac{\partial \dot{q}}{\partial \alpha} \right|_1^2 - \int_1^2 \frac{\partial \dot{q}}{\partial \alpha} \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}} \right) dt$$

y usando la condición de extremo fijo en las  $\dot{q}_i$ , vemos que se elimina el primer término a la derecha. Una segunda integración por partes nos da

$$\begin{aligned} - \int_1^2 \frac{\partial \dot{q}}{\partial \alpha} \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}} \right) \right] dt &= - \int_1^2 \frac{\partial^2 q}{\partial t \partial \alpha} \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}} \right) \right] dt \\ &= - \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}} \right) \right] \frac{\partial q}{\partial \alpha} \Big|_1^2 - \int_1^2 - \frac{\partial q}{\partial \alpha} \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}} \right) dt \end{aligned}$$

el primer término a la derecha se elimina por la condición de extremo fijo en los  $q_i$ . Por tanto

$$\int_1^2 \frac{\partial L}{\partial \ddot{q}} \frac{\partial \ddot{q}}{\partial \alpha} dt = \int_1^2 \frac{\partial q}{\partial \alpha} \left[ \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}} \right) \right] dt \quad (4.48)$$

reemplazando (4.48) en (4.47) y usando la definición de  $\delta q$  se tiene que

$$\delta I = \int_1^2 \left\{ \frac{\partial L}{\partial q} \delta q - \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) \right] \delta q + \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}} \right) \delta q \right\} dt$$

factorizando  $\delta q$ 's, y colocando los índices de nuevo, resulta

$$\delta I = \int_1^2 \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{d^2}{dt^2} \frac{\partial L}{\partial \ddot{q}_i} \right] \delta q_i dt = 0$$

puesto que las  $q_i$  son independientes, las variaciones  $\delta q_i$  son independientes y podemos aplicar el lema fundamental del cálculo variacional y ver que  $\delta I = 0$  requiere que tanto el integrando como los coeficientes de  $\delta q_i$  se anulen separadamente, llegando entonces a las ecuaciones de movimiento:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}_i} \right) = 0 \quad ; \quad i = 1, 2, \dots, n. \quad (4.49)$$

Nótese que para utilizar el principio variacional de Hamilton extendido, el espacio de configuraciones debe ser de dimensión  $2n$ , con  $q_i$  y  $\dot{q}_i$  en los ejes, ya que la condición completa de variación cero en los extremos requiere esta extensión del espacio. Además estas condiciones de variación cero conducen a la invarianza de las ecuaciones de movimiento, bajo el siguiente tipo de transformaciones gauge del Lagrangiano

$$L' = L + \frac{dF(q, \dot{q}, t)}{dt} \quad (4.50)$$

Tomemos como ejemplo sencillo de aplicación el siguiente Lagrangiano

$$L = -\frac{1}{2}mq\ddot{q} - \frac{k}{2}q^2 \quad (4.51)$$

para el cual tenemos

$$\frac{\partial L}{\partial q} = -\frac{1}{2}m\ddot{q} - kq \quad ; \quad -\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) = 0 \quad ; \quad \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}} \right) = \frac{d^2}{dt^2} \left( -\frac{1}{2}mq \right) = -\frac{1}{2}m\ddot{q}$$

al reemplazar estas relaciones en las ecuaciones de movimiento (4.49) tenemos

$$-m\ddot{q} - kq = 0$$

Que corresponde a la ley de Hooke. Nótese que el Lagrangiano (4.51) a pesar de su apariencia extraña, nos da las ecuaciones del oscilador armónico cuando usamos este “formalismo de sacudida” de las ecuaciones de Lagrange. Para entender porqué, utilizaremos la invarianza gauge en (4.50), para obtener el Lagrangiano (4.51) a partir del Lagrangiano usual para el oscilador armónico simple  $L_{SHO}$

$$\begin{aligned} L(q, \dot{q}, \ddot{q}, t) &= L_{SHO}(q, \dot{q}, t) + \frac{d}{dt} \left( -\frac{mq\dot{q}}{2} \right) = \frac{m\dot{q}^2}{2} - \frac{kq^2}{2} - \frac{mq\dot{q}}{2} - \frac{m\dot{q}^2}{2} \\ L(q, \dot{q}, \ddot{q}, t) &= -\frac{1}{2}mq\ddot{q} - \frac{k}{2}q^2 \end{aligned}$$

es muy importante enfatizar que los Lagrangianos  $L$  y  $L_{SHO}$  son equivalentes solo bajo el principio de Hamilton **extendido** con variación cero en  $q, \dot{q}$ . De otra forma el Gauge (4.50) ya no sería válido.

De lo anterior se vé que se requiere un principio variacional específico o equivalentemente, un tipo específico de ecuaciones de movimiento, para que un Lagrangiano dado tenga toda la información física del sistema. Así mismo, las transformaciones gauge posibles para el Lagrangiano dependen del principio variacional que se postule, o de las ecuaciones de movimiento que se asuman.

## 4.7. Ejercicios

- Supongamos un disco de radio  $R$  que rueda sobre un plano horizontal  $XY$ , de manera que permanece siempre vertical. Sean  $(x, y)$  las coordenadas del centro del disco. Definimos también un ángulo de rotación  $\phi$  alrededor del eje del disco, y un ángulo  $\theta$  entre el eje  $X$  y el eje del disco (ver Fig. 4.2). Si el disco rueda sin deslizar, demuestre que la condición de rodadura se manifiesta en una ligadura de la forma

$$dx - R \sin \theta \, d\phi = 0 \quad ; \quad dy + R \cos \theta \, d\phi = 0 \quad (4.52)$$

- Supongamos que tenemos un conjunto de ecuaciones diferenciales lineales de ligadura del tipo

$$\sum_{k=1}^n g_k(x_1, \dots, x_n) \, dx_k = 0 \quad (4.53)$$

este tipo de ligaduras son holónomas solo si existe una función integrante  $f(x_1, \dots, x_n)$  que convierta a estas ecuaciones en diferenciales exactas i.e.

$$\sum_{k=1}^n g_k(x_1, \dots, x_n) \, dx_k = df = \sum_{k=1}^n \frac{\partial f}{\partial x_k} dx_k \quad \Rightarrow \quad g_k = \frac{\partial f}{\partial x_k} \quad (4.54)$$

multiplicando la última ecuación por  $f$  y derivando el producto parcialmente con respecto a  $x_j$  se obtiene

$$\frac{\partial}{\partial x_j} (fg_k) = \frac{\partial f}{\partial x_j} \frac{\partial f}{\partial x_k} + f \frac{\partial^2 f}{\partial x_j \partial x_k} \quad (4.55)$$

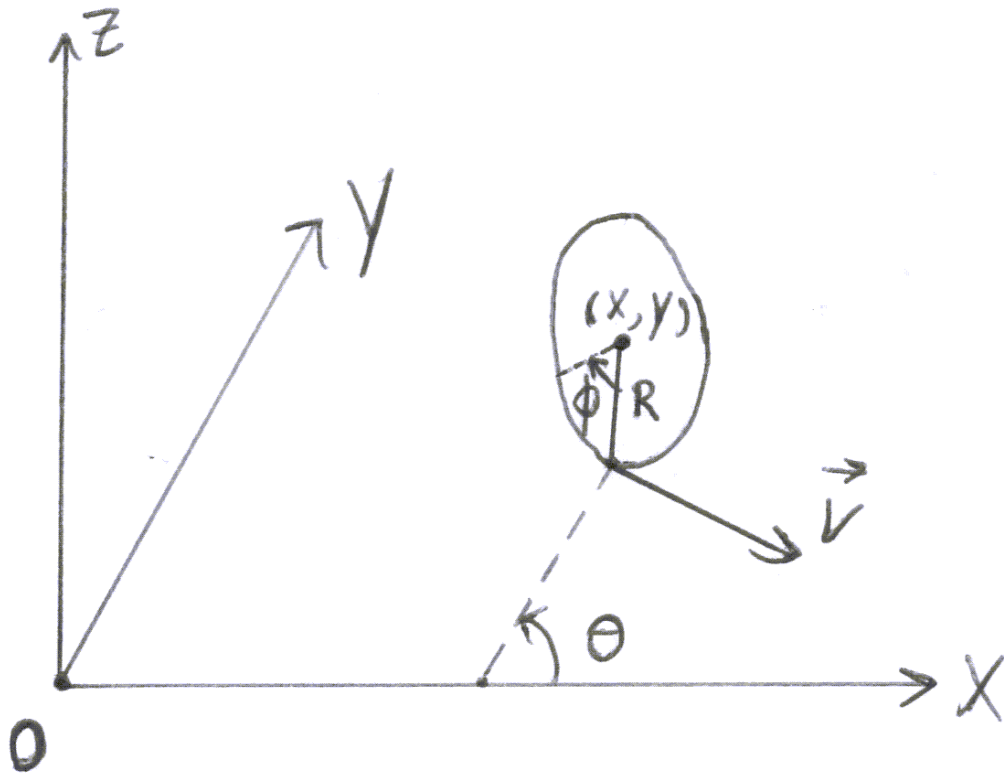


Figura 4.2: Disco vertical que rueda sin deslizar sobre el plano  $XY$ .

invirtiendo el rol de los índices  $j, k$ , se obtiene una ecuación similar, y el miembro derecho de tal ecuación es idéntico al miembro derecho de la Ec. (4.55) si las segundas derivadas parciales son continuas. Por tanto, la función integrante debe cumplir la condición<sup>14</sup>

$$\frac{\partial (fg_k)}{\partial x_j} = \frac{\partial (fg_j)}{\partial x_k} ; \quad j, k = 1, \dots, n \quad (4.56)$$

demuestre que no se puede encontrar una función integrante para la ligadura de rodadura Ec. (4.52). Sin embargo, tal función integrante sí existe cuando  $\theta$  es constante en la Ec. (4.52), como ocurre por ejemplo con un disco que rueda sobre un plano inclinado.

3. Resolver el ejercicio de la sección 4.4.2, sustituyendo el aro por una esfera, un cilindro, un disco y un elipsoide de revolución. Asuma constante la densidad de cada figura.
4. Demuestre que el gauge (4.50), deja invariantes las ecuaciones de movimiento (4.49), que se derivan del principio de Hamilton extendido, asociado al “formalismo de la sacudida”. Demuéstrelo (a) apelando directamente a las ecuaciones diferenciales, (b) apelando al principio variacional de Hamilton extendido.
5. En la sección 2.8.6 Pág. 32, se estudió la dinámica de un aro que rueda sin deslizar sobre una cuña que desliza sin rozamiento sobre el suelo. Trate la ligadura de rodadura utilizando el método de multiplicadores de Lagrange. Contraste los resultados con los obtenidos en la sección 4.4.2, Pág. 61, en la cual la cuña está fija en el suelo.
6. Una esfera de masa  $m$  y radio  $r$  rueda sin deslizar sobre un cilindro fijo de radio  $R$ , como se indica en la Fig. 4.3. Si la esfera comienza a rodar sin deslizar desde el punto más alto del cilindro y partiendo

<sup>14</sup>En las condiciones (4.54, 4.56) podemos agregar la coordenada temporal digamos  $x_{n+1} \equiv t$ , y las Ecs. (4.54, 4.56) serían válidas para  $j, k = 1, \dots, n + 1$ . Nótese que esta es la estructura de las Ecs. (4.4), Pág. (55).

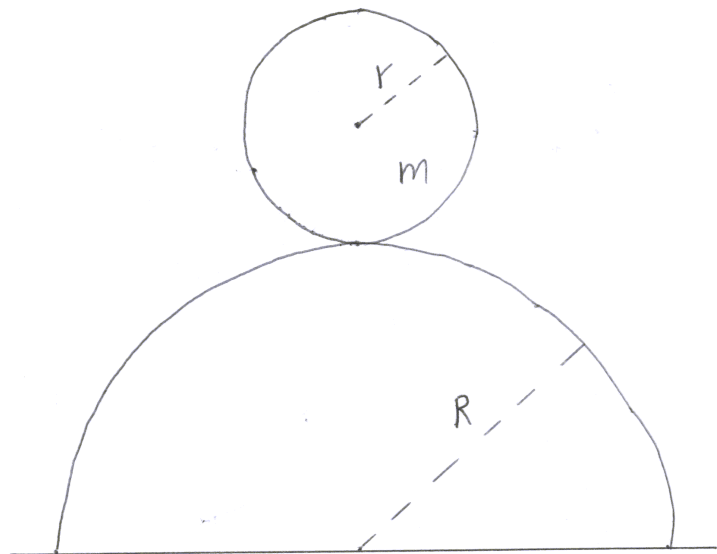


Figura 4.3: Esfera de masa  $m$  y radio  $r$  que rueda sin deslizar sobre un cilindro fijo de radio  $R$ .

del reposo. (a) Plantee las ecuaciones de Lagrange con multiplicadores que den cuenta de la fuerza de ligadura del cilindro sobre la esfera. (b) Encuentre el punto donde la esfera se separa del cilindro. (c) Resuelva el mismo problema reemplazando la esfera por un aro y por un cilindro, cada uno con masa  $m$  y radio  $r$ .

# Simetrías y cantidades conservadas en el formalismo de Lagrange

## 5.1. Teoremas de conservación y propiedades de simetría

Las estrategias planteadas hasta el momento nos permiten encontrar las ecuaciones diferenciales que determinan la dinámica de un sistema con  $n$  grados de libertad. Como estas ecuaciones son de segundo orden en el tiempo, se requieren  $2n$  constantes de integración que usualmente se determinan con las condiciones iniciales en las coordenadas y velocidades generalizadas. Desafortunadamente, la mayoría de problemas no son integrables completamente. No obstante, hemos aprendido del formalismo Newtoniano, que incluso para sistemas no integrables es posible extraer mucha información valiosa del sistema aunque las ecuaciones de movimiento no estén completamente resueltas. En general, esta información se traduce en principios de conservación que debemos reexaminar en este nuevo formalismo.

La solución completa de una ecuación de segundo orden requiere formalmente de dos procesos de integración, con frecuencia ocurre que podemos hacer un “primer proceso de integración” pero el último proceso de integración no se puede llevar a cabo fácilmente. En otras palabras, podemos obtener ecuaciones de la forma

$$f(q, \dot{q}, t) = cte$$

que son ecuaciones diferenciales de primer orden. Estas estructuras se conocen como primeras integrales de las ecuaciones de movimiento. Mucha información se puede extraer de estas primeras integrales, en particular las leyes de conservación.

Consideremos un sistema de partículas puntuales bajo la influencia de fuerzas que se derivan de potenciales que solo dependen de la posición. En este caso podemos escribir

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}_i} &\equiv \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum_k \frac{1}{2} m_k (\dot{x}_k^2 + \dot{y}_k^2 + \dot{z}_k^2) \\ \frac{\partial L}{\partial \dot{x}_i} &= m_i \dot{x}_i = p_{x_i} \end{aligned}$$

que corresponde a la componente  $x$  del momento lineal de la partícula  $i$ -ésima. Esto sugiere la forma de elaborar el concepto de momento generalizado cuando usamos coordenadas generalizadas. El momento generalizado  $p_j$  asociado a la coordenada generalizada  $q_j$  se define como

$$p_j \equiv \frac{\partial L}{\partial \dot{q}_j} \tag{5.1}$$

por razones que veremos posteriormente, a  $p_j$  también se le conoce como momento canónicamente conjugado

a  $q_j$ . Nótese que  $p_j$  no tiene necesariamente dimensiones de momento lineal<sup>1</sup>. Cuando tenemos potenciales dependientes de la velocidad incluso en coordenadas cartesianas el momento generalizado difiere del momento mecánico. Un ejemplo notable es el de un conjunto de partículas en un campo electromagnético, cuyo Lagrangiano es<sup>2</sup>

$$L = \sum_{i=1}^N \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - \sum_{i=1}^N q_i \phi(\mathbf{r}_i) + \sum_{i=1}^N q_i \mathbf{A}(\mathbf{r}_i) \cdot \dot{\mathbf{r}}_i$$

$q_i$  en este caso denota carga eléctrica. El momento generalizado es

$$p_{x_i} = \frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i + q_i A_x(\mathbf{r}_i) \quad (5.2)$$

que es el momento mecánico asociado a la partícula mas un término adicional que depende del campo.

En algunas ocasiones el Lagrangiano no depende de una cierta coordenada  $q_j$ , pero sí depende de su velocidad generalizada asociada  $\dot{q}_j$ . En este caso se dice que  $q_j$  es una coordenada **cíclica** o **ignorable**, y la ecuación de Lagrange asociada a esta coordenada se reduce a

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) = 0 \quad (5.3)$$

de la definición de momento generalizado se observa que

$$\frac{dp_j}{dt} = 0 \Rightarrow p_j = cte \quad (5.4)$$

de lo cual resulta un teorema de conservación: **el momento generalizado canónicamente conjugado a una coordenada cíclica se conserva**. Vale recalcar que la validez de este teorema depende de que las coordenadas generalizadas sean independientes entre sí, ya que de ello dependen las ecuaciones de Lagrange (sin multiplicadores  $\lambda$ ). Por ejemplo, podemos apreciar en la Ec. (4.34), que en el ejemplo del aro que rueda sin deslizar sobre un plano inclinado fijo en el suelo, el Lagrangiano no contiene a la coordenada  $\theta$ , sin embargo la cantidad  $p_\theta$  no es una constante de movimiento debido a que el ángulo aparece en la ecuación de ligadura  $dx - a d\theta = 0$ .

Nótese otra diferencia importante entre el concepto de momento mecánico y el de momento generalizado. El primero se asocia a una partícula y el segundo se asocia a una coordenada generalizada. Por otro lado, ya se ha mencionado que una coordenada generalizada no necesariamente está asociada a una partícula y por tanto el momento conjugado tampoco lo estará.

Los momentos generalizados conservados constituyen primeras integrales de movimiento, pues al pasar de (5.3) a (5.4) hemos pasado de una ecuación diferencial de segundo orden a una de primer orden, lo cual equivale a hacer una integración. Formalmente, la velocidad generalizada de una coordenada cíclica puede ser reemplazada por el momento conservado, de modo que la variable cíclica desaparece completamente del Lagrangiano, este procedimiento desarrollado por Routh, será discutido en el capítulo de formulación Hamiltoniana sección 6.9.

Las condiciones de conservación del momento generalizado son mas generales que las que se derivan en mecánica Newtoniana para el momento lineal y angular. Por ejemplo, un momento conservado aparece incluso cuando la ley de acción y reacción no es válida. En particular, este es el caso cuando están presentes fuerzas electromagnéticas. Supongamos que tenemos una partícula inmersa en un campo electromagnético, tal que los potenciales electrodinámicos no dependen de  $x$ . De esta forma  $x$  no aparece en el Lagrangiano (aunque sí aparece  $\dot{x}$ ) y el momento canónico  $p_x$  viene dado por la Ec. (5.2)

$$p_x = m\dot{x} + qA_x = cte$$

<sup>1</sup>Sin embargo, es claro de la definición (5.1), que  $p_j \dot{q}_j$  tiene dimensiones de energía. Por tanto,  $p_j q_j$  tiene dimensiones de energía por tiempo, es decir de momento angular.

<sup>2</sup>Nótese que cuando al sistema de partículas le añadimos el campo, no es necesario diferenciar para una partícula entre la contribución debida a las fuerzas internas y las externas, ya que ambas contribuciones están contenidas en el campo.

De la teoría electrodinámica clásica, es bien conocido que cuando  $\mathbf{A}$  y  $\phi$  no dependen de  $x$ , la cantidad  $qA_x$  es la componente  $x$  del momento lineal del campo electromagnético asociado con la carga  $q$ , el momento mecánico  $m\dot{x}$  está asociado únicamente a la carga. Por tanto, el momento canónico  $p_x$  es una cantidad mixta que está asociada tanto a la carga como al campo, esto no supone ninguna contradicción ya que como dijimos, el momento canónico está realmente asociado a la coordenada y no a la partícula (ni al campo).

**Example 2** Para el problema de la sección 2.8.6 del aro que rueda sin deslizar sobre un plano inclinado, que a su vez se desliza sobre el suelo sin rozamiento, el Lagrangiano está dado por la Ec. (2.51) Pág. 33

$$L = m\dot{S}^2 + \frac{1}{2}(m + M)\dot{\xi}^2 + m\dot{\xi}\dot{S}\cos\alpha - mg[R\cos\alpha + (l - S)\sin\alpha] \quad (5.5)$$

Se observa que en este Lagrangiano la coordenada  $\xi$  es cíclica, y por tanto su momento canónicamente conjugado es constante. Esto nos conduce a una primera integral de movimiento (ecuación diferencial de primer orden)

$$p_\xi = \frac{\partial L}{\partial \dot{\xi}} = (m + M)\dot{\xi} + m\dot{S}\cos\alpha = K$$

esta expresión se puede encontrar también al integrar una de las ecuaciones de Lagrange que resultaron en el problema, Ec. (2.53), Pág. 33.

Para el anterior ejemplo, es importante mencionar que el Lagrangiano 5.5, está escrito en coordenadas realmente independientes y por esta razón el momento conjugado  $p_\xi$  es realmente una constante de movimiento. Ya mencionamos que en un problema similar, obtuvimos un Lagrangiano Ec. (4.34), que contiene una coordenada cíclica que no conduce a la conservación de su momento canónico, en virtud de que las coordenadas generalizadas en tal problema no eran independientes. Por otro lado, el ejemplo anterior también muestra con claridad que la invarianza de un momento conjugado equivale a un proceso de integración que nos lleva a una ecuación diferencial de primer orden. Sin embargo, en este ejemplo particular este proceso de integración es trivial y se podía ejecutar sin mayores dificultades desde las ecuaciones de Lagrange mismas. Es necesario decir no obstante, que el proceso que nos lleva de las ecuaciones de Lagrange (de segundo orden) a primeras integrales de movimiento (ecuaciones de primer orden), puede ser muy complejo en general, y esto justifica la introducción del concepto de coordenada cíclica y de la conservación de su momento conjugado asociado.

Veremos ahora que los teoremas usuales de conservación están contenidos en la regla general sobre coordenadas cíclicas.

### 5.1.1. Momento lineal y coordenadas globales de traslación

Pensemos primero en una coordenada generalizada  $q_j$  cuyo cambio  $dq_j$  represente una traslación del sistema como un todo en cierta dirección<sup>3</sup>. Esto implica que el corrimiento del sistema como un todo debe ser compatible con las ligaduras. En algunos casos las fuerzas externas podrían imponer ligaduras sobre el sistema que impidan esta clase de movimiento, con lo cual  $q_j$  no sería una coordenada generalizada independiente.

Supondremos entonces que (a)  $q_j$  es una coordenada generalizada independiente. Asumiremos además que (b) La coordenada  $q_j$  no aparece en la energía cinética, (c) el potencial no depende de las velocidades generalizadas, esto excluye de la formulación por ejemplo a las fuerzas electromagnéticas. La ecuación de Lagrange asociada a esta coordenada queda

$$\frac{d}{dt} \left( \frac{\partial(T - V)}{\partial \dot{q}_j} \right) - \frac{\partial(T - V)}{\partial q_j} = \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) + \frac{\partial V}{\partial q_j} = 0$$

<sup>3</sup>Por ejemplo, un conjunto de partículas colineales insertadas en una varilla sin masa. Una coordenada  $q_j$  puede ser una coordenada cartesiana cuyo eje va a lo largo de la varilla. Un desplazamiento  $dq_j$  implica que el sistema se mueve como un todo gracias a la ligadura de cuerpo rígido entre las partículas. El corrimiento de una coordenada cartesiana del centro de masa de un sistema es otro ejemplo.

la derivada total en el tiempo del momento generalizado está dada por

$$\dot{p}_j = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) = \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) = -\frac{\partial V}{\partial q_j} = Q_j \quad (5.6)$$

a continuación veremos que esta es la ecuación de movimiento para el momento lineal total a lo largo de  $q_j$ . Es decir que  $Q_j$  es la componente de la fuerza (en el sentido Newtoniano) en la dirección  $q_j$ , en tanto que  $p_j$  es la componente del momento lineal a lo largo de tal dirección. La fuerza generalizada  $Q_j$  viene dada por

$$Q_j = \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (5.7)$$

nótese que al ser  $q_j$  una variable de traslación, sus unidades sí deben ser de longitud y se puede definir unívocamente un vector unitario de traslación  $\mathbf{n}$ , a lo largo de  $dq_j$ . Como  $dq_j$  representa una traslación del sistema a lo largo de cierto eje, la diferencia entre los vectores  $\mathbf{r}_i(q_j)$  y  $\mathbf{r}_i(q_j + dq_j)$  va a lo largo de  $\mathbf{n}$ , y tiene como magnitud  $dq_j$

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \lim_{\Delta q_j \rightarrow 0} \frac{\mathbf{r}_i(q_j + \Delta q_j) - \mathbf{r}_i(q_j)}{\Delta q_j} = \frac{dq_j \mathbf{n}}{dq_j} = \mathbf{n} \quad (5.8)$$

obsérvese que se ha escrito  $\mathbf{n}$  y no  $\mathbf{n}_i$  ya que al ser una traslación del sistema como un todo, cada partícula se desplaza en la misma dirección que las otras. De aquí resulta

$$Q_j = \sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{n} = \mathbf{n} \cdot \mathbf{F}$$

por lo tanto,  $Q_j$  representa la componente de la fuerza total sobre el sistema a lo largo de la dirección  $\mathbf{n}$  de traslación. Para ver el significado del momento canónico, calculemos  $p_j$

$$\begin{aligned} p_j &= \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} = \frac{\partial}{\partial \dot{q}_j} \left( \frac{1}{2} \sum_{i=1}^N m_i \dot{\mathbf{r}}_i^2 \right) = \sum_{i=1}^N m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} \\ p_j &= \sum_{i=1}^N m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \end{aligned}$$

donde hemos usado (2.12). Teniendo en cuenta (5.8) llegamos a

$$p_j = \mathbf{n} \cdot \sum_i m_i \mathbf{v}_i$$

que corresponde a la componente del momento mecánico lineal total del sistema a lo largo de la dirección de traslación. Vemos que en este caso, el momento generalizado asociado a una sola coordenada  $q_j$  resultó estar asociado al momento lineal de todo el sistema, esto no es de extrañar, ya que  $q_j$  representó una traslación para el sistema como un todo. Esto enfatiza el hecho de que el momento generalizado está asociado a una coordenada y no a una partícula. El mismo comentario vale para la fuerza generalizada. En realidad, todas las variables generalizadas en la formulación Lagrangiana están asociadas a coordenadas en lugar de partículas. A este hecho se debe en parte la facilidad de extender este formalismo a campos y sistemas continuos.

Vale la pena mencionar que en el presente tratamiento solo se asumió que el potencial no dependía de la velocidad generalizada  $\dot{q}_j$  pero podría depender de las velocidades generalizadas de otras coordenadas. Por otro lado, para sistemas aislados cuyo potencial no dependa de  $\dot{q}_j$  las demás condiciones se satisfacen (para una dirección arbitraria de la traslación), ya que al no existir fuerzas externas no hay ligaduras externas que signifiquen una violación de las ligaduras cuando el sistema se traslada como un todo. Adicionalmente, debido a la homogeneidad del espacio, una traslación del sistema como un todo es equivalente para un sistema aislado a un corrimiento del origen, de lo cual es claro que la energía cinética no puede depender de  $q_j$ . Nótese que



si el sistema no es aislado, un corrimiento del origen no necesariamente es equivalente a un corrimiento del sistema, ya que el corrimiento del sistema lo puede estar acercando o alejando de ciertos objetos externos, en tanto que un corrimiento en el origen no. Si por ejemplo ubicamos un plano infinito de masa o de carga uniformemente distribuída por ejemplo sobre el plano  $XY$ , es claro que una traslación en  $x$  e  $y$  cumplirá las condiciones aquí expuestas, pero un corrimiento en  $z$  probablemente no.

Si ahora suponemos (d) que la coordenada de traslación  $q_j$  es cíclica,  $q_j$  no podría aparecer en el potencial de modo que

$$-\frac{\partial V}{\partial q_j} \equiv Q_j = 0 = \dot{p}_j$$

lo cual nos da la conservación de la componente del momento lineal total del sistema a lo largo de la dirección de traslación. De modo que si cierta componente de la fuerza (Newtoniana) total del sistema es nula, la correspondiente componente del momento lineal es constante. Esta condición también la cumplen los sistemas aislados para una dirección arbitraria de la traslación, en virtud de la homogeneidad del espacio, por tanto vemos que el momento total del sistema se debe conservar para sistemas aislados cuyo potencial no dependa de las velocidades generalizadas. En este caso hablamos del teorema de conservación del momento lineal. Sin embargo, es un hecho experimental que el momento lineal total de un sistema aislado se conserva incluso si no se cumple dicha condición (aunque con un concepto extendido de momento), con lo cual la conservación del momento adquiere el carácter de principio.

### 5.1.2. Momento angular y coordenadas globales de rotación

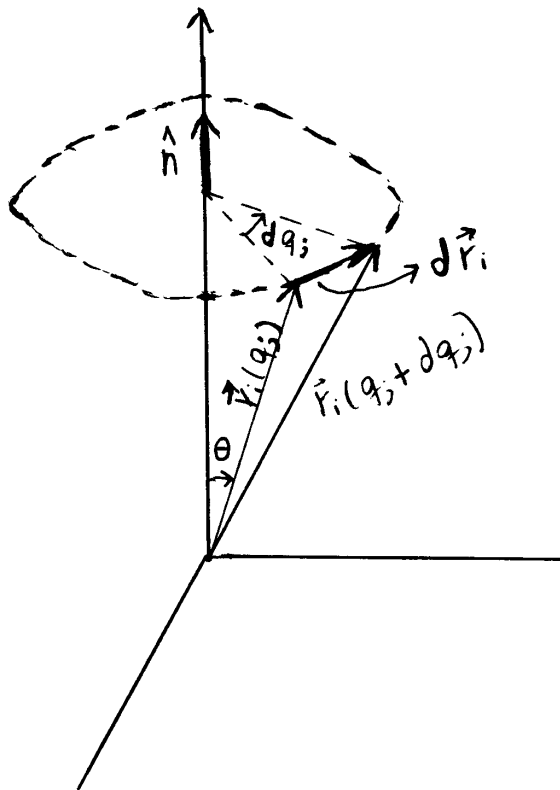


Figura 5.1: Rotación infinitesimal del sistema como un todo, caracterizada por el desplazamiento angular  $dq_j$ .

De manera similar al caso anterior, veremos que si  $dq_j$  corresponde a una rotación del sistema como un todo alrededor de cierto eje, el momento generalizado corresponde al momento angular del sistema a lo largo del

eje de rotación en tanto que  $Q_j$  corresponde a la componente del torque en la dirección de dicho eje. Cuando la variable  $q_j$  se vuelve cíclica se llega a la conservación del momento angular.

Asumiremos las mismas condiciones del caso anterior pero para una coordenada  $q_j$  que produce la rotación del sistema como un todo alrededor de un eje fijo. La Fig. 5.1 muestra la rotación en una cantidad  $dq_j$  de la partícula  $i$ -ésima del sistema, donde por comodidad y sin pérdida de generalidad se ha colocado el eje de rotación a lo largo del eje  $Z$ .  $\mathbf{r}_i(q_j)$  es la posición de esa partícula cuando la coordenada generalizada vale  $q_j$  y  $\mathbf{r}_i(q_j + dq_j)$  es la posición de la partícula  $i$ -ésima cuando la rotación  $dq_j$  se ha realizado,  $\theta$  es el ángulo entre  $\mathbf{r}_i$  y el eje de rotación y  $R_i$  el radio del círculo descrito por la partícula  $i$  alrededor del eje de rotación. Como hemos asumido que  $T$  es independiente de  $q_j$  y que  $V$  es independiente de  $\dot{q}_j$  llegamos de nuevo a (5.6), y la fuerza generalizada  $Q_j$  está dada de nuevo por (5.7) pero en este caso la derivada adquiere un sentido diferente

$$\|d\mathbf{r}_i\| = R_i dq_j = r_i \sin \theta dq_j \Rightarrow \left\| \frac{\partial \mathbf{r}_i}{\partial q_j} \right\| = r_i \sin \theta$$

la dirección de  $d\mathbf{r}_i$  y por lo tanto la de  $\partial \mathbf{r}_i / \partial q_j$  es perpendicular a  $\mathbf{n}$  (vector unitario a lo largo del eje de rotación). Adicionalmente dado que  $\mathbf{r}_i$  no cambia de magnitud, se tiene que  $d\mathbf{r}_i$  (y por tanto  $\partial \mathbf{r}_i / \partial q_j$ ) es también perpendicular a  $\mathbf{r}_i$ . Por tanto,  $\partial \mathbf{r}_i / \partial q_j$  es perpendicular a  $\mathbf{n}$  y a  $\mathbf{r}_i$  y su magnitud es  $r_i \sin \theta = \|\mathbf{n}\| \|\mathbf{r}_i\| \sin \theta$ , de lo cual se deduce que

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \mathbf{n} \times \mathbf{r}_i \quad (5.9)$$

la fuerza generalizada queda entonces

$$Q_j = \sum_{i=1}^N \mathbf{F}_i \cdot (\mathbf{n} \times \mathbf{r}_i) = \sum_{i=1}^N \mathbf{n} \cdot (\mathbf{r}_i \times \mathbf{F}_i)$$

donde hemos usado la identidad  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ . Vemos que  $\mathbf{N}_i \equiv \mathbf{r}_i \times \mathbf{F}_i$  es el torque de la partícula  $i$ -ésima, medido con respecto al origen definido en la Fig. 5.1. De modo que

$$Q_j = \mathbf{n} \cdot \sum_{i=1}^N \mathbf{N}_i = \mathbf{n} \cdot \mathbf{N}$$

la fuerza generalizada corresponde entonces a la componente del torque total del sistema a lo largo del eje de rotación<sup>4</sup>. Veamos el momento conjugado

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^N m_i \mathbf{v}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \sum_{i=1}^N m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_{i=1}^N m_i \mathbf{v}_i \cdot (\mathbf{n} \times \mathbf{r}_i) = \sum_{i=1}^N \mathbf{n} \cdot (\mathbf{r}_i \times m_i \mathbf{v}_i)$$

$$p_j = \mathbf{n} \cdot \sum_{i=1}^N \mathbf{L}_i = \mathbf{n} \cdot \mathbf{L}$$

el momento canónico es entonces la componente del momento angular total en la dirección del eje de rotación. Si la variable  $q_j$  se vuelve ignorable, se llega a la conservación de la componente del momento angular total en la dirección del eje de rotación. La variable  $q_j$  es en este caso una variable angular y por tanto, adimensional.

Vemos que las condiciones anteriores se cumplen para un sistema aislado. La discusión es muy similar al caso traslacional salvo que esta vez apelamos a la isotropía del espacio y no a la homogeneidad. Recordemos que la homogeneidad nos dice que la estructura del espacio es la misma si cambiamos el origen del sistema coordenado, en tanto que la isotropía nos menciona que la estructura del espacio se vé igual si hacemos una reorientación de los ejes coordenados (sin cambiar el origen).

<sup>4</sup>Es importante notar que esta afirmación es válida siempre y cuando el origen con respecto al cual se mide el torque, sea tal que el eje de rotación pase por dicho origen, como se aprecia en la Fig. 5.1.

### 5.1.3. Consideraciones generales sobre simetrías asociadas a coordenadas cíclicas y cantidades conservadas

Es sencillo ver que el hecho de que una variable sea cíclica está asociado con algún tipo de simetría del sistema. Por ejemplo, si la traslación del sistema como un todo no afecta al problema, lo que estamos diciendo es que el sistema es **invariante** bajo traslaciones en cierta dirección, y por lo tanto dicha invarianza ante la operación de traslación conduce a la conservación del momento lineal en la dirección de dicha traslación. Por otra parte si un sistema permanece invariante ante una rotación alrededor de cierto eje, la componente del momento angular a lo largo de dicho eje se conserva. Los teoremas de conservación están fuertemente ligados a las simetrías del sistema. Si el sistema es esféricamente simétrico, el momento angular del sistema se conserva en todas direcciones. Estas consideraciones de simetría para llegar a primeras integrales es aplicable incluso a sistemas muy complejos en los cuales no es posible resolver la dinámica completa.

Veamos una aplicación: supongamos que tenemos un sistema de partículas inmerso en un potencial generado por una distribución homogénea de masa, carga etc. que forma un plano infinito (el plano  $XY$ ). Claramente la traslación de este sistema (si es localizado) a lo largo de  $x$  e  $y$  no afecta la dinámica de éste, pero una traslación en  $z$  puede posiblemente afectarlo. Por otro lado también hay una clara invarianza cuando rotamos al sistema alrededor de  $z$ . Sin conocer los detalles del sistema, deducimos que  $p_x, p_y$  y  $L_z$  se conservan. Si en vez del plano  $XY$ , la distribución externa solo ocupa al semiplano infinito  $x \geq 0$ , entonces solo permanece la invarianza de  $p_y$ . La relación estrecha entre las simetrías y las constantes de movimiento, constituye uno de los principios más profundos y fructíferos en la Física y adquiere una dimensión aún mayor a la luz del Teorema de Noether que discutiremos más adelante.

## 5.2. Función energía y conservación de la energía

Por supuesto, también es de esperarse que el teorema de conservación de la energía se pueda obtener del formalismo Lagrangiano, cuando las fuerzas del sistema son derivables de un potencial que solo depende de la posición. Por otro lado, hemos visto que la ausencia de una coordenada generalizada en el Lagrangiano conduce a la conservación de un momento generalizado, es natural entonces preguntarse si la ausencia explícita de la variable tiempo conduce a algún teorema de conservación. La analogía no es tan directa puesto que no hemos asociado un momento generalizado relacionado con la variable tiempo ni tiene sentido una variable  $\dot{t}$ , en ese sentido no podemos hablar del tiempo como una coordenada generalizada. Más bien, el tiempo se considera un parámetro que además de aparecer explícitamente, regula la evolución de las coordenadas.

Al igual que en el caso de los momentos, veremos que del formalismo Lagrangiano sale un teorema de conservación más general que incluye a la conservación de la energía como caso particular.

Consideremos un Lagrangiano que es función de las coordenadas generalizadas, velocidades generalizadas y el tiempo (la dependencia temporal explícita puede provenir de las variaciones de fuentes exteriores, o de ligaduras dependientes del tiempo), con lo cual la derivada total respecto al tiempo es

$$\frac{dL}{dt} = \sum_{j=1}^n \frac{\partial L}{\partial q_j} \frac{dq_j}{dt} + \sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t} \quad (5.10)$$

a partir de las ecuaciones de Lagrange

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) \quad (5.11)$$

y reemplazándolo en (5.10)

$$\begin{aligned} \frac{dL}{dt} &= \sum_{j=1}^n \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) \right] \frac{dq_j}{dt} + \sum_{j=1}^n \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t} \Rightarrow \\ \frac{dL}{dt} &= \sum_{j=1}^n \left[ \frac{d}{dt} \left( \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) \right] + \frac{\partial L}{\partial t} \end{aligned}$$

lo cual se puede reescribir en la forma

$$\frac{d}{dt} \left[ \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L \right] + \frac{\partial L}{\partial t} = 0 \quad (5.12)$$

la cantidad entre paréntesis se denomina **función energía**, y se denota por  $h$

$$h(q, \dot{q}, t) \equiv \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L = \left( \sum_{j=1}^n \dot{q}_j p_j \right) - L \quad (5.13)$$

la función energía es idéntica en valor al Hamiltoniano, cuya formulación veremos más adelante. Sin embargo, se denota con una letra diferente (el Hamiltoniano se denota como  $H$ ) ya que las dos funciones difieren en cuanto a los argumentos que utilizan,  $h$  es función de  $q_j, \dot{q}_j, t$ ; en tanto que  $H$  es función de  $q_j, p_j, t$  siendo  $p_j$  el momento conjugado a  $q_j$ . Aplicando (5.13), la ecuación (5.12) se escribe

$$\frac{dh}{dt} = - \frac{\partial L}{\partial t} \quad (5.14)$$

de lo cual se ve en forma inmediata que si el Lagrangiano no es función explícita del tiempo, es decir su dependencia temporal aparece solo a través de  $q(t)$ , y  $\dot{q}(t)$ , la función energía es una constante de movimiento.  $h$  es en consecuencia, una primera integral de movimiento y se le denomina integral de Jacobi.

En este punto conviene clarificar que para que una cierta cantidad sea constante de movimiento, es necesario y suficiente que su derivada total con respecto al tiempo (y no necesariamente la parcial) sea nula. Para ver la razón de esto, recordemos el significado de cada una de estas derivadas, la derivada parcial corresponde a dejar las coordenadas y velocidades generalizadas fijas y solo se mueve el parámetro tiempo, el hecho de que la derivada parcial se anule significa entonces que la cantidad en cuestión se mantiene constante en un proceso virtual en el cual las coordenadas y velocidades generalizadas del sistema se mantuvieran fijas y solo variara el parámetro tiempo, es decir solo evolucionan las influencias exteriores al sistema<sup>5</sup>. Una cantidad es constante de movimiento cuando se mantiene constante su valor en un proceso real, y en un proceso real las coordenadas y velocidades generalizadas también evolucionan a medida que transcurre el tiempo, en consecuencia es la derivada total la que describe correctamente la evolución de una cierta cantidad con el tiempo. En particular, la cantidad será constante de movimiento si y solo si su derivada total es cero.

Más adelante veremos que para el Hamiltoniano (que coincide numéricamente con la función energía) las derivadas temporales total y parcial coinciden (ver Ec. 6.38, Pág. 101). Naturalmente esto también será válido para la función energía, de modo que en este caso muy particular la anulación de la derivada parcial nos conduce a que  $h$  sea constante de movimiento. Debemos enfatizar sin embargo, que en general las derivadas total y parcial con respecto al tiempo de una cantidad arbitraria pueden ser muy diferentes.

### 5.2.1. Relación entre energía y función energía

Bajo ciertas circunstancias, la función  $h$  es la energía total del sistema. Para determinar bajo que circunstancias, recordemos que la energía cinética se puede escribir de la forma

$$T = T_0(q) + T_1(q, \dot{q}) + T_2(q, \dot{q})$$

donde  $T_0$  es independiente de las velocidades generalizadas,  $T_1$  es lineal en las velocidades generalizadas, y  $T_2$  es una función cuadrática de éstas (ver Ec. 2.24, Pág. 2.24). Para un amplio número de sistemas, una descomposición similar es posible con el Lagrangiano completo

$$L(q, \dot{q}, t) = L_0(q, t) + L_1(q, \dot{q}, t) + L_2(q, \dot{q}, t) \quad (5.15)$$

<sup>5</sup>Vale decir que en este caso estamos hablando de un proceso virtual muy diferente al definido para el principio de D'Alembert. Pues en el escenario actual es el tiempo el que evoluciona y se fijan las coordenadas de las partículas, en tanto que para el principio de D'Alembert es todo lo contrario.

donde  $L_2$  es una función homogénea de segundo grado (no simplemente cuadrática) en  $\dot{q}$ .  $L_1$  es homogénea de primer grado en  $\dot{q}$ , y  $L_0$  es independiente de  $\dot{q}$  (i.e. homogénea de grado cero en  $\dot{q}$ ). No hay ninguna razón de primeros principios para asumir que el Lagrangiano tenga esta forma, pero esta estructura aparece en una gran cantidad de problemas. Por ejemplo, el Lagrangiano adquiere esta forma cuando el potencial no depende en forma explícita de la velocidad. Sin embargo, aún para ciertos potenciales dependientes de la velocidad, esta separación es posible como se puede ver para el caso más característico del potencial de una carga en un campo electromagnético.

Aplicaremos ahora el teorema de Euler, que nos dice que si  $f(x_1, \dots, x_p)$  es una función homogénea de grado  $n$  en las variables  $x_i$  entonces

$$\sum_{i=1}^p x_i \frac{\partial f}{\partial x_i} = n f \quad (5.16)$$

aplicando la definición de la función  $h$ , Ec. (5.13), para Lagrangianos de la forma (5.15), resulta

$$h = \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L = \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L_0}{\partial \dot{q}_j} \right) - L_0 + \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L_1}{\partial \dot{q}_j} \right) - L_1 + \left( \sum_{j=1}^n \dot{q}_j \frac{\partial L_2}{\partial \dot{q}_j} \right) - L_2$$

y aplicando el teorema de Euler (5.16), para funciones homogéneas

$$\begin{aligned} h &= 0 - L_0 + L_1 - L_1 + 2L_2 - L_2 \\ h &= L_2 - L_0 \end{aligned} \quad (5.17)$$

Si las transformaciones de coordenadas descritas en (2.5) no dependen explícitamente del tiempo, la estructura de la energía cinética descrita en (2.24), resulta  $T = T_2$ . Si adicionalmente, el potencial no depende de las velocidades generalizadas, se tendría que  $L_2 = T$ , y  $L_0 = -V$ , de tal manera que

$$h = T + V = E$$

y la función energía corresponde en este caso a la energía total del sistema. Esta energía no necesariamente se conserva puesto que el potencial puede ser función explícita del tiempo. Cuando se asume adicionalmente a las condiciones anteriores, que el potencial no depende explícitamente del tiempo, la Ec. (5.14) nos conduce a la conservación de  $h$  y en este caso a la conservación de la energía total del sistema. Nótese sin embargo, que las condiciones de conservación de  $h$  son en general, muy diferentes de las condiciones para la conservación de la energía.  $h$  no necesariamente corresponde a la energía, y cuando corresponde a la energía del sistema no necesariamente se conserva. Es particularmente importante el hecho de que mientras el valor numérico del Lagrangiano es independiente de las coordenadas generalizadas empleadas, la función  $h$  depende en valor y en forma funcional del sistema coordinado elegido. Mas aún, esta cantidad puede ser conservada para una cierta escogencia de coordenadas y no serlo para otra, o ser la energía total en un sistema coordinado y no serlo en otro. En realidad, para un mismo sistema físico, diferentes funciones  $h$  pueden ser generadas de acuerdo con el sistema de coordenadas elegido, esto se puede ver de la definición (5.13) ya que  $\dot{q}_j$  y  $p_j$  son muy diferentes cuando cambiamos de sistema coordinado. Volveremos sobre este punto en la sección 6.7.1, en el lenguaje del Hamiltoniano.

Uno de los casos más comunes en mecánica clásica es aquel en el cual la energía cinética es de la forma  $T = m\dot{q}_i^2/2$  ó  $p_i^2/2m$  y la energía potencial depende solo de las coordenadas generalizadas. En este caso,  $h$  es la energía total del sistema y se conserva.

### 5.2.2. Función energía con fuerzas disipativas

Un caso interesante ocurre cuando tenemos fuerzas disipativas que se pueden generar de una función de disipación  $F$ . Tomando de nuevo como punto de partida la Ec. (5.10) y teniendo en cuenta que las ecuaciones de Lagrange con fuerzas disipativas tipo Rayleigh vienen dadas por (2.38) vemos que la Ec. (5.11) debe ser corregida para obtener

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_j} \right) + \frac{\partial F}{\partial \dot{q}_j}$$

con esto la Ec. (5.12) queda en la forma

$$\frac{d}{dt} \left[ \left( \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L \right] + \frac{\partial L}{\partial t} + \sum_j \frac{\partial F}{\partial \dot{q}_j} \dot{q}_j = 0$$

y junto con (5.13) adquieren la forma

$$\frac{dh}{dt} + \frac{\partial L}{\partial t} = - \sum_j \frac{\partial F}{\partial \dot{q}_j} \dot{q}_j$$

la definición de  $F$  dada por la Ec. (2.37), nos muestra que esta función es homogénea de grado 2 en las  $\dot{q}$ 's. Aplicando de nuevo el teorema de Euler, resulta

$$\frac{dh}{dt} = -2F - \frac{\partial L}{\partial t} \quad (5.18)$$

Si  $L$  no es función explícita del tiempo y el sistema es tal que  $h$  es la energía del sistema, la Ec. (5.18) nos dice que  $2F$  es la rata de disipación de energía

$$\frac{dE}{dt} = -2F \quad (5.19)$$

que concuerda con lo demostrado en la sección 2.5, aunque allí fué probado en circunstancias menos generales.

### 5.3. Teorema de Noether para sistemas discretos (opcional)

Veremos a continuación que la relación entre simetrías y leyes de conservación se puede englobar de una forma muy general en el celebrado teorema de Noether, que contiene a los casos ya estudiados como casos particulares. En términos simples este teorema asocia a cada simetría continua del sistema una cantidad conservada. Veamos su enunciado preciso.

**Theorem 3** *Teorema de Noether: Sea un Lagrangiano de la forma  $L = L(\mathbf{q}, \dot{\mathbf{q}}, t)$ . Supongamos que las ecuaciones de movimiento de Lagrange Ecs. (4.3), son invariantes bajo una transformación continua de coordenadas de la forma  $[t, \mathbf{q}] \rightarrow [t'(t), \mathbf{q}'(\mathbf{q}, t)]$ . Entonces existe una integral de movimiento i.e. una cantidad conservada asociada a dicha invarianza.*

**Demostración:** Dado un Lagrangiano  $L(\mathbf{q}, \dot{\mathbf{q}}, t)$  que depende de las coordenadas  $q_i$  ( $i = 1, \dots, n$ ), sus derivadas temporales  $\dot{q}_i$  y el tiempo  $t$ , podemos introducir unas nuevas coordenadas con la transformación

$$t' \equiv t'(t) \quad ; \quad q'_i \equiv q'_i(\mathbf{q}, t) \quad (5.20)$$

esta transformación debe ser invertible ya que de lo contrario, el nuevo conjunto coordenado no sería independiente. Parametrizaremos las nuevas coordenadas en la forma:

$$t' \equiv t + \delta t(t), \quad q'_i \equiv q_i + \delta q_i(\mathbf{q}, t) \quad (5.21)$$

inicialmente las transformaciones  $\delta t$  y  $\delta q_i$  son arbitrarias, por notación se tiene que:

$$\dot{q}_i \equiv \frac{dq_i}{dt} \quad ; \quad \dot{q}'_i \equiv \frac{dq'_i}{dt'}$$

podemos conectar estas cantidades a través de las relaciones

$$\dot{q}'_i = \frac{dq'_i}{dt'} = \frac{dq'_i}{dt} \frac{dt}{dt'} = \frac{d(q_i + \delta q_i)}{dt} \frac{dt}{dt'} \quad (5.22)$$

y teniendo en cuenta que

$$\frac{dt}{dt'} = \left( \frac{dt'}{dt} \right)^{-1} = \left( \frac{d[t + \delta t]}{dt} \right)^{-1} = \left( 1 + \frac{d}{dt} \delta t \right)^{-1} = \frac{1}{1 + \frac{d}{dt} \delta t} \quad (5.23)$$

podemos reemplazar (5.23) en (5.22) y se tiene que

$$\dot{q}'_i = \left( \dot{q}_i + \frac{d}{dt} \delta q_i \right) \frac{1}{\left( 1 + \frac{d}{dt} \delta t \right)}$$

de modo que:

$$\delta \dot{q}_i \equiv \dot{q}'_i - \dot{q}_i = \left( \dot{q}_i + \frac{d}{dt} \delta q_i \right) \frac{1}{\left( 1 + \frac{d}{dt} \delta t \right)} - \dot{q}_i$$

Si consideramos transformaciones infinitesimales de modo que  $\delta q_i$  y  $\delta t$  se convierten en cantidades diferenciales, se tiene que:

$$\begin{aligned} \delta \dot{q}_i &\approx \left( \dot{q}_i + \frac{d}{dt} \delta q_i \right) \left( 1 - \frac{d}{dt} \delta t \right) - \dot{q}_i = \dot{q}_i - \dot{q}_i \frac{d}{dt} \delta t + \left( \frac{d}{dt} \delta q_i \right) - \left( \frac{d}{dt} \delta q_i \right) \left( \frac{d}{dt} \delta t \right) - \dot{q}_i \\ \delta \dot{q}_i &\approx \frac{d}{dt} \delta q_i - \dot{q}_i \frac{d}{dt} \delta t \end{aligned} \quad (5.24)$$

donde hemos despreciado términos de orden cuadrático en  $\delta t, \delta q$ .

Dado que la Física no puede cambiar con esta transformación de coordenadas, la acción debe permanecer invariante:

$$S(t_1, t_2) \equiv \int_{t_1}^{t_2} L[\mathbf{q}(t), \dot{\mathbf{q}}(t), t] dt = S'[t'(t_1), t'(t_2)] \equiv \int_{t'(t_1)}^{t'(t_2)} L'[\mathbf{q}'(t'), \dot{\mathbf{q}}'(t'), t'] dt'$$

para que esto se cumpla, se requiere de la siguiente igualdad

$$L'(\mathbf{q}', \dot{\mathbf{q}}', t') \equiv L[\mathbf{q}(\mathbf{q}', t'), \dot{\mathbf{q}}(\mathbf{q}', \dot{\mathbf{q}}', t'), t(t')] \frac{dt}{dt'} \quad (5.25)$$

para verificarlo basta con hacer una transformación de coordenadas y tiempo a la acción

$$\int_{t_1}^{t_2} L[\mathbf{q}(t), \dot{\mathbf{q}}(t), t] dt \rightarrow \int_{t'(t_1)}^{t'(t_2)} L'[\mathbf{q}'(t'), \dot{\mathbf{q}}'(t'), t'] dt' = \int_{t_1}^{t_2} L[\mathbf{q}, \dot{\mathbf{q}}, t] dt$$

donde en la última igualdad hemos aplicado justamente (5.25) y los límites de integración cambian por el hecho de que cambia el diferencial del cual dependen éstos.

Si la forma de las ecuaciones de movimiento es invariante ante esta transformación de coordenadas, se dice que dicha transformación es simétrica. En el caso más simple, el lagrangiano como tal es invariante:

$$L'(\mathbf{q}', \dot{\mathbf{q}}', t') = L(\mathbf{q}', \dot{\mathbf{q}}', t')$$

esto sin embargo, no es necesario, ya hemos visto que es suficiente que se cumpla la relación.

$$L'(\mathbf{q}', \dot{\mathbf{q}}', t') = L(\mathbf{q}', \dot{\mathbf{q}}', t') + \frac{d}{dt'} \Omega(\mathbf{q}', t') \quad (5.26)$$

es decir, que ambas funciones lagrangianas pueden diferir en una derivada total con respecto al nuevo parámetro de tiempo. Si insertamos la ecuación (5.26) en (5.25), se tiene:

$$L[\mathbf{q}(\mathbf{q}', t'), \dot{\mathbf{q}}(\mathbf{q}', \dot{\mathbf{q}}', t'), t(t')] \frac{dt}{dt'} = L(\mathbf{q}', \dot{\mathbf{q}}', t') + \frac{d}{dt'} \Omega(\mathbf{q}', t') \quad (5.27)$$

de lo cual queda:

$$\begin{aligned} L[\mathbf{q}(\mathbf{q}', t'), \dot{\mathbf{q}}(\mathbf{q}', \dot{\mathbf{q}}', t'), t(t')] &= L(\mathbf{q}', \dot{\mathbf{q}}', t') \frac{dt'}{dt} + \left[ \frac{d}{dt'} \Omega(\mathbf{q}', t') \right] \frac{dt'}{dt} \\ L[\mathbf{q}, \dot{\mathbf{q}}, t] &= L(\mathbf{q}', \dot{\mathbf{q}}', t') \frac{dt'}{dt} + \frac{d}{dt} \Omega(\mathbf{q}', t') \end{aligned}$$

que junto con (5.23) nos da la ecuación:

$$\begin{aligned} L[\mathbf{q}, \dot{\mathbf{q}}, t] - L(\mathbf{q}', \dot{\mathbf{q}}', t') \frac{dt'}{dt} &= \frac{d}{dt} \Omega(\mathbf{q}', t') \\ L[\mathbf{q}, \dot{\mathbf{q}}, t] - L(\mathbf{q}', \dot{\mathbf{q}}', t') \left(1 + \frac{d}{dt} \delta t\right) &= \frac{d}{dt} \Omega(\mathbf{q}', t') \end{aligned}$$

de modo que

$$L[\mathbf{q}, \dot{\mathbf{q}}, t] - L(\mathbf{q}', \dot{\mathbf{q}}', t') = L(\mathbf{q}', \dot{\mathbf{q}}', t') \frac{d}{dt} \delta t + \frac{d}{dt} \Omega(\mathbf{q}', t') \quad (5.28)$$

y dado que la transformación es continua, es posible considerar transformaciones infinitesimales en (5.21). Definiendo

$$L(\mathbf{q}', \dot{\mathbf{q}}', t') - L[\mathbf{q}, \dot{\mathbf{q}}, t] \equiv \delta L$$

y tomando (5.21), la ecuación (5.28) se convierte en:

$$-\delta L \equiv L[\mathbf{q}, \dot{\mathbf{q}}, t] - L(\mathbf{q} + \delta \mathbf{q}, \dot{\mathbf{q}} + \delta \dot{\mathbf{q}}, t + \delta t) = L(\mathbf{q} + \delta \mathbf{q}, \dot{\mathbf{q}} + \delta \dot{\mathbf{q}}, t + \delta t) \frac{d}{dt} \delta t + \frac{d}{dt} \Omega(\mathbf{q} + \delta \mathbf{q}, t + \delta t)$$

pero por expansión de Taylor

$$L(\mathbf{q} + \delta \mathbf{q}, \dot{\mathbf{q}} + \delta \dot{\mathbf{q}}, t + \delta t) \frac{d}{dt} \delta t = L(\mathbf{q}, \dot{\mathbf{q}}, t) \frac{d}{dt} \delta t + \mathcal{O}(\delta^2)$$

despreciando términos cuadráticos en  $\delta \mathbf{q}$ ,  $\delta \dot{\mathbf{q}}$  y/o  $\delta t$ , resulta

$$-\delta L = L(\mathbf{q}, \dot{\mathbf{q}}, t) \frac{d}{dt} \delta t + \frac{d}{dt} \Omega(\mathbf{q} + \delta \mathbf{q}, t + \delta t)$$

En particular, si escogemos  $\delta \mathbf{q} = \delta t = 0$ , se tiene que  $\mathbf{q} = \mathbf{q}'$  y  $t = t'$ , además usando (5.24) se obtiene  $\delta \dot{\mathbf{q}} = 0$ , de modo que  $\dot{\mathbf{q}} = \dot{\mathbf{q}}'$ . Con estas consideraciones y usando la Ec. (5.27) se tendría que  $\frac{d}{dt} \Omega(\mathbf{q}, t) = 0$ . Podemos añadir este cero para reescribir  $-\delta L$  como

$$\begin{aligned} -\delta L &= L(\mathbf{q}, \dot{\mathbf{q}}, t) \frac{d}{dt} \delta t + \frac{d}{dt} [\Omega(\mathbf{q} + \delta \mathbf{q}, t + \delta t) - \Omega(\mathbf{q}, t)] \\ \delta L &= -L(\mathbf{q}, \dot{\mathbf{q}}, t) \frac{d}{dt} \delta t - \frac{d}{dt} \delta \Omega(\mathbf{q}, t) \end{aligned} \quad (5.29)$$

por otro lado, dado que estamos escribiendo  $L$  en función de  $\mathbf{q}$ ,  $\dot{\mathbf{q}}$ ,  $t$  la regla de la cadena para  $\delta L$  nos da

$$\delta L = \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] + \frac{\partial L}{\partial t} \delta t \quad (5.30)$$

igualando (5.29) con (5.30) y usando (5.24) resulta:

$$\begin{aligned} \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] + \frac{\partial L}{\partial t} \delta t + L \frac{d}{dt} \delta t &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \\ \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \left( \frac{d}{dt} \delta q_i - \dot{q}_i \frac{d}{dt} \delta t \right) \right] + \left( \frac{\partial L}{\partial t} + L \frac{d}{dt} \right) \delta t &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \\ \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \right] \delta q_i - \sum_{i=1}^n \left[ \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \frac{d}{dt} \right] \delta t + \left( \frac{\partial L}{\partial t} + L \frac{d}{dt} \right) \delta t &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \\ \sum_{i=1}^n \left[ \frac{\partial L}{\partial q_i} + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \right] \delta q_i + \frac{\partial L}{\partial t} \delta t + \left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \frac{d}{dt} \delta t &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \end{aligned} \quad (5.31)$$



la Ec. (5.31) describe la condición que un  $\delta\Omega$  debe cumplir para un Lagrangiano dado, a fin de que las ecuaciones de movimiento (4.3), permanezcan invariantes ante una transformación infinitesimal dada por (5.21)<sup>6</sup>. El problema se reduce usualmente a la existencia (o no existencia) de una solución para la función  $\delta\Omega$  en la Ec. (5.31) para una transformación específica de la forma (5.21). En particular, si se cumplen las condiciones

$$\frac{d(\delta\Omega)}{dt} = 0 \quad y \quad \frac{d(\delta t)}{dt} = 0 \quad (5.32)$$

entonces la Ec. (5.29) nos lleva a que  $\delta L = 0$ . Por tanto, bajo estas condiciones la función Lagrangiana misma permanecería invariante bajo la transformación de coordenadas. Si la Ec. (5.31) se satisface, entonces al usar las ecuaciones de movimiento

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right)$$

se obtiene:

$$\sum_{i=1}^n \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \right] \delta q_i + \frac{\partial L}{\partial t} \delta t + \left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \frac{d}{dt} \delta t = -\frac{d}{dt} \delta\Omega(\mathbf{q}, t) \quad (5.33)$$

el primer y tercer términos a la izquierda de (5.33) se pueden escribir en la forma

$$\sum_{i=1}^n \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \right] \delta q_i = \frac{d}{dt} \left\{ \sum_{i=1}^n \left[ 2 \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i \right\} - \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{q}_i} \right) \frac{d}{dt} \delta q_i - \sum_{i=1}^n \delta q_i \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \quad (5.34)$$

$$\left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \frac{d}{dt} \delta t = \frac{d}{dt} \left\{ \left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \delta t \right\} - \frac{dL}{dt} \delta t + \delta t \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \quad (5.35)$$

reemplazando (5.34) y (5.35) en (5.33) se obtiene

$$\begin{aligned} & \frac{d}{dt} \left\{ \sum_{i=1}^n \left[ 2 \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i \right\} - \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{q}_i} \right) \frac{d}{dt} \delta q_i - \sum_{i=1}^n \delta q_i \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial t} \delta t + \\ & + \frac{d}{dt} \left\{ \left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \delta t \right\} - \frac{dL}{dt} \delta t + \delta t \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) + \frac{d}{dt} \delta\Omega(\mathbf{q}, t) = 0 \end{aligned}$$

organizando los términos que aparecen bajo la derivada temporal total resulta

$$\begin{aligned} & \frac{d}{dt} \left\{ \sum_{i=1}^n 2 \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta t + \delta\Omega(\mathbf{q}, t) \right\} - \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{q}_i} \right) \frac{d}{dt} \delta q_i - \sum_{i=1}^n \delta q_i \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \\ & + \frac{\partial L}{\partial t} \delta t - \frac{dL}{dt} \delta t + \delta t \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) = 0 \end{aligned}$$

simplificando términos

$$\begin{aligned} & \frac{d}{dt} \left\{ \sum_{i=1}^n 2 \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta t + \delta\Omega(\mathbf{q}, t) \right\} - \frac{d}{dt} \left\{ \sum_{i=1}^n \left[ \left( \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] \right\} + \\ & + \left[ \frac{\partial L}{\partial t} - \frac{dL}{dt} + \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \right] \delta t = 0 \end{aligned}$$

<sup>6</sup>Vale la pena mencionar que al Lagrangiano lo podemos ver en este teorema, como una función arbitraria que depende de un parámetro  $t$ , de unas coordenadas  $q_i$  y de  $\dot{q}_i$ , de tal manera que las ecuaciones que rigen el comportamiento de las coordenadas  $q_i$  con respecto al parámetro  $t$ , sean las Ecs. (4.3). El sistema no tiene que ser mecánico y de hecho no tiene que ser un sistema Físico.

de nuevo agrupamos los términos con derivada temporal total

$$\frac{d}{dt} \left\{ \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta t + \delta \Omega(\mathbf{q}, t) \right\} + \left[ \frac{\partial L}{\partial t} + \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - \frac{dL}{dt} \right] \delta t = 0 \quad (5.36)$$

usando regla de la cadena para  $dL/dt$  y las ecuaciones de Lagrange, evaluamos el término proporcional a  $\delta t$

$$\begin{aligned} \frac{\partial L}{\partial t} + \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - \frac{dL}{dt} &= \frac{\partial L}{\partial t} + \frac{d}{dt} \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - \sum_{i=1}^n \frac{\partial L}{\partial q_i} \dot{q}_i - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^n \dot{q}_i \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) + \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \sum_{i=1}^n \frac{\partial L}{\partial q_i} \dot{q}_i - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \\ &= \sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial q_i} + \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \sum_{i=1}^n \frac{\partial L}{\partial q_i} \dot{q}_i - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i = 0 \end{aligned}$$

por tanto la ecuación (5.36) se reduce a:

$$\frac{d}{dt} \left\{ \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta t + \delta \Omega(\mathbf{q}, t) \right\} = 0$$

es decir que la cantidad

$$\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta t + \delta \Omega(\mathbf{q}, t) = cte \quad (5.37)$$

es una constante de movimiento o cantidad conservada (integral de movimiento). En síntesis, si para una transformación específica de coordenadas de la forma (5.21), existe un valor de  $\delta \Omega(\mathbf{q}, t)$  que satisfaga la Ec. (5.31), dicho valor de  $\delta \Omega$  nos conducirá a una constante de movimiento de la forma (5.37). Recordemos que la condición (5.31) equivale a la invarianza de las ecuaciones de movimiento ante la transformación (5.21).

### 5.3.1. Comentarios sobre el teorema de Noether

1. Es importante enfatizar en que dicho teorema solo es válido para transformaciones continuas del espacio tiempo, ya que en la demostración es fundamental el uso de transformaciones infinitesimales.
2. Aunque  $\Omega(\mathbf{q}, t)$  es una función arbitraria, esta debe ser derivable hasta segundo orden en todas sus variables y esta segunda derivada debe ser continua. Ya que en una parte de la demostración se intercambian las segundas derivadas.
3. El teorema implica la invarianza de la acción (i.e. de las ecuaciones de movimiento). Sin embargo, no implica la invarianza del lagrangiano mismo, esta solo se cumple si  $d(\delta \Omega)/dt = 0$ ,  $d(\delta t)/dt = 0$ .
4. En teoría de campos (ver sección 18.3) se puede hacer una demostración similar con las siguientes correspondencias:  $\mathbf{q} \rightarrow \Phi(\mathbf{q})$ ,  $\dot{\mathbf{q}} \rightarrow \partial_\mu \Phi(\mathbf{q})$ ,  $L \rightarrow \mathcal{L}$  donde  $\Phi(\mathbf{q})$  representa los campos (un arreglo vectorial de ellos),  $\partial_\mu \Phi(\mathbf{q})$  representa sus derivadas con respecto al espacio y el tiempo, y  $\mathcal{L}$  representa una densidad Lagrangiana con

$$L \equiv \int \mathcal{L} d^3q,$$

5. A pesar de que la función  $\Omega$  actúa como un “gauge” para el Lagrangiano, vemos que la cantidad conservada depende del cambio de esta función  $\delta \Omega$  evaluado entre los dos sistemas coordenados. Esto es lógico ya que  $\delta \Omega$  surge de manera natural a partir de la transformación de coordenadas que se hizo. En realidad esta cantidad se debe fijar a través de la Ec. (5.31) con el fin de que la transformación en cuestión deje invariantes a las ecuaciones de movimiento.

6. Nótese que la solución para  $\delta\Omega$  en la Ec. (5.31), no tiene porqué ser única. Por otro lado, puede ocurrir que no exista solución, en este caso la transformación de coordenadas no es una transformación de simetría del sistema y no tendríamos una cantidad conservada.
7. Es necesario distinguir entre la transformación de coordenadas  $[t, \mathbf{q}] \rightarrow [t', \mathbf{q}']$  y la transformación gauge  $L' = L + d\Omega/dt$ . Nótese que esta última solo es válida en un sistema coordenado fijo.
8. Una condición esencial para la validez del teorema, es que el nuevo parámetro  $t'$ , dependa únicamente del antiguo parámetro  $t' = t'(t)$ , y no de las antiguas coordenadas, como se vé en las Ecs. (5.20). Esto es importante para la consistencia de la teoría, puesto que el parámetro debe ser completamente independiente de las coordenadas, de modo que solo él regule la evolución de las coordenadas, conservando la independencia entre éstas.
9. En esta demostración se ha supuesto que el Lagrangiano depende de  $\mathbf{q}, \dot{\mathbf{q}}$  y  $t$ . No se habla sobre Lagrangianos con dependencia de las derivadas de orden superior. Por ejemplo, esta demostración no es válida para un Lagrangiano de la forma  $L(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, t)$  como el discutido en la sección 4.6.
10. El teorema implica que las simetrías del sistema deben reflejarse en su acción  $S$ .
11. Una estrategia muy fructífera para el uso del teorema de Noether es la siguiente: para una transformación de coordenadas específica, buscamos las condiciones requeridas para que la Ec. (5.31) tenga solución, con la solución así obtenida vamos a la Ec. (5.37) para encontrar la constante de movimiento que se genera. Las condiciones que se necesiten para que (5.31) tenga solución, serán entonces las condiciones físicas que debe tener mi sistema para que la transformación de coordenadas sea una transformación de simetría para mi sistema, y por ende para que la cantidad generada en (5.37) sea realmente conservada.

## 5.4. Ejemplos de aplicación del teorema de Noether

### 5.4.1. Invarianza ante traslación temporal y conservación de la energía

Supongamos que las ecuaciones de movimiento de un sistema físico son invariantes ante una traslación temporal. ¿Que condición debe satisfacer el Lagrangiano y cual es la cantidad conservada?.

Una traslación temporal se puede escribir como:  $\delta\mathbf{q} = \delta\dot{\mathbf{q}} = 0$ ,  $\delta t(t) = \delta\tau = cte$ . De la ecuación (5.31) resulta:

$$\begin{aligned} \frac{\partial L}{\partial t}\delta\tau + \left[ L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right] \frac{d}{dt}\delta\tau &= -\frac{d}{dt}\delta\Omega(\mathbf{q}, t) \\ \frac{\partial L}{\partial t}\delta\tau &= -\frac{d}{dt}\delta\Omega(\mathbf{q}, t) \end{aligned} \quad (5.38)$$

si  $L$  no depende explícitamente del tiempo entonces  $\delta\Omega$  es constante y además la Ec. (5.29) nos indica que el Lagrangiano mismo es invariante ante la transformación en cuestión. Por tanto, la cantidad conservada dada por la Ec. (5.37) sería

$$\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \delta q_i + \left( L - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) \delta\tau + \delta\Omega = c_a$$

teniendo en cuenta que  $\delta q_i = 0$  y como  $\delta\tau$  y  $\delta\Omega$  son constantes resulta:

$$h(q_i, \dot{q}_i, t) \equiv \left( \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) = \left( \sum_{i=1}^n p_i \dot{q}_i - L \right) \equiv H = cte$$

es decir, la función energía del sistema (Función Hamiltoniana), es la constante de movimiento. Cuando la función energía (el Hamiltoniano) corresponde a la energía del sistema llegamos a la conservación de la energía.

Vale decir que de acuerdo con nuestras condiciones, la energía se conserva incluso para algunos sistemas no aislados, pues si los campos externos son independientes del tiempo, el Lagrangiano no dependerá del tiempo que es la condición requerida para llegar a la conservación de esta cantidad (además de las condiciones para que la función energía, sea la energía del sistema).

En la discusión anterior asumimos que el Lagrangiano no depende explícitamente del tiempo, con lo cual el Lagrangiano mismo permanece invariante ante la traslación temporal. Vale la pena preguntarse si podemos encontrar una condición más general en la cual las ecuaciones de movimiento permanezcan invariantes ante dicha transformación, pero no necesariamente el Lagrangiano mismo. Asumamos en consecuencia que el Lagrangiano puede depender explícitamente del tiempo, dado que  $L = T - V$ , usualmente la energía cinética no depende explícitamente del tiempo (a menos que la transformación a coordenadas generalizadas dependa explícitamente del tiempo), si asumimos que en cambio la energía potencial es dependiente del tiempo, la Ec. (5.38) queda

$$\frac{\partial V}{\partial t} = \frac{1}{\delta\tau} \frac{d}{dt} \delta\Omega$$

en general no es posible encontrar un  $\delta\Omega$  que satisfaga esta ecuación, ya que  $\frac{\partial V}{\partial t}$  no tiene que ser una derivada total. Efectivamente en este caso la función energía no es constante de movimiento y no hay garantía de que se pueda encontrar alguna función que sí sea constante de movimiento.

En las tres secciones siguientes asumiremos que la función Lagrangiana del sistema en coordenadas cartesianas está dada por

$$L = \frac{1}{2} m \dot{\mathbf{r}}^2 - V(\mathbf{r}) \quad (5.39)$$

y encontraremos las condiciones para que las traslaciones espaciales, las rotaciones y las transformaciones de Galileo sean transformaciones de simetría del sistema. Así mismo encontraremos las cantidades conservadas asociadas a cada simetría y veremos que los resultados son consistentes con los ya obtenidos.

#### 5.4.2. Invarianza ante traslación espacial y conservación del momento lineal

En analogía con la sección anterior, la idea es exigir invarianza ante traslaciones espaciales y ver a que condiciones lleva esta exigencia. Tomaremos como punto de partida el Lagrangiano (5.39).

Por simplicidad asumamos que solo hay traslación a lo largo del eje  $x_3$ . Las transformaciones correspondientes a traslaciones espaciales son:

$$\delta x_1 = \delta x_2 = 0, \quad \delta x_3 = cte, \quad \delta t = 0 \quad (5.40)$$

Para  $\delta t = 0$  la Ec. (5.31) se reduce a:

$$\sum_{i=1}^n \left[ \frac{\partial L}{\partial x_i} + \frac{\partial L}{\partial \dot{x}_i} \frac{d}{dt} \right] \delta x_i = - \frac{d}{dt} \delta\Omega(\mathbf{x}, t) \quad (5.41)$$

si existe un  $\delta\Omega$  que cumpla esta condición se tiene de (5.37) que:

$$\sum_{i=1}^n \frac{\partial L}{\partial \dot{x}_i} \delta x_i + \delta\Omega(\mathbf{q}, t) = c_a \quad (5.42)$$

para la forma de nuestro Lagrangiano Ec. (5.39) se tiene que

$$\frac{\partial L}{\partial x_i} = - \frac{\partial V}{\partial x_i}; \quad \frac{\partial L}{\partial \dot{x}_i} = m \dot{x}_i \quad (5.43)$$

esto es válido para la invarianza traslacional, rotacional y galileana ya que hasta ahora solo hemos usado  $\delta t = 0$ . En el caso de la invarianza traslacional, usando (5.41), (5.40) y (5.43), se tiene:

$$\begin{aligned} \sum_{i=1}^n -\frac{\partial V}{\partial x_i} \delta x_i &= -\frac{d}{dt} \delta \Omega(\mathbf{x}, t) \\ -\frac{\partial V}{\partial x_3} \delta x_3 &= -\sum_{i=1}^n \left[ \frac{\partial}{\partial x_i} \delta \Omega(\mathbf{x}, t) \right] \dot{x}_i - \frac{\partial}{\partial t} \delta \Omega(\mathbf{x}, t) \\ \sum_{i=1}^n \left[ \frac{\partial}{\partial x_i} \delta \Omega(\mathbf{x}, t) \right] \dot{x}_i &= \frac{\partial V}{\partial x_3} \delta x_3 - \frac{\partial}{\partial t} \delta \Omega(\mathbf{x}, t) \end{aligned} \quad (5.44)$$

claramente, la expresión de la derecha no contiene a los  $\dot{x}_i$ , por lo tanto los coeficientes de la izquierda tampoco pueden contenerlos i.e.

$$\frac{\partial}{\partial x_i} \delta \Omega(\mathbf{x}, t) = 0 \Rightarrow \delta \Omega(\mathbf{x}, t) = \delta \Omega(t)$$

de modo que la condición (5.44) se reduce a:

$$\frac{\partial V}{\partial x_3} \delta x_3 = \frac{d}{dt} \delta \Omega(t)$$

debido a la forma de nuestro Lagrangiano, el potencial  $V$  solo es función de la posición de modo que a la izquierda tenemos un término que solo depende de la posición y a la derecha otro que solo depende del tiempo, de lo cual se sigue que  $\frac{\partial V}{\partial x_3}$  debe ser independiente de  $(\mathbf{x}, t)$ , es decir es constante. Integrando<sup>7</sup>:

$$\left( \frac{\partial V}{\partial x_3} \delta x_3 \right) t = \delta \Omega \quad (5.45)$$

con este valor de  $\delta \Omega$  las ecuaciones de movimiento son invariantes en forma ante una transformación espacial de  $x_3$ . Es decir, que cuando se cumple (5.45), la traslación espacial es una transformación de simetría, la constante de movimiento se sigue de (5.42)

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}_3} \delta x_3 + \delta \Omega(t) &= c_a \\ m \dot{x}_3 + \left( \frac{\partial V}{\partial x_3} \right) t &= \frac{c_a}{\delta x_3} \equiv c_b \end{aligned}$$

recordando que la fuerza se escribe como  $F_i = -\frac{\partial V}{\partial x_i}$ . tendremos en general que

$$m \dot{x}_3 - F_3 t = c_b$$

ahora bien, si tenemos una invarianza similar asociada a  $\delta x_1$  y  $\delta x_2$ , la cantidad conservada es:

$$\tilde{\mathbf{P}} = \mathbf{p} - \mathbf{F}t = m\dot{\mathbf{r}} - \mathbf{F}t = \mathbf{c}_b \quad (5.46)$$

es decir que el momento lineal  $\mathbf{p}$ , es una función lineal del tiempo. Recordemos que bajo las condiciones aquí establecidas, las componentes de la fuerza deben ser constantes de modo que  $\mathbf{F}$  es un campo de fuerzas constante y homogéneo.

Volviendo a la invarianza en solo  $\delta x_3$ , vemos que la exigencia de dicha invarianza nos lleva a que  $\partial_3 V = -F_3$  sea constante. Es decir, a que la componente de la fuerza en esa dirección sea constante. Si en particular suponemos que  $F_3 = 0$ , se sigue que la cantidad conservada es justamente la componente del momento en esa dirección.

$$\frac{\partial L}{\partial \dot{x}_3} = m \dot{x}_3 = p_3 = c_b \quad (5.47)$$

<sup>7</sup>Una posible constante de integración se puede absorber en  $\delta \Omega$ .

adicionalmente, se puede ver de (5.45) que cuando  $F_3 = 0$ ,  $\Rightarrow \delta\Omega = 0$  y el Lagrangiano como tal es invariante. Por supuesto, la invarianza del vector momento se seguirá si cada componente de la fuerza se anula.

Por tanto, vemos que el principio de conservación del momento se sigue de la invarianza del Lagrangiano ante una traslación espacial, pero no de la invarianza de las ecuaciones de movimiento. Cuando solo esta última se cumple para todas las coordenadas,  $\mathbf{F}$  es un campo de fuerzas constante y homogéneo.

La existencia de un campo constante y homogéneo ilustra la diferencia entre homogeneidad local y homogeneidad global del espacio. El espacio que ocupa el campo es localmente homogéneo, porque ningún punto de dicho espacio se puede distinguir de otro por una medición local (en nuestro caso, se obtendría el mismo valor de la fuerza en cada punto). Sin embargo, esta fuerza tiene que ser generada por alguna fuente (las placas de un condensador, una masa distante etc.) la existencia de dicha fuente destruye la homogeneidad global del espacio. Por tanto, la homogeneidad local del espacio implica que el momento es una función lineal del tiempo en tanto que la homogeneidad global implica su conservación<sup>8</sup>.

### 5.4.3. Invarianza ante rotaciones espaciales y la conservación del momento angular

Al exigir invarianza ante rotaciones espaciales, supondremos por simplicidad que la rotación se realiza en el plano  $X - Y$ , es decir alrededor del eje  $Z$ . Las transformaciones correspondientes son:

$$\begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

Si consideramos que la rotación está descrita por un ángulo infinitesimal  $\delta\phi$  constante, la transformación a primer orden en  $\delta\phi$  queda:

$$\begin{pmatrix} 1 & -\delta\phi \\ \delta\phi & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

$$\begin{aligned} x'_1 &= x_1 - (\delta\phi) x_2 \\ x'_2 &= x_1 \delta\phi + x_2 \end{aligned}$$

de modo que

$$\begin{aligned} \delta x_1 &= x'_1 - x_1 = -(\delta\phi) x_2 \\ \delta x_2 &= x'_2 - x_2 = (\delta\phi) x_1 \\ \delta x_3 &= \delta t = 0 \end{aligned} \tag{5.48}$$

Recordemos que las ecuaciones (5.41, 5.42, 5.43), solo emplearon la condición  $\delta t = 0$ , y la estructura (5.39) del Lagrangiano, de modo que son aplicables en este contexto. A partir de (5.41) se obtiene:

$$\sum_{i=1}^n \left[ \frac{\partial L}{\partial x_i} + \frac{\partial L}{\partial \dot{x}_i} \frac{d}{dt} \right] \delta x_i = -\frac{d}{dt} \delta\Omega(\mathbf{x}, t) \tag{5.49}$$

para la forma de nuestro Lagrangiano se aplican las Ecs. (5.43). Usando (5.43) y (5.48), los miembros de la izquierda de (5.49) se escriben como

$$\begin{aligned} \sum_{i=1}^n \frac{\partial L}{\partial x_i} \delta x_i &= \frac{\partial L}{\partial x_1} \delta x_1 + \frac{\partial L}{\partial x_2} \delta x_2 + \frac{\partial L}{\partial x_3} \delta x_3 = \left( \frac{\partial V}{\partial x_1} x_2 - \frac{\partial V}{\partial x_2} x_1 \right) \delta\phi \\ \sum_{i=1}^n \left[ \frac{\partial L}{\partial \dot{x}_i} \frac{d}{dt} \right] \delta x_i &= \frac{\partial L}{\partial \dot{x}_1} \frac{d}{dt} \delta x_1 + \frac{\partial L}{\partial \dot{x}_2} \frac{d}{dt} \delta x_2 = -m\dot{x}_1 (\delta\phi) \dot{x}_2 + m\dot{x}_2 (\delta\phi) \dot{x}_1 = 0 \end{aligned}$$

<sup>8</sup>Por supuesto la homogeneidad global absoluta no existe, ya que toda fuerza debe tener sus fuentes en alguna parte. Sin embargo, si las fuentes están a distancias mucho mayores que todas las distancias típicas de mi problema, podemos pensar que las fuentes están “en el infinito”, en cuyo caso adquiere sentido la homogeneidad global como una buena aproximación.

aplicando estas relaciones en (5.49)

$$\left( \frac{\partial V}{\partial x_1} x_2 - \frac{\partial V}{\partial x_2} x_1 \right) \delta\phi = -\frac{d}{dt} \delta\Omega(\mathbf{x}, t) \quad (5.50)$$

$$(\mathbf{r} \times \nabla V)_3 \delta\phi = \frac{d}{dt} \delta\Omega(\mathbf{x}, t) \quad (5.51)$$

$$(\mathbf{r} \times \nabla V)_3 \delta\phi = \dot{x}_i \frac{\partial}{\partial x_i} \delta\Omega(\mathbf{x}, t) + \frac{\partial}{\partial t} \delta\Omega(\mathbf{x}, t) \quad (5.52)$$

$$(\mathbf{r} \times \nabla V)_3 \delta\phi - \frac{\partial}{\partial t} \delta\Omega(\mathbf{x}, t) = \dot{x}_i \frac{\partial}{\partial x_i} \delta\Omega(\mathbf{x}, t) \quad (5.53)$$

pero dado que los términos de la izquierda no dependen de  $\dot{x}_i$  los coeficientes de  $\dot{x}_i$  a la derecha deben anularse con lo cual  $\partial_{x_i} \delta\Omega = 0$  i.e.  $\delta\Omega(\mathbf{x}, t) = \delta\Omega(t)$ . Con lo cual la Ec. (5.53) queda

$$(\mathbf{r} \times \nabla V)_3 \delta\phi = \frac{d}{dt} \delta\Omega(t) \quad (5.54)$$

Por tanto, el miembro de la izquierda depende de coordenadas espaciales y el de la derecha solo del tiempo, de tal forma que cada miembro debe ser constante, lo cual nos lleva a:

$$(\mathbf{r} \times \nabla V)_3 = cte \quad (5.55)$$

La Ec. (5.55), es la condición que se requiere para que las ecuaciones de Lagrange sean invariantes bajo una rotación sobre el plano  $X - Y$ , este término corresponde a menos la tercera componente del torque. Es decir que la invarianza de las ecuaciones de Lagrange ante las rotaciones espaciales alrededor de un eje, requiere que la componente del torque a lo largo de dicho eje sea uniforme y constante. Esto es análogo a la condición de que la fuerza sea constante y uniforme para que las ecuaciones de Lagrange sean invariantes ante traslaciones espaciales. Integrando (5.54) obtenemos el valor de  $\delta\Omega(t)$

$$\delta\Omega(t) = -(\delta\phi) (\mathbf{r} \times \mathbf{F})_3 t = -(\delta\phi) \tau_3 t \quad (5.56)$$

reemplazando (5.43, 5.48, 5.56) en (5.37), la constante de movimiento queda

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}_1} \delta x_1 + \frac{\partial L}{\partial \dot{x}_2} \delta x_2 - (\delta\phi) \tau_3 t &= c_a \\ -m\dot{x}_1 (\delta\phi) x_2 + m\dot{x}_2 (\delta\phi) x_1 - (\delta\phi) \tau_3 t &= c_a \\ -m\dot{x}_1 x_2 + m\dot{x}_2 x_1 - \tau_3 t &= c_b \\ (\mathbf{r} \times \mathbf{p})_3 - \tau_3 t &= c_b \end{aligned}$$

si asumimos que hay invarianza ante rotaciones en tres ejes mutuamente perpendiculares, la constante de movimiento es

$$\mathbf{L} - \vec{\tau} t = \mathbf{c}_b \quad (5.57)$$

que es el análogo de (5.46) en el caso de invarianza traslacional. En este caso el momento angular es función lineal del tiempo.

Volviendo a la invarianza ante rotaciones con respecto a  $x_3$ , vemos que si la tercera componente del torque es cero la Ec. (5.54) nos dice que

$$(\mathbf{r} \times \nabla V)_3 = 0 \Rightarrow \frac{d}{dt} \delta\Omega(\mathbf{x}, t) = 0$$

y  $\delta\Omega$  es una constante, esto nos lleva a la invarianza del Lagrangiano puesto que se cumplen las condiciones (5.32). Reemplazando  $\tau_3 = 0$  en (5.57) la constante de movimiento es

$$\mathbf{L}_3 = c_b \quad (5.58)$$

que es el análogo de (5.47) para traslaciones espaciales. Luego la conservación del momento angular total se sigue de la invarianza del Lagrangiano ante rotaciones en tres ejes independientes.

Nótese la analogía con las traslaciones espaciales, la invarianza de las ecuaciones de movimiento ante traslaciones tenía como condición que las fuerzas fueran constantes y la invarianza del Lagrangiano mismo nos llevaba a la conservación del momento lineal. En este caso la invarianza rotacional de las ecuaciones de movimiento requiere torques constantes y uniformes, y la invarianza del Lagrangiano mismo nos lleva a la conservación del momento angular. Es de anotar que en ambos casos cuando las ecuaciones quedan invariantes pero no el Lagrangiano, las cantidades conservadas no son el momento lineal o el momento angular sino las cantidades definidas en (5.46, 5.57).

#### 5.4.4. Transformaciones de Galileo

Cuando los fenómenos físicos se ven desde dos sistemas de referencia inerciales  $S$  y  $S'$ , es necesario conciliar las observaciones de cada uno, para lo cual es necesario conocer la forma en que se conectan las coordenadas cartesianas de espacio y tiempo de los dos sistemas en cuestión. Las transformaciones de Galileo nos proveen esa conexión.

Asumamos una partícula de masa  $m$  con coordenadas  $x_i, t$  con respecto al sistema de referencia  $S$ , y coordenadas  $x'_i, t'$  con respecto al sistema de referencia  $S'$ . Por simplicidad asumiremos que el sistema  $S'$  viaja con velocidad constante  $v_3$  a lo largo del eje  $X_3$  del sistema  $S$ , con lo cual las transformaciones de Galileo se escriben en la forma siguiente

$$\begin{aligned} x'_1 &= x_1 & ; & & x'_2 &= x_2 & ; & & x'_3 &= x_3 + v_3 t & ; & & t' &= t \\ \delta x_1 &= \delta x_2 = \delta t = 0 & ; & & \delta x_3 &= x'_3 - x_3 = v_3 t \end{aligned}$$

para transformaciones infinitesimales (i.e. velocidades relativas infinitesimales) tenemos:

$$\begin{aligned} x'_3 &= x_3 + (\delta v_3) t & ; & & (\delta v_3) &= cte \\ \delta x_1 &= \delta x_2 = \delta t = 0 & ; & & \delta x_3 &= (\delta v_3) t \end{aligned} \quad (5.59)$$

la condición para que esta sea una transformación de simetría, se encuentra tomando de nuevo la Ec. (5.31) y aplicándole las Ecs. (5.59) y (5.43)

$$\begin{aligned} \sum_{i=1}^n \left[ \frac{\partial L}{\partial x_i} + \frac{\partial L}{\partial \dot{x}_i} \frac{d}{dt} \right] \delta x_i &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \\ \left[ \frac{\partial L}{\partial x_3} + \frac{\partial L}{\partial \dot{x}_3} \frac{d}{dt} \right] (\delta v_3) t &= -\frac{d}{dt} \delta \Omega(\mathbf{q}, t) \end{aligned} \quad (5.60)$$

$$\left[ -\frac{\partial V}{\partial x_3} t + m \dot{x}_3 \right] (\delta v_3) = -\sum_{i=1}^n \frac{\partial \delta \Omega}{\partial x_i} \dot{x}_i - \frac{\partial}{\partial t} \delta \Omega \quad (5.61)$$

escribiendo explícitamente la sumatoria y reorganizando términos

$$-\frac{\partial V}{\partial x_3} (\delta v_3) t + \frac{\partial}{\partial t} \delta \Omega + \left[ m (\delta v_3) + \frac{\partial \delta \Omega}{\partial x_3} \right] \dot{x}_3 + \frac{\partial \delta \Omega}{\partial x_1} \dot{x}_1 + \frac{\partial \delta \Omega}{\partial x_2} \dot{x}_2 = 0 \quad (5.62)$$

en virtud de la independencia de los  $\dot{x}_i$ , los coeficientes de  $\dot{x}_1$ ,  $\dot{x}_2$ ,  $\dot{x}_3$  y los términos independientes de  $\dot{x}_i$  en (5.62) deben ser nulos, por lo tanto.

$$\frac{\partial \delta \Omega}{\partial x_1} = \frac{\partial \delta \Omega}{\partial x_2} = 0 & ; & \frac{\partial \delta \Omega}{\partial x_3} = -m (\delta v_3) & ; & \frac{\partial V}{\partial x_3} t (\delta v_3) = \frac{\partial}{\partial t} \delta \Omega \quad (5.63)$$

de modo que  $\delta \Omega$  solo es función de  $x_3$  y  $t$ . Integrando la segunda de las Ecs. (5.63) resulta

$$\begin{aligned} \delta \Omega &= -m x_3 (\delta v_3) + f(t) \\ \Rightarrow \frac{\partial}{\partial t} \delta \Omega &= \frac{d}{dt} f(t) \end{aligned} \quad (5.64)$$



y la tercera de las Ecs. (5.63) queda entonces

$$\frac{\partial V}{\partial x_3} t (\delta v_3) = \frac{d}{dt} f(t) \Rightarrow \frac{\partial V}{\partial x_3} (\delta v_3) = \frac{1}{t} \frac{d}{dt} f(t) \quad (5.65)$$

y como  $\frac{1}{t} \frac{d}{dt} f(t)$  solo depende del tiempo y  $\frac{\partial V}{\partial x_3} (\delta v_3)$  solo del espacio, se tiene que

$$\frac{\partial V}{\partial x_3} = cte$$

ya que  $\delta v_3$  es constante. La condición Física requerida para la existencia de  $\delta\Omega$  es entonces que la componente  $F_3$  de la fuerza sobre la partícula sea constante y uniforme. Con esta condición, podemos integrar fácilmente (5.65) y se obtiene

$$f(t) = \frac{\partial V}{\partial x_3} \frac{t^2}{2} (\delta v_3) + C \quad (5.66)$$

reemplazando (5.66) en (5.64)

$$\delta\Omega = \left( -mx_3 + \frac{1}{2} \frac{\partial V}{\partial x_3} t^2 \right) (\delta v_3) + C$$

la constante se puede absorber en la cantidad conservada. Nuevamente usamos (5.37) para encontrar la cantidad conservada y le aplicamos las Ecs. (5.43, 5.59)

$$\begin{aligned} \sum_{i=1}^n \frac{\partial L}{\partial \dot{x}_i} \delta x_i + \delta\Omega(\mathbf{q}, t) &= c_a \\ \frac{\partial L}{\partial x_3} \delta x_3 + \delta\Omega(\mathbf{q}, t) &= c_a \\ m\dot{x}_3 (\delta v_3) t + \left( -mx_3 + \frac{1}{2} \frac{\partial V}{\partial x_3} t^2 \right) (\delta v_3) &= c_b \\ m\dot{x}_3 t - mx_3 + \frac{1}{2} \frac{\partial V}{\partial x_3} t^2 &= c_1 \end{aligned}$$

y la cantidad conservada queda finalmente

$$m\dot{x}_3 t - mx_3 - \frac{1}{2} F_3 t^2 = c_1 \quad (5.67)$$

De nuevo enfatizamos que  $F_3$  debe ser constante. Si adicionalmente, el Lagrangiano  $L$  también tiene invarianza traslacional a lo largo de  $X_3$ , tendremos que  $\partial_3 V = F_3 = 0$ , y la cantidad conservada se reduce a:

$$\begin{aligned} m\dot{x}_3 t - mx_3 &= c_1 \Rightarrow \\ x_3 - \dot{x}_3 t &= c_2 = x_3 - \frac{p_3}{m} t \end{aligned}$$

el valor de la constante en la última ecuación se determina fácilmente haciendo  $t = 0$  y se obtiene

$$c_2 = x_3(0) = x_3 - \frac{p_3}{m} t \quad (5.68)$$

y como  $F_3 = 0$ , la partícula se mueve con velocidad constante en la dirección  $x_3$ . Efectivamente, la Ec. (5.68) describe un movimiento uniforme en la dirección de  $X_3$ , ya que  $p_3$  es constante en virtud de la invarianza traslacional del Lagrangiano a lo largo de  $X_3$ . Recordemos que  $\dot{x}_3$  se refiere a la velocidad de la partícula, en tanto que  $v_3$  ó  $\delta v_3$  se refiere a la velocidad del sistema  $S'$  con respecto a  $S$ .

## 5.5. Ejercicios

1. Sea  $L(q_i, \dot{q}_i, \ddot{q}_i)$  un Lagrangiano asociado al “formalismo de la sacudida” (ver sección 4.6). Para este Lagrangiano definamos una coordenada cíclica  $q_k$  como una coordenada que no aparece en el Lagrangiano, pero aparece  $\dot{q}_k$  y/o  $\ddot{q}_k$ . Defina una cantidad adecuada (un momento canónico **extendido** conjugado a  $q_k$ ) que sea constante de movimiento cuando  $q_k$  es cíclica. **Sugerencia:** Observe la ecuación de movimiento (4.49), cuando está asociada a una coordenada cíclica.
2. Supongamos que el potencial  $U$  de un sistema físico depende de las velocidades generalizadas. (a) Demuestre que el momento  $p_\theta$  canónicamente conjugado a una coordenada global de rotación  $\theta$  del sistema como un todo, viene dada por

$$p_\theta = L_\theta - \sum_{i=1}^N \mathbf{n} \cdot (\mathbf{r}_i \times \nabla_{v_i} U)$$

donde  $\mathbf{n}$  es un vector unitario en la dirección del eje de rotación y  $\nabla_{v_i}$  es el operador diferencial definido como

$$\nabla_{v_i} \equiv \mathbf{u}_x \frac{\partial}{\partial v_{ix}} + \mathbf{u}_y \frac{\partial}{\partial v_{iy}} + \mathbf{u}_z \frac{\partial}{\partial v_{iz}}$$

siendo  $v_{ix}$  la componente  $x$  de la velocidad de la  $i$ -ésima partícula, y lo mismo para las otras componentes. Finalmente,  $L_\theta$  es el momento angular mecánico total a lo largo del eje de rotación i.e.

$$L_\theta = \mathbf{n} \cdot \sum_{i=1}^N (\mathbf{r}_i \times \mathbf{p}_i)$$

- (b) Aplique estos resultados a un sistema de partículas inmerso en un campo electromagnético.
3. Una partícula se mueve sobre un aro sin masa de radio  $R$ . El aro está siempre vertical y gira alrededor de un eje vertical que pasa por su diámetro, con velocidad angular constante  $\omega$ . Supondremos que las únicas fuerzas exteriores son las de gravedad. (a) Encuentre un Lagrangiano y las ecuaciones de movimiento. (b) Encuentre posibles constantes de movimiento. (c) Demuestre que si  $\omega > \omega_0$  para un cierto valor crítico  $\omega_0$ , existe una solución tal que la partícula permanece fija en el aro en un punto que no es el más bajo. Pero si  $\omega < \omega_0$ , el único punto estacionario para la partícula es el punto más bajo del aro. Encuentre el valor de  $\omega_0$ .
4. Para el péndulo esférico del problema 4 Pág. 4. (a) Encuentre los momentos canónicamente conjugados a las coordenadas generalizadas que utilizó. ¿Alguno de ellos es constante de movimiento?. (b) Encuentre la función energía, ¿se conserva?, ¿es igual a la energía del sistema?. Recuerde que sus respuestas pueden depender de las coordenadas generalizadas utilizadas.
5. Demuestre que si existe una función  $G_i$  tal que

$$\frac{\partial L}{\partial q_i} = \frac{dG_i}{dt} \tag{5.69}$$

para una coordenada generalizada  $q_i$ , entonces  $p_i - G_i$  es una constante de movimiento, siendo  $p_i$  el momento canónicamente conjugado a  $q_i$ . Si  $G_i$  existe, ¿Es única?.

6. Sea una carga puntual  $q$  no relativista de masa  $m$  inmersa en un campo eléctrico constante y homogéneo  $\mathbf{E}$ . Un campo eléctrico constante y homogéneo  $\mathbf{E}$ , se puede describir por cualquiera de los siguientes conjuntos de potenciales

$$\phi = -\mathbf{E} \cdot \mathbf{r} ; \quad \mathbf{A} = 0 \tag{5.70}$$

$$\phi' = 0 ; \quad \mathbf{A}' = -\mathbf{E}t \tag{5.71}$$

correspondientes a diferentes gauges. (a) Tomando la convención (5.70), demuestre que se conserva la cantidad  $\mathbf{p} - q\mathbf{E}t$ , donde  $\mathbf{p}$  es el vector cuyas componentes son los momentos canónicamente conjugados a las coordenadas cartesianas. (b) Demuestre que si tomamos la convención (5.71), la cantidad conservada es el momento canónicamente conjugado  $\mathbf{p}$ . (c) Demuestre que en ambos casos la cantidad conservada se reduce a  $m\dot{\mathbf{x}} - q\mathbf{E}t$ . Este ejercicio muestra que el momento canónicamente conjugado a las coordenadas de una partícula en un campo electromagnético, así como su contenido físico pueden depender del gauge.

7. Sea una carga puntual  $q$  no relativista de masa  $m$  inmersa en un campo magnético constante y homogéneo  $\mathbf{B}$ . Un campo magnético constante y homogéneo  $\mathbf{B}$ , se puede describir con los potenciales

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r} \quad ; \quad \phi = 0 \quad (5.72)$$

demuestre que se conserva la cantidad

$$\mathbf{p} - \frac{q}{2}\mathbf{r} \times \mathbf{B} \quad (5.73)$$

siendo  $\mathbf{p}$  las componentes de los momentos canónicamente conjugados. Demuestre que en términos de la velocidad, esta cantidad conservada se puede escribir como

$$m\dot{\mathbf{x}} - q\mathbf{r} \times \mathbf{B} \quad (5.74)$$

8. Una partícula de masa  $m$ , se mueve en una dimensión sometida a la fuerza

$$F(x, t) = \frac{k}{x^2} e^{-t/\tau}$$

donde  $k$  y  $\tau$  son constantes positivas. Encuentre el Lagrangiano, los momentos conjugados y la función energía. Compare la función energía con la energía total del sistema y encuentre las cantidades conservadas si las hay.

9. Obtenga el valor de la constante de movimiento asociada a la invarianza galileana Ec. (5.67), en términos de las condiciones iniciales de dos maneras: (a) haciendo  $t = 0$  en la Ec. (5.67). (b) Teniendo en cuenta que  $F_3$  es constante de modo que hay un movimiento uniformemente acelerado en  $X_3$ , y reemplazando las expresiones de  $x_3(t)$  y  $\dot{x}_3(t)$  para un movimiento uniformemente acelerado, en la Ec. (5.67).

## Ecuaciones de Movimiento de Hamilton

La formulación Hamiltoniana no tiene un contenido Físico nuevo, y no es particularmente ventajosa para la solución de problemas concretos en mecánica. El poder de este formalismo consiste en que es fácilmente extendible a otras áreas de la Física, en Mecánica Clásica este formalismo permite desarrollos posteriores como son la teoría de Hamilton Jacobi, la teoría de perturbaciones y el caos. Asumiremos ligaduras holónomas y sistemas monogénicos.

### 6.1. Consideraciones generales

Dado que las ecuaciones de Lagrange constituyen un conjunto de  $n$  ecuaciones diferenciales de segundo orden,  $2n$  constantes de movimiento se requieren para su completa solución, ellas pueden ser por ejemplo los valores de las coordenadas y velocidades generalizadas en un cierto tiempo, o el valor de las coordenadas en dos tiempos distintos. La dinámica del sistema se puede describir con una trayectoria en el espacio  $n$  dimensional de configuraciones, donde el tiempo actúa como un parámetro que traza la curva. En este formalismo, hay una ecuación asociada a cada una de las  $n$  coordenadas y su correspondiente velocidad generalizada. En su forma más simple, las ecuaciones requieren que las  $n$  coordenadas sean independientes. Para el formalismo Hamiltoniano, la independencia de las  $n$  coordenadas es esencial.

La idea fundamental detrás del formalismo Hamiltoniano, es la de convertir las  $n$  ecuaciones de segundo orden, en  $2n$  ecuaciones de primer orden. Naturalmente, el número de ecuaciones diferenciales debe duplicarse puesto que las ecuaciones de primer orden solo requieren una condición inicial y el número de condiciones iniciales totales ( $2n$ ) debe conservarse. Las  $2n$  ecuaciones diferenciales parciales deben escribirse en términos de  $2n$  variables independientes. De esta forma, la descripción de la dinámica del sistema se hará ahora en un sistema coordinado  $2n$  dimensional (espacio de fase) cuyas coordenadas serán las  $2n$  variables independientes. Es natural aunque no obligatorio, pensar que las primeras  $n$  coordenadas sean las coordenadas generalizadas  $q_k$ , veremos adicionalmente que las ecuaciones de movimiento quedan altamente simétricas si se eligen como las otras  $n$  coordenadas los momentos canónicamente conjugados definidos a través de la ecuación (5.1).

$$p_j \equiv \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_j} \quad (6.1)$$

a las cantidades  $(q_j, p_j)$  se les conoce como **variables canónicas**. La combinación de ecuaciones de primer orden y variables  $(q, p)$  resulta particularmente motivante, puesto que ya se discutió que muchas primeras integrales de movimiento conducen a cantidades conservadas, esto implica que al ser las ecuaciones de Hamilton de primer orden tales primeras integrales deben aparecer de manera más directa<sup>1</sup>. Además si una variable  $q_i$  es cíclica, su momento canónicamente conjugado es constante y dado que los momentos conjugados son parte del conjunto de variables independientes, estas constantes aparecen de manera más directa en el formalismo.

<sup>1</sup>Al ser las ecuaciones de Lagrange de segundo orden, se requiere realizar un proceso de integración para llegar a primeras integrales, las cuales son de primer orden.

En las ecuaciones de Lagrange las variables  $q, \dot{q}$  fueron tratadas todas como independientes, pero cada ecuación involucraba a  $q$  y  $\dot{q}$ . Dado que ahora queremos escribir un formalismo en términos de las variables  $q_k, p_k, t$ , debemos realizar un cambio de variables del conjunto  $(q_k, \dot{q}_k, t)$  al conjunto  $(q_k, p_k, t)$ . Por simplicidad consideremos primero un Lagrangiano que depende de una sola coordenada generalizada y una sola velocidad generalizada  $L(q, \dot{q})$ . Un diferencial de esta función se escribe como

$$dL = u dq + v d\dot{q} ; \quad u \equiv \frac{\partial L}{\partial q} ; \quad v \equiv \frac{\partial L}{\partial \dot{q}}$$

donde  $u$  y  $v$  son funciones de  $q$  y  $\dot{q}$ . Para poder escribir con base en esto una función  $H(q, p)$ , requerimos escribir un diferencial de esta función en la forma

$$dH = u' dq + v' dp ; \quad u' \equiv \frac{\partial H}{\partial q} ; \quad v' \equiv \frac{\partial H}{\partial p}$$

donde  $u', v'$  son funciones de  $q$  y  $p$ . El procedimiento matemático para realizar la transformación de la función  $L$  (Lagrangiano) a la función  $H$  (Hamiltoniano), se denomina una transformación de Legendre.

## 6.2. Transformaciones de Legendre

Escribiremos el problema en una notación más general. Sea  $f(x, y)$  de modo que el diferencial de  $f$  se escribe como

$$df = u dx + v dy ; \quad u \equiv \frac{\partial f}{\partial x} ; \quad v \equiv \frac{\partial f}{\partial y} \quad (6.2)$$

queremos cambiar la base de variables de  $(x, y)$  a  $(u, y)$ . Por tanto un diferencial de una función  $g(u, y)$  se escribe en términos de los diferenciales  $du$  y  $dy$ . Definamos  $g$  como una función de  $u$  y de  $y$  definida por

$$g = f - ux \quad (6.3)$$

escribamos el diferencial de  $g$

$$\begin{aligned} dg &= df - u dx - x du = u dx + v dy - u dx - x du \\ dg &= v dy - x du \end{aligned}$$

este diferencial tiene entonces la forma deseada. Las variables  $x$  y  $v$  son ahora funciones de  $u$  y  $y$  de la forma

$$x = -\frac{\partial g}{\partial u} ; \quad v = \frac{\partial g}{\partial y}$$

que son los análogos de (6.2). Un ejemplo de uso frecuente de la transformación de Legendre en Física aparece en la termodinámica. Para un gas experimentando un proceso reversible, se puede demostrar que el cambio diferencial de energía  $dU$  se puede escribir como

$$dU = T dS - P dV ; \quad T = \frac{\partial U}{\partial S} ; \quad P = -\frac{\partial U}{\partial V}$$

siendo  $T, S, P, V$  la temperatura, la entropía, la presión y el volumen respectivamente. A partir de la función energía interna  $U(S, V)$  se puede generar la entalpía  $H(S, P)$ , a través de una transformación de Legendre

$$\begin{aligned} H &= U + PV \Rightarrow dH = dU + P dV + V dP = T dS - P dV + P dV + V dP \\ dH &= T dS + V dP ; \quad T = \frac{\partial H}{\partial S} ; \quad V = \frac{\partial H}{\partial P} \end{aligned}$$

las energías libres de Helmholtz y de Gibbs, están dadas por otras transformaciones de Legendre

$$F \equiv U - TS ; \quad G \equiv H - TS$$

### 6.3. Generación del Hamiltoniano y Ecuaciones de Hamilton

La única diferencia importante entre las transformaciones de Legendre arriba descritas y las requeridas para generar una función de  $(q_k, p_k, t)$ , con base en una función de  $(q_k, \dot{q}_k, t)$ , es que en el último caso un número  $n$  de variables debe ser transformado. Comencemos por escribir el diferencial total de  $L$

$$dL = \frac{\partial L}{\partial q_k} dq_k + \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k + \frac{\partial L}{\partial t} dt \quad (6.4)$$

donde emplearemos de aquí en adelante, convención de suma sobre índices repetidos. De las ecuaciones de Lagrange y la definición de momento conjugado se tiene

$$\dot{p}_k = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) = \frac{\partial L}{\partial q_k} \quad (6.5)$$

de modo que (6.4) se escribe como

$$dL = \dot{p}_k dq_k + p_k d\dot{q}_k + \frac{\partial L}{\partial t} dt \quad (6.6)$$

La función Hamiltoniana o Hamiltoniano se genera por medio de la siguiente transformación de Legendre

$$H(q, p, t) \equiv \dot{q}_k p_k - L(q, \dot{q}, t) \quad (6.7)$$

la forma diferencial de  $H$  se obtiene de su definición (6.7) y de la forma diferencial del Lagrangiano (6.6)

$$\begin{aligned} dH &= p_k d\dot{q}_k + \dot{q}_k dp_k - dL \\ dH &= p_k d\dot{q}_k + \dot{q}_k dp_k - \dot{p}_k dq_k - p_k d\dot{q}_k - \frac{\partial L}{\partial t} dt \\ dH &= \dot{q}_k dp_k - \dot{p}_k dq_k - \frac{\partial L}{\partial t} dt \end{aligned} \quad (6.8)$$

y el diferencial adquiere la forma deseada, puesto que los diferenciales  $d\dot{q}_k$  han sido removidos por la transformación de Legendre, y en su lugar aparecen los diferenciales  $dp_k$ . Puesto que exigiremos que  $H$  sea función exclusiva de las variables  $(q_k, p_k, t)$ , el diferencial de  $H$  también se puede escribir de la forma

$$dH = \frac{\partial H}{\partial q_k} dq_k + \frac{\partial H}{\partial p_k} dp_k + \frac{\partial H}{\partial t} dt \quad (6.9)$$

comparando (6.8) con (6.9), y teniendo en cuenta que todas las variables  $q_k, p_k$  son todas independientes entre sí, se obtiene un conjunto de  $2n + 1$  ecuaciones

$$\begin{aligned} \dot{q}_k &= \frac{\partial H}{\partial p_k} ; \quad k = 1, \dots, n \\ -\dot{p}_k &= \frac{\partial H}{\partial q_k} ; \quad k = 1, \dots, n \end{aligned} \quad (6.10)$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \quad (6.11)$$

Las Ecs. (6.10), son conocidas como ecuaciones de movimiento de Hamilton, y constituyen el conjunto de  $2n$  ecuaciones de primer orden que se buscaba. El primer conjunto de ecuaciones se puede considerar como el inverso de las Ecs. (6.1) que definen al momento conjugado, con lo cual se puede pensar que no dan ninguna información nueva. Esto es cierto desde el punto de vista de la resolución de problemas, pero dentro del formalismo ambos conjuntos de ecuaciones tienen gran significado si el Hamiltoniano puede ser conocido de alguna manera.

Como se puede ver comparando (5.13) con (6.7), la función energía  $h$  y el hamiltoniano  $H$  son numéricamente idénticos, pero se usa un símbolo diferente para cada una puesto que  $h$  es función de  $q, \dot{q}, t$  en tanto que el Hamiltoniano debe ser función exclusiva de  $q, p, t$ . El procedimiento para utilizar las ecuaciones de Hamilton se ve más bien laborioso ya que comprende las siguientes etapas

1. Con un conjunto de coordenadas generalizadas **independientes** se construye el Lagrangiano  $L(q, \dot{q}, t) = T - V$
2. Se definen los momentos conjugados a través de (6.1) como funciones de  $q, \dot{q}, t$
3. Se usa (6.7) para construir el Hamiltoniano. Sin embargo, la Ec. (6.7) deja al Hamiltoniano como función mixta de  $q, \dot{q}, p, t$
4. Las ecuaciones (6.1) se invierten para obtener  $\dot{q}$  en función de  $q, p, t$ . Este proceso de inversión presenta varias dificultades que veremos más adelante.
5. Los resultados anteriores se aplican para eliminar las  $\dot{q}$  de  $H$  con el fin de expresar esta función únicamente en términos de  $q, p, t$

## 6.4. Algoritmo matricial para la obtención del Hamiltoniano

Afortunadamente, en muchos casos reales es posible abreviar los pasos descritos en la sección anterior, gracias a la característica ya mencionada para muchos Lagrangianos de ser separables en tres términos homogéneos de orden cero, uno y dos en las  $\dot{q}$  (ver Ec. 5.15). Recordando que  $h$  es numéricamente igual a  $H$ , retomaremos el análisis realizado en la sección 5.2.1, de modo que la Ec. (5.15) conduce a (5.17)

$$h = H = L_2 - L_0 \quad (6.12)$$

También aprendimos en la sección 5.2.1, que si además las transformaciones que llevan a las coordenadas generalizadas no dependen explícitamente del tiempo, y las fuerzas derivan de un potencial que no depende de las  $\dot{q}$ , la función energía (y por tanto el Hamiltoniano), será la energía total

$$h = H = T + V \quad (6.13)$$

Ahora haremos una suposición un tanto más restrictiva (pero suficientemente general) que la que se asume en (5.15). Asumiremos que el Lagrangiano tiene la estructura

$$L = L_0(q, t) + \dot{q}_i a_i(q, t) + \frac{1}{2} \dot{q}_i \dot{q}_k T_{ik}(q, t) \quad (6.14)$$

En tal caso y teniendo en cuenta la Ec. (6.7), el Hamiltoniano  $H$  viene dado por la siguiente prescripción

$$H = p_n \dot{q}_n - L_0(q, t) - \dot{q}_i a_i(q, t) - \frac{1}{2} \dot{q}_i \dot{q}_k T_{ik}(q, t) \quad (6.15)$$

si se cumple cualquiera de las relaciones (6.12, 6.13, 6.15), los pasos 3 y 4 arriba indicados se abreviarían. En particular, bajo la suposición (6.15), los pasos del 2 al 5 se pueden realizar de una vez al menos formalmente. Para ello escribiremos (6.14) en forma matricial

$$L(q, \dot{q}, t) = L_0(q, t) + \tilde{\mathbf{q}} \mathbf{a} + \frac{1}{2} \tilde{\mathbf{q}} \mathbf{T} \dot{\mathbf{q}} \quad (6.16)$$

donde  $\dot{\mathbf{q}}$ ,  $\mathbf{a}$  son matrices columna (no vectores Euclidianos!),  $\mathbf{T}$  es una matriz  $n \times n$ , que sin pérdida de generalidad se puede tomar como simétrica<sup>2</sup> y  $\tilde{\mathbf{q}}$  es la traspuesta de  $\dot{\mathbf{q}}$ . Los elementos de las matrices son en

<sup>2</sup>Un término cuadrático en  $\dot{q}$  típico del Lagrangiano es de la forma  $M_{12} \dot{q}_1 \dot{q}_2$ . En la suma sobre índices de  $\dot{q}_i T_{ij} \dot{q}_j$  aparecen dos términos relacionados con este coeficiente de modo que  $M_{12} \dot{q}_1 \dot{q}_2 = (\dot{q}_1 T_{12} \dot{q}_2 + \dot{q}_2 T_{21} \dot{q}_1) / 2$  (teniendo en cuenta el factor 1/2 en la Ec. 6.16). Esto implica que la única restricción sobre  $\mathbf{T}$  es que  $M_{12} = (T_{12} + T_{21}) / 2$ . En consecuencia, los elementos de la matriz se pueden definir de muchas maneras, en particular podemos elegir  $T_{12} = T_{21} = M_{12}$ , y lo mismo para los otros coeficientes, en cuyo caso la matriz será simétrica.

general funciones de  $q$  y  $t$ . La transformación de Legendre que define al Hamiltoniano Ec. (6.15), se puede escribir matricialmente como

$$\begin{aligned} H &= \tilde{\mathbf{q}}\mathbf{p} - L = \tilde{\mathbf{q}}\mathbf{p} - \left[ L_0(q, t) + \tilde{\mathbf{q}}\mathbf{a} + \frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\dot{\mathbf{q}} \right] \\ H &= \tilde{\mathbf{q}}(\mathbf{p} - \mathbf{a}) - \frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\dot{\mathbf{q}} - L_0(q, t) \end{aligned} \quad (6.17)$$

derivando (6.14) se obtienen los momentos conjugados

$$\begin{aligned} p_n &= \frac{\partial L}{\partial \dot{q}_n} = \frac{\partial}{\partial \dot{q}_n} \left[ L_0(q, t) + \dot{q}_i a_i(q, t) + \frac{1}{2}\dot{q}_i \dot{q}_k T_{ik}(q, t) \right] \\ p_n &= a_n(q, t) + \frac{1}{2}\dot{q}_k T_{nk}(q, t) + \frac{1}{2}\dot{q}_i T_{in}(q, t) \end{aligned}$$

teniendo en cuenta que la matriz  $\mathbf{T}$  se eligió como simétrica y que los índices  $k, i$  son mudos, se obtiene

$$p_n = a_n(q, t) + T_{nk}\dot{q}_k$$

de modo que los momentos conjugados escritos en forma de matriz columna  $\mathbf{p}$  están dados por

$$\mathbf{p} = \mathbf{T}\dot{\mathbf{q}} + \mathbf{a} \Rightarrow \mathbf{p} - \mathbf{a} = \mathbf{T}\dot{\mathbf{q}}$$

invirtiendo esta ecuación

$$\dot{\mathbf{q}} = \mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) \quad (6.18)$$

este paso presupone que existe el inverso de  $\mathbf{T}$ , lo cual usualmente está garantizado por la positividad de la energía cinética<sup>3</sup>. Transponiendo la ecuación anterior se cumple

$$\tilde{\mathbf{q}} = (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} \quad (6.19)$$

donde hemos usado el hecho de que  $\mathbf{T}^{-1}$  también es simétrica. Obsérvese que las expresiones (6.18, 6.19) permiten reemplazar las  $\dot{q}$  en términos de los  $p, q, t$  ya que  $\mathbf{a}$  y  $\mathbf{T}$  solo son funciones de  $q, t$ . Es decir hemos logrado el proceso descrito en el paso 4 de invertir las ecuaciones (6.1). Reemplazando (6.18, 6.19) en el Hamiltoniano (6.17)

$$\begin{aligned} H &= [(\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1}] (\mathbf{p} - \mathbf{a}) - \frac{1}{2} [(\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1}] \mathbf{T} [\mathbf{T}^{-1} (\mathbf{p} - \mathbf{a})] - L_0(q, t) \\ H &= (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - \frac{1}{2} (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - L_0(q, t) \\ H &= \frac{1}{2} (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - L_0(q, t) \end{aligned} \quad (6.20)$$

Por lo tanto, si el Lagrangiano se puede escribir en la forma (6.16), el Hamiltoniano se escribe directamente en la forma (6.20). En el caso más usual, la matriz  $\mathbf{T}$  es diagonal, en cuyo caso el inverso es también diagonal donde  $(\mathbf{T}^{-1})_{ii} = T_{ii}^{-1}$ .

#### 6.4.1. Hamiltoniano para un cuerpo sometido a una fuerza central en coordenadas esféricas

A manera de ejemplo, consideremos el movimiento de un cuerpo bajo una fuerza central en coordenadas esféricas, calculemos primero la energía cinética

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

<sup>3</sup>Técnicamente, la matriz  $\mathbf{T}$  es simétrica real y definida positiva. Esto implica que sus valores propios y su determinante son estrictamente positivos con lo cual la matriz es invertible (ver sección 12.14).



para las coordenadas esféricas tenemos

$$\begin{aligned}x &= r \sin \theta \cos \phi ; \quad y = r \sin \theta \sin \phi ; \quad z = r \cos \theta \\ \dot{x} &= \dot{r} \sin \theta \cos \phi + r \dot{\theta} \cos \theta \cos \phi - r \dot{\phi} \sin \theta \sin \phi \\ \dot{y} &= \dot{r} \sin \theta \sin \phi + r \dot{\theta} \cos \theta \sin \phi + r \dot{\phi} \sin \theta \cos \phi \\ \dot{z} &= \dot{r} \cos \theta - r \dot{\theta} \sin \theta\end{aligned}$$

$$\begin{aligned}v^2 &= \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \left( \dot{r} \sin \theta \cos \phi + r \dot{\theta} \cos \theta \cos \phi - r \dot{\phi} \sin \theta \sin \phi \right)^2 \\ &\quad + \left( \dot{r} \sin \theta \sin \phi + r \dot{\theta} \cos \theta \sin \phi + r \dot{\phi} \sin \theta \cos \phi \right)^2 + \left( \dot{r} \cos \theta - r \dot{\theta} \sin \theta \right)^2\end{aligned}$$

$$\begin{aligned}v^2 &= \dot{r}^2 \sin^2 \theta \cos^2 \phi + r^2 \dot{\theta}^2 \cos^2 \theta \cos^2 \phi + r^2 \dot{\phi}^2 \sin^2 \theta \sin^2 \phi \\ &\quad + 2\dot{r}r\dot{\theta} \sin \theta \cos \phi \cos \theta \cos \phi - 2\dot{r}r\dot{\phi} \sin \theta \cos \phi \sin \theta \sin \phi - 2r^2 \dot{\phi} \dot{\theta} \cos \theta \cos \phi \sin \theta \sin \phi \\ &\quad + \dot{r}^2 \sin^2 \theta \sin^2 \phi + r^2 \dot{\theta}^2 \cos^2 \theta \sin^2 \phi + r^2 \dot{\phi}^2 \sin^2 \theta \cos^2 \phi \\ &\quad + 2\dot{r}r\dot{\theta} \sin \theta \sin \phi \cos \theta \sin \phi + 2\dot{r}r\dot{\phi} \sin \theta \sin \phi \sin \theta \cos \phi + 2r^2 \dot{\phi} \dot{\theta} \cos \theta \sin \phi \sin \theta \cos \phi \\ &\quad + \dot{r}^2 \cos^2 \theta + r^2 \dot{\theta}^2 \sin^2 \theta - 2\dot{r}r\dot{\theta} \cos \theta \sin \theta\end{aligned}$$

reorganizando los términos cuadráticos y cruzados se obtiene

$$\begin{aligned}v^2 &= \dot{r}^2 \sin^2 \theta \cos^2 \phi + \dot{r}^2 \sin^2 \theta \sin^2 \phi + \dot{r}^2 \cos^2 \theta + r^2 \dot{\theta}^2 \cos^2 \theta \cos^2 \phi + r^2 \dot{\theta}^2 \cos^2 \theta \sin^2 \phi \\ &\quad + r^2 \dot{\phi}^2 \sin^2 \theta + r^2 \dot{\phi}^2 \sin^2 \theta \sin^2 \phi + r^2 \dot{\phi}^2 \sin^2 \theta \cos^2 \phi \\ &\quad + 2\dot{r}r\dot{\theta} \sin \theta \cos \phi \cos \theta \cos \phi + 2\dot{r}r\dot{\theta} \sin \theta \sin \phi \cos \theta \sin \phi - 2\dot{r}r\dot{\theta} \cos \theta \sin \theta \\ &\quad - 2\dot{r}r\dot{\phi} \sin \theta \cos \phi \sin \theta \sin \phi + 2\dot{r}r\dot{\phi} \sin \theta \sin \phi \sin \theta \cos \phi \\ &\quad - 2r^2 \dot{\phi} \dot{\theta} \cos \theta \cos \phi \sin \theta \sin \phi + 2r^2 \dot{\phi} \dot{\theta} \cos \theta \sin \phi \sin \theta \cos \phi\end{aligned}$$

$$\begin{aligned}v^2 &= \dot{r}^2 \sin^2 \theta (\cos^2 \phi + \sin^2 \phi) + \dot{r}^2 \cos^2 \theta + r^2 \dot{\theta}^2 \cos^2 \theta (\cos^2 \phi + \sin^2 \phi) + r^2 \dot{\theta}^2 \sin^2 \theta \\ &\quad + r^2 \dot{\phi}^2 \sin^2 \theta (\sin^2 \phi + \cos^2 \phi) + 2\dot{r}r\dot{\theta} \sin \theta \cos \theta (\cos^2 \phi + \sin^2 \phi) - 2\dot{r}r\dot{\theta} \cos \theta \sin \theta\end{aligned}$$

$$v^2 = \dot{r}^2 (\sin^2 \theta + \cos^2 \theta) + r^2 \dot{\theta}^2 (\cos^2 \theta + \sin^2 \theta) + r^2 \dot{\phi}^2 \sin^2 \theta + 2\dot{r}r\dot{\theta} \sin \theta \cos \theta - 2\dot{r}r\dot{\theta} \cos \theta \sin \theta$$

$$v^2 = \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \dot{\phi}^2 \sin^2 \theta$$

y la energía cinética final es

$$\frac{1}{2}mv^2 = \frac{1}{2}m \left[ \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \dot{\phi}^2 \sin^2 \theta \right] \quad (6.21)$$

la energía potencial se escribe  $V(r)$  de modo que el Lagrangiano tiene la forma

$$L = T - V = \frac{m}{2} \left( \dot{r}^2 + r^2 \sin^2 \theta \dot{\phi}^2 + r^2 \dot{\theta}^2 \right) - V(r) \quad (6.22)$$

este Lagrangiano se puede descomponer en la forma  $L = L_0 + L_1 + L_2$ , adicionalmente la transformación de coordenadas cartesianas a esféricas no depende explícitamente del tiempo y el potencial no depende de las  $\dot{q}$  ni del tiempo, por tanto el Hamiltoniano es la energía total y además se conserva. Claramente, el Lagrangiano (6.22), posee la siguiente estructura

$$L_2 = T_2(\dot{r}, \dot{\phi}, \dot{\theta}) = T \quad ; \quad L_1 = 0 \quad ; \quad L_0 = -V(r) \quad (6.23)$$

por lo tanto el Hamiltoniano tiene la forma (6.12). En arreglo matricial, el Lagrangiano (6.22) se escribe

$$L = -V(r) + \begin{pmatrix} \dot{r} & \dot{\phi} & \dot{\theta} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \dot{r} & \dot{\phi} & \dot{\theta} \end{pmatrix} \begin{pmatrix} m & 0 & 0 \\ 0 & mr^2 \sin^2 \theta & 0 \\ 0 & 0 & mr^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\phi} \\ \dot{\theta} \end{pmatrix}$$

comparando con (6.16) resulta

$$L_0 = -V ; \quad \mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ; \quad \mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & mr^2 \sin^2 \theta & 0 \\ 0 & 0 & mr^2 \end{pmatrix} ; \quad \mathbf{T}^{-1} = \begin{pmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{mr^2 \sin^2 \theta} & 0 \\ 0 & 0 & \frac{1}{mr^2} \end{pmatrix} \quad (6.24)$$

reemplazando estas expresiones en (6.20)

$$H = \frac{1}{2} \tilde{\mathbf{p}} \mathbf{T}^{-1} \mathbf{p} + V(r) = \frac{1}{2} \begin{pmatrix} p_r & p_\phi & p_\theta \end{pmatrix} \begin{pmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{mr^2 \sin^2 \theta} & 0 \\ 0 & 0 & \frac{1}{mr^2} \end{pmatrix} \begin{pmatrix} p_r \\ p_\phi \\ p_\theta \end{pmatrix} + V(r)$$

Finalmente, el Hamiltoniano de una partícula sometida a una interacción central definida por el potencial  $V(r)$ , se escribe en coordenadas esféricas de la forma

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_\phi^2}{r^2 \sin^2 \theta} + \frac{p_\theta^2}{r^2} \right) + V(r) \quad (6.25)$$

si reescribimos el problema en coordenadas cartesianas, el Hamiltoniano tiene otra forma funcional

$$T = \frac{m}{2} \dot{x}_i \dot{x}_i \Rightarrow H(x_i, p_i) = \frac{p_i p_i}{2m} + V(\sqrt{x_i x_i})$$

escribiendo los momentos conjugados  $p_i$  en un arreglo en forma de matriz columna queda

$$H(x_i, p_i) = \frac{\mathbf{p} \cdot \mathbf{p}}{2m} + V(\sqrt{x_i x_i})$$

las componentes de  $\mathbf{p}$  se pueden tomar relativos a cualquier sistema coordenado que deseemos. Pero es importante no confundir  $p_k$  con la componente  $k$ -ésima de  $\mathbf{p}$ . Por ejemplo  $p_\theta \neq (\mathbf{p})_\theta$ , el primero tiene dimensiones de momento angular y el segundo tiene dimensiones de momento lineal. De aquí en adelante, cuando se use un vector para representar a un momento conjugado, se referirá a momentos conjugados a coordenadas cartesianas de posición, a menos que se indique lo contrario (recordemos que en general las  $n$ -uplas de  $p_k$  no forman vectores euclidianos al igual que en el caso de las  $q_k$ ).

### 6.4.2. Hamiltoniano de una carga no relativista inmersa en un campo electromagnético

Como un segundo ejemplo consideremos una carga no relativista inmersa en un campo electromagnético, el Lagrangiano se escribe

$$\begin{aligned} L &= \frac{1}{2} m v^2 - q\phi + q\mathbf{A} \cdot \mathbf{v} = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + qA_x \dot{x} + qA_y \dot{y} + qA_z \dot{z} - q\phi \\ &= -q\phi + q \begin{pmatrix} \dot{x} & \dot{y} & \dot{z} \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \dot{x} & \dot{y} & \dot{z} \end{pmatrix} \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} \end{aligned}$$

de nuevo comparando con (6.16) resulta

$$\begin{aligned} L_0 &= -q\phi ; \quad \mathbf{a} = \begin{pmatrix} qA_x \\ qA_y \\ qA_z \end{pmatrix} ; \quad \mathbf{T} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} = m\mathbf{1} \\ \mathbf{p} - \mathbf{a} &= \begin{pmatrix} p_x - qA_x \\ p_y - qA_y \\ p_z - qA_z \end{pmatrix} ; \quad \mathbf{T}^{-1} = \begin{bmatrix} 1/m & 0 & 0 \\ 0 & 1/m & 0 \\ 0 & 0 & 1/m \end{bmatrix} = \frac{1}{m} \end{aligned}$$

y reemplazando en (6.20)

$$\begin{aligned}
 H &= \frac{1}{2} (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - L_0 = \frac{1}{2} \begin{pmatrix} p_x - qA_x & p_y - qA_y & p_z - qA_z \end{pmatrix} \frac{1}{m} \begin{pmatrix} p_x - qA_x \\ p_y - qA_y \\ p_z - qA_z \end{pmatrix} + q\phi \\
 H &= \frac{1}{2m} (p_i - qA_i) (p_i - qA_i) + q\phi \\
 H &= \frac{1}{2m} (\mathbf{p} - q\mathbf{A})^2 + q\phi
 \end{aligned} \tag{6.26}$$

donde el momento canónicamente conjugado  $\mathbf{p}$  ya había sido calculado en la sección 5.1, Ec. (5.2), Pág. 70

$$p_i = m\dot{x}_i + qA_i \tag{6.27}$$

de modo que el Hamiltoniano en términos de la velocidad (o más bien la función energía), queda

$$h = H = \frac{m\dot{\mathbf{x}}^2}{2} + q\phi = T + q\phi$$

nótese que en este caso el potencial tiene un término lineal en las velocidades de modo que el Hamiltoniano no corresponde a  $T + U$ . Sin embargo, en este caso particular el Hamiltoniano aún corresponde a la energía total, ya que el término magnético, que es el que rompe la condición para que  $H = T + U$ , no produce trabajo<sup>4</sup> y la energía potencial (entendida como la capacidad para realizar trabajo) depende solo de  $\phi$ . De lo anterior se deduce que la condición  $H = T + U$  es suficiente pero no necesaria para que  $H$  sea la energía del sistema.

## 6.5. Forma Simpléctica de las Ecuaciones de Hamilton

Las ecuaciones de Hamilton

$$\begin{aligned}
 \dot{q}_k &= \frac{\partial H}{\partial p_k} ; \quad k = 1, \dots, n \\
 -\dot{p}_k &= \frac{\partial H}{\partial q_k} ; \quad k = 1, \dots, n
 \end{aligned} \tag{6.28}$$

son muy simétricas excepto por un cambio de signo al intercambiar el papel de las variables canónicas  $q_k \leftrightarrow p_k$ . Una estrategia muy fructífera para llegar a ecuaciones más simétricas es el uso de las ecuaciones de Hamilton en notación simpléctica. Para un sistema con  $n$  grados de libertad, construimos una matriz columna  $\boldsymbol{\eta}$  con  $2n$  elementos y definida por

$$\eta_i = q_i, \quad \eta_{i+n} = p_i ; \quad i \leq n \tag{6.29}$$

similarmente la matriz columna  $\frac{\partial H}{\partial \boldsymbol{\eta}}$  tiene elementos

$$\left( \frac{\partial H}{\partial \boldsymbol{\eta}} \right)_i = \frac{\partial H}{\partial q_i} ; \quad \left( \frac{\partial H}{\partial \boldsymbol{\eta}} \right)_{i+n} = \frac{\partial H}{\partial p_i} ; \quad i \leq n \tag{6.30}$$

finalmente definamos una matriz  $\mathbf{J}$  de dimension  $2n \times 2n$

$$\mathbf{J} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{bmatrix} ; \quad \tilde{\mathbf{J}} \equiv \begin{bmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{6.31}$$

donde cada elemento representa una submatriz  $n \times n$ . Es fácil verificar las siguientes propiedades para  $\mathbf{J}$

$$\begin{aligned}
 \mathbf{J}\tilde{\mathbf{J}} &= \tilde{\mathbf{J}}\mathbf{J} = \mathbf{1} ; \quad \tilde{\tilde{\mathbf{J}}} = -\mathbf{J} = \mathbf{J}^{-1} \\
 \mathbf{J}^2 &= -\mathbf{1} ; \quad \det \mathbf{J} = +1
 \end{aligned} \tag{6.32}$$

<sup>4</sup>Nótese que la noción de potencial generalizado se creó como un artificio para poder absorber en el Lagrangiano la información física contenida en ciertas fuerzas, pero no como artificio para calcular trabajo que es la forma en que surgen los potenciales que solo dependen de la posición y están asociados a la energía de la partícula.

las ecuaciones de Hamilton (6.10) se pueden escribir entonces de forma compacta, así

$$\dot{\eta} = \mathbf{J} \frac{\partial H}{\partial \eta} \quad (6.33)$$

por ejemplo, para dos coordenadas las ecuaciones de Hamilton son

$$\dot{q}_1 = \frac{\partial H}{\partial p_1}, \quad \dot{q}_2 = \frac{\partial H}{\partial p_2}; \quad -\dot{p}_1 = \frac{\partial H}{\partial q_1}; \quad -\dot{p}_2 = \frac{\partial H}{\partial q_2} \quad (6.34)$$

la matriz columna  $\eta$ , y la matriz  $\mathbf{J}$  serían

$$\eta = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix}; \quad \mathbf{J} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

es claro que

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\dot{p}_1 \\ -\dot{p}_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} \quad (6.35)$$

sustituyendo las Ecs. (6.34) en (6.35), nos queda

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \partial H / \partial q_1 \\ \partial H / \partial q_2 \\ \partial H / \partial p_1 \\ \partial H / \partial p_2 \end{bmatrix} \quad (6.36)$$

y teniendo en cuenta la notación definida en (6.30), es claro que las Ecs. (6.36), son la forma explícita de las ecuaciones de Hamilton (6.33) en estructura simpléctica, para dos coordenadas generalizadas. Esta forma matricial simpléctica de las ecuaciones de Hamilton, será muy utilizada en lo que sigue. La palabra simpléctico significa “entrelazado”, efectivamente las ecuaciones de Hamilton entrelazan a  $q$  y  $p$  y tal entrelazamiento en la parametrización (6.33) se produce gracias a los elementos fuera de la diagonal de la matriz  $\mathbf{J}$  en la Ec. (6.31). Nótese además que la matriz  $\mathbf{J}$  definida en la Ec. (6.31), “absorbe” el cambio de signo que se produce en las ecuaciones de Hamilton al intercambiar  $q_k \leftrightarrow p_k$ . Esto lo hace gracias a la diferencia de signo entre las submatrices de  $\mathbf{J}$ . Esta “absorción” del cambio de signo es lo que le da una apariencia más simétrica a las ecuaciones de Hamilton simplécticas, con respecto a las ecuaciones de Hamilton normales.

## 6.6. Coordenadas cíclicas y teoremas de conservación

En el formalismo Lagrangiano hemos definido una coordenada cíclica o ignorable como aquella coordenada generalizada que no aparece en el Lagrangiano, pero sí aparece su velocidad generalizada correspondiente. Vimos que para este tipo de coordenadas su momento canónicamente conjugado es una constante de movimiento. De la transformación de Legendre que define a  $H$ , Ec. (6.7), tenemos que

$$H = \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L(q, \dot{q}, t)$$

y claramente una coordenada cíclica no aparece al lado derecho de esta ecuación, mostrando que una coordenada cíclica tampoco aparece en el Hamiltoniano<sup>5</sup>. Por otro lado, a partir de las Ecs. (6.5) y (6.28) se encuentra que

$$\dot{p}_j = \frac{\partial L}{\partial q_j} = -\frac{\partial H}{\partial q_j} \quad (6.37)$$

<sup>5</sup>Debe añadirse que al escribir el Hamiltoniano en sus argumentos finales  $q, p, t$ ; una coordenada cíclica tampoco aparece, ya que al invertir las Ecs. (6.1), las  $\dot{q}_i$  tampoco dependerán de una variable cíclica.

de aquí se deduce que una coordenada generalizada está ausente en el Hamiltoniano si y solo si está ausente en el Lagrangiano. En consecuencia, si cierta coordenada generalizada está ausente de  $H$ , su momento canónicamente conjugado será constante de movimiento. Por tanto, las constantes de movimiento asociadas a variables cíclicas aparecen de inmediato como se anticipó debido a que las ecuaciones son de primer orden y que los momentos conjugados se tomaron como coordenadas independientes. Esto significa que los teoremas de conservación que se derivaron en la sección 5.1 se pueden derivar de la formulación Hamiltoniana, así como la conexión entre simetrías del sistema y constantes de movimiento. En particular, la invarianza ante traslaciones del sistema como un todo en cierta dirección conduce a la conservación de la componente del momento lineal del sistema en esa dirección, esta coordenada de traslación no aparece en  $H$ . La invarianza ante una rotación con respecto a cierto eje hace que dicha coordenada generalizada de rotación (ángulo), no aparezca en el Hamiltoniano y lleva a la conservación de la componente del momento angular total a lo largo del eje de rotación.

Cuando trabajamos la función energía en la sección 5.2, vimos que si  $L$  no es función explícita del tiempo,  $h$  es una constante de movimiento. Naturalmente, esto es también válido para  $H$ . Con el fin de probar la consistencia de esta aseveración llegaremos a la misma conclusión por otro camino: Tomando la derivada total de  $H$  con respecto al tiempo

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial t}$$

usando las Ecs. de Hamilton (6.10), se eliminan los dos primeros términos de la derecha, y utilizando (6.11) resulta

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad (6.38)$$

de aquí resulta que  $t$  no aparece explícitamente en el Lagrangiano ( $\partial_t L = 0$ ) si y solo si, no aparece explícitamente en el Hamiltoniano ( $\partial_t H = 0$ ). En consecuencia, cuando  $t$  no aparece explícitamente en el Hamiltoniano, tenemos que  $H$  será una constante de movimiento ( $dH/dt = 0$ )<sup>6</sup>. Recordemos que la conservación de  $h$  (y  $H$ ) no significa necesariamente que esta función coincida con la energía total del sistema.

Por otro lado, también se vió en la sección 5.2 que la función  $h$  (y por tanto  $H$ ) es la energía total del sistema si se cumplen las siguientes condiciones: (a) El Lagrangiano es expresable en la forma  $L = L_0 + L_1 + L_2$  siendo cada sumando una función homogénea de grado cero uno y dos en las  $\dot{q}_i$  (b) las transformaciones que definen las coordenadas generalizadas Ecs. (2.5) no dependen explícitamente del tiempo, y (c) el potencial es independiente de las velocidades generalizadas. Si el potencial depende explícitamente del tiempo, entonces la energía del sistema (el Hamiltoniano) no es una constante de movimiento, pero si además de las anteriores condiciones ocurre que el potencial no depende explícitamente del tiempo, la energía del sistema (el Hamiltoniano) es una constante de movimiento.

## 6.7. El Hamiltoniano en diferentes sistemas coordenados

La anterior discusión muestra que la identificación de  $H$  como la energía del sistema o como constante de movimiento son dos cosas aparte, aunque no necesariamente excluyentes. Otro aspecto discutido en la sección 5.2 es la dependencia tanto funcional como numérica de  $h$  con respecto a la escogencia del sistema de coordenadas generalizadas. Lo mismo ocurre para  $H$ . Es posible por ejemplo, que para una escogencia de coordenadas  $H$  se conserve, en tanto que con otra escogencia el nuevo  $H'$  no se conserve. Por otro lado, es plausible que el Hamiltoniano en cierto sistema coordenado corresponda a la energía del sistema, en tanto que en otro sistema coordenado no lo sea. Esta es una de las diferencias mas fundamentales entre  $H$  y  $L$ , ya que este último cambia en forma funcional pero no en magnitud, cuando se hace un cambio a otras coordenadas generalizadas. Ilustraremos las anteriores consideraciones con un ejemplo.

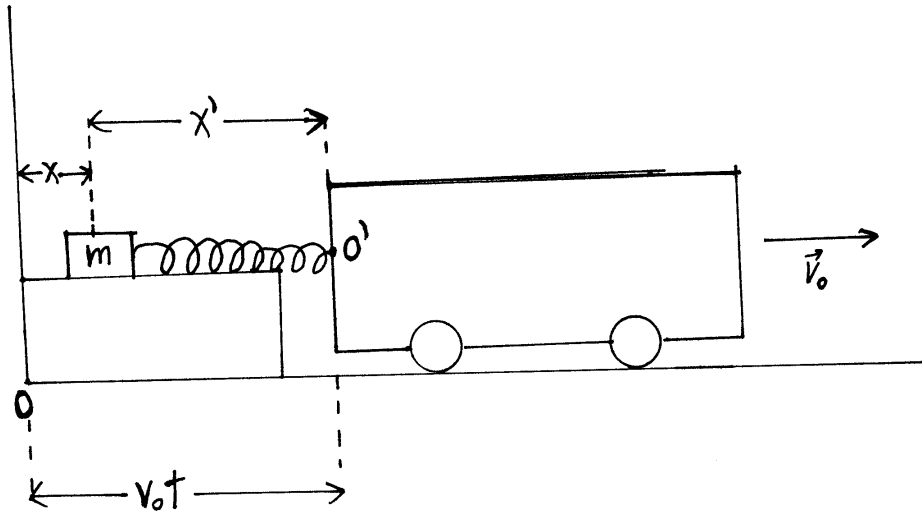


Figura 6.1: Sistema masa resorte donde el resorte se ata a un carro que viaja a velocidad constante.

### 6.7.1. Hamiltoniano de un sistema masa resorte en diferentes sistemas coordenados

Veamos un ejemplo un tanto académico que muestra claramente la dependencia del Hamiltoniano con las coordenadas generalizadas elegidas: Sea una partícula atada a un resorte cuyo extremo está sujeto a un carro que se mueve con velocidad constante  $v_0$  como muestra la figura 6.1. La masa descansa sobre una superficie sin rozamiento. Tomemos como coordenada generalizada la posición  $x$  de la partícula en cualquier instante. Por simplicidad asumimos que en  $t = 0$  el resorte tiene longitud cero y que la masa (y el punto  $O'$  de unión del resorte con el carro) pasan por el origen en ese instante. El Lagrangiano se escribe

$$L(x, \dot{x}, t) = T - V = \frac{m\dot{x}^2}{2} - \frac{k}{2}(v_0 t - x)^2 \Rightarrow m\ddot{x} = -k(x - v_0 t)$$

una forma de resolver el problema es cambiar a una variable  $x'$

$$x' = x - v_0 t \quad (6.39)$$

$$\Rightarrow m\ddot{x}' = -kx' \quad (6.40)$$

donde  $x'$  es el desplazamiento de la partícula medido desde el punto  $O'$ . La Ec. (6.40) nos dice que el sistema de referencia asociado al carro móvil vé un movimiento armónico simple como se espera del principio de equivalencia de Galileo<sup>7</sup>.

Veamos ahora la formulación Hamiltoniana. En términos de la variable  $x$  vemos que la transformación (identidad) a coordenadas generalizadas no depende explícitamente del tiempo, el Lagrangiano se puede descomponer en la forma  $L = L_0 + L_1 + L_2$  y además el potencial no dependen de la velocidad generalizada  $\dot{x}$ . Por tanto, el Hamiltoniano es la energía total del sistema y queda

$$H(x, p, t) = T + V = \frac{p^2}{2m} + \frac{k}{2}(x - v_0 t)^2$$

sin embargo  $H$  no es una cantidad conservada puesto que depende explícitamente del tiempo. Esto es físicamente entendible ya que un agente externo debe proveer de energía al sistema para que el carro viaje a velocidad constante en contra de la reacción de la partícula.

<sup>6</sup>Esto se debe a su vez a la propiedad particular de  $H$  (o de  $h$ ), de que su derivada parcial temporal coincide con su derivada temporal total. En este caso, hemos probado esta aseveración como una consecuencia de las ecuaciones de Hamilton.

<sup>7</sup>De la Ec. (6.39) vemos que para el caso particular de la Fig. 6.1, la variable  $x'$  es negativa, puesto que  $v_0 t \geq x \geq 0$ .

Ahora formulemos el Lagrangiano haciendo una transformación de la coordenada  $x$  a la coordenada  $x'$  Ec. (6.39). El Lagrangiano queda

$$\begin{aligned} L &= \frac{m \left(\frac{dx}{dt}\right)^2}{2} - \frac{k}{2} (x - v_0 t)^2 = \frac{m \left[\frac{d}{dt} (x' + v_0 t)\right]^2}{2} - \frac{k}{2} [(x' + v_0 t) - v_0 t]^2 \\ L(x', \dot{x}', t) &= \frac{m [\dot{x}' + v_0]^2}{2} - \frac{k}{2} x'^2 = \frac{m \dot{x}'^2}{2} + m v_0 \dot{x}' + \frac{m v_0^2}{2} - \frac{k}{2} x'^2 \end{aligned} \quad (6.41)$$

escribiendo el Lagrangiano de acuerdo con la estructura (6.16) en donde las matrices son  $1 \times 1$  tenemos

$$\begin{aligned} L(x', \dot{x}', t) &= \frac{m v_0^2}{2} - \frac{k}{2} x'^2 + m v_0 \dot{x}' + \frac{1}{2} \dot{x}' m \dot{x}' \\ L_0 &= \frac{m v_0^2}{2} - \frac{k}{2} x'^2 \quad ; \quad \mathbf{a} = m v_0 \quad ; \quad \mathbf{T} = m \end{aligned} \quad (6.42)$$

el nuevo Hamiltoniano se obtiene reemplazando (6.42) en (6.20)

$$\begin{aligned} H'(x', p', t) &= \frac{1}{2} (p' - a) T^{-1} (p' - a) - L_0 = \frac{1}{2} (p' - m v_0) \frac{1}{m} (p' - m v_0) + \frac{k}{2} x'^2 - \frac{m v_0^2}{2} \\ H'(x', p', t) &= \frac{(p' - m v_0)^2}{2m} + \frac{k x'^2}{2} - \frac{m v_0^2}{2} \end{aligned}$$

La Ec. (6.41) nos indica que el Lagrangiano escrito en el nuevo sistema coordenado, posee un término lineal en  $\dot{x}'$ . Por otro lado, la transformación a coordenadas generalizadas depende explícitamente del tiempo según se vé en la Ec.(6.39). Esto indica que  $H'$  ya no es la energía total del sistema. En cambio, sí es una constante de movimiento ya que no depende explícitamente del tiempo y  $\partial H'/\partial t = dH'/dt = 0$ . Por otro lado, se puede notar que el término  $m v_0^2/2$  es una constante que se puede remover tanto del Lagrangiano como del Hamiltoniano, sin afectar las ecuaciones de movimiento. Finalmente, se puede ver que excepto por tal término constante,  $H'$  se puede identificar con la energía total de movimiento de la partícula relativa al carro móvil. Ambos Hamiltonianos difieren en magnitud, forma funcional y dependencia temporal. Sin embargo, se puede verificar que ambos conducen al mismo movimiento de la partícula.

Con lo anterior podría quedar la sensación de que hemos cambiado de sistema de referencia en virtud de que  $x'$  es la coordenada que mediría el sistema del carro móvil (lo llamaremos sistema  $S'$ ) y  $H'$  sería la energía total relativa a dicho sistema también. No obstante, la coordenada  $x'$  es medida por el sistema original fijo a tierra (lo llamaremos  $S$ ), esto se puede ver en la forma en que se construye el Lagrangiano Ec. (6.41), tanto la energía cinética como la potencial se siguen midiendo con respecto al sistema  $S$  aunque se escriban en términos de la coordenada  $x'$ . Efectivamente, ante un cambio de sistema de coordenadas el Lagrangiano preserva su magnitud (aunque no su forma funcional). En contraste, ante un cambio de sistema de referencia tanto la magnitud como la forma funcional del Lagrangiano pueden cambiar. El Hamiltoniano  $H'$  se construyó usando el Lagrangiano (6.41) de modo que también está asociado al sistema  $S$ . Las interpretaciones como coordenada relativa a  $S'$  y energía relativa a  $S'$  son una forma de ver a estas cantidades, pero no han sido construídas en este sistema de referencia. Nuevamente se insiste en no confundir un cambio de sistema de referencia con un cambio de sistema de coordenadas generalizadas.

## 6.8. Problemas de aplicación de las ecuaciones de Hamilton

### 6.8.1. Partícula sobre superficie cilíndrica

Una partícula está sujeta a moverse sobre una superficie cilíndrica definida por  $x^2 + y^2 = R^2$ , y está sujeta a una fuerza que siempre apunta hacia el origen y proporcional a la distancia de la partícula al origen, i.e.  $\mathbf{F} = -kr$ .

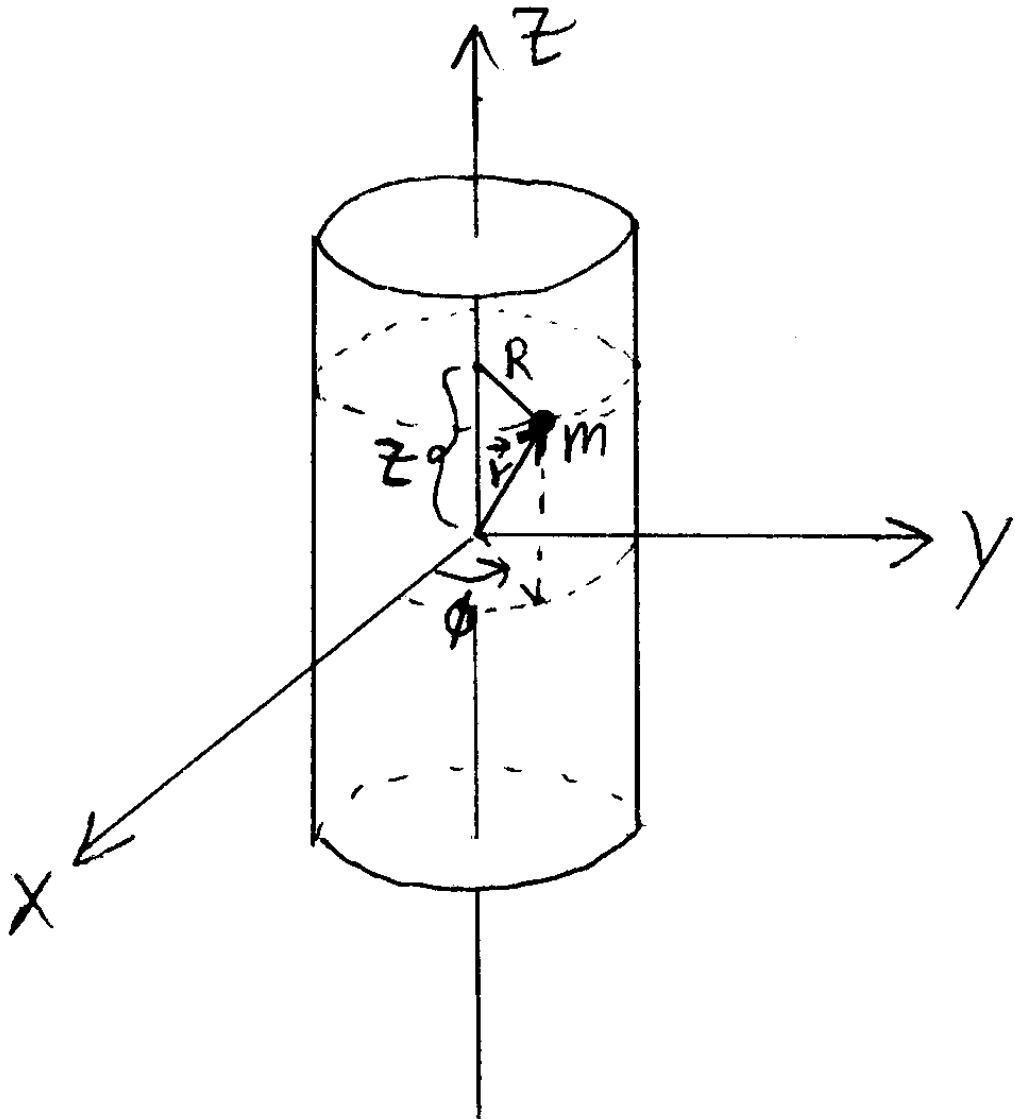


Figura 6.2: Partícula restringida a moverse sobre una superficie cilíndrica.

Tomaremos las coordenadas generalizadas  $z$  y  $\phi$  indicadas en la figura 6.2. Su relación con las coordenadas cartesianas es

$$x = R \cos \phi \quad , \quad y = R \sin \phi \quad , \quad z = z \quad (6.43)$$

$$\dot{x} = -\dot{\phi} R \sin \phi \quad , \quad \dot{y} = \dot{\phi} R \cos \phi \quad , \quad \dot{z} = \dot{z} \quad (6.44)$$

La partícula está sometida a la fuerza restauradora y la fuerza de ligadura que la mantiene sobre la superficie cilíndrica. Por supuesto, solo la fuerza restauradora contribuye a la energía potencial, la cual viene dada por

$$V' = \frac{1}{2} k r^2 = \frac{1}{2} k (x^2 + y^2 + z^2) = \frac{1}{2} k (R^2 + z^2)$$

$$V = \frac{1}{2} k z^2$$

donde hemos suprimido el término constante  $(1/2) k R^2$  obteniendo una energía potencial equivalente. La



energía cinética es

$$\begin{aligned} T &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m(\dot{\phi}^2 R^2 \sin^2 \phi + \dot{\phi}^2 R^2 \cos^2 \phi + \dot{z}^2) \\ T &= \frac{1}{2}m(R^2 \dot{\phi}^2 + \dot{z}^2) \end{aligned}$$

el Lagrangiano queda en la forma

$$\begin{aligned} L &= T - V = \frac{1}{2}m(R^2 \dot{\phi}^2 + \dot{z}^2) - \frac{1}{2}kz^2 = L_0 + L_1 + L_2 \\ L_0 &= -\frac{1}{2}kz^2, \quad L_1 = 0, \quad L_2 = \frac{1}{2}m(R^2 \dot{\phi}^2 + \dot{z}^2) \end{aligned}$$

la transformación de coordenadas (6.43) no depende explícitamente del tiempo, el Lagrangiano se puede descomponer en la forma  $L_0 + L_1 + L_2$ , y el potencial no depende de las velocidades generalizadas. Por tanto la función energía es la energía del sistema

$$h = T + V = \frac{1}{2}m(R^2 \dot{\phi}^2 + \dot{z}^2) + \frac{1}{2}kz^2$$

para encontrar el Hamiltoniano en los argumentos adecuados, podemos en este caso invertir directamente las ecuaciones (6.1), ya que

$$\begin{aligned} p_\phi &= \frac{\partial L}{\partial \dot{\phi}} = mR^2 \dot{\phi}, \quad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} \Rightarrow \\ \dot{\phi} &= \frac{p_\phi}{mR^2}; \quad \dot{z} = \frac{p_z}{m} \end{aligned} \quad (6.45)$$

sustituyendo estas expresiones en la función energía, obtenemos el Hamiltoniano

$$\begin{aligned} H &= \frac{1}{2}m \left[ R^2 \left( \frac{p_\phi}{mR^2} \right)^2 + \left( \frac{p_z}{m} \right)^2 \right] + \frac{1}{2}kz^2 \\ H &= \frac{1}{2m} \left[ \frac{p_\phi^2}{R^2} + p_z^2 \right] + \frac{1}{2}kz^2 \end{aligned}$$

el Hamiltoniano no depende explícitamente del tiempo y por tanto la energía de la partícula se conserva. Las ecuaciones de Hamilton nos dan

$$\dot{p}_\phi = -\frac{\partial H}{\partial \phi} = 0, \quad \dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{mR^2} \quad (6.46)$$

$$\dot{p}_z = -\frac{\partial H}{\partial z} = -kz, \quad \dot{z} = \frac{\partial H}{\partial p_z} = \frac{p_z}{m} \quad (6.47)$$

Nótese que las ecuaciones de Hamilton para  $\dot{\phi}$  y  $\dot{z}$  son las mismas ecuaciones (6.45) de definición de los momentos conjugados. Las dos ecuaciones (6.46) asociadas a  $p_\phi$  y  $\dot{\phi}$  tienen solución inmediata, debido al carácter cíclico de la coordenada  $\phi$

$$p_\phi = mR^2 \dot{\phi} = c_0 \quad (6.48)$$

de modo que el momento angular con respecto al eje  $Z$  se conserva, esto a su vez proviene del hecho de que el problema es invariante ante una rotación alrededor de dicho eje, puesto que es un eje de simetría del sistema. Equivalentemente, la velocidad angular de giro alrededor del eje  $z$  es constante. Integrando la ecuación (6.48) se obtiene  $\phi(t)$

$$\phi(t) = \frac{c_0}{mR^2}t + \phi_0; \quad c_0 = p_\phi = \frac{\dot{\phi}}{mR^2} \quad (6.49)$$

Derivando la segunda de las Ecs. (6.47) obtenemos  $\dot{p}_z = m\ddot{z}$ , y reemplazando esto en la primera de las Ecs. (6.47) queda

$$\ddot{z} + \omega_0^2 z = 0 \quad ; \quad \omega_0^2 \equiv \frac{k}{m} \quad (6.50)$$

la proyección del movimiento sobre  $Z$  es entonces armónica simple con frecuencia angular  $\omega_0$ . Es claro de las ecuaciones (6.49, 6.50) que las constantes de integración se obtienen conociendo condiciones iniciales tales como  $\phi(0)$ ,  $\dot{\phi}(0)$ ,  $z(0)$ ,  $\dot{z}(0)$ . Equivalentemente puede ser el conjunto  $\phi(0)$ ,  $p_\phi$ ,  $z(0)$ ,  $p_z$  etc.

### 6.8.2. Ejemplo de aplicación del algoritmo matricial

Supongamos que cierto sistema físico viene descrito por el Lagrangiano

$$L = k\dot{q}_1^2 + \frac{\dot{q}_2^2}{a + bq_1^2} + k_1q_1^2 + k_2\dot{q}_1\dot{q}_2 \quad (6.51)$$

donde  $a, b, k_1, k_2$  y  $k$  son constantes. Encontraremos el Hamiltoniano asociado. Para ello, emplearemos el algoritmo matricial descrito en la sección 6.4, observando primero que el Lagrangiano tiene la estructura matricial dada por la Ec. (6.16)

$$L(q, \dot{q}, t) = L_0(q, t) + \tilde{\mathbf{q}}\mathbf{a} + \frac{1}{2}\tilde{\mathbf{q}}\mathbf{T}\dot{\mathbf{q}} \quad (6.52)$$

como se puede ver al escribir el Lagrangiano (6.51) en la forma

$$L = k_1\dot{q}_1^2 + \frac{1}{2} \begin{pmatrix} \dot{q}_1 & \dot{q}_2 \end{pmatrix} \begin{pmatrix} 2k & k_2 \\ k_2 & \frac{2}{a+bq_1^2} \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} \quad (6.53)$$

comparando las ecuaciones (6.52, 6.53) se tiene que<sup>8</sup>

$$L_0 = k_1\dot{q}_1^2 \quad ; \quad \mathbf{a} = 0 \quad ; \quad \mathbf{T} = \begin{pmatrix} 2k & k_2 \\ k_2 & \frac{2}{a+bq_1^2} \end{pmatrix} \quad (6.54)$$

para un Lagrangiano con la estructura (6.52), el Hamiltoniano viene dado por (6.20)

$$H(q, p, t) = \frac{1}{2}(\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) - L_0(q, t) \quad (6.55)$$

por tanto, debemos calcular  $\mathbf{T}^{-1}$

$$\mathbf{T}^{-1} = \begin{pmatrix} \frac{2}{4k - bk_2^2q_1^2 - ak_2^2} & \frac{bk_2q_1^2 + ak_2}{bk_2^2q_1^2 + ak_2^2 - 4k} \\ \frac{bk_2q_1^2 + ak_2}{bk_2^2q_1^2 + ak_2^2 - 4k} & \frac{2k(bq_1^2 + a)}{4k - bk_2^2q_1^2 - ak_2^2} \end{pmatrix} \quad (6.56)$$

sustituyendo (6.54, 6.56) en la Ec. (6.55), el Hamiltoniano queda

$$\begin{aligned} H(q, p, t) &= \frac{1}{2}\tilde{\mathbf{p}}\mathbf{T}^{-1}\mathbf{p} - k_1q_1^2 = \frac{1}{2} \begin{pmatrix} p_1 & p_2 \end{pmatrix} \begin{pmatrix} \frac{2}{4k - bk_2^2q_1^2 - ak_2^2} & \frac{bk_2q_1^2 + ak_2}{bk_2^2q_1^2 + ak_2^2 - 4k} \\ \frac{bk_2q_1^2 + ak_2}{bk_2^2q_1^2 + ak_2^2 - 4k} & \frac{2k(bq_1^2 + a)}{4k - bk_2^2q_1^2 - ak_2^2} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} - k_1q_1^2 \\ H(q, p, t) &= \frac{p_1^2}{4k - ak_2^2 - bk_2^2q_1^2} + \frac{k(a + bq_1^2)}{4k - ak_2^2 - bk_2^2q_1^2} \frac{p_2^2}{2} + \frac{(ak_2 + bk_2q_1^2)}{ak_2^2 + bk_2^2q_1^2 - 4k} \frac{p_1p_2}{2} - k_1q_1^2 \end{aligned} \quad (6.57)$$

se deja como ejercicio al lector escribir las ecuaciones de Hamilton.

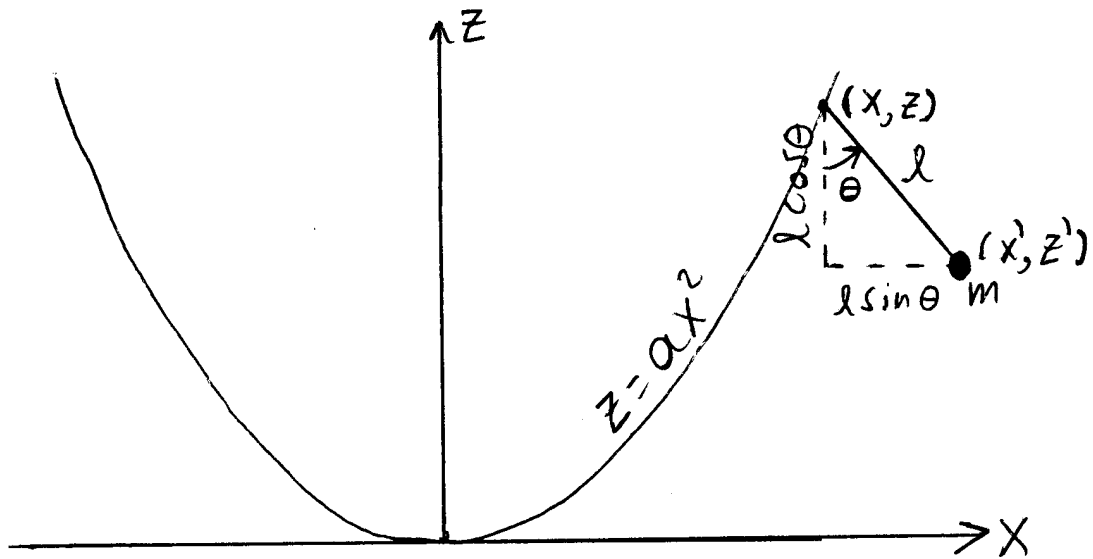


Figura 6.3: Péndulo simple cuyo punto de suspensión está restringido a moverse sobre una parábola  $z = ax^2$ . La posición del punto de suspensión del péndulo está dada por las coordenadas  $(x, z)$ , en tanto que la posición de la lenteja está dada por las coordenadas  $(x', z')$ .

### 6.8.3. Péndulo sujeto a una recorrido parabólico

El punto de suspensión de un péndulo simple de longitud  $l$  y masa  $m$  está restringido a moverse sobre una parábola  $z = ax^2$  en el plano vertical, como lo indica la figura 6.3. Encontraremos el Hamiltoniano de dicho péndulo y las ecuaciones de movimiento de Hamilton asociadas.

El punto de suspensión del péndulo se denota por las coordenadas  $(x, z)$ , y la posición de la lenteja por las coordenadas  $(x', z')$ . De la Figura 6.3 es claro que

$$x' = x + l \sin \theta \quad ; \quad z' = z - l \cos \theta = ax^2 - l \cos \theta \quad (6.58)$$

con lo cual

$$\dot{x}' = \dot{x} + l \dot{\theta} \cos \theta \quad ; \quad \dot{z}' = 2ax\dot{x} + l \dot{\theta} \sin \theta \quad (6.59)$$

las Ecs. (6.58) nos sugieren utilizar a  $x$  y  $\theta$ , como coordenadas generalizadas<sup>9</sup>. La energía cinética y potencial de la lenteja están dadas por

$$\begin{aligned} T &= \frac{1}{2}m(\dot{x}'^2 + \dot{z}'^2) = \frac{1}{2}m \left[ (\dot{x} + l \dot{\theta} \cos \theta)^2 + (2ax\dot{x} + l \dot{\theta} \sin \theta)^2 \right] \\ V &= mgz' = mg(ax^2 - l \cos \theta) \end{aligned}$$

donde hemos usado las Ecs. (6.59, 6.58). El Lagrangiano queda

<sup>8</sup>Nótese que la elección

$$\mathbf{T} = \begin{pmatrix} 2k & \frac{2k_2}{a+bq_1^2} \\ 0 & \frac{2k_2}{a+bq_1^2} \end{pmatrix}$$

en las Ecs. (6.54, 6.53), hubiese conducido al mismo Lagrangiano. Sin embargo, en tal caso la matriz  $\mathbf{T}$  no sería simétrica, lo cual fué una hipótesis fundamental en los desarrollos de la sección 6.4.

<sup>9</sup>De la Figura 6.3, se observa que se puede mover  $\theta$  sin mover  $x$  y sin violar las ligaduras (moviendo  $\theta$  con el punto de suspensión fijo). Así mismo, se puede mover la coordenada  $x$  manteniendo fijo  $\theta$  y respetando las ligaduras. Esto se logra moviendo el punto de suspensión a lo largo de la parábola manteniendo  $\theta$  fijo, es decir con una traslación paralela de la cuerda.

$$L = \frac{1}{2}m \left( \dot{x}^2 + 2l\dot{x}\dot{\theta} \cos \theta + 4a^2x^2\dot{x}^2 + 4alx\dot{x}\dot{\theta} \sin \theta + l^2\dot{\theta}^2 \right) - mg(ax^2 - l \cos \theta)$$

$$L = -mg(ax^2 - l \cos \theta) + \frac{1}{2}m \left[ \dot{x}^2 (1 + 4a^2x^2) + 2\dot{x}\dot{\theta}l (\cos \theta + 2ax \sin \theta) + l^2\dot{\theta}^2 \right]$$

este Lagrangiano se puede escribir en estructura matricial de la forma

$$L = -mg(ax^2 - l \cos \theta) + \frac{1}{2} \begin{pmatrix} \dot{x} & \dot{\theta} \end{pmatrix} \begin{pmatrix} m(1 + 4a^2x^2) & ml(\cos \theta + 2ax \sin \theta) \\ ml(\cos \theta + 2ax \sin \theta) & ml^2 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix}$$

que al comparar con (6.52) nos da

$$L_0 = -mg(ax^2 - l \cos \theta) ; \quad \mathbf{a} = \mathbf{0} ; \quad \mathbf{T} = \begin{pmatrix} m(1 + 4a^2x^2) & ml(\cos \theta + 2ax \sin \theta) \\ ml(\cos \theta + 2ax \sin \theta) & ml^2 \end{pmatrix} \quad (6.60)$$

ahora debemos calcular  $\mathbf{T}^{-1}$  para lo cual usamos

$$\mathbf{T}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

tenemos que

$$\begin{aligned} ad - bc &= m^2l^2(1 + 4a^2x^2) - m^2l^2(\cos \theta + 2ax \sin \theta)^2 \\ &= m^2l^2(\sin^2 \theta + 4a^2x^2 - 4ax \cos \theta \sin \theta - 4a^2x^2 \sin^2 \theta) \\ ad - bc &= m^2l^2(\sin^2 \theta - 4ax \sin \theta \cos \theta + 4a^2x^2 \cos^2 \theta) = m^2l^2(\sin \theta - 2ax \cos \theta)^2 \equiv m^2l^2Y \end{aligned}$$

con lo cual la inversa de  $\mathbf{T}$  queda

$$\mathbf{T}^{-1} = \frac{1}{m^2l^2Y} \begin{pmatrix} ml^2 & -ml(\cos \theta + 2ax \sin \theta) \\ -ml(\cos \theta + 2ax \sin \theta) & m(1 + 4a^2x^2) \end{pmatrix}$$

$$\mathbf{T}^{-1} = \frac{1}{mY} \begin{pmatrix} 1 & -(\cos \theta + 2ax \sin \theta)/l \\ -(\cos \theta + 2ax \sin \theta)/l & (1 + 4a^2x^2)/l^2 \end{pmatrix}$$

escribiremos  $\mathbf{T}^{-1}$  como

$$\mathbf{T}^{-1} = \frac{1}{mY} \begin{pmatrix} 1 & -J \\ -J & (1 + 4a^2x^2)/l^2 \end{pmatrix}$$

$$J \equiv (\cos \theta + 2ax \sin \theta)/l ; \quad Y \equiv (\sin \theta - 2ax \cos \theta)^2 \quad (6.61)$$

reemplazando (6.60, 6.61) en (6.55) obtenemos el Hamiltoniano

$$\begin{aligned} H(q, p, t) &= \frac{1}{2} \tilde{\mathbf{p}} \mathbf{T}^{-1} \mathbf{p} - L_0(q, t) = \frac{1}{2mY} \begin{pmatrix} p_x & p_\theta \end{pmatrix} \begin{pmatrix} 1 & -J \\ -J & (1 + 4a^2x^2)/l^2 \end{pmatrix} \begin{pmatrix} p_x \\ p_\theta \end{pmatrix} + \\ &+ mg(ax^2 - l \cos \theta) \\ &= \frac{1}{2mY} \begin{pmatrix} p_x & p_\theta \end{pmatrix} \begin{pmatrix} p_x - Jp_\theta & \\ -Jp_x + \frac{(1+4a^2x^2)p_\theta}{l^2} \end{pmatrix} + mg(ax^2 - l \cos \theta) \\ &= \frac{1}{2mY} \left[ p_x^2 - 2Jp_\theta p_x + \left( \frac{1 + 4a^2x^2}{l^2} \right) p_\theta^2 \right] + mg(ax^2 - l \cos \theta) \end{aligned}$$

reemplazando las definiciones de  $J$  y  $Y$  el Hamiltoniano queda

$$H(x, \theta, p_x, p_\theta) = \frac{1}{2m(\sin \theta - 2ax \cos \theta)^2} \left[ p_x^2 - 2 \left( \frac{\cos \theta + 2ax \sin \theta}{l} \right) p_\theta p_x + \left( \frac{1 + 4a^2 x^2}{l^2} \right) p_\theta^2 \right] + mg(ax^2 - l \cos \theta) \quad (6.62)$$

Es fácil ver que este Hamiltoniano es la energía del sistema y se conserva. Este Hamiltoniano no tiene coordenadas cíclicas para el sistema coordinado elegido. Escribiremos ahora las ecuaciones de Hamilton dadas por

$$\dot{x} = \frac{\partial H}{\partial p_x} ; \quad \dot{\theta} = \frac{\partial H}{\partial p_\theta} ; \quad \dot{p}_x = -\frac{\partial H}{\partial x} ; \quad \dot{p}_\theta = -\frac{\partial H}{\partial \theta}$$

las dos primeras nos dan

$$\dot{x} = \frac{1}{m(\sin \theta - 2ax \cos \theta)^2} \left[ p_x - \frac{\cos \theta + 2ax \sin \theta}{l} p_\theta \right] \quad (6.63)$$

$$\dot{\theta} = \frac{1}{ml(\sin \theta - 2ax \cos \theta)^2} \left[ -(\cos \theta + 2ax \sin \theta) p_x + \left( \frac{1 + 4a^2 x^2}{l} \right) p_\theta \right] \quad (6.64)$$

las otras son un poco más extensas de calcular comencemos con

$$\begin{aligned} -\dot{p}_x &= \frac{\partial H}{\partial x} = \frac{1}{2m(\sin \theta - 2ax \cos \theta)^2} \left[ \frac{-4a \sin \theta}{l} p_\theta p_x + \frac{8a^2 x}{l^2} p_\theta^2 \right] \\ &+ \frac{-2(-2a \cos \theta)}{2m(\sin \theta - 2ax \cos \theta)^3} \left[ p_x^2 - 2 \left( \frac{\cos \theta + 2ax \sin \theta}{l} \right) p_\theta p_x + \left( \frac{1 + 4a^2 x^2}{l^2} \right) p_\theta^2 \right] + 2mgax \end{aligned}$$

agrupando los términos de  $p_x^2$ ,  $p_\theta^2$  y  $p_x p_\theta$  nos queda

$$\begin{aligned} -\dot{p}_x &= \frac{4a^2 x}{ml^2(\sin \theta - 2ax \cos \theta)^2} p_\theta^2 + \frac{2a \cos \theta (1 + 4a^2 x^2)}{ml^2(\sin \theta - 2ax \cos \theta)^3} p_\theta^2 + \frac{2a \cos \theta}{m(\sin \theta - 2ax \cos \theta)^3} p_x^2 \\ &- \frac{2a \sin \theta}{ml(\sin \theta - 2ax \cos \theta)^2} p_\theta p_x - \frac{4a \cos \theta (\cos \theta + 2ax \sin \theta)}{ml(\sin \theta - 2ax \cos \theta)^3} p_\theta p_x + 2mgax \\ &= \left[ \frac{4a^2 x (\sin \theta - 2ax \cos \theta) + 2a \cos \theta (1 + 4a^2 x^2)}{ml^2(\sin \theta - 2ax \cos \theta)^3} \right] p_\theta^2 + \frac{2a \cos \theta}{m(\sin \theta - 2ax \cos \theta)^3} p_x^2 \\ &- \left[ \frac{2a \sin \theta (\sin \theta - 2ax \cos \theta) + 4a \cos \theta (\cos \theta + 2ax \sin \theta)}{ml(\sin \theta - 2ax \cos \theta)^3} \right] p_\theta p_x + 2mgax \\ -\dot{p}_x &= \frac{2a(\cos \theta + 2ax \sin \theta)}{ml^2(\sin \theta - 2ax \cos \theta)^3} p_\theta^2 + \frac{2a \cos \theta}{m(\sin \theta - 2ax \cos \theta)^3} p_x^2 \\ &+ \frac{2a(\sin^2 \theta - 2 - 2ax \cos \theta \sin \theta)}{ml(\sin \theta - 2ax \cos \theta)^3} p_\theta p_x + 2mgax \end{aligned}$$

resultando finalmente

$$\dot{p}_x = \frac{2a [l^2 \cos \theta p_x^2 + (\cos \theta + 2ax \sin \theta) p_\theta^2 + l(\sin^2 \theta - 2 - 2ax \sin \theta \cos \theta) p_x p_\theta]}{ml^2(2ax \cos \theta - \sin \theta)^3} - 2mgax \quad (6.65)$$

de una manera similar obtenemos la cuarta ecuación de Hamilton

$$\begin{aligned} -\dot{p}_\theta &= \frac{\partial H}{\partial \theta} = \frac{1}{2m(\sin \theta - 2ax \cos \theta)^2} \left( \frac{2 \sin \theta - 4ax \cos \theta}{l} \right) p_\theta p_x \\ &+ \left[ p_x^2 - 2 \left( \frac{\cos \theta + 2ax \sin \theta}{l} \right) p_\theta p_x + \left( \frac{1 + 4a^2 x^2}{l^2} \right) p_\theta^2 \right] \left[ \frac{-2(\cos \theta + 2ax \sin \theta)}{2m(\sin \theta - 2ax \cos \theta)^3} \right] + mgl \sin \theta \end{aligned}$$

$$\begin{aligned}
-\dot{p}_\theta &= \frac{-(\cos \theta + 2ax \sin \theta)}{m(\sin \theta - 2ax \cos \theta)^3} p_x^2 + \left\{ \left[ \frac{1 + 4a^2 x^2}{l^2} \right] \left[ \frac{-(\cos \theta + 2ax \sin \theta)}{m(\sin \theta - 2ax \cos \theta)^3} \right] \right\} p_\theta^2 \\
&+ \left\{ \frac{\sin \theta - 2ax \cos \theta}{ml(\sin \theta - 2ax \cos \theta)^2} + 2 \left[ \frac{\cos \theta + 2ax \sin \theta}{l} \right] \left[ \frac{(\cos \theta + 2ax \sin \theta)}{m(\sin \theta - 2ax \cos \theta)^3} \right] \right\} p_\theta p_x + mgl \sin \theta \\
\dot{p}_\theta &= \frac{1}{ml^2(\sin \theta - 2ax \cos \theta)^3} \left\{ l^2(\cos \theta + 2ax \sin \theta) p_x^2 + (\cos \theta + 2ax \sin \theta)(1 + 4a^2 x^2) p_\theta^2 \right. \\
&\quad \left. - l \left[ (\sin \theta - 2ax \cos \theta)^2 + 2(\cos \theta + 2ax \sin \theta)^2 \right] p_\theta p_x \right\} - mgl \sin \theta
\end{aligned}$$

y se obtiene finalmente

$$\begin{aligned}
\dot{p}_\theta &= \frac{1}{ml^2(\sin \theta - 2ax \cos \theta)^3} \left\{ l^2(\cos \theta + 2ax \sin \theta) p_x^2 + (\cos \theta + 2ax \sin \theta)(1 + 4a^2 x^2) p_\theta^2 \right. \\
&\quad \left. - l \left[ (2ax \sin \theta + \cos \theta)^2 + 1 + 4a^2 x^2 \right] p_\theta p_x \right\} - mgl \sin \theta
\end{aligned} \tag{6.66}$$

reuniendo las ecuaciones (6.63, 6.64, 6.65, 6.66) obtenemos las ecuaciones de Hamilton para el péndulo cuyo punto de suspensión se mueve a lo largo de la parábola  $z = ax^2$

$$\dot{x} = \frac{1}{m(\sin \theta - 2ax \cos \theta)^2} \left[ p_x - \frac{\cos \theta + 2ax \sin \theta}{l} p_\theta \right] \tag{6.67}$$

$$\dot{\theta} = \frac{1}{ml(\sin \theta - 2ax \cos \theta)^2} \left[ -(\cos \theta + 2ax \sin \theta) p_x + \left( \frac{1 + 4a^2 x^2}{l} \right) p_\theta \right] \tag{6.68}$$

$$\dot{p}_x = \frac{2a \left[ l^2 \cos \theta p_x^2 + (\cos \theta + 2ax \sin \theta) p_\theta^2 + l(\sin^2 \theta - 2 - 2ax \sin \theta \cos \theta) p_x p_\theta \right]}{ml^2(2ax \cos \theta - \sin \theta)^3} - 2mga x \tag{6.69}$$

$$\begin{aligned}
\dot{p}_\theta &= \frac{1}{ml^2(\sin \theta - 2ax \cos \theta)^3} \left\{ l^2(\cos \theta + 2ax \sin \theta) p_x^2 + (\cos \theta + 2ax \sin \theta)(1 + 4a^2 x^2) p_\theta^2 \right. \\
&\quad \left. - l \left[ (2ax \sin \theta + \cos \theta)^2 + 1 + 4a^2 x^2 \right] p_\theta p_x \right\} - mgl \sin \theta
\end{aligned} \tag{6.70}$$

## 6.9. Procedimiento de Routh

Hemos mencionado que la formulación Hamiltoniana no es en general adecuada para la solución directa de problemas mecánicos. Sin embargo, hay una importante excepción a este hecho: la formulación Hamiltoniana es muy adecuada para el tratamiento de problemas que involucran variables cíclicas.

Pensemos en un sistema cuyo Lagrangiano se puede escribir como

$$L(q_1, \dots, q_s; \dot{q}_1, \dots, \dot{q}_n; t)$$

es decir las últimas  $n-s$  coordenadas son cíclicas (naturalmente  $s < n$ ). Dado que en el formalismo Lagrangiano todas las velocidades generalizadas aparecen y pueden ser en general funciones del tiempo, el problema tiene aún  $n$  grados de libertad a pesar de la existencia de las coordenadas cíclicas. En contraste, en el formalismo Hamiltoniano las coordenadas cíclicas se vuelven realmente ignorables, ya que los correspondientes momentos conjugados se vuelven constantes de movimiento que denotaremos por  $\alpha_k$ . El Hamiltoniano se puede escribir en la forma

$$H = H(q_1, \dots, q_s; p_1, \dots, p_s; \alpha_{s+1}, \dots, \alpha_n; t)$$

el Hamiltoniano describe un problema de solo  $s$  grados de libertad, y las coordenadas cíclicas han sido completamente ignoradas, excepto que se manifiestan como constantes de integración  $\alpha_k$  que se determinan con las condiciones iniciales. El comportamiento de la coordenada cíclica como tal se encuentra integrando la siguiente ecuación de movimiento

$$\dot{q}_n = \frac{\partial H}{\partial \alpha_n}$$

esto sugiere utilizar un método que permita aprovechar las ventajas del formalismo Hamiltoniano con las coordenadas cíclicas junto con el formalismo Lagrangiano para las coordenadas no cíclicas. Esto se logra usando una transformación matemática que convierta la base de las  $q, \dot{q}$  en la base  $q, p$  solo para las coordenadas cíclicas. De esta forma se obtienen  $s$  ecuaciones de Lagrange para las coordenadas no cíclicas, y  $2(n - s)$  ecuaciones de Hamilton para las coordenadas cíclicas. Asumiendo que las coordenadas  $q_{s+1}, \dots, q_n$  son cíclicas, se propone una nueva función  $R$  (conocida como el Routhiano) que defina una “transformación de Legendre parcial” en donde solo se transformen los términos asociados a las variables cíclicas, por tanto  $R$  está definida por

$$R(q_1, \dots, q_s; \dot{q}_1, \dots, \dot{q}_s; p_{s+1}, \dots, p_n; t) \equiv \sum_{i=s+1}^n p_i \dot{q}_i - L \quad (6.71)$$

esto equivale a escribir

$$R(q_1, \dots, q_s; \dot{q}_1, \dots, \dot{q}_s; p_{s+1}, \dots, p_n; t) \equiv H_{cicl}(p_{s+1}, \dots, p_n; t) - L_{no\ cicl}(q_1, \dots, q_s; \dot{q}_1, \dots, \dot{q}_s; t)$$

para las  $s$  coordenadas no ignorables i.e.  $k = 1, \dots, s$  podemos calcular

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial R}{\partial \dot{q}_k} \right) &= \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{q}_k} \left( \sum_{i=s+1}^n p_i \dot{q}_i \right) - \frac{\partial L}{\partial \dot{q}_k} \right] = -\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}_k} \right] \\ \frac{\partial R}{\partial q_k} &= \frac{\partial}{\partial q_k} \left( \sum_{i=s+1}^n p_i \dot{q}_i \right) - \frac{\partial L}{\partial q_k} = -\frac{\partial L}{\partial q_k} \end{aligned}$$

y teniendo en cuenta las ecuaciones de Lagrange resulta

$$\frac{d}{dt} \left( \frac{\partial R}{\partial \dot{q}_k} \right) - \frac{\partial R}{\partial q_k} = 0, \quad k = 1, \dots, s \quad (6.72)$$

en tanto que para las  $n - s$  coordenadas ignorables i.e.  $k = s + 1, \dots, n$  tenemos que

$$\begin{aligned} \frac{\partial R}{\partial q_k} &= \frac{\partial}{\partial q_k} \left( \sum_{i=s+1}^n p_i \dot{q}_i \right) - \frac{\partial L}{\partial q_k} = -\frac{\partial L}{\partial q_k} = -\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) \\ \frac{\partial R}{\partial q_k} &= -\dot{p}_k = 0 \end{aligned}$$

donde hemos tenido en cuenta las ecuaciones de Lagrange (2.23), la definición de momento conjugado (5.1), y el hecho de que los momentos conjugados a variables cíclicas son constantes. Por otro lado para  $k = s + 1, \dots, n$  también tenemos

$$\frac{\partial R}{\partial p_k} = \frac{\partial}{\partial p_k} \left( \sum_{i=s+1}^n p_i \dot{q}_i \right) - \frac{\partial L}{\partial p_k} = \dot{q}_k$$

con lo que queda finalmente

$$\frac{\partial R}{\partial q_k} = -\dot{p}_k = 0 \quad ; \quad \frac{\partial R}{\partial p_k} = \dot{q}_k \quad ; \quad k = s + 1, \dots, n \quad (6.73)$$

de modo que se logra el objetivo de tener ecuaciones tipo Lagrange para las variables no cíclicas, Ecs. (6.72) y tipo Hamilton para las variables cíclicas Ecs. (6.73). Este método se conoce como procedimiento de Routh. Es necesario escribir el Routhiano con los argumentos dados en la Ec. (6.71).

### 6.9.1. Partícula sometida a un potencial central atractivo por el método de Routh

Para ilustrar el método tomemos como ejemplo el problema de una partícula sujeta a un potencial central bajo la influencia de una fuerza atractiva proporcional a  $r^{-n-1}$ . Este movimiento siempre se realiza en un plano en virtud de la conservación del momento angular. El Lagrangiano en coordenadas polares se escribe

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2) + \frac{k}{r^n}$$

la coordenada ignorable es  $\theta$  y su momento conjugado se denotará por  $p_\theta$ . El Routhiano en coordenadas polares se escribe

$$R = p_\theta \dot{\theta} - L = p_\theta \dot{\theta} - \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2) - \frac{k}{r^n} \quad (6.74)$$

este es el valor del Routhiano que se obtiene por reemplazo directo en (6.71). No obstante, esta expresión no está escrita con los argumentos adecuados, ya que según (6.71) el Routhiano se debe escribir en términos de las coordenadas no cíclicas (en este caso  $r$ ), las velocidades generalizadas de las coordenadas no cíclicas ( $\dot{r}$ ), los momentos conjugados asociados a las variables cíclicas ( $p_\theta$ ) y el tiempo (que aquí no aparece), es decir  $R = R(r, \dot{r}, p_\theta)$ . Ahora bien en (6.74) el Routhiano está escrito en términos de  $R = R(r, \dot{r}, p_\theta, \dot{\theta})$ , por tanto es necesario encontrar a  $\dot{\theta} = \dot{\theta}(r, \dot{r}, p_\theta)$  para que el Routhiano quede con los argumentos adecuados. Esta relación la sacamos a partir del momento conjugado  $p_\theta$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2 \dot{\theta} \quad \Rightarrow \quad \dot{\theta} = \frac{p_\theta}{mr^2}$$

reemplazando esta expresión en el Routhiano queda

$$\begin{aligned} R(r, \dot{r}, p_\theta) &= p_\theta \left( \frac{p_\theta}{mr^2} \right) - \frac{m}{2} \left[ \dot{r}^2 + r^2 \left( \frac{p_\theta}{mr^2} \right)^2 \right] - \frac{k}{r^n} \\ R(r, \dot{r}, p_\theta) &= \frac{p_\theta^2}{mr^2} - \frac{1}{2} m \dot{r}^2 - \frac{1}{2} \left( \frac{p_\theta^2}{mr^2} \right) - \frac{k}{r^n} \\ R(r, \dot{r}, p_\theta) &= \frac{p_\theta^2}{2mr^2} - \frac{1}{2} m \dot{r}^2 - \frac{k}{r^n} \end{aligned}$$

debemos aplicar las ecuaciones (6.72) tipo Lagrange para la coordenada no cíclica  $r$ , para lo cual calculamos

$$\begin{aligned} \frac{\partial R}{\partial \dot{r}} &= -m\dot{r} \quad ; \quad \frac{d}{dt} \left( \frac{\partial R}{\partial \dot{r}} \right) = -m\ddot{r} \quad ; \quad \frac{\partial R}{\partial r} = -\frac{p_\theta^2}{mr^3} + \frac{nk}{r^{n+1}} \\ -m\ddot{r} + \frac{p_\theta^2}{mr^3} - \frac{nk}{r^{n+1}} &= 0 \\ \ddot{r} - \frac{p_\theta^2}{m^2 r^3} + \frac{nk}{mr^{n+1}} &= 0 \end{aligned} \quad (6.75)$$

ahora debemos aplicar las ecuaciones (6.73) tipo Hamilton, a la coordenada cíclica  $\theta$  y su momento conjugado  $p_\theta$

$$\begin{aligned} \frac{\partial R}{\partial \theta} &= 0 \quad ; \quad \frac{\partial R}{\partial p_\theta} = \frac{p_\theta}{mr^2} \\ \dot{p}_\theta &= 0 \quad ; \quad \frac{p_\theta}{mr^2} = \dot{\theta} \end{aligned} \quad (6.76)$$

cuya solución es

$$p_\theta = mr^2 \dot{\theta} \equiv l = cte \quad (6.77)$$

Las Ecs. (6.75, 6.77), junto con las condiciones iniciales, nos dan la solución a la dinámica de la partícula. Un conjunto de condiciones iniciales puede ser  $r(0)$ ,  $\dot{r}(0)$ ,  $\theta(0)$  y  $l$ .

El formalismo de Routh es frecuentemente útil para problemas prácticos en Física y en Ingeniería. Sin embargo, como instrumento formal para posteriores construcciones en mecánica clásica, el formalismo puramente Hamiltoniano es mucho más fructífero.



## 6.10. Derivación de las ecuaciones de Hamilton a partir del principio variacional de Hamilton

La derivación hasta aquí desarrollada de las ecuaciones de Hamilton está basada en un principio diferencial que hace uso directo de las Ecuaciones de Lagrange combinado con una transformación de Legendre. La pregunta natural es entonces de que manera pueden surgir estas ecuaciones a partir de un formalismo integral. Vimos que el formalismo Lagrangiano puede extraerse de una formulación integral via principio variacional de Hamilton. Estos principios variacionales serán útiles en contextos mas allá de la mecánica. Es natural entonces tratar de derivar las ecuaciones de Hamilton partiendo del principio variacional de Hamilton como fuente de una formulación integral de estas ecuaciones

$$\delta I \equiv \delta \int_{t_1}^{t_2} L dt = 0 \quad (6.78)$$

En la formulación Lagrangiana las integrales se evalúan con base en las trayectorias en el espacio de configuraciones (adaptado para la formulación Lagrangiana). La filosofía del formalismo Hamiltoniano requiere tratar a las coordenadas  $q, p$  como independientes entre sí. Por tanto, el espacio adecuado para trabajar trayectorias en este formalismo consiste en tener  $2n$  ejes coordenados uno para cada  $q_i$  y uno para cada  $p_i$ , un punto en este espacio (espacio de fase) traza una curva a través del parámetro tiempo. En consecuencia, las integrales deben ser evaluadas con base en las trayectorias en el espacio de fase. Dado que el conjunto  $q, p$  se considera independiente, el integrando será, en general función de  $q, \dot{q}, p, \dot{p}$ . Usando la Ec. (6.7) que define al Hamiltoniano, en la Ec. (6.78) resulta

$$\delta I = \delta \int_{t_1}^{t_2} [p_i \dot{q}_i - H(q, p, t)] dt = 0 \quad (6.79)$$

A la Ec. (6.79) se le suele denominar principio **modificado** de Hamilton en virtud de que la integral de línea se está realizando en un espacio diferente (el espacio de fase). Teniendo en cuenta que las técnicas desarrolladas en la sección 3.2 son válidas para cualquier conjunto de variables y sus derivadas con respecto a un parámetro (en nuestro caso el tiempo)<sup>10</sup>, tenemos que genéricamente nuestro principio variacional se escribe

$$\delta I = \delta \int_{t_1}^{t_2} f(q, \dot{q}, p, \dot{p}) dt = 0$$

de lo cual resultan  $2n$  ecuaciones de Euler-Lagrange

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial f}{\partial \dot{q}_j} \right) - \frac{\partial f}{\partial q_j} &= 0 ; j = 1, \dots, n \\ \frac{d}{dt} \left( \frac{\partial f}{\partial \dot{p}_j} \right) - \frac{\partial f}{\partial p_j} &= 0 ; j = 1, \dots, n \end{aligned} \quad (6.80)$$

por tanto, las ecuaciones que se derivan del principio de Hamilton modificado, se obtienen aplicando las relaciones (6.80) al integrando de la Ec. (6.79)

$$f(q, \dot{q}, p) = p_i \dot{q}_i - H(q, p, t) \quad (6.81)$$

resultando

$$\dot{p}_j + \frac{\partial H}{\partial q_j} = 0 ; -\dot{q}_j + \frac{\partial H}{\partial p_j} = 0$$

De modo que hemos llegado a las ecuaciones de Hamilton a través del principio variacional modificado de Hamilton. Se puede argumentar que los momentos no pueden considerarse independientes en virtud de su definición (5.1) que nos muestra que  $p$  es en general función de  $q$  y  $\dot{q}$ , con lo cual no se puede variar  $q$

<sup>10</sup>Desde el punto de vista matemático el espacio de fase es simplemente un espacio de configuración de dimensión  $2n$ , ya que lo que se hace es cartesianizar las variables independientes  $q$  y  $p$  colocando un eje para cada variable independiente.

ó  $\dot{q}$ , sin variar  $p$ . Recordemos sin embargo, que la filosofía misma del formalismo Hamiltoniano requiere de considerar a  $q$  y  $p$  independientes, es decir que una vez establecido el formalismo Hamiltoniano la definición (5.1) no constituye parte del formalismo. Los momentos conjugados son elevados a la categoría de variables independientes, con un papel análogo a las coordenadas y conectados con ellas y con el tiempo solo a través de las ecuaciones de movimiento y no a través de definiciones. Visto desde el punto de vista de un conteo de grados de libertad, al introducir los momentos “olvidando” su definición, estamos aumentando el número de grados de libertad, pero también estamos aumentando el número de ecuaciones de modo que el número de grados de libertad real se conserva.

Un aspecto similar ocurre en la formulación Lagrangiana, estrictamente  $q$  y  $\dot{q}$  no son independientes en virtud de la definición  $\dot{q} = dq/dt$ , pero cuando escribimos las ecuaciones de Lagrange “olvidamos” la definición de  $\dot{q}$  y la elevamos al estatus de independiente, la dependencia original se recupera a través de las  $n$  ecuaciones de movimiento.

Podemos preguntarnos si el principio modificado de Hamilton contiene nueva Física respecto al original, sin embargo este principio fué establecido con el fin de obtener las ecuaciones de Hamilton, al igual que el principio variacional original fué construido con el fin de reproducir las ecuaciones de Lagrange. Una vez que el Hamiltoniano es construido, la transformación de Legendre muestra que las formulaciones de Lagrange y Hamilton así como los respectivos principios variacionales de donde vienen, contienen todos la misma Física.

Vale la pena mirar si las restricciones de punto fijo en el principio de Hamilton sufren algún cambio en el principio de Hamilton modificado. Por ejemplo en el espacio de configuraciones, la condición de extremo fijo significa  $\delta q_i = 0$ , en tanto que en el espacio de fase significa  $\delta q_i = \delta p_i = 0$ . Para esto nos debemos preguntar si es necesaria la condición de extremo fijo en el espacio de fase para llegar a las ecuaciones de Euler Lagrange dadas por (6.80). Si volvemos a la demostración de las ecuaciones de Euler Lagrange en la sección 3.2, vemos que la condición de extremo fijo se requiere para eliminar el primer término de la derecha en la Ec. (3.22). No obstante, teniendo en cuenta que nuestro integrando definido por (6.81) no es función explícita de  $\dot{p}_j$  se observa que la eliminación del término ya descrito, es automática para las variables  $p_j$  y no requiere de condición de punto fijo en estas variables. Por tanto, el principio de Hamilton modificado conduce a las ecuaciones de Hamilton bajo las mismas condiciones variacionales que el principio de Hamilton original, es decir bajo la condición de extremo fijo solo en las  $q_i$ .

La anterior discusión muestra que la condición  $\delta p_j = 0$  no es una condición necesaria pero sí sería una condición suficiente para llegar a las ecuaciones de Hamilton. Existen en realidad grandes ventajas al imponer extremos fijos en el espacio de fase, ya que en este caso obtenemos una simetría Gauge para el Hamiltoniano similar a la que posee el Lagrangiano, es decir se puede adicionar al integrando una función de la forma  $dF(q, p, t)/dt$ , siendo  $F$  dos veces diferenciable, sin afectar la validez del principio variacional. Como ejemplo sencillo si adicionamos al integrando en la Ec. (6.79) un término de la forma  $-\frac{d}{dt}(q_i p_i)$  el principio modificado de Hamilton quedaría

$$\delta \int_{t_1}^{t_2} [-\dot{p}_i q_i - H(q, p, t)] dt = 0 \quad (6.82)$$

nótese que ahora el integrando sí depende de  $\dot{p}_j$ . Por tanto, si queremos que esta nueva expresión aún nos conduzca a las ecuaciones de Hamilton, es necesario imponer  $\delta p_j = 0$  (con esta nueva expresión no necesitamos que  $\delta q_j = 0$ , pero al hacer la transformación inversa y volver al integrando original, se ve que dicha condición sí es necesaria). De modo que si requerimos una simetría gauge que permita adicionar un término de la forma  $\frac{dF(q, p, t)}{dt}$  se requiere de condición de extremo fijo en el espacio de fase<sup>11</sup>. El integrando en (6.82) no es el Lagrangiano ni se puede relacionar fácilmente con el Lagrangiano a través de una transformación puntual. La condición de extremo fijo en el espacio de fase provee un método para llegar a la formulación Hamiltoniana sin pasar por una formulación Lagrangiana. Esto elimina la necesidad de ligar las  $q$ ,  $\dot{q}$  de un Lagrangiano, con las  $q$ ,  $p$  de un Hamiltoniano. Esto será importante cuando estudiemos transformaciones de variables en el espacio de fase que preservan la forma de las ecuaciones de Hamilton.

La independencia entre  $q$  y  $p$  en la formulación Hamiltoniana es una de las grandes diferencias con las

<sup>11</sup>Nótese que en ausencia de la condición  $\delta p_j = 0$ , aún tenemos un gauge aunque más débil, ya que podríamos adicionar un factor de la forma  $\frac{dF(q, t)}{dt}$ . Esto no es lo deseable ya que la idea es que  $F$  contenga todas las variables independientes.

formulación Lagrangiana en la cual los momentos aparecen como una definición Ec. (5.1). Podemos pensar como si tuviéramos  $2n$  coordenadas generalizadas y sus derivadas temporales. Solo ampliando el número de grados de libertad de  $n$  a  $2n$  se pueden obtener ecuaciones de primer orden.

### 6.11. El principio de mínima acción (opcional)

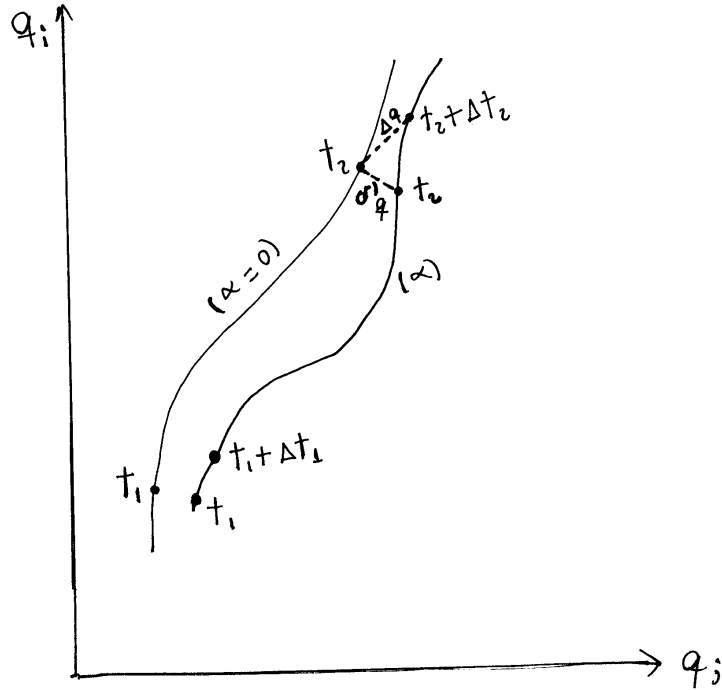


Figura 6.4: Ilustración de la  $\delta$ -variación y de la  $\Delta$ -variación en el espacio de configuraciones. La “curva correcta” corresponde a  $(\alpha = 0)$ , en tanto que la curva variada corresponde a un valor no nulo de  $\alpha$ .

Existe otro principio variacional muy útil asociado con el formalismo Hamiltoniano conocido como principio de mínima acción. Para esto planteamos otro tipo de variacional  $\Delta$ . Recordemos que teniendo la curva correcta definida entre los tiempos  $t_1$  y  $t_2$ ; el variacional  $\delta$  define la familia de curvas que poseen extremo fijo tanto en el tiempo como en el espacio de configuraciones, es decir todas las curvas están definidas en este intervalo de tiempo y de modo que  $\delta q_i(t_1) = \delta q_i(t_2) = 0$ . La nueva variación  $\Delta$  relaja ambas condiciones, las curvas vecinas pueden estar definidas en un intervalo temporal diferente  $[t_1 + \Delta t_1, t_2 + \Delta t_2]$ , y sus coordenadas generalizadas no tienen que coincidir con las del camino correcto en los extremos. Es posible por ejemplo, que una curva vecina no se interseque con la curva correcta en ningún punto. La Fig. 6.4 ilustra la  $\Delta$ -variación en el espacio de configuraciones.

Usaremos no obstante la misma parametrización que se usó para la variación  $\delta$  (ver Ec. 3.21)

$$q_i(t, \alpha) = q_i(t, 0) + \alpha \eta_i(t) \tag{6.83}$$

donde  $\alpha$  es el parámetro infinitesimal que conduce al camino correcto cuando se anula. Naturalmente, ya no es necesaria la condición de que las  $\eta_i(t)$  se anulen en los extremos  $t_1$  y  $t_2$  (ni en los extremos  $t_1 + \Delta t_1$  y  $t_2 + \Delta t_2$  de un camino variado). Solo requerimos que sean funciones continuas y diferenciables en cualquier intervalo real o variado.

Evaluemos la variación  $\Delta$  de la integral de acción.

$$\Delta \int_{t_1}^{t_2} L dt \equiv \int_{t_1 + \Delta t_1}^{t_2 + \Delta t_2} L(\alpha) dt - \int_{t_1}^{t_2} L(0) dt \tag{6.84}$$

donde  $L(\alpha)$  representa el valor de  $L$  cuando tomamos el camino variado siendo  $L(0)$  el camino correcto. La variación se descompone en dos partes, una surge del cambio en los límites de integración y la otra del parámetro  $\alpha$  i.e. la variación de la curva como tal. Escribiremos la primera integral como

$$\int_{t_1+\Delta t_1}^{t_2+\Delta t_2} = \int_{t_1}^{t_2} - \int_{t_1}^{t_1+\Delta t_1} + \int_{t_2}^{t_2+\Delta t_2}$$

con lo cual (6.84) queda

$$\Delta \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} L(\alpha) dt - \int_{t_1}^{t_2} L(0) dt - \int_{t_1}^{t_1+\Delta t_1} L(\alpha) dt + \int_{t_2}^{t_2+\Delta t_2} L(\alpha) dt \quad (6.85)$$

es útil parametrizar  $L(\alpha)$  como

$$L(\alpha) \equiv L(0) + F(\alpha)$$

donde  $F(\alpha)$  nos da cuenta de la desviación en el Lagrangiano debida a la evaluación en la curva vecina, con respecto a su valor en la curva real. Claramente  $F(\alpha)$  es una cantidad infinitesimal. Definiremos ahora la  $\delta'$ -variación, la cual mantiene fijo el intervalo temporal, pero no tiene extremos fijos en las  $q_i$  (ver Fig. 6.4). Con estas consideraciones (6.85) queda de la forma

$$\begin{aligned} \Delta \int_{t_1}^{t_2} L dt &= \delta' \int_{t_1}^{t_2} L dt - \int_{t_1}^{t_1+\Delta t_1} [L(0) + F(\alpha)] dt + \int_{t_2}^{t_2+\Delta t_2} [L(0) + F(\alpha)] dt \\ \Delta \int_{t_1}^{t_2} L dt &= \delta' \int_{t_1}^{t_2} L dt - \int_{t_1}^{t_1+\Delta t_1} L(0) dt + \int_{t_2}^{t_2+\Delta t_2} L(0) dt \\ &\quad - \int_{t_1}^{t_1+\Delta t_1} F(\alpha) dt + \int_{t_2}^{t_2+\Delta t_2} F(\alpha) dt \end{aligned}$$

y dado que las cantidades  $\Delta t_1$ ,  $\Delta t_2$  y  $F(\alpha)$  son infinitesimales, las dos últimas integrales nos dan diferenciales de segundo orden y por tanto se desprecian. Similarmente,  $L(0)$  se puede considerar constante dentro de los intervalos de integración  $[t_i, t_i + \Delta t_i]$ , puesto que una variación  $\Delta L$  (infinitesimal) del Lagrangiano dentro de uno de estos intervalos, daría una contribución del orden de  $\Delta L \cdot \Delta t_i$ . Con estas consideraciones, la  $\Delta$ -variación a primer orden queda finalmente

$$\Delta \int_{t_1}^{t_2} L dt = L(t_2, \alpha = 0) \Delta t_2 - L(t_1, \alpha = 0) \Delta t_1 + \delta' \int_{t_1}^{t_2} L dt \quad (6.86)$$

Los dos primeros términos en el miembro derecho de la Ec. (6.86), se deben a la variación en los límites de integración, tomados sobre el “camino correcto”. El tercer término a la derecha de (6.86), se origina en la variación de la curva como tal a través del parámetro  $\alpha$ , tomando los mismos límites de integración que en la curva original.

La variación  $\delta'$  en la integral de la derecha puede realizarse a través de una parametrización del camino variado, como en la sección 3.2, pero la variación de  $q_i$  no se anula en los extremos. Retomando la ecuación (3.22), el primer término de la derecha en dicha ecuación ya no se anula, con lo cual la Ec. (3.23) se modifica en la forma

$$\delta' \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \left[ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \right] \delta' q_i dt + \left. \frac{\partial L}{\partial \dot{q}_i} \delta' q_i \right|_1^2$$

donde  $\delta' q_i$  se define en forma análoga a la Ec. (3.24). Usando las ecuaciones de Lagrange<sup>12</sup>, esta nueva variación se escribe

$$\delta' \int_{t_1}^{t_2} L dt = \left. \frac{\partial L}{\partial \dot{q}_i} \delta' q_i \right|_1^2$$

<sup>12</sup>Al tomar ecuaciones de Lagrange, estamos asumiendo la validez del principio de Hamilton.

y la variación  $\Delta$  queda

$$\Delta \int_{t_1}^{t_2} L dt = [L(t) \Delta t + p_i \delta' q_i] \Big|_1^2 \quad (6.87)$$

$\delta' q_i(1)$  es la variación en  $q_i$  con respecto a la coordenada en la curva correcta tomada en el tiempo  $t_1$ , otro tanto ocurre con  $\delta' q_i(2)$ . La idea es escribir la  $\Delta$ -variación de la acción en términos de las  $\Delta$ -variaciones de las coordenadas, calculemos entonces  $\Delta q_i(2)$

$$\begin{aligned} \Delta q_i(2) &= q_i(t_2 + \Delta t_2, \alpha) - q_i(t_2, 0) = q_i(t_2 + \Delta t_2, 0) - q_i(t_2, 0) + \alpha \eta_i(t_2 + \Delta t_2) \\ \Delta q_i(2) &= \left[ \frac{q_i(t_2 + \Delta t_2, 0) - q_i(t_2, 0)}{\Delta t_2} \right] \Delta t_2 + \alpha \eta_i(t_2) \end{aligned}$$

donde hemos usado la parametrización (6.83) y eliminado contribuciones de segundo orden. A primer orden en las cantidades infinitesimales  $\alpha$  y  $\Delta t_2$

$$\Delta q_i(2) = \dot{q}_i(2) \Delta t_2 + \delta' q_i(2)$$

nótese que  $\alpha \eta_i(t_2)$  nos da una  $\delta'$  variación solo si no se impone que  $\eta_i$  se anule en los extremos con  $t_1$  y  $t_2$  fijos.

De manera similar se calcula para el otro extremo quedando

$$\Delta q_i = \delta' q_i + \dot{q}_i \Delta t \quad (6.88)$$

y reemplazando (6.88) en (6.87) resulta

$$\begin{aligned} \Delta \int_{t_1}^{t_2} L dt &= (L \Delta t - p_i \dot{q}_i \Delta t + p_i \Delta q_i) \Big|_1^2 \Rightarrow \\ \Delta \int_{t_1}^{t_2} L dt &= (p_i \Delta q_i - H \Delta t) \Big|_1^2 \end{aligned} \quad (6.89)$$

para obtener el principio de mínima acción, realizaremos algunas suposiciones adicionales

1. Consideraremos sistemas en donde  $L$ , y por tanto  $H$  no dependen explícitamente del tiempo, de modo que  $H$  es una cantidad conservada.
2. Restringiremos las variaciones de tal manera que para los caminos vecinos  $H$  también sea conservada al igual que en el camino correcto.
3. Exigiremos además que los caminos variados cumplan el requerimiento de que  $\Delta q_i$  se anule en los extremos (pero no  $\Delta t$ ). Esto implica entonces que  $q_i(t_1, 0)$  para el camino real debe coincidir con  $q_i(t_1 + \Delta t_1, \alpha)$  para el camino variado, y similarmente para  $t_2$ .

Un ejemplo particular de trayectoria vecina que podría satisfacer los tres requisitos, es un camino en donde la “línea” que describe la trayectoria variada en el espacio de configuraciones coincida con la “línea” de la trayectoria real, la diferencia consiste en la rapidez con que se desplaza el punto que describe la trayectoria en ambos casos, es decir que las funciones  $q_i(t)$  están alteradas en el camino variado. En este caso es necesario cambiar los tiempos en los extremos de la curva variada a fin de que  $H$  se mantenga constante en todos los puntos de la trayectoria variada.

Teniendo en cuenta las condiciones adicionales impuestas, la  $\Delta$ -variación de la acción dada por la Ec. (6.89), queda

$$\Delta \int_{t_1}^{t_2} L dt = -H (\Delta t_2 - \Delta t_1) \quad (6.90)$$

por otro lado, bajo las mismas condiciones, la acción como tal queda

$$\int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} p_i \dot{q}_i dt - H (t_2 - t_1)$$

escribiendo la  $\Delta$  variación de esta acción, resulta

$$\Delta \int_{t_1}^{t_2} L dt = \Delta \int_{t_1}^{t_2} p_i \dot{q}_i dt - H (\Delta t_2 - \Delta t_1) \quad (6.91)$$

comparando (6.90) con (6.91) se obtiene el principio de mínima acción

$$\Delta \int_{t_1}^{t_2} p_i \dot{q}_i dt = 0 \quad (6.92)$$

el nombre de este principio proviene del hecho de que la integral en (6.92), se conocía antiguamente como la acción. La literatura moderna utiliza el término acción para la integral involucrada en el principio variacional de Hamilton. Para la integral en (6.92) se utiliza con frecuencia el término **acción abreviada**.

### 6.11.1. Algunas aplicaciones del principio de mínima acción

El principio de mínima acción se puede escribir de diversas maneras. Por ejemplo, si la ecuación que define a las coordenadas generalizadas no depende explícitamente del tiempo, la energía cinética solo contiene factores homogéneos de segundo grado en  $\dot{q}$ .

$$T = \frac{1}{2} M_{jk}(q) \dot{q}_j \dot{q}_k$$

si adicionalmente, el potencial no depende de las velocidades generalizadas (es decir el Hamiltoniano es la energía del sistema), los momentos canónicos se derivan solo de  $T$

$$\begin{aligned} p_i &= \frac{\partial T}{\partial \dot{q}_i} = \frac{1}{2} (M_{ik} \dot{q}_k + M_{ji} \dot{q}_j) \Rightarrow \\ p_i \dot{q}_i &= \frac{1}{2} (M_{ik} \dot{q}_k \dot{q}_i + M_{ji} \dot{q}_j \dot{q}_i) = 2T \end{aligned}$$

empleando esta última igualdad en (6.92) resulta

$$\Delta \int_{t_1}^{t_2} T dt = 0 \quad (6.93)$$

si además  $T$  se conserva al igual que  $H$  (por ejemplo es el caso de un cuerpo rígido aislado puesto que en tal caso la energía potencial es solo interna y no cambia con el tiempo). El principio de mínima acción adquiere entonces la forma

$$\Delta (t_2 - t_1) = 0$$

esta ecuación nos dice que de todos los caminos posibles entre dos puntos en el espacio de configuraciones, consistente con la conservación de  $H$  y  $T$ , el sistema se moverá a lo largo de un camino particular para el cual el tiempo de transito es el mínimo posible (mas estrictamente estacionario). En esta forma el principio de mínima acción nos recuerda al principio de Fermat en óptica geométrica, en donde un rayo de luz atraviesa el camino que minimiza el tiempo de viaje. Nótese que para llegar a esta conclusión por este camino, fué necesario que el intervalo temporal de integración no fuera fijo, como efectivamente ocurre en la  $\Delta$ -variación.

Pensemos ahora el caso especial de una partícula asumiendo que  $T$  no es necesariamente constante, la energía cinética se escribe

$$T = \frac{m}{2} \left( \frac{ds}{dt} \right)^2$$

donde  $s$  denota la longitud de arco. De esta ecuación se puede despejar  $dt$

$$dt = \frac{ds}{\sqrt{2T/m}} \quad (6.94)$$

de modo que podemos reemplazar el  $dt$  de la acción abreviada (6.93)

$$\Delta \int_{t_1}^{t_2} T dt = 0 \Rightarrow \Delta \int_{p_1}^{p_2} T \frac{ds}{\sqrt{2T}} = 0 \Rightarrow \Delta \int_{p_1}^{p_2} \sqrt{T} ds = 0$$

finalmente se puede escribir

$$\Delta \int_{p_1}^{p_2} \sqrt{H - V(q)} ds = 0 \quad (6.95)$$

la Ec. (6.95) se conoce usualmente como la forma de Jacobi para el principio de mínima acción. Este resultado es extendible a caminos en espacios de configuraciones con curvatura que describen a un sistema de partículas (ver Ref. [1]). La Ec. (6.94) nos muestra que el punto atraviesa el camino en el espacio de configuraciones a una rapidez  $\sqrt{2T/m}$ . Si  $T$  es constante (e.g. en un sistema aislado), la forma de Jacobi del principio de mínima acción nos dice que el punto del sistema viaja por el camino mas corto en el espacio de configuraciones (es decir una geodésica). Tal resultado tiene muchas aplicaciones en óptica geométrica y en óptica electrónica. Es importante mencionar que en la forma de Jacobi del principio de mínima acción, nos referimos al camino (y no a la curva) que describe el punto figurativo del sistema. Es decir, no es relevante la dependencia de este camino con el parámetro tiempo, ya que en la integral (6.95)  $H$  es constante,  $V$  solo depende de las coordenadas, y solo interviene el elemento de camino  $ds$ .

Existen una gran variedad de principios variacionales. Por ejemplo, un principio variacional derivable del principio de mínima acción abreviada es el principio de mínima curvatura de Hertz, según el cual, una partícula libre se mueve a lo largo del camino de mínima curvatura, lo cual conduce a su vez al movimiento a lo largo de geodésicas.

Los principios variacionales no contienen física nueva y son en general poco ventajosos para resolver problemas prácticos, su utilidad yace en la facilidad para desarrollar nuevas formulaciones de la estructura teórica de la mecánica y otros campos de la Física mas allá de la mecánica.

## 6.12. Ejercicios

1. Demuestre explícitamente que dado un Hamiltoniano  $H$ , el nuevo Hamiltoniano dado por

$$H' = H + \frac{dF(q, p, t)}{dt}$$

conduce a las mismas ecuaciones de Hamilton que  $H$ . ¿Qué condiciones se necesitan sobre  $F$  para garantizar la equivalencia de ambos Hamiltonianos?.

2. Use las ecuaciones de Hamilton para encontrar las ecuaciones de movimiento de un péndulo esférico, usando coordenadas generalizadas apropiadas.
3. En el problema de la sección 6.8.1, de una partícula restringida a moverse en una superficie cilíndrica, encuentre la energía del sistema en términos de las condiciones iniciales.
4. Encuentre las ecuaciones de Hamilton asociadas al Hamiltoniano de la Ec. (6.54), Pág. 106.
5. Un Hamiltoniano de un solo grado de libertad tiene la forma

$$H(q, p) = \frac{p^2}{2\alpha} - bqpe^{-\alpha t} + \frac{ba}{2}q^2e^{-\alpha t}(\alpha + be^{-\alpha t}) + \frac{kq^2}{2} \quad (6.96)$$

donde  $a, b, \alpha$ , y  $k$  son constantes (a) Encuentre el Lagrangiano asociado a este Hamiltoniano, (b) Encuentre un Lagrangiano equivalente que no dependa explícitamente del tiempo. (c) Encuentre el Hamiltoniano asociado al nuevo Lagrangiano, así como su relación con el Hamiltoniano (6.96).

6. Encuentre el Hamiltoniano y las ecuaciones de Hamilton para el péndulo doble descrito por la Fig. 2.1, Pág. 16.

## Transformaciones canónicas

Hay un tipo de problema que sería de muy fácil solución con el formalismo Hamiltoniano, el caso en el cual  $H$  es una constante de movimiento y todas las coordenadas generalizadas son cíclicas, de modo que todos los momentos conjugados son constantes

$$p_i = \alpha_i \quad ; \quad i = 1, \dots, n$$

y dado que el Hamiltoniano no puede ser función ni del tiempo ni de las coordenadas cíclicas, este se escribe como

$$H = H(\alpha_1, \dots, \alpha_n)$$

y las ecuaciones de Hamilton para  $q_i$  son simplemente

$$\dot{q}_i = \frac{\partial H}{\partial \alpha_i} = \omega_i \quad ; \quad i = 1, \dots, n \quad (7.1)$$

dado que las  $\omega_i$  son funciones de las  $\alpha_i$  únicamente, son también constantes en el tiempo. Las Ecs. (7.1) tienen solución inmediata

$$q_i = \omega_i t + \beta_i \quad ; \quad i = 1, \dots, n$$

donde las  $\beta_i$  así como las  $\alpha_i$  se determinan con las  $2n$  condiciones iniciales. Podría pensarse que esta clase de problemas son de interés solo académico. No obstante, cabe recordar que existe en general mas de un sistema coordinado generalizado para un sistema físico dado. Por ejemplo, para fuerzas centrales el uso de coordenadas cartesianas no conduce a variables cíclicas en tanto que las coordenadas polares producen automáticamente una variable cíclica. El número de variables cíclicas depende en general del sistema de coordenadas generalizadas utilizado. Es posible entonces que encontremos un conjunto de coordenadas en las que todas las coordenadas sean cíclicas, una vez encontrado el resto del problema es muy sencillo. Dado que en general el sistema coordinado mas obvio no va a ser normalmente cíclico, debemos desarrollar un método para transformar este conjunto coordinado en otro que sea más adecuado. Las transformaciones de este tipo que hemos usado hasta el momento, son de un conjunto  $\{q_i\}$  de coordenadas generalizadas a otro conjunto  $\{Q_i\}$

$$Q_i = Q_i(q_i, t) \quad ; \quad i = 1, \dots, n \quad (7.2)$$

por ejemplo, la transformación de coordenadas cartesianas a polares tiene esta forma<sup>1</sup>. Estas transformaciones se conocen como transformaciones puntuales.

En el formalismo Hamiltoniano, los momentos conjugados están al mismo nivel que las coordenadas de tal modo que las transformaciones puntuales deben realizarse en el espacio de fase y no en el de configuraciones, por tanto debemos estudiar transformaciones del tipo

$$Q_i = Q_i(q, p, t) \quad ; \quad P_i = P_i(q, p, t) \quad (7.3)$$

<sup>1</sup>No toda transformación de coordenadas cartesianas a otro conjunto de coordenadas es del tipo (7.2), ya que éstas últimas se refieren a transformaciones entre dos conjuntos de coordenadas independientes. Por ejemplo, las transformaciones descritas por (2.2), relacionan en general a un conjunto de coordenadas dependientes (las cartesianas) con otro conjunto de coordenadas independientes (las  $q_i$ ).



En el formalismo Hamiltoniano, solo interesan aquellas transformaciones para las cuales exista una función  $K(Q, P, t)$  tal que las ecuaciones de movimiento en el nuevo sistema de coordenadas  $Q, P$ , tengan la estructura de las ecuaciones de Hamilton, es decir

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} ; \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} \quad (7.4)$$

Esta clase de transformaciones se denominan **transformaciones canónicas**. La función  $K(Q, P, t)$  juega el papel del Hamiltoniano en el nuevo sistema coordinado  $Q, P$ . Esta función se denomina Hamiltoniano transformado o Kamiltoniano. Dado que estas transformaciones deben ser invertibles, resulta arbitrario decir cual es el “nuevo Hamiltoniano o Kamiltoniano” y el “antiguo Hamiltoniano”.

Es importante mencionar que la transformación de coordenadas debe ser independiente del problema, lo cual significa que dadas las ecuaciones de transformación (7.3) entonces las expresiones (7.4) deben ser las ecuaciones de movimiento sin importar la forma del Hamiltoniano original. Aunque la transformación pudo ser inspirada en un problema específico, las mismas transformaciones deben conducir a las ecuaciones de Kamilton, cuando se aplican a otro problema con el mismo número de grados de libertad, aunque naturalmente el valor específico numérico y funcional de  $H$  y  $K$  sí sería específico para cada problema.

## 7.1. Caracterización de las transformaciones canónicas a través del principio de Hamilton modificado

Si  $Q_i, P_i$  son coordenadas canónicas, ellas deben satisfacer el principio variacional modificado de Hamilton en la forma

$$\delta \int_{t_1}^{t_2} [P_i \dot{Q}_i - K(Q, P, t)] dt = 0 \quad (7.5)$$

al mismo tiempo, las antiguas variables canónicas deben satisfacer el mismo principio

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t)) dt = 0 \quad (7.6)$$

teniendo en cuenta que el principio modificado de Hamilton tiene variación cero en los extremos<sup>2</sup>, la validez simultánea de (7.5) y (7.6) conduce a que los dos integrandos estén relacionados de la siguiente forma

$$\lambda [p_i \dot{q}_i - H(q, p, t)] = P_i \dot{Q}_i - K(Q, P, t) + \frac{dF}{dt} \quad (7.7)$$

donde  $F$  es una función de las coordenadas del espacio de fase y del tiempo con derivadas continuas hasta segundo orden.  $\lambda$  es una constante independiente de las coordenadas y el tiempo, que produce una transformación de escala.

**Example 4** *Veamos un caso de transformación de escala. Pasemos de  $\{q, p\}$  a otro conjunto  $\{Q', P'\}$  con una transformación de la forma*

$$Q'_i = \mu q_i ; \quad P'_i = \nu p_i \quad (7.8)$$

*las ecuaciones de Hamilton (7.4) serán satisfechas si elegimos el kamiltoniano*

$$K'(Q', P') \equiv \mu\nu H(q, p) \quad (7.9)$$

*Los integrandos correspondientes al principio de Hamilton modificado están relacionados por*

$$\mu\nu (p_i \dot{q}_i - H) = P'_i \dot{Q}'_i - K' \quad (7.10)$$

<sup>2</sup>Recordemos que para llegar a las ecuaciones de Hamilton solo requerimos que  $\delta q_i = 0$ , pero también podemos imponer  $\delta p_i = 0$ , con el fin de enriquecer el gauge del Hamiltoniano.

que es de la forma (7.7) con  $\lambda = \mu\nu$ , y  $dF/dt = 0$ . Sin embargo, si tenemos una transformación  $\{q, p\} \rightarrow \{Q', P'\}$  con factor de escala  $\alpha \neq 1$

$$\alpha [p_i \dot{q}_i - H(q, p, t)] = P'_i \dot{Q}'_i - K'(Q', P', t) + \frac{dF'}{dt} \quad (7.11)$$

siempre podemos encontrar un conjunto intermedio de coordenadas  $\{Q, P\}$  relacionadas con  $\{Q', P'\}$  por una transformación de escala de la forma (7.8, 7.10)

$$\begin{aligned} Q'_i &= \mu Q_i \quad ; \quad P'_i = \nu P_i \\ \mu\nu (P_i \dot{Q}_i - K) &= P'_i \dot{Q}'_i - K' \end{aligned} \quad (7.12)$$

sustituyendo (7.12) en (7.11) y ajustando  $\mu\nu \equiv \alpha$ , obtenemos

$$\alpha [p_i \dot{q}_i - H(q, p, t)] = \alpha (P_i \dot{Q}_i - K) + \frac{dF'}{dt}$$

y definiendo  $F \equiv F'/\alpha$  vemos que la transformación entre los dos conjuntos  $\{q, p\}$  y  $\{Q, P\}$  satisface entonces la Ec. (7.7) pero con  $\lambda = 1$

$$p_i \dot{q}_i - H(q, p, t) = P_i \dot{Q}_i - K(Q, P, t) + \frac{dF}{dt} \quad (7.13)$$

El ejemplo anterior nos muestra que la transformación de escala es básicamente trivial y no aporta significativamente al formalismo de las transformaciones canónicas, de modo que trabajaremos las transformaciones dadas por (7.13). Cuando  $\lambda \neq 1$  se habla de transformaciones canónicas extendidas. Cuando  $\lambda = 1$  hablamos simplemente de una transformación canónica. Lo que hemos visto es que una transformación canónica extendida siempre se puede separar en una transformación canónica seguida de una transformación de escala. Cuando la transformación canónica no depende explícitamente del tiempo se habla de transformación canónica restringida. De aquí en adelante se trabaja con transformaciones canónicas (no extendidas) a menos que se indique lo contrario.

El último término a la derecha de (7.13) solo contribuye a la variación de la acción en los extremos y por lo tanto se anula si  $F$  es una función de  $(q, p, t)$  o de  $(Q, P, t)$  o cualquier mezcla de las coordenadas de los dos espacios de fase  $\{q, p\}$  y  $\{Q, P\}$ , dado que todas ellas tienen variación cero en los extremos. Además, por medio de las ecuaciones de transformación (7.3) y sus inversas, es posible expresar a  $F$  en términos de coordenadas nuevas y viejas. En realidad estas funciones adquieren mayor utilidad cuando se expresan con un número igual de variables viejas y nuevas (además del tiempo). En ese sentido  $F$  actúa como puente entre los dos sistemas de coordenadas y se denomina la función generadora de la transformación.

## 7.2. Caracterización de las funciones generadoras de una transformación canónica

Veamos de que manera  $F$  puede generar la transformación. Asumamos que  $F$  es función de las  $n$  primeras coordenadas generalizadas, nuevas y viejas i.e. de  $q, Q$

$$F \equiv F_1(q, Q, t) \quad (7.14)$$

la ecuación (7.13) toma la forma

$$\begin{aligned} p_i \dot{q}_i - H &= P_i \dot{Q}_i - K + \frac{dF_1}{dt} \\ p_i \dot{q}_i - H &= P_i \dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i} \dot{q}_i + \frac{\partial F_1}{\partial Q_i} \dot{Q}_i \end{aligned} \quad (7.15)$$

pasando todo a un solo lado

$$0 = H + \left( \frac{\partial F_1}{\partial q_i} - p_i \right) \dot{q}_i + \left( P_i + \frac{\partial F_1}{\partial Q_i} \right) \dot{Q}_i - K + \frac{\partial F_1}{\partial t}$$

y dado que las coordenadas nuevas y viejas son separadamente independientes y  $H, K, F_1$  no son funciones de  $\dot{q}_i$  ni de  $\dot{Q}_i$ , se llega a que los coeficientes que acompañan a  $\dot{q}_i, \dot{Q}_i$  deben anularse

$$p_i = \frac{\partial F_1}{\partial q_i} \quad ; \quad P_i = -\frac{\partial F_1}{\partial Q_i} \quad ; \quad i = 1, \dots, n \quad (7.16)$$

quedando finalmente

$$K = H + \frac{\partial F_1}{\partial t} \quad (7.17)$$

las primeras  $n$  de las Ecs. (7.16) definen a  $p_i$  como funciones de  $q_j, Q_j$  y  $t$ . Asumiendo que son invertibles, se pueden despejar las  $Q_i$  en términos de  $q_j, p_j, t$ , lo cual nos daría la primera mitad de las transformaciones (7.3). Una vez establecidas, las  $Q_i(q_j, p_j, t)$  se pueden sustituir en la segunda serie de Ecs. (7.16) de tal manera que se obtienen las  $P_i$  como funciones de  $q_j, p_j, t$ ; y por tanto la segunda mitad de las transformaciones (7.3). Finalmente, la relación (7.17) nos da la conexión entre el nuevo Hamiltoniano  $K$  y el antiguo  $H$ , de esta relación se vé que  $K$  difiere numéricamente de  $H$  si y solo si  $F_1$  depende explícitamente del tiempo. Es necesario tener en cuenta que  $K$  debe ser función de  $Q, P, t$ , para lo cual los argumentos de  $H$  es decir  $q, p, t$  deben ser convertidos en  $Q, P, t$  a través de la inversa de las relaciones (7.3), de la misma forma, los argumentos  $q_i, t$  de la función  $\partial_t F_1$  deben expresarse en términos de  $Q, P, t$ ; con lo cual la función  $K$  ya queda en términos de los argumentos correctos.

El procedimiento anterior describe formalmente como se pueden obtener las transformaciones canónicas partiendo de una función generatriz  $F_1$ . Este proceso puede ser revertido para obtener la función generatriz comenzando con las transformaciones canónicas (7.3). Se comienza por invertir el primer conjunto de transformaciones (7.3) con el fin de expresar los  $p_i$  como funciones de  $q, Q, t$ , tal expresión se reemplaza entonces en el segundo conjunto de transformaciones (7.3) con lo cual se obtienen los  $P_i$  también como funciones de  $q, Q, t$ . Se obtienen entonces dos funciones específicas

$$p_i = p_i(q, Q, t) \quad ; \quad P_i = P_i(q, Q, t)$$

reemplazando estas funciones específicas en las ecuaciones (7.16), se obtiene un conjunto de  $2n$  ecuaciones diferenciales con  $2n$  coordenadas  $(q_i, Q_i)$ , para  $F_1(q, Q, t)$ . Estas ecuaciones permiten en principio encontrar a  $F_1$  siempre que las transformaciones sean realmente canónicas. Se puede ver que existe una arbitrariedad en la solución para  $F_1$ . En particular, se le puede adicionar una función arbitraria cuyo argumento es el tiempo solamente<sup>3</sup>

$$F_1'(q, Q, t) = F_1(q, Q, t) + g(t)$$

Esta función aditiva claramente no afecta las ecuaciones de transformación. Por otro lado, se aprecia en la Ec. (7.17), que el Hamiltoniano con  $F_1'$  sí cambia con respecto al Hamiltoniano con  $F_1$ , pero en una derivada total con respecto al tiempo, que no afecta a las ecuaciones de Hamilton. En algunas ocasiones  $F_1$  puede tener otras ambigüedades.

En ocasiones no es adecuado escribir la función generatriz en términos de los argumentos  $q, Q, t$ . Por ejemplo, la transformación puede ser tal que los  $p_i$  no se puedan escribir como funciones de  $q, Q, t$ , sino más bien como funciones de  $q, P, t$ . En este caso se debe buscar una función generatriz que sea función de las antiguas coordenadas  $q$  y los nuevos momentos  $P$ . Claramente la Ec. (7.15) debe reemplazarse por una relación similar pero que involucre a  $\dot{P}_i$  en lugar de  $\dot{Q}_i$ . Esto se logra escribiendo a  $F$  en (7.13) como

$$F \equiv F_2(q, P, t) - Q_i P_i \quad (7.18)$$

<sup>3</sup>Este es un conjunto de ecuaciones diferenciales sin condiciones iniciales ni de frontera. Por esta razón no se espera unicidad en su solución. Además,  $F_1$  no corresponde a un observable físico, de modo que tampoco es necesaria dicha unicidad.

Función generatriz	Derivadas de las $F_i$	Caso especial trivial
$F = F_1(q, Q, t)$	$p_i = \frac{\partial F_1}{\partial q_i} ; P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i, Q_i = p_i, P_i = -q_i$
$F = F_2(q, P, t) - Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i} ; Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i, Q_i = q_i, P_i = p_i$
$F = F_3(p, Q, t) + q_i p_i$	$q_i = -\frac{\partial F_3}{\partial p_i} ; P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i, Q_i = -q_i, P_i = -p_i$
$F = F_4(p, P, t) + q_i p_i - Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i} ; Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i, Q_i = p_i, P_i = -q_i$

Cuadro 7.1: Funciones generatrices de los 4 tipos, sus ecuaciones diferenciales asociadas y algunos ejemplos simples.

sustituyendo (7.18) en (7.13) y usando regla de la cadena se tiene que

$$\begin{aligned}
 p_i \dot{q}_i - H(q, p, t) &= P_i \dot{Q}_i - K(Q, P, t) + \frac{dF_2}{dt} - \frac{d}{dt}(Q_i P_i) \\
 p_i \dot{q}_i - H(q, p, t) &= P_i \dot{Q}_i - K(Q, P, t) + \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial q_i} \dot{q}_i + \frac{\partial F_2}{\partial P_i} \dot{P}_i - \dot{Q}_i P_i - Q_i \dot{P}_i \\
 p_i \dot{q}_i - H(q, p, t) &= -K(Q, P, t) + \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial q_i} \dot{q}_i + \frac{\partial F_2}{\partial P_i} \dot{P}_i - Q_i \dot{P}_i
 \end{aligned}$$

pasando todo a un extremo

$$\left( \frac{\partial F_2}{\partial q_i} - p_i \right) \dot{q}_i + \left( \frac{\partial F_2}{\partial P_i} - Q_i \right) \dot{P}_i + H - K + \frac{\partial F_2}{\partial t} = 0$$

por razones similares al caso de  $F_1$ , los términos proporcionales a  $\dot{q}_i$  y  $\dot{P}_i$  deben anularse y por lo tanto los tres últimos términos de la izquierda deben cancelarse. De estas cancelaciones se obtienen las ecuaciones fundamentales para  $F_2$  y para  $K$

$$p_i = \frac{\partial F_2}{\partial q_i} ; Q_i = \frac{\partial F_2}{\partial P_i} \tag{7.19}$$

$$K = H + \frac{\partial F_2}{\partial t} \tag{7.20}$$

al igual que con el caso anterior, el primer conjunto de ecuaciones (7.19) debe resolverse para  $P_i$  en función de  $q_j, p_j, t$ , con lo cual se tiene la segunda mitad de las transformaciones (7.3). La primera mitad de las transformaciones (7.3) se obtiene sustituyendo  $P(q, p, t)$  en la segunda mitad de las Ecs. (7.19).

Obviamente, existen otros dos tipos básicos de funciones generatrices  $F_3(p, Q, t)$  y  $F_4(p, P, t)$ , para las cuales el procedimiento es análogo y se sintetiza en la tabla 7.1. Se puede observar que los cuatro tipos básicos de funciones generatrices están conectados a través de transformaciones de Legendre. Por ejemplo, la transición entre  $F_1$  y  $F_2$  equivale a ir de las variables  $q, Q$  a las variables  $q, P$  con el segundo conjunto de relaciones (7.16)

$$-P_i = \frac{\partial F_1(q, Q, t)}{\partial Q_i}$$

ya que de esta relación se puede obtener  $P_i$  en términos de  $q, Q, t$  la cual formalmente se puede despejar para obtener  $Q_i(q, P, t)$ , que permitiría reemplazar el conjunto  $q, Q$  por el conjunto  $q, P$ . Esta es precisamente la forma requerida para una transformación de Legendre de la base de variables  $(q, Q)$  a la base de variables  $(q, P)$ , como se discutió en la sección 6.2. En analogía con la Ec. (6.3), escribimos

$$F_2(q, P, t) = F_1(q, Q, t) + P_i Q_i \tag{7.21}$$

lo cual es equivalente a igualar (7.14) con (7.18). De manera similar a partir de (7.14), y las definiciones de las otras funciones generatrices dadas en la tabla 7.1, se puede ver que todas las otras funciones generatrices se

pueden ver como transformaciones de Legendre de  $F_1$ . En particular,  $F_4$  se obtiene con dos transformaciones de Legendre de  $F_1$  dado que requiere el cambio de los dos conjuntos de argumentos. Sin embargo, no toda transformación canónica puede ser generada por medio de alguna de las cuatro funciones generatrices básicas. De modo que en algunos casos, estas transformaciones de Legendre pueden conducir a funciones generatrices nulas o indeterminadas. Por esta razón es preferible definir cada tipo de función generadora relativa a la función  $F$  de la Ec. (7.13), la cual es una función de  $2n$  coordenadas y/o momentos independientes, pero que no tiene sus argumentos bien definidos.

En particular, la función generatriz apropiada no necesariamente es del tipo 1,2,3 ó 4 para todas sus coordenadas y/o momentos. Es posible y en algunos casos necesario para ciertas transformaciones canónicas, usar una función generatriz que sea mezcla de los cuatro tipos. A manera de ejemplo, podría ser conveniente que para cierta transformación canónica con dos grados de libertad, definamos una función generatriz de la forma

$$F'(q_1, p_2, P_1, Q_2, t)$$

esta función generatriz estaría relacionada con  $F$  de (7.13) a través de

$$F = F'(q_1, p_2, P_1, Q_2, t) - Q_1 P_1 + q_2 p_2$$

esto define una transformación de Legendre adecuada para cambiar  $Q_1$  por  $P_1$  y para cambiar  $q_2$  por  $p_2$ . Las ecuaciones de transformación se obtienen a través de las relaciones

$$\begin{aligned} p_1 &= \frac{\partial F'}{\partial q_1} ; & Q_1 &= \frac{\partial F'}{\partial P_1} \\ q_2 &= -\frac{\partial F'}{\partial p_2} ; & P_2 &= -\frac{\partial F'}{\partial Q_2} \\ K &= H + \frac{\partial F'}{\partial t} \end{aligned}$$

Se puede ver que  $F'$  es una mezcla entre funciones del tipo  $F_2$  y  $F_3$ .

### 7.3. Ejemplos de transformaciones canónicas

Consideremos primero una función del segundo tipo de la forma

$$F_2 = q_i P_i \tag{7.22}$$

a partir de las Ecs. (7.19, 7.20), las ecuaciones de transformación serán

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i ; \quad Q_i = \frac{\partial F_2}{\partial P_i} = q_i ; \quad K = H$$

las coordenadas nuevas coinciden con las antiguas, de modo que  $F_2$  es una función generatriz de la transformación identidad. Tomemos ahora  $F_3 = p_i Q_i$ , usando las ecuaciones de la tabla 7.1, vemos que esta función genera la transformación canónica  $-I$ , i.e.  $Q_i = -q_i$ ,  $P_i = -p_i$ .

Un tipo más general de transformación del tipo  $F_2$  que incluye a (7.22) como caso particular, se define por

$$F_2 = f_i(q_1, \dots, q_n; t) P_i + g(q_1, \dots, q_n; t) \tag{7.23}$$

donde las  $f_i$  son un conjunto de funciones independientes, y  $g$  es una función diferenciable en  $q, t$ , ambas tienen como argumentos las coordenadas antiguas y el tiempo. Usando el segundo conjunto de Ecs. (7.19) las nuevas coordenadas  $Q_i$  vienen dadas por

$$Q_k = \frac{\partial F_2}{\partial P_k} = f_k(q_1, \dots, q_n; t) \tag{7.24}$$

de modo que con esta transformación las coordenadas nuevas solo dependen de las coordenadas antiguas y el tiempo pero no de los antiguos momentos. Esto define una transformación puntual como la expresada en

(7.2). Para definir una transformación puntual las funciones  $f_i$  deben ser independientes e invertibles, de modo que las  $q_j$  se puedan expresar en términos de las  $Q_k$ . Dado que las  $f_i$  son por lo demás arbitrarias, se concluye que toda transformación puntual puede ser parte de una transformación canónica (recuérdese que una transformación canónica requiere conocer también las transformaciones en los  $p_i$ ). Las Ecs. (7.20, 7.23) nos dan el nuevo Hamiltoniano

$$K = H + \frac{\partial f_i}{\partial t} P_i + \frac{\partial g}{\partial t}$$

Veamos ahora las ecuaciones de transformación para el momento conjugado que induce (7.23), el primer conjunto de Ecs. (7.19) lleva a

$$p_j = \frac{\partial F_2}{\partial q_j} = \frac{\partial f_i}{\partial q_j} P_i + \frac{\partial g}{\partial q_j} \quad (7.25)$$

estas ecuaciones se pueden invertir para dar  $P$  en función de  $(q, p)$ . Este procedimiento es más sencillo si se usa una formulación matricial, definiremos entonces arreglos matriciales de la forma

$$(\mathbf{p})_j \equiv p_j ; (\mathbf{P})_i \equiv P_i ; \left( \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \right)_{ji} \equiv \frac{\partial f_i}{\partial q_j} ; \left( \frac{\partial g}{\partial \mathbf{q}} \right)_j \equiv \frac{\partial g}{\partial q_j}$$

con lo cual la transformación (7.25) se escribe

$$\mathbf{p} = \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \mathbf{P} + \frac{\partial g}{\partial \mathbf{q}} \quad (7.26)$$

de aquí se puede despejar  $\mathbf{P}$

$$\mathbf{P} = \left( \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \right)^{-1} \left( \mathbf{p} - \frac{\partial g}{\partial \mathbf{q}} \right) \quad (7.27)$$

obsérvese que a diferencia de las  $Q_i$  que solo dependían de las  $q_j$  y el tiempo, las  $P_i$  dependen de  $q_j, p_j, t$ .

Otra función generadora interesante es la siguiente

$$F_1 = q_k Q_k \quad (7.28)$$

las ecuaciones de transformación (7.16) en este caso quedan

$$p_i = \frac{\partial F_1}{\partial q_i} = Q_i ; P_i = -\frac{\partial F_1}{\partial Q_i} = -q_i$$

esta función generatriz intercambia los momentos y las coordenadas excepto por un signo. La función  $F_4 = p_i P_i$  produce la misma transformación. Este hecho enfatiza el estatus independiente otorgado a los momentos, ya que tanto los momentos como las coordenadas se necesitan para describir el movimiento en la formulación Hamiltoniana y la distinción entre ambos es asunto de nomenclatura, puesto que podemos intercambiar los nombres con un costo no mayor a un cambio de signo. No hay ningún remanente en la teoría de  $q$  como coordenada espacial o de  $p$  como masa por velocidad. En realidad, de las propias ecuaciones de Hamilton

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} , \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

se puede ver que este intercambio es canónico, con el cambio de signo apropiado. Si  $q_i$  se sustituye por  $p_i$ , las ecuaciones permanecen en la forma canónica solo si  $-p_i$  es sustituido por  $q_i$ .

Podemos ver adicionalmente que la transformación de permutación no se puede generar a partir de una función generatriz del tipo  $F_2(q, P, t)$ , ya que el sistema de ecuaciones parciales que se deriva de (7.19) nos dice que

$$p_i = \frac{\partial F_2(q, P, t)}{\partial q_i}$$

esto implica que  $p_i$  debe escribirse como función de  $q, P, t$  pero en la transformación de permutación  $p_i$  no es función de ninguna de estas variables, en realidad solo es función de las nuevas coordenadas  $Q$ . Por tanto,  $F_2$  no es una función generatriz adecuada para la transformación de permutación.

Similarmente, una función tipo  $F_1(q, Q, t)$  no puede generar la transformación identidad, ya que las Ecs. (7.16) nos llevan a

$$p_i = \frac{\partial F_1(q, Q, t)}{\partial q_i}$$

y definen a  $p_i$  en función de  $q, Q, t$ , pero en la transformación identidad,  $p_i$  solo es función de  $P_i$ . Adicionalmente, no es posible obviar esta dificultad definiendo una función  $F_1$  mediante una transformada de Legendre del tipo (7.21), aplicada a la generadora  $F_2$  de la identidad.

$$F_1(q, Q, t) = F_2(q, P, t) - P_i Q_i = q_i P_i - P_i Q_i$$

pero en la transformación identidad es obvio que  $q = Q$  de modo que

$$F_1(q, Q, t) = Q_i P_i - P_i Q_i = 0$$

que no es una función generatriz. Similar resultado se obtiene si se intenta generar la transformación de permutación a partir de una función  $F_2$  que provenga de la transformación de Legendre del generador  $F_1$  de permutación. No obstante, se puede demostrar que una función  $F_3$  puede generar la transformación identidad y que una función tipo  $F_4$  puede generar la permutación.

Finalmente, una transformación que deja algunas parejas  $(q, p)$  inalteradas, e intercambia el resto (con un cambio de signo), es una transformación canónica de tipo mixto. En un sistema de dos grados de libertad, la transformación

$$\begin{aligned} Q_1 &= q_1, & P_1 &= p_1 \\ Q_2 &= p_2, & P_2 &= -q_2 \end{aligned} \quad (7.29)$$

es generada por la función

$$F = q_1 P_1 + q_2 Q_2 \quad (7.30)$$

que es una mezcla entre funciones de tipo  $F_1$  y  $F_2$ . El término tipo  $F_1$  produce permutación y el término tipo  $F_2$  es del tipo identidad.

## 7.4. Transformaciones canónicas para el oscilador armónico

El Hamiltoniano para el oscilador armónico se escribe como

$$H = \frac{1}{2m} (p^2 + m^2 \omega^2 q^2) \quad ; \quad \omega^2 \equiv \frac{k}{m}$$

esta forma en suma de cuadrados en  $q$  y  $p$  sugiere que para que la nueva coordenada sea cíclica, podemos hallar una transformación de la forma

$$p = f(P) \cos Q \quad ; \quad q = \frac{f(P)}{m\omega} \sin Q \quad (7.31)$$

Dado que la transformación (7.31) no depende explícitamente del tiempo, el Hamiltoniano no cambia numéricamente y en las nuevas coordenadas queda

$$K = H = \frac{f^2(P)}{2m} (\cos^2 Q + \sin^2 Q) = \frac{f^2(P)}{2m} \quad (7.32)$$

de modo que  $Q$  resulta cíclica. Solo nos queda hallar la forma de  $f(P)$  tal que la transformación (7.31) sea canónica. El cociente entre la Ecs. (7.31) nos lleva a

$$p = m\omega q \cot Q \quad (7.33)$$

que es independiente de  $f(P)$ . Esta transformación es de la forma  $p = p(q, Q)$  lo cual sugiere buscar una función generatriz tipo  $F_1$ , pues el primer conjunto de Ecs. (7.16) nos da  $p_i = p_i(q, Q, t)$  en la forma

$$p = \frac{\partial F_1(q, Q, t)}{\partial q} \quad (7.34)$$

y la solución más sencilla de  $F_1$  para la Ec. (7.33) es

$$F_1(q, Q, t) = \frac{m\omega q^2}{2} \cot Q \quad (7.35)$$

naturalmente,  $F_1$  debe satisfacer la otra mitad de las ecuaciones de transformación (7.16)

$$P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2 \sin^2 Q} \quad (7.36)$$

despejando  $q$  se obtiene

$$q = \sqrt{\frac{2P}{m\omega}} \sin Q \quad (7.37)$$

que al compararla con la segunda de las Ecs. (7.31) nos da

$$f(P) = \sqrt{2m\omega P} \quad (7.38)$$

al reemplazar este factor en el nuevo Hamiltoniano (7.32) queda

$$K = H = \omega P$$

que claramente es cíclico en  $Q$ , con lo cual  $P$  es una constante de movimiento. Como  $H$  es constante y es la energía del sistema, el nuevo momento canónico viene dado por

$$P = \frac{E}{\omega} \quad (7.39)$$

y la ecuación de Hamilton para  $Q$  se reduce a

$$\dot{Q} = \frac{\partial K}{\partial P} = \omega$$

cuya solución es

$$Q = \omega t + \alpha \quad (7.40)$$

siendo  $\alpha$  una constante de integración fijada por las condiciones iniciales. Sustituyendo (7.39) y (7.40) en (7.37) resulta

$$q = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \alpha)$$

que es la solución conocida para el oscilador armónico.

Cabe ahora la pregunta, ¿qué nos asegura que la transformación definida por (7.31) y por (7.38) es canónica?. Lo asegura el hecho de que tal transformación es generada por una función generatriz tipo 1 Ec. (7.35) que cumple con sus ecuaciones diferenciales Ecs. (7.34) y (7.36).



## 7.5. Transformaciones canónicas usando la forma simpléctica de las ecuaciones de Hamilton

El método de las funciones generatrices es formalmente completo para tratar el problema de caracterizar todas las transformaciones canónicas, puesto que el principio de Hamilton modificado me garantiza la existencia de una función generatriz para cualquier transformación que deje invariante la forma de las ecuaciones de Hamilton. Sin embargo, en la práctica puede ser muy difícil encontrar una función generatriz<sup>4</sup> y si no estamos seguros de si la transformación es canónica puede ser muy difícil demostrar la existencia o no existencia de la función generatriz. Por esta razón, se hace necesario encontrar métodos que sean más sistemáticos para verificar si una transformación es o no es canónica.

Otro método para tratar las transformaciones canónicas, consiste en la utilización de las ecuaciones de Hamilton simplécticas descritas en la sección 6.5 a través de las Ecs. (6.33). Consideraremos primero el caso de transformaciones canónicas restringidas, es decir que no dependen explícitamente del tiempo

$$Q_i = Q_i(q, p) \quad ; \quad P_i = P_i(q, p) \quad (7.41)$$

la primera característica notable es que el Hamiltoniano no cambia ante una transformación canónica restringida, como se vé de (7.14, 7.17). Su forma funcional puede ser diferente pero se conserva su valor numérico. La idea es que la transformación sea canónica, lo cual es equivalente a exigir que las nuevas coordenadas  $Q, P$  cumplan ecuaciones de Hamilton para el nuevo Hamiltoniano  $K = H$

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} \quad ; \quad -\dot{P}_i = \frac{\partial H}{\partial Q_i} \quad (7.42)$$

Estudiaremos el primer conjunto de Ecs. (7.42), evaluemos primero  $\dot{Q}_i$  el cual se calcula con base en (7.41), y en las ecuaciones de Hamilton para las coordenadas antiguas (es decir asumiendo que las coordenadas originales también son canónicas)

$$\dot{Q}_i = \frac{\partial Q_i}{\partial q_j} \dot{q}_j + \frac{\partial Q_i}{\partial p_j} \dot{p}_j = \frac{\partial Q_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j} \frac{\partial H}{\partial q_j} \quad (7.43)$$

ahora calculamos el término de la derecha en el primer conjunto de Ecs. (7.42), para lo cual se requieren las inversas de (7.41)

$$q_j = q_j(Q, P) \quad ; \quad p_j = p_j(Q, P) \quad (7.44)$$

las cuales nos permiten calcular  $H(q, p, t)$  en términos de  $Q, P, t$ , así como el término en cuestión

$$\frac{\partial H}{\partial P_i} = \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial P_i} + \frac{\partial H}{\partial q_j} \frac{\partial q_j}{\partial P_i} \quad (7.45)$$

El primer conjunto de Ecs. (7.42), nos conduce a igualar los términos de la derecha en (7.43) y (7.45) con lo cual resulta

$$\left( \frac{\partial Q_i}{\partial q_j} \right)_{q,p} = \left( \frac{\partial p_j}{\partial P_i} \right)_{Q,P} \quad ; \quad \left( \frac{\partial Q_i}{\partial p_j} \right)_{q,p} = - \left( \frac{\partial q_j}{\partial P_i} \right)_{Q,P} \quad (7.46)$$

donde los subíndices enfatizan los argumentos en los cuales queda evaluada cada expresión. Con un procedimiento similar, el segundo conjunto de Ecs. (7.42) nos lleva a las siguientes condiciones

$$\left( \frac{\partial P_i}{\partial q_j} \right)_{q,p} = - \left( \frac{\partial p_j}{\partial Q_i} \right)_{Q,P} \quad ; \quad \left( \frac{\partial P_i}{\partial p_j} \right)_{q,p} = \left( \frac{\partial q_j}{\partial Q_i} \right)_{Q,P} \quad (7.47)$$

el conjunto de condiciones (7.46) y (7.47) se conoce usualmente como **condiciones directas** para una transformación canónica restringida.

<sup>4</sup>Vale la pena mencionar que los cuatro tipos básicos de funciones generatrices que se han estudiado no cubren el espectro de todas las transformaciones canónicas.

El procedimiento anterior se puede realizar en forma matricial usando las ecuaciones de Hamilton simplécticas descritas en la sección 6.5. Partimos de la formulación simpléctica descrita en las Ecs. (6.33).

$$\dot{\eta} = \mathbf{J} \frac{\partial H}{\partial \eta} \quad (7.48)$$

con  $\mathbf{J}$  y  $\eta$  definidos por (6.29, 6.31). Similarmente, el nuevo conjunto coordenado  $Q_i, P_i$  define una matriz columna de dimensión  $2n$  análoga a  $\eta$ , denotamos esta matriz por  $\zeta$ , y para una transformación canónica restringida, las Ecs. (7.41) se pueden escribir como

$$\zeta = \zeta(\eta) \quad (7.49)$$

ahora bien, la condición canónica exige que se cumplan las Ecs. (7.42), que matricialmente se escriben en la forma

$$\dot{\zeta} = \mathbf{J} \frac{\partial H}{\partial \zeta} \quad (7.50)$$

donde hemos tenido en cuenta que una transformación canónica restringida no cambia el valor numérico del Hamiltoniano como se vió en el formalismo de la función generatriz. Evaluemos primero el término de la izquierda en (7.50) teniendo en cuenta (7.49)

$$\dot{\zeta}_i = \frac{\partial \zeta_i}{\partial \eta_j} \dot{\eta}_j \quad i, j = 1, \dots, 2n$$

definimos entonces la matriz jacobiana de la transformación

$$M_{ij} \equiv \frac{\partial \zeta_i}{\partial \eta_j} \Rightarrow \dot{\zeta}_i = M_{ij} \dot{\eta}_j \quad (7.51)$$

matricialmente queda

$$\dot{\zeta} = \mathbf{M} \dot{\eta} \quad (7.52)$$

reemplazando (7.48) en (7.52) resulta

$$\dot{\zeta} = \mathbf{M} \mathbf{J} \frac{\partial H}{\partial \eta} \quad (7.53)$$

por otro lado, a través de la relación inversa de (7.49),  $H$  se puede escribir como función de  $\zeta$ , y podemos evaluar también el término de la derecha en (7.50)

$$\frac{\partial H}{\partial \eta_i} = \frac{\partial H}{\partial \zeta_j} \frac{\partial \zeta_j}{\partial \eta_i} \Rightarrow \frac{\partial H}{\partial \eta_i} = M_{ji} \frac{\partial H}{\partial \zeta_j} = \widetilde{M}_{ij} \frac{\partial H}{\partial \zeta_j}$$

matricialmente esto se escribe

$$\frac{\partial H}{\partial \eta} = \widetilde{\mathbf{M}} \frac{\partial H}{\partial \zeta} \quad (7.54)$$

reemplazando (7.54) en (7.53)

$$\dot{\zeta} = \mathbf{M} \mathbf{J} \widetilde{\mathbf{M}} \frac{\partial H}{\partial \zeta} \quad (7.55)$$

y teniendo en cuenta la condición (7.50), en comparación con (7.55), se concluye que la transformación (7.49) es canónica si  $\mathbf{M}$  satisface la condición

$$\mathbf{M} \mathbf{J} \widetilde{\mathbf{M}} = \mathbf{J} \quad (7.56)$$

esta condición también es necesaria para una transformación canónica restringida, como se puede ver invirtiendo los pasos anteriores. La condición (7.56), se puede escribir de varias maneras. Por ejemplo, multiplicando por  $\widetilde{\mathbf{M}}^{-1}$  a la derecha en ambos miembros

$$\mathbf{M} \mathbf{J} = \mathbf{J} \widetilde{\mathbf{M}}^{-1} \quad (7.57)$$

la transposición y la inversión se pueden tomar en cualquier orden. Si se comparan explícitamente las condiciones que se derivan de (7.57), se observa que coinciden con las expresadas en las Ecs. (7.46) y (7.47). Adicionalmente, usando (6.32) vemos que la Ec. (7.57) se puede transformar en

$$\begin{aligned}\mathbf{JMJ}(-\mathbf{J}) &= \mathbf{J}^2\widetilde{\mathbf{M}}^{-1}(-\mathbf{J}) \Rightarrow \\ \mathbf{JM} &= \widetilde{\mathbf{M}}^{-1}\mathbf{J}\end{aligned}$$

o equivalentemente

$$\widetilde{\mathbf{M}}\mathbf{JM} = \mathbf{J} \quad (7.58)$$

La ecuación (7.56), o la versión equivalente (7.58) se conocen como **condiciones simplécticas** para una transformación canónica y la matriz  $\mathbf{M}$  que satisface esta condición, se denomina **matriz simpléctica**.

Para una transformación canónica extendida restringida (independiente del tiempo e incluyendo una transformación de escala), donde  $K = \lambda H$ , la condición (7.56) resulta

$$\mathbf{MJ}\widetilde{\mathbf{M}} = \lambda\mathbf{J} \quad (7.59)$$

### 7.5.1. Ejemplos de transformaciones canónicas con método matricial

Veamos un ejemplo explícito. Tomemos la función mixta definida en la Ec. (7.30)  $F = F_2(q_1, P_1) + F_1(q_2, Q_2)$ . Las matrices  $\boldsymbol{\eta}$ ,  $\boldsymbol{\zeta}$  y  $\mathbf{J}$  son

$$\boldsymbol{\eta} = \begin{pmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{pmatrix} ; \quad \boldsymbol{\zeta} = \begin{pmatrix} Q_1 \\ Q_2 \\ P_1 \\ P_2 \end{pmatrix} ; \quad \mathbf{J} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad (7.60)$$

tomando las ecuaciones de transformación (7.29) y derivándolas con respecto al tiempo, se pueden comparar estas transformaciones con las descritas por  $\dot{\boldsymbol{\zeta}} = \mathbf{M}\dot{\boldsymbol{\eta}}$  (Ec. 7.52) con el fin de obtener la matriz  $\mathbf{M}$

$$\begin{pmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} = \begin{pmatrix} \dot{q}_1 \\ \dot{p}_2 \\ \dot{p}_1 \\ -\dot{q}_2 \end{pmatrix}$$

por lo tanto

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad (7.61)$$

Naturalmente, el hecho de que esta transformación haya sido obtenida con una función generatriz y sus ecuaciones diferenciales asociadas, garantiza el carácter canónico de la transformación. No obstante, el lector puede verificar el carácter canónico de la transformación observando que las matrices  $\mathbf{J}$  y  $\mathbf{M}$  de las Ecs. (7.60, 7.61), cumplen con la condición simpléctica (7.58).

Por otro lado, las ecuaciones de Hamilton para las nuevas variables i.e.  $\dot{\boldsymbol{\zeta}} = \mathbf{J}(\partial H/\partial\boldsymbol{\zeta})$  (Ec. 7.50), se expresan sin dependencia con la función generatriz

$$\begin{aligned}\begin{pmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \partial H/\partial Q_1 \\ \partial H/\partial Q_2 \\ \partial H/\partial P_1 \\ \partial H/\partial P_2 \end{pmatrix} \\ \begin{pmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\dot{P}_1 \\ -\dot{P}_2 \\ \dot{Q}_1 \\ \dot{Q}_2 \end{pmatrix}\end{aligned}$$

$\mathbf{M}$  es función de  $F$  (ya que depende de la transformación canónica específica), pero no así  $\mathbf{J}$  la cual es una matriz constante (ver Ec. 6.31). Este formalismo no se puede aplicar en todos los casos. En particular, no se puede construir una matriz simple  $\mathbf{M}$  para generar la transformación canónica usada para el oscilador armónico en la sección 7.4.

## 7.6. Acercamiento simpléctico para transformaciones canónicas dependientes del tiempo

Cuando la transformación a las nuevas coordenadas depende explícitamente del tiempo, ya no son válidos los argumentos anteriores. Sin embargo, se sigue cumpliendo que la condición simpléctica es necesaria y suficiente para que la transformación sea canónica. Esto se puede probar por argumentos similares a los arriba expuestos, pero es más fácil a través de otro camino que explota el hecho de que el tiempo actúa como un parámetro en la transformación canónica. Una transformación canónica de la forma

$$\zeta = \zeta(\eta, t)$$

evoluciona continuamente con el tiempo a partir de un cierto  $t_0$ . Es una transformación continua uniparamétrica. Si la transformación

$$\eta \rightarrow \zeta(t) \tag{7.62}$$

es canónica, se tiene que en particular la transformación

$$\eta \rightarrow \zeta(t_0) \tag{7.63}$$

es canónica siendo  $t_0$  un tiempo fijo, se sigue entonces de la definición de transformación canónica, que la transformación caracterizada por

$$\zeta(t_0) \rightarrow \zeta(t) \tag{7.64}$$

es también canónica. Puesto que en (7.63) el tiempo es fijo, esta transformación canónica debe satisfacer la condición simpléctica (7.58). Si podemos probar que (7.64) cumple la condición simpléctica, es fácil mostrar que la transformación general (7.62) también la cumple.

Para demostrar que la condición simpléctica es aún necesaria y suficiente para transformaciones de la forma (7.64) usaremos la noción de transformación canónica infinitesimal (TCI) en la cual todas las coordenadas nuevas difieren de las antiguas por cantidades infinitesimales

$$Q_i = q_i + \delta q_i \tag{7.65}$$

$$P_i = p_i + \delta p_i \tag{7.66}$$

que matricialmente se puede escribir como

$$\zeta = \eta + \delta\eta \tag{7.67}$$

la notación  $\delta q_i, \delta p_i$  no se refiere a desplazamientos virtuales, simplemente cambios infinitesimales en las coordenadas (cambio en coordenadas no implica movimiento del sistema ni real ni virtual). Como una transformación canónica infinitesimal difiere solo infinitesimalmente del operador identidad, es natural comenzar con una función generatriz de la identidad colocándole una desviación infinitesimal

$$F_2 = q_i P_i + \varepsilon G(q, P, t) \tag{7.68}$$

$\varepsilon$  es un parámetro infinitesimal y  $G$  es una función con segundas derivadas continuas en sus  $2n+1$  argumentos. Dado que esta función generatriz es de tipo 2, podemos emplear las Ecs. (7.19) de transformación para los momentos

$$p_j = \frac{\partial F_2}{\partial q_j} = P_j + \varepsilon \frac{\partial G}{\partial q_j}$$

con lo cual

$$\delta p_j \equiv P_j - p_j = -\varepsilon \frac{\partial G}{\partial q_j} \quad (7.69)$$

adicionalmente, usando las Ecs. (7.19) para transformación de coordenadas

$$Q_j = \frac{\partial F_2}{\partial P_j} = q_j + \varepsilon \frac{\partial G}{\partial P_j}$$

dado que el segundo término a la derecha es lineal en  $\varepsilon$ , y que  $P_j$  solo difiere de  $p_j$  infinitesimalmente, entonces es consistente a primer orden cambiar la variable de derivación  $P_j$  por  $p_j$ . De esta forma  $G$  se escribe en términos de  $q, p, t$  únicamente. El cambio infinitesimal  $\delta q_j$  se puede escribir a primer orden como

$$\delta q_j \equiv Q_j - q_j = \varepsilon \frac{\partial G}{\partial p_j} \quad (7.70)$$

las ecuaciones de transformación (7.69, 7.70), se pueden sintetizar en una ecuación matricial

$$\delta \boldsymbol{\eta} = \varepsilon \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\eta}} \quad (7.71)$$

una transformación infinitesimal que nos interesa directamente, es la transformación (7.64) cuando  $t$  difiere de  $t_0$  en una cantidad infinitesimal

$$\boldsymbol{\zeta}(t_0) \rightarrow \boldsymbol{\zeta}(t_0 + dt) \quad (7.72)$$

$dt$  hace las veces del  $\varepsilon$ . Dado que entre  $\boldsymbol{\zeta}(\boldsymbol{\eta}, t_0)$  y  $\boldsymbol{\zeta}(\boldsymbol{\eta}, t)$  con  $t_0$  y  $t$  arbitrarios, la transformación es continua, se puede construir como una sucesión de transformaciones infinitesimales en pasos de  $dt$ . Por lo tanto, bastará con probar que la transformación infinitesimal (7.72) satisface la condición simpléctica (7.58). La TCI (7.71), tiene asociada una matriz jacobiana, y se puede ver fácilmente que esta última es simpléctica. Tomando la definición de matriz jacobiana (7.51) para una transformación infinitesimal como (7.67) se obtiene

$$\mathbf{M} \equiv \frac{\partial \boldsymbol{\zeta}}{\partial \boldsymbol{\eta}} = \mathbf{1} + \frac{\partial \delta \boldsymbol{\eta}}{\partial \boldsymbol{\eta}}$$

y usando (7.71), resulta

$$\begin{aligned} \mathbf{M} &= \mathbf{1} + \frac{\partial}{\partial \boldsymbol{\eta}} \left( \varepsilon \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\eta}} \right) \\ \mathbf{M} &= \mathbf{1} + \varepsilon \mathbf{J} \left( \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right) \end{aligned} \quad (7.73)$$

definimos la matriz

$$\left( \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right)_{ij} \equiv \frac{\partial^2 G}{\partial \eta_i \partial \eta_j} \quad (7.74)$$

la cual es claramente simétrica, en virtud de la continuidad de las segundas derivadas de  $G$ . Teniendo además en cuenta que  $\mathbf{J}$  es antisimétrico, y trasponiendo  $\mathbf{M}$  en (7.73) resulta

$$\begin{aligned} \widetilde{\mathbf{M}} &= \mathbf{1} + \varepsilon \left( \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right) \widetilde{\mathbf{J}} \\ \widetilde{\mathbf{M}} &= \mathbf{1} - \varepsilon \left( \frac{\partial^2 G}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}} \right) \mathbf{J} \end{aligned}$$

para chequear la condición simpléctica (7.58) debemos hacer el producto  $\widetilde{\mathbf{M}}\mathbf{J}\mathbf{M}$

$$\begin{aligned}\widetilde{\mathbf{M}}\mathbf{J}\mathbf{M} &= \left[ \mathbf{1} - \varepsilon \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) \mathbf{J} \right] \mathbf{J} \left[ \mathbf{1} + \varepsilon \mathbf{J} \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) \right] \\ &= \mathbf{J} + \varepsilon \mathbf{J}^2 \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) - \varepsilon \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) \mathbf{J}^2 + O(\varepsilon^2) \\ \widetilde{\mathbf{M}}\mathbf{J}\mathbf{M} &= \mathbf{J} - \varepsilon \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) + \varepsilon \left( \frac{\partial^2 G}{\partial \eta \partial \eta} \right) + O(\varepsilon^2) \\ \widetilde{\mathbf{M}}\mathbf{J}\mathbf{M} &= \mathbf{J} + O(\varepsilon^2)\end{aligned}$$

donde hemos usado la propiedad  $\mathbf{J}^2 = -\mathbf{1}$ , de modo que a primer orden en  $\varepsilon$ , se cumple la condición simpléctica para la transformación infinitesimal (7.72). Por tanto cualquier transformación canónica obedece la condición simpléctica aún cuando sea explícitamente dependiente del tiempo. Nótese que es esencial que no haya contribuciones a primer orden en  $\varepsilon$ , ya que una transformación canónica finita, se obtiene con un proceso de integración sobre el parámetro infinitesimal de las TCI. En consecuencia, los términos de primer orden darían una contribución finita para una transformación canónica finita, y solo los términos de segundo orden en adelante tienden a cero después de llevar a cabo la integración.

A pesar de que en este acercamiento poco se ha usado el formalismo de la función generatriz, tanto el formalismo simpléctico como el de la función generatriz están conectados (ver sección 8.5). Se puede demostrar por ejemplo que la condición simpléctica implica la existencia de una función generatriz. Cualquiera de los dos formalismos se puede usar indistintamente de acuerdo con las necesidades y conveniencias. En particular, cualquiera de los dos formalismos sirve para demostrar que las transformaciones canónicas forman un grupo matemático.

## 7.7. Ejemplos de transformaciones canónicas

En esta sección estudiaremos el paso de un conjunto canónico de coordenadas  $\{q_i, p_i\}$  a otro conjunto de coordenadas  $\{Q_i, P_i\}$ . A menos que se indique lo contrario, se considerará que el conjunto de coordenadas  $\{q_i, p_i\}$  es de hecho canónico, y se asumirá además como el conjunto canónico “original” (ya hemos dicho que esto es cuestión de convención), y en algunos casos se procurará averiguar si el conjunto “final”  $\{Q_i, P_i\}$  es canónico o no.

### 7.7.1. Transformación canónica por conjugación compleja

Una forma natural de intentar combinar los dos conjuntos de ecuaciones de Hamilton en uno solo, es formar una cantidad compleja con las coordenadas  $\{q, p\}$ , y convertir los dos conjuntos de ecuaciones en un solo conjunto de ecuaciones complejas. Para un solo grado de libertad, definiremos una transformación de la forma

$$Q = \mu(q + ip) \quad ; \quad P = \nu(q - ip) \quad (7.75)$$

y veremos si esta transformación es canónica, o bajo que condiciones lo es. La transformación inversa es

$$q = \frac{1}{2} \left( \frac{1}{\mu} Q + \frac{1}{\nu} P \right) \quad ; \quad p = \frac{1}{2i} \left( \frac{1}{\mu} Q - \frac{1}{\nu} P \right) \quad (7.76)$$

En este caso utilizaremos las condiciones **directas** para transformaciones canónicas restringidas (independientes del tiempo) dadas por las ecuaciones (7.46) y (7.47)

$$\frac{\partial Q}{\partial q} = \frac{\partial p}{\partial P} \quad ; \quad \frac{\partial Q}{\partial p} = -\frac{\partial q}{\partial P} \quad ; \quad \frac{\partial P}{\partial q} = -\frac{\partial p}{\partial Q} \quad ; \quad \frac{\partial P}{\partial p} = \frac{\partial q}{\partial Q} \quad (7.77)$$

al aplicar (7.77) a las transformaciones (7.75) y sus inversas (7.76), tenemos que

$$\mu = -\frac{1}{2i\nu} \quad ; \quad i\mu = -\frac{1}{2\nu} \quad ; \quad \nu = -\frac{1}{2i\mu} \quad ; \quad -i\nu = \frac{1}{2\mu}$$

las cuatro ecuaciones son equivalentes y nos dan

$$\mu = -\frac{1}{2i\nu} \quad (7.78)$$

en particular las transformaciones con  $\mu = \nu = 1$ , no cumplen la condición (7.78) y por tanto no son canónicas. De hecho, la condición (7.78), indica que  $|\mu| \neq |\nu|$ , y nos muestra que las transformaciones (7.75) solo son canónicas si incluyen una transformación de escala no trivial. El lector puede comprobar de una forma similar al ejemplo 4 Pág. 121, que esta transformación altera el valor del Hamiltoniano aunque sea independiente del tiempo. Esto se debe a la inclusión de la transformación de escala. Vemos entonces que una transformación canónica induce un cambio en el valor del Hamiltoniano si depende explícitamente del tiempo, o si incluye una transformación de escala.

### 7.7.2. Transformación canónica de rotación

Para un sistema de un grado de libertad, es natural pensar en una transformación de “rotación” para el sistema  $\{q, p\}$  en un cierto ángulo  $\alpha$ . Tendremos entonces la transformación

$$\begin{aligned} Q &= q \cos \alpha - p \sin \alpha \\ P &= q \sin \alpha + p \cos \alpha \end{aligned} \quad (7.79)$$

para averiguar si esta transformación es canónica, utilizaremos la condición simpléctica Ec. (7.56)

$$\mathbf{M}\widetilde{\mathbf{J}}\mathbf{M} = \mathbf{J} \quad (7.80)$$

Podemos hallar  $\mathbf{M}$  a través de la ecuación (7.52)

$$\dot{\zeta} = \mathbf{M}\dot{\eta} \quad (7.81)$$

para lo cual derivamos las Ecs. (7.79) con el tiempo, teniendo en cuenta que  $\alpha$  es fijo

$$\begin{pmatrix} \dot{Q} \\ \dot{P} \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \mathbf{M} \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix}$$

evaluando explícitamente la condición (7.80) tenemos

$$\begin{aligned} \mathbf{M}(\widetilde{\mathbf{J}}) &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \left[ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \right] \\ \mathbf{M}\widetilde{\mathbf{J}}\mathbf{M} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} -\sin \alpha & \cos \alpha \\ -\cos \alpha & -\sin \alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathbf{J} \end{aligned}$$

vemos que la transformación es canónica para cada valor de  $\alpha$ .

Vamos a intentar encontrar una función generadora de esta transformación canónica. Primero intentaremos encontrar una función del tipo  $F_1(q, Q)$ , que obedece las ecuaciones diferenciales (7.16)

$$p = \frac{\partial F_1(q, Q)}{\partial q} \quad ; \quad P = -\frac{\partial F_1(q, Q)}{\partial Q} \quad (7.82)$$

La primera de las ecuaciones (7.82), requiere conocer  $p$  como función de  $q, Q$ . Podemos lograr este despeje en la primera de las Ecs. (7.79)

$$p(q, Q) = -\frac{Q}{\sin \alpha} + q \cot \alpha \quad (7.83)$$

con lo cual la primera de las ecuaciones diferenciales (7.82) queda

$$\frac{\partial F_1(q, Q)}{\partial q} = -\frac{Q}{\sin \alpha} + q \cot \alpha$$

que al integrar nos da

$$F_1(q, Q) = -\frac{Qq}{\sin \alpha} + \frac{q^2}{2} \cot \alpha + g(Q) \quad (7.84)$$

para poder emplear la segunda de las ecuaciones diferenciales (7.82), debemos escribir  $P = P(q, Q)$ , lo cual se logra insertando (7.83) en la segunda de las ecuaciones de transformación (7.79)

$$P = q \sin \alpha + p \cos \alpha = q \sin \alpha + \left[ -\frac{Q}{\sin \alpha} + \frac{q \cos \alpha}{\sin \alpha} \right] \cos \alpha$$

$$P(q, Q) = q \sin \alpha + \frac{q \cos^2 \alpha}{\sin \alpha} - \frac{Q \cos \alpha}{\sin \alpha} \quad (7.85)$$

$$P(q, Q) = \frac{q}{\sin \alpha} - \frac{Q \cos \alpha}{\sin \alpha} \quad (7.86)$$

utilizando (7.86) en la segunda de las Ecs. (7.82) resulta

$$\frac{\partial F_1(q, Q)}{\partial Q} = -\frac{q}{\sin \alpha} + \frac{Q \cos \alpha}{\sin \alpha}$$

que se puede integrar para obtener

$$F_1 = -\frac{qQ}{\sin \alpha} + \frac{Q^2}{2} \cot \alpha + h(q) \quad (7.87)$$

comparando las  $F_1$  de las ecuaciones (7.84, 7.87) encontramos una solución haciendo

$$h(q) = \frac{q^2}{2} \cot \alpha \quad ; \quad g(Q) = \frac{Q^2}{2} \cot \alpha$$

quedando finalmente

$$F_1 = -\frac{Qq}{\sin \alpha} + \frac{1}{2}(q^2 + Q^2) \cot \alpha \quad ; \quad \alpha \neq n\pi \quad (7.88)$$

esta solución es válida para todo  $\alpha$ , excepto para  $\alpha = n\pi$  siendo  $n$  un entero, puntos en los cuales esta función diverge. Veremos ahora si existe una solución tipo  $F_2(q, P)$  que pueda cubrir los “huecos” dejados por la solución de  $F_1(q, Q)$ . Las ecuaciones diferenciales (7.19), se escriben

$$p = \frac{\partial F_2(q, P)}{\partial q} \quad ; \quad Q = \frac{\partial F_2(q, P)}{\partial P} \quad (7.89)$$

la primera requiere conocer  $p = p(q, P)$ , y este despeje es directo de la segunda de las ecuaciones (7.79)

$$p = \frac{P - q \sin \alpha}{\cos \alpha} \quad ; \quad p = \frac{\partial F_2}{\partial q} \quad (7.90)$$

la solución para  $F_2(q, P)$  es de la forma

$$F_2 = \frac{qP}{\cos \alpha} - \frac{q^2}{2} \tan \alpha + f(P) \quad (7.91)$$

para resolver la segunda ecuación (7.89) necesitamos  $Q = Q(q, P)$ , lo cual se obtiene insertando (7.90) en la primera de las ecuaciones (7.79)

$$Q = q \cos \alpha - (P - q \sin \alpha) \tan \alpha \quad ; \quad Q = \frac{\partial F_2}{\partial P}$$



la solución para  $F_2(q, P)$  en este caso es

$$\begin{aligned} F_2 &= qP \cos \alpha - \frac{P^2}{2} \tan \alpha + qP \frac{\sin^2 \alpha}{\cos \alpha} + g(q) \\ F_2 &= \frac{qP}{\cos \alpha} - \frac{P^2}{2} \tan \alpha + g(q) \end{aligned} \quad (7.92)$$

y comparando (7.91, 7.92), encontramos una solución para  $F_2(q, P)$

$$F_2(q, P) = -\frac{1}{2}(q^2 + P^2) \tan \alpha + \frac{qP}{\cos \alpha} \quad ; \quad \alpha \neq \left(n + \frac{1}{2}\right) \pi \quad (7.93)$$

esta función converge para  $\alpha = n\pi$ , de modo que llena los “huecos” dejados por la función  $F_1(q, Q)$  de la Ec. (7.88). La función (7.93) diverge para  $\alpha = (n + 1/2)\pi$ , pero estos huecos son llenados por  $F_1$ .

La interpretación de la transformación (7.79) como una rotación bidimensional nos hace intuir que  $\alpha = 0$  corresponde a la identidad, como se observa al hacer  $\alpha = 0$  en (7.79). Similarmente, haciendo  $\alpha = 0$  en (7.93) obtenemos  $F_2(q, P) = qP$ , en concordancia con el generador de la identidad dado por la Ec. (7.22).

Ahora aplicando  $\alpha = \pi/2$  en las ecuaciones (7.79) tenemos

$$Q = -p \quad ; \quad P = q$$

así mismo haciendo  $\alpha = \pi/2$  en (7.88), encontramos  $F_1 = qQ$ , y coincide con la función generadora de la transformación canónica de permutación o intercambio dada por la Ec. (7.28). Geométricamente, es claro que una rotación positiva de  $\pi/2$  convierte el eje  $Y$  en el eje  $X$  y el eje  $X$  en el eje  $Y$  negativo. Finalmente, las ecuaciones (7.88, 7.93) muestran que la transformación de intercambio no se puede generar con una función del tipo  $F_1$  y la identidad no se puede generar con una función tipo  $F_2$ , como ya se mencionó.

### 7.7.3. Un sistema con dos grados de libertad

Para un sistema con dos grados de libertad, determinaremos si la transformación

$$\begin{aligned} Q_1 &= q_1 q_2 & P_1 &= \frac{p_1 - p_2}{q_2 - q_1} + 1 \\ Q_2 &= q_1 + q_2 & P_2 &= \frac{q_2 p_2 - q_1 p_1}{q_2 - q_1} - (q_2 + q_1) \end{aligned}$$

es canónica. Para ello recurriremos a la condición simpléctica. La matriz jacobiana de transformación está dada por la Ec. (7.51)

$$\begin{aligned} M_{11} &= \frac{\partial Q_1}{\partial q_1} = q_2 \quad ; \quad M_{12} = \frac{\partial Q_1}{\partial q_2} = q_1 \quad ; \quad M_{13} = \frac{\partial Q_1}{\partial p_1} = 0 \quad ; \quad M_{14} = \frac{\partial Q_1}{\partial p_2} = 0 \\ M_{21} &= \frac{\partial Q_2}{\partial q_1} = 1 \quad ; \quad M_{22} = \frac{\partial Q_2}{\partial q_2} = 1 \quad ; \quad M_{23} = \frac{\partial Q_2}{\partial p_1} = 0 \quad ; \quad M_{24} = \frac{\partial Q_2}{\partial p_2} = 0 \\ M_{31} &= \frac{\partial P_1}{\partial q_1} = \frac{p_1 - p_2}{(q_2 - q_1)^2} \quad ; \quad M_{32} = \frac{\partial P_1}{\partial q_2} = -\frac{p_1 - p_2}{(q_2 - q_1)^2} \\ M_{33} &= \frac{\partial P_1}{\partial p_1} = \frac{1}{(q_2 - q_1)} \quad ; \quad M_{34} = \frac{\partial P_1}{\partial p_2} = -\frac{1}{(q_2 - q_1)} \end{aligned}$$

$$\begin{aligned}
M_{41} &= \frac{\partial P_2}{\partial q_1} = \frac{p_2 q_2 - p_1 q_1}{(q_2 - q_1)^2} - \frac{p_1}{q_2 - q_1} - 1 = \frac{p_2 q_2 - p_1 q_1 - (q_2 - q_1) p_1 - (q_2 - q_1)^2}{(q_2 - q_1)^2} \\
&= \frac{p_2 q_2 - p_1 q_2 + 2q_1 q_2 - q_1^2 - q_2^2}{(q_2 - q_1)^2} \\
M_{42} &= \frac{\partial P_2}{\partial q_2} = \frac{p_2}{q_2 - q_1} + \frac{p_1 q_1 - p_2 q_2}{(q_2 - q_1)^2} - 1 \\
&= \frac{(q_2 - q_1) p_2 + p_1 q_1 - p_2 q_2 - (q_2 - q_1)^2}{(q_2 - q_1)^2} = \frac{p_1 q_1 - p_2 q_1 + 2q_1 q_2 - q_1^2 - q_2^2}{(q_2 - q_1)^2} \\
M_{43} &= \frac{\partial P_2}{\partial p_1} = -\frac{q_1}{q_2 - q_1} \quad ; \quad M_{44} = \frac{\partial P_2}{\partial p_2} = \frac{q_2}{q_2 - q_1}
\end{aligned}$$

la matriz jacobina y su transpuesta quedan entonces

$$\begin{aligned}
\mathbf{M} &= \begin{pmatrix} q_2 & q_1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \frac{(p_1 - p_2)}{(q_2 - q_1)^2} & -\frac{(p_1 - p_2)}{(q_2 - q_1)^2} & \frac{1}{(q_2 - q_1)} & -\frac{1}{(q_2 - q_1)} \\ \frac{p_2 q_2 - p_1 q_2 + 2q_1 q_2 - q_1^2 - q_2^2}{(q_2 - q_1)^2} & \frac{p_1 q_1 - p_2 q_1 + 2q_1 q_2 - q_1^2 - q_2^2}{(q_2 - q_1)^2} & -\frac{q_1}{(q_2 - q_1)} & \frac{q_2}{(q_2 - q_1)} \end{pmatrix} \\
\widetilde{\mathbf{M}} &= \begin{pmatrix} q_2 & 1 & \frac{p_1 - p_2}{(q_1 - q_2)^2} & -\frac{(q_1^2 + q_2^2 + p_1 q_2 - p_2 q_2 - 2q_1 q_2)}{(q_1 - q_2)^2} \\ q_1 & 1 & -\frac{p_1 - p_2}{(q_1 - q_2)^2} & -\frac{(q_1^2 + q_2^2 - p_1 q_1 + p_2 q_1 - 2q_1 q_2)}{(q_1 - q_2)^2} \\ 0 & 0 & -\frac{1}{q_1 - q_2} & \frac{q_1}{q_1 - q_2} \\ 0 & 0 & \frac{1}{q_1 - q_2} & -\frac{q_2}{q_1 - q_2} \end{pmatrix}
\end{aligned}$$

y la matriz  $\mathbf{J}$  en cuatro dimensiones está dada por

$$\mathbf{J} = \begin{pmatrix} \mathbf{0}_{2 \times 2} & \mathbf{1}_{2 \times 2} \\ -\mathbf{1}_{2 \times 2} & \mathbf{0}_{2 \times 2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

ahora debemos realizar el producto  $\mathbf{M}\widetilde{\mathbf{M}}$  y verificar si coincide con  $\mathbf{J}$ . Con una dosis de paciencia y una buena taza de café, el lector puede comprobar que se cumple la condición simpléctica. Por tanto, esta transformación es canónica.

## 7.8. Ejercicios

1. Demostrar que una función del tipo  $F_3(p, Q, t)$  puede generar la transformación canónica identidad, y una función del tipo  $F_4(p, P, t)$  puede generar la transformación canónica de intercambio o permutación.
2. Sea  $q, p$  un conjunto canónico para un grado de libertad. Demuestre que el conjunto  $Q, P$  dado por

$$Q = \log(1 + \sqrt{q} \cos p) \quad ; \quad P = 2(1 + \sqrt{q} \cos p) \sqrt{q} \sin p \quad (7.94)$$

es también un conjunto canónico y que una función generadora para esta transformación está dada por

$$F_3(p, Q) = -(e^Q - 1)^2 \tan p$$

3. Para un sistema con dos grados de libertad, sean  $\{q_i, p_i\}$  variables canónicas, y una transformación puntual de la forma

$$Q_1 = q_1^2 \quad , \quad Q_2 = q_1 + q_2$$

encuentre la transformación más general para  $P_1$  y  $P_2$  que genere una transformación canónica. Ahora sea el Hamiltoniano

$$H = \left( \frac{p_1 - p_2}{2q_1} \right)^2 + p_2 + (q_1 + q_2)^2$$

Encuentre una transformación particular para  $P_1$  y  $P_2$  de tal forma que  $Q_1$  y  $Q_2$  sean ambas cíclicas. Resuelva las ecuaciones de Hamilton y obtenga la solución para  $q_1, q_2, p_1$  y  $p_2$  en función del tiempo, y en términos de las condiciones iniciales.

4. La Ec. (2.49) define una transformación gauge para los campos electromagnéticos. Dicha transformación cambia al Hamiltoniano (6.26) y al momento canónico (6.27). (a) Muestre que este cambio se puede ver como una transformación canónica en donde  $q$  permanece inalterada. (b) Encuentre una función generadora del tipo 2, que genera dicha transformación canónica. Una transformación canónica para  $n$  grados de libertad que deja invariantes los  $q_i$  se denomina una transformación canónica gauge.

5. Demuestre que la transformación

$$\begin{aligned} Q_1 &= q_1 \cos \alpha - \frac{p_2}{\beta} \sin \alpha & ; & & Q_2 &= q_2 \cos \alpha - \frac{p_1}{\beta} \sin \alpha \\ P_1 &= \beta q_2 \sin \alpha + p_1 \cos \alpha & ; & & P_2 &= \beta q_1 \sin \alpha + p_2 \cos \alpha \end{aligned} \quad (7.95)$$

es canónica.

6. Una partícula de masa  $m$  y carga  $q$  se mueve en un plano bajo la influencia combinada de un potencial armónico  $(m/2)\omega_0^2 [q_1^2 + q_2^2]$  y un campo magnético descrito por el potencial vectorial  $\mathbf{A} = (0, hq_1, 0)$ , siendo  $h$  una constante. Use la transformación canónica (7.95) con  $\tan 2\alpha = m\omega_0 (qh)^{-1}$  y escoja  $\beta$  adecuadamente, para encontrar el movimiento general así como el movimiento en los límites  $h = 0$  y  $h \rightarrow \infty$ .

7. (a) Demuestre que la transformación

$$Q = p + iaq \quad , \quad P = \frac{p - iaq}{2ia}$$

es canónica. (b) Encuentre una función generadora. (c) Use esta transformación para resolver el oscilador armónico lineal.

8. Utilizando el procedimiento de la sección 7.2, complete la tabla 7.1 de la Pág. 124, para las funciones del tipo  $F_3(p, Q, t)$  y  $F_4(p, P, t)$ .

## Corchetes de Poisson y otros invariantes canónicos

Hemos visto que las transformaciones canónicas son aquellas que preservan la forma de las ecuaciones de Hamilton. Esto nos motiva a buscar estructuras que sean invariantes canónicas, ya que si una cierta cantidad preserva su forma ante transformaciones canónicas, entonces es posible que dicha cantidad pueda enlazarse fácilmente con las ecuaciones de movimiento en cualquier base canónica. Por otro lado, los invariantes canónicos tienen el potencial de expresar ecuaciones de movimiento equivalentes a las de Hamilton. Teniendo esto en mente estudiaremos un invariante canónico muy importante: los corchetes de Poisson. Adicionalmente, veremos como las ecuaciones de movimiento y las cantidades conservadas se pueden escribir en el lenguaje de los corchetes de Poisson. Haremos además una breve mención de otros invariantes canónicos.

### 8.1. Corchetes de Poisson

Los corchetes de Poisson de dos funciones  $u, v$  con respecto a las variables canónicas  $q, p$  se definen por

$$[u, v]_{q,p} \equiv \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \quad (8.1)$$

vemos que esta forma bilineal tiene la misma estructura simpléctica de las ecuaciones de Hamilton en donde  $q$  se acopla con  $p$ , y  $p$  con  $-q$ . Por esta razón, esta definición se puede escribir en una estructura matricial también simpléctica

$$[u, v]_{\boldsymbol{\eta}} = \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \mathbf{J} \frac{\partial v}{\partial \boldsymbol{\eta}} \quad ; \quad \eta_i \equiv q_i, \quad \eta_{i+n} \equiv p_i \quad \text{with} \quad i \leq n \quad (8.2)$$

en particular podemos tomar como funciones  $u, v$  elementos del conjunto de variables canónicas  $(q, p)$ . A partir de la definición (8.1) ó (8.2) estos corchetes tienen los valores

$$\begin{aligned} [q_j, q_k]_{q,p} &= 0 = [p_j, p_k]_{q,p} \\ [q_j, p_k]_{q,p} &= \delta_{jk} = -[p_k, q_j]_{q,p} \end{aligned} \quad (8.3)$$

en notación matricial

$$[\eta_i, \eta_j]_{q,p} \equiv \left( [\eta, \eta]_{q,p} \right)_{ij} \equiv [\eta, \eta]_{\boldsymbol{\eta}}$$

todas estas propiedades se pueden escribir en una sola ecuación

$$[\boldsymbol{\eta}, \boldsymbol{\eta}]_{\boldsymbol{\eta}} = \mathbf{J} \quad (8.4)$$

otro conjunto de funciones  $u, v$  interesantes serían los elementos del conjunto de nuevas variables canónicas  $(Q, P)$  ó  $\zeta$ , con respecto a las antiguas variables  $(q, p)$  ó  $\boldsymbol{\eta}$ . Escribiendo estos corchetes en notación matricial usando (8.2), resulta

$$[\zeta, \zeta]_{\boldsymbol{\eta}} = \left( \widetilde{\frac{\partial \zeta}{\partial \boldsymbol{\eta}}} \right) \mathbf{J} \frac{\partial \zeta}{\partial \boldsymbol{\eta}}$$

pero las derivadas parciales definen precisamente el Jacobiano de la transformación, con lo cual tenemos

$$[\zeta, \zeta]_{\eta} = \widetilde{\mathbf{M}}\mathbf{J}\mathbf{M} \quad (8.5)$$

y dado que la transformación  $\eta \rightarrow \zeta$  es canónica, la condición simpléctica (7.58) nos lleva a escribir

$$[\zeta, \zeta]_{\eta} = \mathbf{J} \quad (8.6)$$

se pueden revertir los pasos para mostrar que si (8.6) es válida, la transformación es canónica.

Los corchetes de Poisson de variables canónicas en sí tales como (8.4, 8.6), se denominan **corchetes fundamentales de Poisson**. Las Ecs. (8.4) escritas en las nuevas variables canónicas, dan

$$[\zeta, \zeta]_{\zeta} = \mathbf{J} \quad (8.7)$$

y comparando (8.6) con (8.7) vemos que ambos tienen el mismo valor cuando se evalúan con respecto a cualquier conjunto de coordenadas canónicas. En otras palabras, los corchetes fundamentales de Poisson son invariantes bajo transformaciones canónicas. La ecuación (8.5), muestra que la invarianza es una condición necesaria y suficiente para que la matriz de transformación sea simpléctica. La invarianza de los corchetes fundamentales de Poisson ante el cambio de base canónica, es entonces equivalente en todos los sentidos a la condición simpléctica.

Veremos ahora que todos los corchetes de Poisson son invariantes ante transformaciones canónicas. Retomemos el corchete de Poisson para funciones arbitrarias  $u, v$  Ec. (8.2), la derivada de  $v$  con respecto a  $\eta$  se puede escribir en analogía con la Ec. (7.54)

$$\begin{aligned} \frac{\partial v}{\partial \eta_k} &= \frac{\partial v}{\partial \zeta_i} \frac{\partial \zeta_i}{\partial \eta_k} = \frac{\partial v}{\partial \zeta_i} M_{ik} = \widetilde{M}_{ki} \frac{\partial v}{\partial \zeta_i} \\ \frac{\partial v}{\partial \eta} &= \widetilde{\mathbf{M}} \frac{\partial v}{\partial \zeta} \end{aligned}$$

similarmente

$$\widetilde{\left( \frac{\partial u}{\partial \eta} \right)} = \widetilde{\left( \widetilde{\mathbf{M}} \frac{\partial u}{\partial \zeta} \right)} = \left( \frac{\partial u}{\partial \zeta} \right) \mathbf{M}$$

reemplazando estas expresiones en (8.2)

$$[u, v]_{\eta} = \left( \frac{\partial u}{\partial \eta} \right) \mathbf{J} \frac{\partial v}{\partial \eta} = \left( \frac{\partial u}{\partial \zeta} \right) \mathbf{M} \widetilde{\mathbf{J}} \widetilde{\mathbf{M}} \frac{\partial v}{\partial \zeta}$$

si la transformación  $\eta \rightarrow \zeta$  es canónica, la condición simpléctica (7.56) nos lleva a

$$[u, v]_{\eta} = \left( \frac{\partial u}{\partial \zeta} \right) \mathbf{J} \frac{\partial v}{\partial \zeta} \equiv [u, v]_{\zeta} \quad (8.8)$$

por tanto **todos los corchetes de Poisson son invariantes canónicos**. Este resultado nos permite omitir el subíndice que denota el conjunto de variables que se usan para evaluar el corchete, siempre y cuando las variables que se usen sean canónicas. Vale la pena anotar sin embargo, que las transformaciones de escala, así como las transformaciones canónicas extendidas, en donde la condición simpléctica toma la forma (7.59), **NO** dejan invariantes a los corchetes de Poisson.

Recordemos que la definición original de transformaciones canónicas surge de la necesidad de conservar la forma de las ecuaciones de Hamilton ante un cambio de coordenadas generalizadas y momentos conjugados. Esto sugiere que, dado que los corchetes de Poisson son invariantes canónicos, podemos construir ecuaciones de movimiento en términos de corchetes de Poisson que sean invariantes en forma ante transformaciones canónicas. Desarrollaremos entonces un formalismo paralelo al formalismo de Hamilton basado en los corchetes de Poisson.

## 8.2. Propiedades de los corchetes de Poisson

Los corchetes de Poisson poseen una serie de propiedades algebraicas de fundamental importancia, las cuales resumimos en la siguiente forma

$$[u, u] = 0 \quad ; \quad [u, v] = -[v, u] \quad (\text{antisimetría}) \quad (8.9)$$

$$[au + bv, w] = a[u, w] + b[v, w] \quad (\text{linealidad}) \quad (8.10)$$

$$[uv, w] = [u, w]v + u[v, w] \quad (8.11)$$

$$0 = [u, [v, w]] + [v, [w, u]] + [w, [u, v]] \quad (\text{no asociatividad}) \quad (8.12)$$

donde  $u, v, w$  son funciones arbitrarias de las variables canónicas y del tiempo en tanto que  $a$  y  $b$  son constantes. Todas estas propiedades son directas a partir de la definición (8.1), excepto quizás (8.12) conocida como la identidad de Jacobi. A manera de ejemplo demostraremos la propiedad (8.11) directamente con la definición

$$\begin{aligned} [uv, w] &= \frac{\partial(uv)}{\partial q_i} \frac{\partial w}{\partial p_i} - \frac{\partial(uv)}{\partial p_i} \frac{\partial w}{\partial q_i} = \left[ u \frac{\partial v}{\partial q_i} + \frac{\partial u}{\partial q_i} v \right] \frac{\partial w}{\partial p_i} - \left[ u \frac{\partial v}{\partial p_i} + \frac{\partial u}{\partial p_i} v \right] \frac{\partial w}{\partial q_i} \\ [uv, w] &= \left[ \frac{\partial u}{\partial q_i} \frac{\partial w}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial w}{\partial q_i} \right] v + u \left[ \frac{\partial v}{\partial q_i} \frac{\partial w}{\partial p_i} - \frac{\partial v}{\partial p_i} \frac{\partial w}{\partial q_i} \right] \\ [uv, w] &= [u, w]v + u[v, w] \end{aligned}$$

Veamos ahora la demostración de la identidad de Jacobi (8.12). Esta propiedad muestra que la suma de corchetes de Poisson dobles construídos con permutaciones cíclicas de tres funciones es cero. Estrictamente, esta propiedad se puede demostrar con un reemplazo explícito a partir de la definición, sin embargo, un cambio de nomenclatura puede abreviar notablemente la demostración, las derivadas parciales las escribiremos como

$$u_i \equiv \frac{\partial u}{\partial \eta_i} \quad ; \quad v_{ij} \equiv \frac{\partial^2 v}{\partial \eta_i \partial \eta_j} = v_{ji} \quad (8.13)$$

con esta notación el corchete de Poisson (8.2), se escribe

$$[u, v] = u_i J_{ij} v_j$$

esta notación resulta particularmente práctica para expresar los corchetes dobles de Poisson. Escribamos el primer corchete doble de (8.12) en esta notación

$$[u, [v, w]] = u_i J_{ij} [v, w]_j = u_i J_{ij} (v_k J_{kl} w_l)_j$$

ahora calculamos la derivada parcial con respecto a  $\eta_j$  usando regla del producto, teniendo en cuenta que las  $J_{kl}$  son constantes,

$$[u, [v, w]] = u_i J_{ij} (v_k J_{kl} w_{lj} + v_{kj} J_{kl} w_l)$$

y el mismo procedimiento se realiza para los otros dos corchetes dobles en (8.12), con lo cual se obtiene otros cuatro términos, para un total de 6 términos. Los otros términos se pueden obtener por permutación circular de  $u, v, w$

$$\begin{aligned} [w, [u, v]] &= w_i J_{ij} (u_k J_{kl} v_{lj} + u_{kj} J_{kl} v_l) \\ [v, [w, u]] &= v_i J_{ij} (w_k J_{kl} u_{lj} + w_{kj} J_{kl} u_l) \end{aligned}$$

Todos los 6 términos tienen una segunda derivada en alguna de las funciones  $u, v, w$ . Tomemos los dos términos que contienen segunda derivada en  $w$

$$W_1 \equiv u_i J_{ij} v_k J_{kl} w_{lj} \quad ; \quad W_2 \equiv v_i J_{ij} w_{kj} J_{kl} u_l$$

el primero proviene de  $[u, [v, w]]$  y el segundo de  $[v, [w, u]]$ . Reescribamos estos términos teniendo en cuenta que las segundas derivadas parciales se pueden intercambiar y que la matriz  $\mathbf{J}$  es antisimétrica

$$\begin{aligned} W_1 &\equiv u_i J_{ij} v_k J_{kl} w_{lj} = v_k J_{kl} w_{lj} J_{ij} u_i = -v_k J_{kl} w_{lj} J_{ji} u_i \\ W_2 &= v_i J_{ij} w_{kj} J_{kl} u_l = v_i J_{ij} w_{jk} J_{kl} u_l = -W_1 \end{aligned}$$

donde hemos tenido en cuenta que todos los índices se suman y por tanto son mudos. Es claro entonces que  $W_1 + W_2 = 0$ , de modo que los dos pares de términos que contienen segundas derivadas en  $w$  se anulan. Por argumentos idénticos, los otros dos pares que contienen segundas derivadas en  $u$  y  $v$  se anulan, con lo cual se demuestra la identidad de Jacobi Ec. (8.12). Nótese que la única propiedad que se usó de  $\mathbf{J}$  fué su antisimetría.

### Corchetes de Poisson como estructuras algebraicas

De aquí en adelante abandonamos la notación definida por (8.13). Si imaginamos al corchete de Poisson como una ley de combinación u operación “producto” entre los elementos  $u, v$ , la identidad de Jacobi nos dice que este producto o ley de combinación no es asociativo, y nos dice cual es el efecto cuando cambia la secuencia de “multiplicaciones”. Con frecuencia ocurre que tenemos una ley de combinación o producto entre elementos  $u_i$  y  $u_j$  que satisface las propiedades (8.9-8.12) además de la propiedad

$$[u_i, u_j] = c_{ij}^k u_k \quad (8.14)$$

siendo  $c_{ij}^k$  cantidades denominadas constantes de estructura<sup>1</sup>. En este caso, el conjunto de todos los elementos  $u, v$  junto con la operación producto, forman un álgebra no conmutativa conocida como álgebra de Lie. Nótese que en este contexto la notación  $[u_i, u_j]$  se refiere a una ley de combinación cualquiera y no a los corchetes de Poisson. En el espacio tridimensional, los corchetes de Poisson cumplen la propiedad (8.14) de una manera bien particular. O bien, todas las constantes de estructura son cero, o bien solo hay un término en el lado derecho de la ecuación (8.14), para cada par de índices  $i, j$ .

Un comentario final sobre álgebras de Lie, nombraremos dos conjuntos de álgebras de Lie particularmente útiles en Física, la primera es el álgebra definida sobre los vectores Euclidianos en  $\mathbb{R}^3$  con la ley de combinación

$${}_V [\mathbf{A}, \mathbf{B}] \equiv \mathbf{A} \times \mathbf{B}$$

la segunda es el álgebra definida sobre el conjunto de matrices  $n \times n$  con la operación producto definida por el conmutador entre ellas

$${}_M [\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA}$$

lo interesante es que muchos resultados solo dependen de la estructura de álgebra y no de la forma explícita de la ley de combinación. Profundizar en estos tópicos va más allá de los propósitos de este texto, el lector interesado puede consultar las Refs. [9, 10, 11].

## 8.3. Corchetes de Lagrange

Sean  $u, v$  dos funciones pertenecientes a un conjunto de  $2n$  funciones de las variables canónicas, y que son independientes entre sí. Invertiendo las ecuaciones se pueden escribir las variables canónicas en función de este conjunto de  $2n$  funciones. El corchete de Lagrange de  $u$  y  $v$  con respecto a las variables  $(q, p)$  se define como

$$\begin{aligned} \{u, v\}_{q,p} &= \frac{\partial q_i}{\partial u} \frac{\partial p_i}{\partial v} - \frac{\partial p_i}{\partial u} \frac{\partial q_i}{\partial v} \\ \{u, v\}_{\boldsymbol{\eta}} &= \left( \frac{\partial \boldsymbol{\eta}}{\partial u} \right) \mathbf{J} \frac{\partial \boldsymbol{\eta}}{\partial v} \end{aligned}$$

<sup>1</sup>Nuevamente insistimos en que los índices  $i, j, k$  en la Ec. (8.14) son rótulos para un conjunto dado de elementos, y no denota derivadas como en la Ec. (8.13).

la prueba de su invarianza canónica es muy similar a la de los corchetes de Poisson. Hay una especie de relación inversa entre los corchetes de Poisson y Lagrange, que se manifiesta en la siguiente propiedad: Sea  $u_i$  un conjunto de  $2n$  funciones de las variables canónicas e independientes entre sí. Con este conjunto se puede formar el vector  $\mathbf{u}$  de dimensión  $2n$ , y la matriz  $2n \times 2n$  definida por  $\{\mathbf{u}, \mathbf{u}\}_{ij} \equiv \{u_i, u_j\}$ . Similarmente se define la matriz  $[\mathbf{u}, \mathbf{u}]$  para corchetes de Poisson, se deja como ejercicio verificar que

$$\{\mathbf{u}, \mathbf{u}\} [\mathbf{u}, \mathbf{u}] = -1 \quad (8.15)$$

sin embargo, los corchetes de Lagrange no son de tan amplio uso como los de Poisson. La razón estriba en el hecho de que esta nueva ley de combinación posee las mismas propiedades fundamentales de los corchetes de Poisson, pero con una importante excepción: los corchetes de Lagrange **no** obedecen la identidad de Jacobi y por tanto no forman un álgebra de Lie.

## 8.4. Otros invariantes canónicos

Un invariante canónico muy importante es la magnitud de un elemento de volumen en el espacio de fase. Una transformación canónica  $\boldsymbol{\eta} \rightarrow \boldsymbol{\zeta}$  transforma el espacio de fase  $2n$  dimensional con coordenadas  $\eta_i$  en otro espacio de fase con coordenadas  $\zeta_i$ . El elemento de volumen

$$(d\eta) = dq_1 \dots dq_n dp_1 \dots dp_n$$

se transforma a un nuevo elemento de volumen  $(d\zeta)$

$$(d\zeta) = dQ_1 \dots dQ_n dP_1 \dots dP_n$$

como es bien sabido, las magnitudes de los dos elementos de volumen están enlazadas a través del valor absoluto del determinante de la matriz jacobiana

$$(d\zeta) = \|\mathbf{M}\| (d\eta)$$

Pero tomando el determinante en la condición simpléctica (7.56), se obtiene

$$\|\mathbf{M}\mathbf{J}\widetilde{\mathbf{M}}\| = \|\mathbf{J}\| \Rightarrow \|\mathbf{M}\|^2 \|\mathbf{J}\| = \|\mathbf{J}\| \Rightarrow \|\mathbf{M}\| = 1$$

de modo que el valor absoluto del determinante de  $\mathbf{M}$  es uno (si la transformación canónica es real, el determinante es  $\pm 1$ ). Esto muestra entonces que el elemento diferencial de volumen en el espacio de fase es un invariante canónico. Como corolario resulta que el volumen de una región arbitraria en el espacio de fase

$$J_n = \int \dots \int (d\eta)$$

es un invariante canónico. Esta integral de volumen es el miembro final de una secuencia de invariantes canónicos que constan de integrales sobre subespacios del espacio de fase de diferentes dimensiones. Tal secuencia es conocida como **integrales invariantes de Poincaré**. Una vez más, remitimos al lector interesado en profundizar en estos temas a las Refs. [9, 10, 11].

## 8.5. Relación entre la condición simpléctica y las funciones generatrices (opcional)

La relación entre los corchetes de Poisson y el formalismo de las funciones generatrices se puede sintetizar en el teorema de Carathéodory, el cual enuncia que la condición simpléctica implica la existencia de funciones generatrices.

Por simplicidad examinaremos un sistema con espacio de fase dos dimensional. La extensión a espacios de fase de dimensión arbitraria es directa. Comenzamos con el conjunto de transformaciones de la forma

$$Q = Q(q, p) \quad ; \quad P = P(q, p) \quad (8.16)$$



si suponemos que la primera de estas transformaciones es invertible de modo que  $p$  se puede escribir como función de  $q, Q$  entonces

$$p \equiv \phi(q, Q) \quad (8.17)$$

y si a su vez sustituimos esta expresión en la segunda ecuación de transformación, obtenemos  $P$  también en función de  $q, Q$

$$P = P(q, \phi(q, Q)) \equiv \psi(q, Q) \quad (8.18)$$

queremos ahora enlazar estas transformaciones con las funciones generatrices. Dado que las Ecs. (8.17, 8.18) tienen como argumentos a  $q, Q$  es natural que estas transformaciones se construyan con funciones generatrices de tipo 1. En algunos casos, cuando la primera de las Ecs. (8.16) no es invertible (como en la transformación identidad) podemos comenzar invirtiendo la segunda de las Ecs. (8.16) y reemplazando en la primera, con lo cual llegamos a funciones generatrices del tipo 2.

Volviendo a nuestro caso, asumiendo que existen funciones generatrices del tipo 1 que pueden generar a las Ecs. (8.17, 8.18), éstas se escriben según la prescripción dada por (7.16)

$$p = \frac{\partial F_1(q, Q)}{\partial q} ; \quad P = -\frac{\partial F_1(q, Q)}{\partial Q} \quad (8.19)$$

si las ecuaciones (8.19) se cumplen (es decir, si es consistente nuestra suposición de que existe una función generatriz del tipo 1) entonces se debe cumplir que

$$\begin{aligned} \frac{\partial}{\partial Q} \left( \frac{\partial F_1}{\partial q} \right) &= \frac{\partial}{\partial q} \left( \frac{\partial F_1}{\partial Q} \right) \Rightarrow \frac{\partial}{\partial Q} [p(q, Q)] = \frac{\partial}{\partial q} [-P(q, Q)] \Rightarrow \\ \frac{\partial \phi}{\partial Q} &= -\frac{\partial \psi}{\partial q} \end{aligned} \quad (8.20)$$

donde hemos usado (8.19), (8.17), y (8.18). Recíprocamente, si se cumple (8.20) debe existir una función de tipo 1 que satisfaga (8.19).

Para demostrar la validez de (8.20), escribiremos todas las cantidades en términos de  $q, Q$ . Comenzamos con la identidad

$$\frac{\partial Q}{\partial Q} = 1$$

sustituyendo la Ec. (8.17) en la primera de las Ecs. (8.16)

$$Q = Q(q, \phi(q, Q))$$

la derivada parcial anterior se puede escribir como

$$\frac{\partial Q}{\partial Q} = \frac{\partial Q}{\partial p} \frac{\partial \phi}{\partial Q} = 1 \quad (8.21)$$

ahora evaluamos el corchete de Poisson

$$[Q, P] \equiv \frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial P}{\partial q} \frac{\partial Q}{\partial p} = 1 \quad (8.22)$$

donde hemos usado las propiedades (8.3).

Las derivadas de  $P$  son derivadas de  $\psi$  (ver Ec. 8.18) consideradas como funciones de  $q$  y  $Q(q, p)$  con lo cual se tiene

$$\frac{\partial P}{\partial q} = \frac{\partial \psi(q, Q(q, p))}{\partial q} = \frac{\partial \psi}{\partial q} + \frac{\partial \psi}{\partial Q} \frac{\partial Q}{\partial q}$$

Por lo tanto, el corchete de Poisson (8.22) se escribe

$$[Q, P] = \frac{\partial Q}{\partial q} \frac{\partial \psi}{\partial Q} \frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p} \left( \frac{\partial \psi}{\partial q} + \frac{\partial \psi}{\partial Q} \frac{\partial Q}{\partial q} \right) = 1$$

reorganizando términos

$$\begin{aligned} [Q, P] &= \frac{\partial \psi}{\partial Q} \left( \frac{\partial Q}{\partial q} \frac{\partial Q}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial Q}{\partial q} \right) - \frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q} = 1 \\ [Q, P] &= -\frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q} = 1 \end{aligned} \quad (8.23)$$

Combinando (8.23) con (8.21) resulta

$$\frac{\partial Q}{\partial p} \frac{\partial \phi}{\partial Q} = -\frac{\partial Q}{\partial p} \frac{\partial \psi}{\partial q} \quad (8.24)$$

dado que la derivada parcial  $\frac{\partial Q}{\partial p}$  está evaluada a ambos lados para los mismos argumentos (variando  $Q$  y manteniendo fijo  $q$ ), se tiene que su valor a ambos lados de la igualdad es el mismo. Por otro lado esta derivada es diferente de cero si asumimos que la primera de las Ecs. (8.16) es invertible. Por tanto la ecuación (8.24) nos lleva a la Ec. (8.20) que nos lleva a la existencia de una función generatriz para la transformación (8.16). Para llegar a (8.20) hemos partido de la Ec. (8.22), que nos dice que  $[Q, P] = 1$ , pero recordemos que esta propiedad es equivalente a la condición simpléctica. Por tanto, la condición simpléctica nos lleva a la existencia de una función generatriz y viceversa. Los dos formalismos son entonces equivalentes.

## 8.6. Ecuaciones de movimiento con corchetes de Poisson

Dado que los corchetes de Poisson son invariantes canónicos, las ecuaciones de movimiento que involucren a estas cantidades van a conservar su forma ante transformaciones canónicas. Por ejemplo, veamos la ecuación de movimiento asociada a una función  $u$  de las variables canónicas  $q, p, t$ . Usando las ecuaciones de Hamilton, la derivada temporal total de esta función se puede escribir como

$$\frac{du}{dt} = \frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial u}{\partial t}$$

lo cual se escribe como

$$\frac{du}{dt} = [u, H] + \frac{\partial u}{\partial t} \quad (8.25)$$

la derivación de (8.25), también se puede realizar en notación simpléctica, la regla de la cadena la escribimos como

$$\frac{du}{dt} = \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \dot{\boldsymbol{\eta}} + \frac{\partial u}{\partial t}$$

y usando las Ecs. de Hamilton y la definición de los corchetes de Poisson en forma simpléctica Ecs. (6.33) y (8.2) se tiene

$$\frac{du}{dt} = \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \mathbf{J} \frac{\partial H}{\partial \boldsymbol{\eta}} + \frac{\partial u}{\partial t} = [u, H] + \frac{\partial u}{\partial t}$$

La expresión (8.25) se puede ver como una ecuación de movimiento generalizada para una función arbitraria  $u(q, p, t)$  en la formulación de los corchetes de Poisson. Una escogencia muy natural es  $u = q_i$  ó  $u = p_i$  con lo cual resulta

$$\begin{aligned} \dot{q}_i &= [q_i, H] \quad ; \quad \dot{p}_i = [p_i, H] \\ \dot{\boldsymbol{\eta}} &= [\boldsymbol{\eta}, H] \end{aligned} \quad (8.26)$$

nótese que estas ecuaciones dependen de que  $\partial q / \partial t = 0$ . Esto se puede ver del hecho de que  $\partial u(q, p, t) / \partial t$  consiste en evaluar la razón de cambio de  $u$  manteniendo fijo  $q$  y  $p$  moviendo solo el tiempo, y esto es válido en particular cuando  $u = q$ . Se observa que esto es diferente al caso en el cual se toma  $q = q(t)$  en cuyo caso

$\partial q/\partial t = dq/dt = \dot{q}$ . Argumento similar se sigue para ver que  $\partial p/\partial t = 0$ . Usando la definición simpléctica de los corchetes de Poisson (8.2) queda

$$[\boldsymbol{\eta}, u] = \left( \widetilde{\frac{\partial \eta}{\partial \boldsymbol{\eta}}} \right) \mathbf{J} \frac{\partial u}{\partial \boldsymbol{\eta}} = \mathbf{I} \mathbf{J} \frac{\partial u}{\partial \boldsymbol{\eta}} = \mathbf{J} \frac{\partial u}{\partial \boldsymbol{\eta}} \quad (8.27)$$

tomando  $u = H$  y reemplazando en (8.26), resulta

$$\dot{\boldsymbol{\eta}} = \mathbf{J} \frac{\partial H}{\partial \boldsymbol{\eta}}$$

que son las ecuaciones de Hamilton en forma simpléctica. Vemos entonces que tomando la expresión general (8.25), con  $u = q, p$  resultan las ecuaciones de movimiento para  $q, p$  que coinciden con las ecuaciones de Hamilton como era de esperarse. Otra propiedad familiar resulta cuando se toma  $u = H$  en (8.25), lo cual nos arroja

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}$$

que coincide con (6.38). Nótese que la ecuación de movimiento (8.25) es invariante en forma ante una transformación canónica. La ecuación es válida cualquiera que sea el conjunto de coordenadas canónicas que se use para expresar a  $u$  y para evaluar el corchete de Poisson. No obstante, debe tenerse en cuenta que el Hamiltoniano que se use debe ser apropiado para el conjunto de variables canónicas elegido, cuando se pasa a otro conjunto de variables canónica debemos cambiar al Hamiltoniano transformado o Kamiltoniano.

Un comentario final, la ecuación de movimiento (8.25) solo será válida si  $u$  solo es función explícita de las coordenadas  $q, p$  del sistema y del tiempo, es decir  $u$  no puede ser función explícita de una variable externa (por ejemplo no podría ser función explícita de un campo eléctrico externo) toda la dependencia de  $u$  con respecto al exterior debe estar en el parámetro tiempo.

## 8.7. Constantes de movimiento con corchetes de Poisson

Otro caso de gran interés surge cuando  $u$  es una constante de movimiento i.e.  $du/dt = 0$ , con lo cual (8.25) se reduce a

$$[H, u] = \frac{\partial u}{\partial t} \quad (8.28)$$

e inversamente, todas las funciones que obedezcan la ecuación de movimiento (8.28), son constantes de movimiento. Esta ecuación nos provee de un método sistemático para encontrar constantes de movimiento. Adicionalmente, si la constante de movimiento no involucra al tiempo explícitamente, la condición sobre  $u$  se reduce a

$$[H, u] = 0 \quad (8.29)$$

Si conocemos dos constantes de movimiento, la identidad de Jacobi nos da la posibilidad de obtener más constantes de movimiento. Si  $u, v$  son constantes de movimiento que no dependen explícitamente del tiempo, podemos usar la identidad de Jacobi (8.12) para escribir

$$[u, [v, H]] + [v, [H, u]] + [H, [u, v]] = 0 \quad (8.30)$$

los dos primeros miembros de la izquierda son cero y queda

$$[H, [u, v]] = 0$$

lo cual nos dice que  $[u, v]$  es una constante de movimiento. Cuando  $u, v$  dependen explícitamente del tiempo, el corchete de Poisson entre ellos sigue siendo constante de movimiento, aunque la demostración es un tanto

más elaborada. Al ser  $u$  y  $v$  constantes de movimiento se cumple para ambos la Ec. (8.28), al reemplazar dicha ecuación en la identidad de Jacobi (8.30) se obtiene

$$\begin{aligned} \left[ u, \left( -\frac{\partial v}{\partial t} \right) \right] + \left[ v, \left( \frac{\partial u}{\partial t} \right) \right] + [H, [u, v]] &= 0 \\ \left[ \left( \frac{\partial v}{\partial t} \right), u \right] + \left[ v, \left( \frac{\partial u}{\partial t} \right) \right] &= -[H, [u, v]] \end{aligned}$$

usando la definición de los corchetes de Poisson obtenemos

$$\begin{aligned} \left\{ \left[ \frac{\partial}{\partial q_i} \left( \frac{\partial v}{\partial t} \right) \right] \frac{\partial u}{\partial p_i} - \left[ \frac{\partial}{\partial p_i} \left( \frac{\partial v}{\partial t} \right) \right] \frac{\partial u}{\partial q_i} \right\} + \left\{ \frac{\partial v}{\partial q_i} \left[ \frac{\partial}{\partial p_i} \left( \frac{\partial u}{\partial t} \right) \right] - \frac{\partial v}{\partial p_i} \left[ \frac{\partial}{\partial q_i} \left( \frac{\partial u}{\partial t} \right) \right] \right\} &= [H, [v, u]] \\ \left[ \frac{\partial}{\partial t} \left( \frac{\partial v}{\partial q_i} \right) \right] \frac{\partial u}{\partial p_i} - \left[ \frac{\partial}{\partial t} \left( \frac{\partial v}{\partial p_i} \right) \right] \frac{\partial u}{\partial q_i} + \frac{\partial v}{\partial q_i} \left[ \frac{\partial}{\partial t} \left( \frac{\partial u}{\partial p_i} \right) \right] - \frac{\partial v}{\partial p_i} \left[ \frac{\partial}{\partial t} \left( \frac{\partial u}{\partial q_i} \right) \right] &= [H, [v, u]] \\ \left[ \frac{\partial}{\partial t} \left( \frac{\partial v}{\partial q_i} \right) \right] \frac{\partial u}{\partial p_i} + \frac{\partial v}{\partial q_i} \left[ \frac{\partial}{\partial t} \left( \frac{\partial u}{\partial p_i} \right) \right] - \frac{\partial v}{\partial p_i} \left[ \frac{\partial}{\partial t} \left( \frac{\partial u}{\partial q_i} \right) \right] - \left[ \frac{\partial}{\partial t} \left( \frac{\partial v}{\partial p_i} \right) \right] \frac{\partial u}{\partial q_i} &= [H, [v, u]] \end{aligned}$$

reagrupando términos

$$\begin{aligned} \frac{\partial}{\partial t} \left[ \frac{\partial v}{\partial q_i} \frac{\partial u}{\partial p_i} \right] - \frac{\partial}{\partial t} \left[ \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} \right] &= [H, [v, u]] \\ \frac{\partial}{\partial t} \left[ \frac{\partial v}{\partial q_i} \frac{\partial u}{\partial p_i} - \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} \right] &= [H, [v, u]] \\ \frac{\partial}{\partial t} [v, u] &= [H, [v, u]] \end{aligned}$$

al comparar con (8.28) se obtiene que  $[v, u] = -[u, v]$  es una constante de movimiento. Llegamos entonces al

**Theorem 5** *Teorema de Poisson: El corchete de Poisson de dos constantes de movimiento es también una constante de movimiento.*

En consecuencia, la aplicación reiterada del teorema de Poisson nos puede proveer de un conjunto de constantes de movimiento. Desafortunadamente, este algoritmo nos lleva con frecuencia a constantes de movimiento triviales o claramente dependientes de las anteriores. Sin embargo, el método debe ser tenido en cuenta para múltiples aplicaciones.

Vale la pena mencionar que la condición (8.28) es necesaria y suficiente para que una cierta función de las variables  $q, p$  del sistema y el tiempo sea constante de movimiento. Sin embargo, tal expresión no nos da un algoritmo para encontrar constantes de movimiento, más bien es un método para chequear si una cantidad dada es o no es una constante de movimiento del sistema. Esta forma de evaluación posee no obstante la ventaja de que podemos verificar si una cantidad es o no es constante de movimiento sin resolver completamente la evolución del sistema. Finalmente, enfatizamos de nuevo que para que esta formulación sea válida,  $u$  no puede ser función explícita de ninguna variable externa al sistema.

## 8.8. Ejemplos de constantes de movimiento evaluadas por corchetes de Poisson

### 8.8.1. Sistema con dos grados de libertad

Sea un sistema de dos grados de libertad descrito por el Hamiltoniano

$$H = q_1 p_1 - q_2 p_2 - a q_1^2 + b q_2^2$$

mostraremos que

$$F_1 = \frac{p_1 - a q_1}{q_2} \quad F_2 = q_1 q_2$$

son constantes de movimiento. Puesto que  $F_1$  y  $F_2$  no dependen explícitamente del tiempo, solo debemos chequear que los corchetes de Poisson de estas cantidades con el Hamiltoniano se anulen

$$\begin{aligned} [H, F_1] &= \left( \frac{\partial H}{\partial q_1} \frac{\partial F_1}{\partial p_1} - \frac{\partial F_1}{\partial q_1} \frac{\partial H}{\partial p_1} \right) + \left( \frac{\partial H}{\partial q_2} \frac{\partial F_1}{\partial p_2} - \frac{\partial F_1}{\partial q_2} \frac{\partial H}{\partial p_2} \right) \\ [H, F_1] &= \left[ (p_1 - 2aq_1) \frac{1}{q_2} - \left( -\frac{a}{q_2} \right) q_1 \right] + \left[ (-p_2 + 2bq_2) \times 0 - \left( -\frac{(p_1 - aq_1)}{q_2^2} \right) (-q_2) \right] \\ [H, F_1] &= \frac{(p_1 - 2aq_1)}{q_2} + \frac{aq_1}{q_2} - \frac{(p_1 - aq_1)}{q_2} = 0 \end{aligned}$$

por tanto  $F_1$  es constante de movimiento, veamos a  $F_2$

$$\begin{aligned} [H, F_2] &= [(p_1 - 2aq_1) \times 0 - (q_2) q_1] + [(-p_2 + 2bq_2) \times 0 - (q_1) (-q_2)] \\ [H, F_2] &= -q_2 q_1 + q_1 q_2 = 0 \end{aligned}$$

luego  $F_2$  también es constante de movimiento. Dado que tenemos dos constantes de movimiento, es inmediato pensar que el teorema de Poisson podría proveernos de otras constantes de movimiento independientes. Para ello debemos evaluar el corchete de Poisson entre estas constantes

$$\begin{aligned} [F_2, F_1] &= \left( \frac{\partial F_2}{\partial q_1} \frac{\partial F_1}{\partial p_1} - \frac{\partial F_1}{\partial q_1} \frac{\partial F_2}{\partial p_1} \right) + \left( \frac{\partial F_2}{\partial q_2} \frac{\partial F_1}{\partial p_2} - \frac{\partial F_1}{\partial q_2} \frac{\partial F_2}{\partial p_2} \right) \\ [F_2, F_1] &= \left[ (q_2) \left( \frac{1}{q_2} \right) - \left( -\frac{a}{q_2} \right) (\times 0) \right] + \left[ (q_1) (\times 0) - \left( -\frac{(p_1 - aq_1)}{q_2^2} \right) \times 0 \right] \\ [F_2, F_1] &= 1 \end{aligned}$$

se concluye que la unidad es una constante de movimiento, lo cual es cierto pero trivial. En este caso el teorema de Poisson no es útil para generar constantes de movimiento independientes. Nótese que en general lo único que afirma el teorema es que el corchete entre dos constantes de movimiento es otra constante de movimiento, pero no nos dice si esa constante es no trivial o si es independiente de las dos anteriores. Más adelante veremos algunos ejemplos en donde el teorema de Poisson nos genera constantes de movimiento no triviales e independientes de las anteriores.

### 8.8.2. Otro sistema con dos grados de libertad

Un sistema físico que se mueve en un plano, está descrito por el Hamiltoniano

$$H = \frac{|\mathbf{p}|^n}{b} - ar^{-n} \quad (8.31)$$

donde  $\mathbf{p}$  es el vector de los momentos conjugados a las coordenadas cartesianas. Usaremos los corchetes de Poisson para demostrar que la cantidad

$$D = \frac{\mathbf{p} \cdot \mathbf{r}}{n} - Ht \quad (8.32)$$

es una constante de movimiento. Comenzaremos escribiendo el Hamiltoniano (8.31) en componentes

$$\begin{aligned} H &= \frac{(p_k p_k)^{n/2}}{b} - a (x_k x_k)^{-n/2} \\ \frac{\partial H}{\partial x_i} &= an x_i (x_k x_k)^{-1-n/2} \quad ; \quad \frac{\partial H}{\partial p_i} = \frac{n}{b} p_i (p_k p_k)^{-1+n/2} \end{aligned}$$

evaluaremos el siguiente corchete de Poisson

$$\begin{aligned} [\mathbf{p} \cdot \mathbf{r}, H] &= [p_k x_k, H] = \frac{\partial (p_k x_k)}{\partial x_i} \frac{\partial H}{\partial p_i} - \frac{\partial (p_k x_k)}{\partial p_i} \frac{\partial H}{\partial x_i} = \frac{n}{b} (p_i p_i) (p_k p_k)^{-1+\frac{n}{2}} - an (x_i x_i) (x_k x_k)^{-1-\frac{n}{2}} \\ [\mathbf{p} \cdot \mathbf{r}, H] &= \frac{n}{b} (p_k p_k)^{n/2} - an (x_k x_k)^{-n/2} = \frac{n}{b} (\mathbf{p}^2)^{n/2} - an (\mathbf{r}^2)^{-n/2} \\ [\mathbf{p} \cdot \mathbf{r}, H] &= n \left( \frac{|\mathbf{p}|^n}{b} - a |\mathbf{r}|^{-n} \right) = nH \end{aligned} \quad (8.33)$$

la ecuación de movimiento para la cantidad  $D$  definida en (8.32), será entonces

$$\begin{aligned}\frac{dD}{dt} &= [D, H] + \frac{\partial D}{\partial t} = \left[ \frac{\mathbf{p} \cdot \mathbf{r}}{n} - Ht, H \right] + \frac{\partial}{\partial t} \left( \frac{\mathbf{p} \cdot \mathbf{r}}{n} - Ht \right) = \frac{[\mathbf{p} \cdot \mathbf{r}, H]}{n} - [H, H]t - H = \frac{nH}{n} - 0 - H \\ \frac{dD}{dt} &= 0\end{aligned}$$

donde hemos usado (8.33), y el hecho de que  $H$  no depende explícitamente del tiempo. Hemos tenido en cuenta además que  $\partial_t(\mathbf{p} \cdot \mathbf{r}) = 0$  (¿porqué?). Concluimos entonces que  $D$  es una cantidad conservada.

Podemos aplicar este ejemplo al caso particular de un solo grado de libertad en el cual  $n = a^{-1} = b = 2$ , el Hamiltoniano y la cantidad conservada quedan

$$H = \frac{p^2}{2} - \frac{1}{2q^2} \quad ; \quad D = \frac{pq}{2} - Ht \quad (8.34)$$

para este Hamiltoniano la transformación  $Q = \lambda q$ ,  $p = \lambda P$  es una transformación canónica de escala. Consideremos adicionalmente, una transformación de escala en el tiempo

$$Q = \lambda q, \quad p = \lambda P, \quad t' = \lambda^2 t \quad (8.35)$$

las nuevas ecuaciones de Hamilton las expresaremos en términos de  $Q(t')$  y  $P(t')$

$$Q(t') = \lambda q(t) = \lambda q \left( \frac{t'}{\lambda^2} \right) \quad ; \quad P(t') = \frac{1}{\lambda} p(t) = \frac{1}{\lambda} p \left( \frac{t'}{\lambda^2} \right) \quad (8.36)$$

puesto que  $q$  y  $p$  son canónicas, ellas satisfacen las ecuaciones de Hamilton con el Hamiltoniano (8.34)<sup>2</sup>

$$\frac{dq(t)}{dt} = \frac{\partial H}{\partial p} = p(t) \quad ; \quad \frac{dp(t)}{dt} = -\frac{\partial H}{\partial q} = -\frac{1}{q^3(t)} \quad \Rightarrow \quad (8.37)$$

$$\frac{dQ(t)}{dt} = p(t) \quad ; \quad \frac{dP(t)}{dt} = -\frac{1}{q^3(t)} \quad (8.38)$$

Por otro lado, teniendo en cuenta las ecuaciones (8.35, 8.36, 8.38) tenemos

$$\begin{aligned}\frac{dQ(t')}{dt'} &= \frac{1}{\lambda^2} \frac{dQ(t)}{dt} = \frac{\lambda}{\lambda^2} \frac{dq(t)}{dt} = \frac{1}{\lambda} p(t) = P(t') \\ \frac{dP(t')}{dt'} &= \frac{1}{\lambda^2} \frac{dP(t)}{dt} = \frac{1}{\lambda^3} \frac{dp(t)}{dt} = -\frac{1}{\lambda^3} \frac{1}{q^3(t)} = -\frac{1}{[\lambda q(t)]^3} = -\frac{1}{Q^3(t')}\end{aligned}$$

hemos llegado entonces a las ecuaciones

$$\frac{dQ(t')}{dt'} = P(t') \quad ; \quad \frac{dP(t')}{dt'} = -\frac{1}{Q^3(t')} \quad (8.39)$$

las cuales son idénticas en forma a las ecuaciones (8.38). Es notable que las ecuaciones de movimiento para  $q$  y  $p$  con el Hamiltoniano (8.34) son invariantes bajo la transformación (8.35), a pesar de que dicha transformación no es canónica. Esto se puede ver del hecho de que las transformaciones canónicas son las transformaciones de  $q, p$  más generales que dejan invariantes las ecuaciones de movimiento, pero no contemplan transformaciones en el parámetro tiempo. La constante de movimiento  $D$ , en la ecuación (8.34) está asociada a esta invarianza<sup>3</sup>.

<sup>2</sup>en este caso escribimos  $dq/dt$  y  $dp/dt$  en lugar de  $\dot{q}$  y  $\dot{p}$ , ya que debemos diferenciar bien los dos parámetros temporales.

<sup>3</sup>Más adelante veremos que la relación entre  $H$  y  $t$ , es muy similar a la relación entre  $q$  y  $p$ , a pesar de que  $H$  y  $t$  no son variables canónicamente conjugadas. En la cantidad conservada  $D$  de la Ec. (8.34), podemos apreciar un producto  $qp$  asociado a la transformación de escala de las variables canónicas y un producto  $Ht$  asociado a la transformación de escala del Hamiltoniano y el tiempo.

### 8.8.3. Constante de movimiento del oscilador armónico

Los problemas considerados hasta aquí son un tanto académicos. Consideremos el problema más Físico del oscilador armónico unidimensional con Hamiltoniano

$$H(q, p) = \frac{p^2}{2m} + \frac{kq^2}{2} \quad (8.40)$$

Demostraremos que para este sistema la cantidad Física definida por

$$u(q, p, t) = \ln(p + im\omega q) - i\omega t \quad ; \quad \omega \equiv \sqrt{\frac{k}{m}} \quad (8.41)$$

es una constante de movimiento y le daremos una interpretación Física. En primer lugar, evaluamos la ecuación de movimiento para  $u(q, p, t)$

$$\frac{du}{dt} = [u, H] + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial u}{\partial p} \frac{\partial H}{\partial q} + \frac{\partial u}{\partial t}$$

y utilizando el Hamiltoniano (8.40), resulta

$$\begin{aligned} \frac{du}{dt} &= \frac{im\omega}{p + im\omega q} \left( \frac{p}{m} \right) - \frac{1}{p + im\omega q} (kq) - i\omega = \frac{i\omega p - kq}{p + im\omega q} - i\omega \\ \frac{du}{dt} &= \frac{i\omega p - m\omega^2 q}{p + im\omega q} - i\omega = i\omega \frac{p + im\omega q}{p + im\omega q} - i\omega = i\omega - i\omega \\ \frac{du}{dt} &= 0 \end{aligned}$$

para ver el significado Físico de esta constante de movimiento exponenciamos la ecuación (8.41)

$$\begin{aligned} e^u &= (p + im\omega q) e^{-i\omega t} = \sqrt{p^2 + (m\omega q)^2} e^{i\theta} e^{-i\omega t} \quad ; \quad \theta = \arctan\left(\frac{m\omega q}{p}\right) \\ e^u &= \sqrt{2m\bar{H}} e^{i(\theta - \omega t)} \quad ; \quad \theta = \arctan\left(\frac{m\omega q}{p}\right) \end{aligned} \quad (8.42)$$

las soluciones para  $q$  y  $p = m\dot{q}$  están dadas por

$$\begin{aligned} q &= A \cos(\omega t + \delta) = A \sin\left(\omega t + \delta + \frac{\pi}{2}\right) \\ p &= m\omega A \cos\left(\omega t + \delta + \frac{\pi}{2}\right) \end{aligned}$$

el cociente nos da

$$\begin{aligned} \frac{m\omega q}{p} &= \tan\left(\omega t + \delta + \frac{\pi}{2}\right) \Rightarrow \\ \arctan\left(\frac{m\omega q}{p}\right) &= \omega t + \delta + \frac{\pi}{2} \end{aligned} \quad (8.43)$$

sustituyendo (8.43) en (8.42) y teniendo en cuenta que  $H$  es la energía del oscilador armónico, tenemos

$$\begin{aligned} e^u &= \sqrt{2m\bar{H}} e^{i(\delta + \frac{\pi}{2})} \Rightarrow u = \ln\left[\sqrt{2m\bar{H}} e^{i(\delta + \frac{\pi}{2})}\right] = \ln(2mE)^{1/2} + i\left(\delta + \frac{\pi}{2}\right) \\ u &= \frac{1}{2} \ln(2mE) + i\left(\delta + \frac{\pi}{2}\right) \end{aligned}$$

por tanto la parte real de  $u$  está relacionada con la energía y la parte imaginaria con la fase inicial.

## 8.9. Transformaciones canónicas infinitesimales en el formalismo de los corchetes de Poisson

Las transformaciones canónicas infinitesimales (TCI) forman parte de un tipo especial de transformaciones que son una función continua de un parámetro y que comienzan con la transformación identidad para algún valor inicial del parámetro. Estas transformaciones adquieren particular interés ya que una transformación canónica arbitraria se puede construir como una sucesión de transformaciones infinitesimales. De nuevo el carácter de invariante canónico de los corchetes de Poisson, hace que exista un especial interés en escribir las TCI en este formalismo.

En las transformaciones canónicas infinitesimales, las nuevas variables difieren de las antiguas en cantidades infinitesimales

$$\zeta = \eta + \delta\eta$$

esta ecuación solo me caracteriza una transformación infinitesimal arbitraria (no necesariamente canónica). Cuando esta transformación es canónica, el cambio se especifica con el parámetro infinitesimal  $\varepsilon$  y la función generatriz de las transformaciones infinitesimales  $G$  (ver Ec. 7.71)<sup>4</sup>

$$\delta\eta = \varepsilon \mathbf{J} \frac{\partial G(\eta)}{\partial \eta} \quad (8.44)$$

tomando la Ec. (8.27)

$$[\eta, u] = \mathbf{J} \frac{\partial u}{\partial \eta} \quad (8.45)$$

esta expresión es válida independientemente del conjunto de variables canónicas usadas para evaluar el corchete. Haciendo  $u = G$  y reemplazando (8.45) en (8.44)

$$\delta\eta = \varepsilon [\eta, G] \quad (8.46)$$

con lo cual hemos logrado el propósito de escribir la TCI en términos de corchetes de Poisson, en donde el corchete involucra a las variables canónicas en cuestión, la función generatriz de la TCI y el parámetro infinitesimal de evolución. Consideremos el caso en el cual el parámetro continuo es el tiempo de modo que  $\varepsilon = dt$ . Tomemos como función generadora el Hamiltoniano<sup>5</sup>. Las ecuaciones de transformación para las TCI se obtienen de (8.46)

$$\delta\eta = [\eta, H] dt = \dot{\eta} dt = d\eta \quad (8.47)$$

donde hemos usado las Ecs. de Hamilton en corchetes de Poisson Ecs. (8.26). Para ver el significado de (8.47) recordemos el significado de  $\delta\eta$  y de  $d\eta$ . La cantidad  $\delta\eta$  se refiere al cambio infinitesimal en las coordenadas debido a una transformación canónica, recordemos que un cambio de coordenadas no se refiere a la evolución del sistema. En contraste  $d\eta$  se refiere al cambio de las coordenadas desde el tiempo  $t$  a sus valores en el tiempo  $t + dt$  como producto de la evolución del sistema. En consecuencia, la igualdad entre estas dos cantidades infinitesimales nos muestra que la TCI generada por el Hamiltoniano y usando el tiempo como parámetro continuo, cambia las coordenadas y momentos en la misma forma que lo haría la evolución del sistema. En otras palabras, el movimiento del sistema Físico en un intervalo de tiempo  $dt$  se puede describir a través de una transformación canónica infinitesimal generada por el Hamiltoniano, siendo  $dt$  el parámetro infinitesimal que modula la TCI. Por otro lado, una transformación canónica arbitraria se puede obtener por medio de TCI's sucesivas (esto es formalmente un proceso de integración en el parámetro infinitesimal). Es decir, que el

<sup>4</sup>Estrictamente la función generatriz de las transformaciones infinitesimales es la función  $F_2$  dada por la Ec. (7.68), Pág. 132. Sin embargo, dado que el término  $q_i P_i$  de esta función es el generador de la identidad, la función  $G$  en esta ecuación es la parte no trivial de esta transformación. Por tanto, nos referiremos a  $G$  de aquí en adelante como la función generadora de una transformación infinitesimal modulada por el parámetro  $\varepsilon$ .

<sup>5</sup>Nótese que en el paso desde la Ec. (7.68) hasta la Ec. (7.71), solo hemos exigido que las segundas derivadas de  $G$  sean continuas en sus  $2n+1$  argumentos. Esto con el fin de garantizar que la matriz (7.74) sea simétrica, lo cual a su vez conduce a que la condición simpléctica sea necesaria y suficiente para llegar a la condición canónica en la transformación. Por lo demás,  $G$  y su parámetro infinitesimal  $\varepsilon$  de modulación son arbitrarios.



movimiento del sistema en un intervalo finito de tiempo se obtiene con sucesivos corrimientos  $dt$ . Los valores de  $q, p$  en cierto tiempo se pueden obtener a partir de sus valores iniciales por una transformación canónica que es función continua del tiempo. De acuerdo a este punto de vista, el movimiento de un sistema mecánico corresponde a una evolución continua de las transformaciones canónicas. Todo este razonamiento nos lleva a concluir que el Hamiltoniano es el generador del movimiento del sistema con el tiempo.

Inversamente, debe existir una transformación canónica que a partir de los valores de las variables  $q, p$  en un tiempo  $t$ , nos lleve a sus valores iniciales. Encontrar esta transformación canónica es equivalente a resolver la ecuación de movimiento. Habíamos sugerido previamente la posibilidad de obtener un Hamiltoniano con todas las coordenadas cíclicas que conducía a que todos los momentos eran constantes de movimiento. La presente estrategia, si es posible, nos llevaría a una transformación canónica en donde  $q_i$  y  $p_i$  serían todas constantes de movimiento. Volveremos sobre estas consideraciones en el siguiente capítulo para obtener soluciones formales para sistemas mecánicos.

Lo anterior nos lleva a contemplar la posibilidad de ver a las transformaciones canónicas desde otro punto de vista, así como a los efectos que ésta produce. La noción de transformación canónica se introdujo como un cambio en las coordenadas (pero no en la configuración del sistema) para describir el mismo sistema usando otro espacio de fase. En esta visión, cambiamos de las coordenadas  $(q, p)$  de un espacio de fase  $\eta$ , a otro espacio de fase  $\zeta$  con coordenadas  $(Q, P)$ . Si el estado de un sistema en un cierto tiempo está descrito por un punto  $A$  en el espacio de fase  $\eta$ , también puede describirse equivalentemente a través del punto transformado  $A'$  del espacio de fase  $\zeta$  (en el mismo instante de tiempo). Cualquier función del sistema de variables tendrá el mismo valor para una configuración dada del sistema, bien sea que la describamos con el conjunto  $(q, p)$  o con  $(Q, P)$ . Es decir, la función tendrá el mismo valor en  $A$  que en  $A'$  (aunque diferente forma funcional). Este se llama un punto de vista **pasivo** de la TC. Desde el punto de vista matemático, esto corresponde a un mapeo desde el espacio  $\eta$  al espacio  $\zeta$  (con inversa). La Fig. 8.1a ilustra el concepto de TC desde el punto de vista pasivo. Nótese que con este enfoque, la transformación de coordenadas está totalmente desligada de la evolución del sistema.

La TC que genera el Hamiltoniano usando al tiempo como parámetro, sugiere otra interpretación para esta TC. Cuando movemos el parámetro tiempo desde  $t$  hasta  $\Delta t$ , esta TC cambia las coordenadas y momentos desde sus valores en el tiempo  $t$  hasta los valores que estas tendrían en el tiempo  $t + \Delta t$ . Esta transformación canónica se puede interpretar consistentemente de la siguiente manera: Dicha TC nos relaciona las coordenadas  $(q, p)$  de un punto en el espacio de fase (definido con las coordenadas  $q, p$ ) con las coordenadas  $(Q, P)$  de otro punto **en el mismo espacio de fase**. Esto corresponde a un mapeo del espacio de fase en sí mismo. Lo anterior nos conduce a una **interpretación activa** de la TC como la generadora de “movimiento” del punto en el espacio de fase de una posición  $(q, p)$  a otra posición  $(Q, P)$ , como se ilustra en la Fig. 8.1b. Por supuesto, la transformación canónica como tal no puede mover o cambiar la configuración del sistema. Lo que ocurre es que el cambio de coordenadas producido por la TC emula el cambio dinámico de las coordenadas producido por la evolución del sistema. En otras palabras, el paso de  $(q, p)$  a  $(Q, P)$  bajo una TC significa que estoy cambiando el sistema coordinado que uso para describir al sistema, el mismo cambio interpretado en forma dinámica se refiere al cambio que **el mismo sistema coordinado** sufre por efecto de la evolución del sistema. El punto es que cuando las dos transformaciones coinciden numéricamente, puedo atribuir la dinámica del sistema (desde un punto de vista práctico) a la transformación canónica.

La interpretación activa no es siempre útil. Por ejemplo, la TC que nos lleva de las coordenadas cartesianas a las coordenadas polares esféricas, es una TC de tipo pasivo, y una interpretación activa sería absurda. Por ejemplo, los ejes de ambos espacios de fase no poseen las mismas unidades físicas pues en coordenadas cartesianas ambos ejes de coordenadas tienen unidades de longitud y los ejes de momentos tienen unidades de masa por velocidad, en contraste se tiene que en coordenadas polares hay un eje angular adimensional en la coordenada  $\theta$  y el eje de  $p_\theta$  tiene unidades de momento angular (aunque sí son iguales las dimensiones en cuanto al número de coordenadas i.e. ejes independientes, es decir son iguales las dimensiones de los dos espacios de fase como espacios vectoriales).

El punto de vista activo es particularmente útil en TC's que dependen en forma continua de un solo

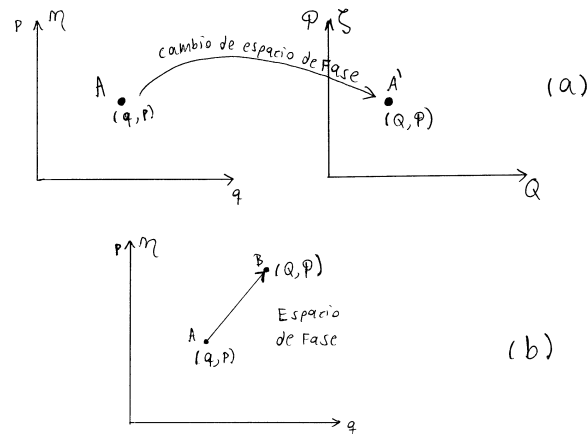


Figura 8.1: Descripción de una transformación canónica (a) desde el punto de vista pasivo y (b) desde el punto de vista activo.

parámetro<sup>6</sup>. En la interpretación activa, el efecto de la TC es el de “mover” el punto del sistema de manera continua trazando así una curva en el espacio de fase, a medida que el parámetro cambia en forma continua. Cuando la TC es la generada por el Hamiltoniano usando al tiempo como el parámetro continuo, la curva sobre la cual se mueve el punto del sistema coincide con la trayectoria del sistema en el espacio de fase.

## 8.10. Cambio de una función del sistema bajo una transformación canónica en los enfoques pasivo y activo

Nos preguntamos ahora por el cambio de una función del sistema  $u = u(q, p, t)$  bajo una TC. Veremos que el cambio en esta función depende de si tomamos el punto de vista activo o pasivo.

Bajo el punto de vista pasivo, debemos tener en cuenta que si en un instante dado  $t_0$  la transformación canónica se describe por  $Q_0 = Q(q_0, p_0, t_0)$ ,  $P_0 = P(q_0, p_0, t_0)$  entonces el conjunto  $(q_0, p_0)$  describe la misma configuración del sistema que el conjunto  $(Q_0, P_0)$ . Ahora bien, el valor de una función del sistema solo puede depender de la configuración de éste. Por lo tanto, el cambio del conjunto coordenado  $(q_0, p_0)$  al conjunto

<sup>6</sup>Nótese que en transformaciones canónicas continuas uniparamétricas no es de esperarse que las unidades de las coordenadas y momentos cambien con la transformación. A manera de ejemplo, asumamos por un momento que para un sistema unidimensional partimos de variables  $(q, p)$  donde  $q$  tiene unidades de longitud y  $p$  de momento lineal, y terminamos con un sistema  $(Q, P)$  donde  $Q$  es adimensional y  $P$  tiene unidades de momento angular. Es claro que el cambio de unidades es discreto y no es de esperarse que una TC que se puede generar por TCI's sucesivas me lleve a un cambio en las unidades de las coordenadas, a menos que se presente el caso bastante atípico de una TC que varíe de manera continua las unidades de las variables. Por tanto las TC continuas uniparamétricas permiten interpretar consistentemente el hecho de que la transformación siempre esté en el mismo espacio de fase.

$(Q_0, P_0)$  debe dejar inalterado el valor de una función del sistema, aunque la función puede cambiar de forma o dependencia funcional con las nuevas variables. Podemos escribir entonces que  $u(q_0, p_0, t_0) = u'(Q_0, P_0, t_0)$ . En otras palabras, el valor numérico de una función del sistema no puede depender del espacio de fase que utilice para describir a dicha función.

En contraste, una visión activa de la TC nos habla de una traslación del sistema del punto  $A$  al punto  $B$ , de la posición  $(q_A, p_A)$  a la posición  $(q_B, p_B)$  en el mismo espacio de fase. Desde este punto de vista, la función  $u(q, p)$  no cambia su dependencia funcional con respecto a  $q, p$  ya que no estamos cambiando de espacio de fase. En cambio, la función sí cambia su valor como resultado del cambio en el valor de los argumentos  $u(q_A, p_A) \neq u(q_B, p_B)$ . Esto tiene que ver con el hecho de que en la visión activa  $A$  y  $B$  me están describiendo diferentes configuraciones del sistema.

Usaremos el símbolo  $\partial$  para denotar el cambio en el valor de una función bajo una TCI **activa**

$$\partial u \equiv u(B) - u(A) \quad (8.48)$$

donde  $A$  y  $B$  están infinitesimalmente cerca, de nuevo usando notación matricial escribimos

$$\partial u = u(\boldsymbol{\eta} + \delta\boldsymbol{\eta}) - u(\boldsymbol{\eta})$$

expandiendo en serie de Taylor hasta primer orden en infinitesimales y usando (8.44) se obtiene

$$\begin{aligned} \partial u &= \left[ u(\boldsymbol{\eta}) + \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \delta\boldsymbol{\eta} + \dots \right] - u(\boldsymbol{\eta}) \\ \partial u &= \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \delta\boldsymbol{\eta} = \varepsilon \left( \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} \right) \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\eta}} \end{aligned}$$

recordemos que el uso de la Ec. (8.44) es lo que nos garantiza que la transformación sea canónica. Ahora utilizamos la definición de corchetes de Poisson (8.2) para obtener:

$$\partial u = \varepsilon [u, G] \quad (8.49)$$

de nuevo la aplicación mas inmediata consiste en usar (8.49) cuando  $u$  es una coordenada del espacio de fase, teniendo en cuenta además la Ec. (8.46), resulta

$$\partial \boldsymbol{\eta} = \varepsilon [\boldsymbol{\eta}, G] = \delta \boldsymbol{\eta} \quad (8.50)$$

Este resultado es obvio a partir de la definición del punto  $B$  con respecto al punto  $A$ , el “cambio” en las coordenadas desde  $A$  hasta  $B$  es precisamente la diferencia infinitesimal entre las coordenadas viejas y nuevas.

## 8.11. Cambio del Hamiltoniano bajo una transformación canónica

Las consideraciones anteriores son mucho más delicadas cuando estudiamos el cambio en el Hamiltoniano. La razón estriba en el hecho de que **el Hamiltoniano no es una función del sistema**. Para ver la razón, recordemos que el Hamiltoniano puede diferir tanto funcional como numéricamente cuando pasamos de un conjunto canónico a otro, la expresión

$$H(q, p, t) \neq K(Q, P, t)$$

nos indica que el Hamiltoniano puede cambiar tanto funcional como numéricamente, incluso si los dos sistemas canónicos describen la misma configuración del sistema. Esto viola claramente nuestra definición de función del sistema. El Hamiltoniano se refiere mas bien a la función que en un determinado espacio de fase define las ecuaciones canónicas de movimiento. Cuando la transformación canónica depende del tiempo, el significado mismo de Hamiltoniano se transforma. De esta forma  $H(A)$  no se transforma en  $H'(A')$ <sup>7</sup> sino en  $K(A')$ , y

<sup>7</sup>Usaremos la convención de que  $u'$  es la forma funcional que hace que  $u(A) = u'(A')$ , es decir es la forma en que debe transformar la forma funcional de  $u$  ante un cambio de espacio de fase a fin de que dicha cantidad defina una función del sistema.

$H(A)$  no tendrá necesariamente el mismo valor que  $K(A')$ . A manera de ejemplo, en cierto espacio de fase el Hamiltoniano puede ser la energía del sistema y en otro espacio de fase puede no serlo. En virtud de lo anterior, designaremos por  $\partial H$  a la diferencia en el valor final del Hamiltoniano bajo las dos interpretaciones.

$$\partial H = H(B) - K(A') \quad (8.51)$$

en los casos en que la función misma no cambia con la transformación canónica, las dos formas de cambio descritas por (8.48) y (8.51) son idénticas puesto que  $u(A) = u'(A')$  (recordemos que el valor numérico de las funciones coincide en  $A$  y  $A'$  en una interpretación pasiva). En general,  $K$  está relacionado con  $H$  a través de la ecuación

$$K = H + \frac{\partial F}{\partial t} \quad (8.52)$$

Una vez que definimos la relación entre los Hamiltonianos en ambos espacios de fase a través de la relación (8.52), las funciones  $K$  y  $H$  ya se pueden tomar como funciones del sistema, es decir como funciones bien definidas para una configuración dada del sistema. Por ejemplo  $H(A) = H'(A')$  ya que se trata de la misma función del sistema definida en puntos  $A$  y  $A'$  de diferentes espacios de fase pero que describen la misma configuración del sistema. Para una TCI, la función generatriz está dada por (7.68) en términos de  $G$ . Dado que en (7.68) solo  $G$  puede ser función explícita del tiempo, el valor del nuevo Hamiltoniano es

$$K(A') = H'(A') + \varepsilon \frac{\partial G}{\partial t} = H(A) + \varepsilon \frac{\partial G}{\partial t}$$

y el cambio en el Hamiltoniano definido en (8.51) es

$$\partial H = H(B) - H(A) - \varepsilon \frac{\partial G}{\partial t} \quad (8.53)$$

siguiendo un camino similar al que nos llevó desde (8.48) hasta (8.49), vemos que  $\partial H$  está dado por

$$\partial H = \varepsilon [H, G] - \varepsilon \frac{\partial G}{\partial t}$$

tomando la ecuación generalizada de movimiento (8.25) con  $u = G$ , resulta

$$\begin{aligned} \partial H &= \varepsilon \left\{ \frac{\partial G}{\partial t} - \frac{dG}{dt} \right\} - \varepsilon \frac{\partial G}{\partial t} \Rightarrow \\ \partial H &= -\varepsilon \frac{dG}{dt} \end{aligned} \quad (8.54)$$

## 8.12. Cantidades conservadas e invarianzas del Hamiltoniano

Nos restringiremos al caso en el cual  $G$  no es función explícita del tiempo. En este caso, la Ec. (8.53) nos lleva a

$$\partial H = H(B) - H(A) \quad \text{si} \quad G = G(q_i, P_i)$$

de modo que el cambio  $\partial H$  que hemos definido para el Hamiltoniano, corresponde en este caso al cambio en el valor de éste cuando el sistema se mueve de la configuración  $A$  a la configuración  $B$ . Si adicionalmente  $G$  es una constante de movimiento, la Ec. (8.54) nos dice que  $G$  genera una TCI que no cambia el valor del Hamiltoniano. De esta forma, podemos ver que **muchas constantes de movimiento son funciones generadoras de aquellas TCI que dejan el Hamiltoniano invariante**. Recíprocamente, si el Hamiltoniano es invariante ante una TCI inducida por una función  $G(q_i, P_i)$ , la Ec. (8.54) nos indica que  $G$  es una constante de movimiento.

Nuevamente, la aseveración anterior nos muestra una conexión entre propiedades de simetría y cantidades conservadas<sup>8</sup>. Si una cierta función del sistema  $G(q_i, P_i)$ , es constante de movimiento, entonces la TCI que

<sup>8</sup>Debe tenerse en cuenta sin embargo, que aquí estamos hablando de simetrías del Hamiltoniano ante TCI's y no ante transformaciones arbitrarias.

dicha función induce es tal que el Hamiltoniano permanece invariante ante dicha TCI. Es esencial insistir en que estas conclusiones solo son válidas si  $G$  no es función explícita del tiempo.

Veremos enseguida que un momento  $p_i$  canónicamente conjugado a una variable  $q_i$ , se conserva si y solo si la coordenada  $q_i$  es cíclica. No obstante, este escenario abarca muchas constantes de movimiento independientes y no solo los momentos generalizados conservados. Sin embargo, el presente formalismo no abarca todas las constantes de movimiento (recordemos que la Ec. 8.25 abarca formalmente todas las constantes de movimiento). Esto por dos razones (1) Estas simetrías están restringidas solo a TCI's y no a transformaciones arbitrarias. (2) Las transformaciones de simetría más generales son las que dejan invariantes las ecuaciones de movimiento, incluso si no dejan invariante al Hamiltoniano<sup>9</sup>. Ya hemos visto en la sección 5.3 que con el formalismo Lagrangiano la invarianza de las ecuaciones de movimiento nos lleva a constantes de movimiento que no están incluídas en escenarios donde el Lagrangiano mismo es invariante. Lo esencial es tener en cuenta que ni el Lagrangiano ni el Hamiltoniano son observables del sistema, y por tanto pueden ser modificados de una manera especial sin cambiar el contenido físico de éstos. En todo caso, la relación entre la constante de movimiento  $G$  y la invarianza del Hamiltoniano nos muestra una vez más el fuerte nexo entre simetrías y cantidades conservadas.

Veremos a continuación que los teoremas de conservación de los momentos generalizados son casos especiales de lo anterior. Si una coordenada  $q_j$  es cíclica, es obvio que el Hamiltoniano es invariante ante una transformación infinitesimal que involucre un desplazamiento de  $q_j$  únicamente. Por supuesto, es necesario asegurarse que dicha transformación infinitesimal es de naturaleza canónica, de ser así, debe existir una función generatriz  $G$  que me genere la hipotética TCI a través de las ecuaciones (7.69, 7.70), puede verificarse que la función<sup>10</sup>

$$G(q, p) = p_i \quad (8.55)$$

genera una transformación canónica infinitesimal a través de (7.69, 7.70) descrita por

$$\delta q_j = \varepsilon \delta_{ij} \quad ; \quad \delta p_i = 0 \quad (8.56)$$

es decir  $p_i$  es una función generatriz de una transformación canónica infinitesimal que desplaza a  $q_i$  únicamente. Dado que solo  $q_i$  se desplaza, es obvio que si dicha coordenada es cíclica el Hamiltoniano queda invariante, lo cual nos lleva a la conservación de  $G$  según la Ec. (8.54), pero  $G$  en nuestro caso es precisamente  $p_i$  como se vé en (8.55). Por tanto, se observa que si  $q_i$  es cíclica llegamos a la conservación de su momento conjugado como se esperaba. Recíprocamente si  $p_i$  es constante entonces la función  $G(q, P)$  en (8.55) es constante de movimiento y esto nos lleva a la invarianza del Hamiltoniano ante la TCI (8.56), que desplaza solo a  $q_j$  dejando todas las demás coordenadas y momentos fijos, esto implica entonces que  $\partial H / \partial q_i = 0$  y por tanto que la coordenada es cíclica.

En virtud de la simetría entre  $q$  y  $p$  en la formulación Hamiltoniana, es natural pensar que  $G(q, P) = q_i$  genere una TCI en donde solo se mueva  $p_i$  (tal vez con un cambio de signo debido a la estructura simpléctica). Ambas consideraciones se pueden escribir en un contexto mas general. Tomemos como generador de una TCI a la función

$$G_l = (\mathbf{J}\eta)_l = J_{lr}\eta_r \quad (8.57)$$

la ecuaciones de transformación (7.71) aplicadas a (8.57) dan

$$\delta \eta_k = \varepsilon J_{ks} \frac{\partial G_l}{\partial \eta_s} = \varepsilon J_{ks} J_{lr} \frac{\partial \eta_r}{\partial \eta_s} = \varepsilon J_{ks} J_{lr} \delta_{rs} = \varepsilon J_{ks} J_{ls} = \varepsilon J_{ks} \tilde{J}_{sl}$$

lo cual en virtud de la ortogonalidad de  $\mathbf{J}$  (ver Ec. 6.32, Pág. 99), resulta

$$\delta \eta_k = \varepsilon \delta_{kl}$$

esto muestra claramente que una transformación que cambie solo a  $\eta_l$  es generada por su variable conjugada. Si  $\eta_l$  es  $q_i$  entonces  $G$  es  $p_i$ . Si  $\eta_l$  es  $p_i$ ,  $G$  es entonces  $-q_i$ .

<sup>9</sup>Nótese que una transformación que cambie al Hamiltoniano en la forma  $H = H + \frac{dF(q,p,t)}{dt}$ , deja invariantes las ecuaciones de movimiento pero no al Hamiltoniano.

<sup>10</sup>Estrictamente  $G$  es de tipo 2, de modo que sus argumentos son  $G(q_i, P_i)$ . Sin embargo, en la interpretación activa  $P_i$  está en el mismo espacio de fase que el momento original, razón por la cual escribiremos simplemente  $G(q_i, p_i)$ , al menos cuando se trabaje con la interpretación activa de la TC.

### 8.12.1. El momento lineal total como generador de TCI's que generan traslaciones

Examinemos una vez más el caso en el cual realizamos la traslación de un sistema como un todo pero ahora en el marco del formalismo de las TCI's. En primer lugar, es importante enfatizar que el significado físico de la función generatriz  $G$  no puede depender del conjunto canónico empleado para describirla. Para ver esto, observemos que en la Ec. (8.46) el cambio de una cierta variable canónica  $\eta_i$  es el mismo independiente del sistema canónico que usemos para expresar  $G$ , ello en virtud de la invarianza canónica de los corchetes de Poisson. Por lo anterior, podemos utilizar en particular las coordenadas cartesianas de las partículas del sistema. Definamos la función generatriz

$$G = \sum_{k=1}^N p_{kx} \quad ; \quad \varepsilon \equiv dx \quad (8.58)$$

en donde  $p_{kx}$  es la componente  $x$  del momento canónico de la  $k$ -ésima partícula. Utilizando las Ecs. (7.69, 7.70) vemos que la TCI inducida por esta función es

$$\begin{aligned} \delta x_i &= dx \frac{\partial G}{\partial p_{ix}} = dx \frac{\partial}{\partial p_{ix}} \left( \sum_{k=1}^N p_{kx} \right) = dx \quad , \quad \delta p_{ix} = -dx \frac{\partial G}{\partial q_{ix}} = 0 \\ \delta y_i &= dx \frac{\partial G}{\partial p_{iy}} = 0 \quad , \quad \delta p_{iy} = -dx \frac{\partial G}{\partial q_{iy}} = 0 \quad ; \quad \delta z_i = \delta p_{iz} = 0 \end{aligned}$$

de modo que

$$\delta x_i = dx \quad ; \quad \delta y_i = \delta z_i = \delta p_{ix} = \delta p_{iy} = \delta p_{iz} = 0$$

es decir la función  $G$  definida en (8.58) genera una traslación del sistema como un todo en una cantidad infinitesimal  $dx$  en la dirección  $x$ , sin traslación en las otras coordenadas ni momentos. Vemos pues que la traslación del sistema como un todo a lo largo de cierto eje, corresponde efectivamente a una transformación canónica.

Puesto que  $G$  no depende explícitamente del tiempo, es constante de movimiento si y solo si la TCI que genera deja invariante el Hamiltoniano. Es decir, el momento lineal total en la dirección  $x$  (que corresponde a  $G$  en este caso), se conserva si y solo si la traslación del sistema como un todo en la dirección  $x$ , deja invariante al Hamiltoniano. Nótese que la dirección  $x$  se puede escoger arbitrariamente. Si denotamos  $\mathbf{n}$  como el vector unitario en la dirección de la traslación y  $\mathbf{P}$  como el momento lineal total del sistema, podemos decir que

$$G = \mathbf{P} \cdot \mathbf{n}$$

se conserva si y solo si el Hamiltoniano es invariante ante una TCI que genera una traslación del sistema como un todo en la dirección  $\mathbf{n}$ . Recordemos que si las fuerzas que actúan sobre el sistema se derivan de potenciales que dependen de la velocidad, el momento canónico asociado a las coordenadas cartesianas no es el momento lineal de la forma  $m\dot{x}$ . Por tanto nuestros resultados son más generales que los obtenidos en la sección 5.1.1 en donde se supuso explícitamente que los potenciales son independientes de la velocidad.

### 8.12.2. El momento angular total como generador de TCI's que generan rotaciones

Ahora analizaremos la rotación de un sistema Físico como un todo en el marco de las TCI's. Por los mismos argumentos de la sección anterior, podemos emplear las coordenadas cartesianas. Adicionalmente, y sin pérdida de generalidad elegiremos al eje de rotación como nuestro eje  $z$ . Para una rotación infinitesimal antihoraria en un ángulo  $d\theta$  la matriz de rotación a primer orden se escribe

$$\begin{pmatrix} \cos d\theta & -\sin d\theta \\ \sin d\theta & \cos d\theta \end{pmatrix} \simeq \begin{pmatrix} 1 & -d\theta \\ d\theta & 1 \end{pmatrix}$$

ahora tendremos en cuenta que en coordenadas cartesianas tanto las coordenadas  $(x_i, y_i, z_i)$  como los momentos conjugados  $(p_{ix}, p_{iy}, p_{iz})$  forman vectores euclidianos, de modo que se comportan de la misma forma ante rotaciones<sup>11</sup> con lo cual

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} 1 & -d\theta \\ d\theta & 1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \Rightarrow x'_i = x_i - y_i d\theta ; y'_i = y_i + x_i d\theta \\ x'_i - x_i = -y_i d\theta ; y'_i - y_i = x_i d\theta$$

dado que los momentos poseen las mismas transformaciones, basta con hacer un reemplazo de la forma  $x_i \rightarrow p_{ix}$  para obtener la transformación de los momentos, las transformaciones infinitesimales quedan entonces

$$\begin{aligned} \delta x_i &= -y_i d\theta ; \delta y_i = x_i d\theta ; \delta z_i = 0 \\ \delta p_{ix} &= -p_{iy} d\theta ; \delta p_{iy} = p_{ix} d\theta ; \delta p_{iz} = 0 \end{aligned} \quad (8.59)$$

estrictamente, aún no hemos demostrado que esta transformación es canónica. Esto equivale a demostrar que existe una función generatriz  $G$ , tal que las Ecs. (7.69, 7.70) reproducen adecuadamente a las ecuaciones (8.59). Podemos ver que una función generatriz adecuada para esta TCI es

$$G = x_k p_{ky} - y_k p_{kx} ; \varepsilon \equiv d\theta \quad (8.60)$$

Esto se puede ver explícitamente reemplazando (8.60) en las Ecs. (7.69, 7.70)

$$\begin{aligned} \delta x_i &= d\theta \frac{\partial G}{\partial p_{ix}} = -y_i d\theta ; \delta y_i = d\theta \frac{\partial G}{\partial p_{iy}} = x_i d\theta ; \delta z_i = d\theta \frac{\partial G}{\partial p_{iz}} = 0 \\ \delta p_{ix} &= -d\theta \frac{\partial G}{\partial x_i} = -p_{iy} d\theta ; \delta p_{iy} = -d\theta \frac{\partial G}{\partial y_i} = p_{ix} d\theta ; \delta p_{iz} = -d\theta \frac{\partial G}{\partial z_i} = 0 \end{aligned}$$

que reproduce correctamente las Ecs. (8.59). Es inmediato ver que la función  $G$  dada por (8.60) corresponde a la componente  $z$  del momento angular canónico total

$$G = L_z \equiv (\mathbf{r}_i \times \mathbf{p}_i)_z \quad (8.61)$$

y dado que el eje  $z$  se eligió en la dirección del eje de rotación cuya orientación es arbitraria, se concluye que la función generatriz  $G$  corresponde a la componente del momento angular canónico total a lo largo del eje de rotación. Denotando  $\mathbf{n}$  al vector unitario a lo largo del eje de rotación se tiene que

$$G = \mathbf{L} \cdot \mathbf{n} \quad (8.62)$$

es importante notar que el momento angular canónico que hemos definido puede diferir del momento angular mecánico, ya que si las fuerzas que actúan sobre el sistema se derivan de potenciales que dependen de la velocidad, las cantidades  $\mathbf{p}_i$  no necesariamente corresponden con el momento lineal (aunque siguen siendo vectores Euclidianos que fué nuestra única suposición). En consecuencia, la expresión (8.61) no corresponde necesariamente al momento angular mecánico. Por lo tanto, nuestros resultados son más generales que los obtenidos en la sección 5.1.2 en donde se supuso explícitamente que los potenciales eran independientes de la velocidad. Podemos ver entonces que el momento conjugado asociado a una coordenada generalizada de rotación del sistema como un todo alrededor de un eje  $\mathbf{n}$ , es la componente del momento angular canónico total según este eje, incluso si el potencial depende de la velocidad. Este resultado se puede derivar también de las Ecs. (8.55, 8.56).

Concluimos que el Hamiltoniano es invariante bajo la rotación del sistema como un todo alrededor de la dirección  $\mathbf{n}$ , si y solo si se conserva la función generatriz  $G$  definida en la Ec. (8.62), es decir si se conserva el momento angular canónico total en la dirección  $\mathbf{n}$ . Una vez más vemos como la invarianza del Hamiltoniano ante rotaciones alrededor de un eje conduce a la conservación del momento angular **canónico** total alrededor de dicho eje.

En síntesis, así como el Hamiltoniano es un generador de desplazamiento del sistema en el tiempo, el momento lineal es un generador de desplazamiento lineal espacial (traslación) del sistema, y el momento angular es un generador de desplazamiento angular espacial (rotación) del sistema. Recordemos que aquí la palabra **generador** indica que éste induce una TCI que nos lleva a la transformación en cuestión.

<sup>11</sup>Para más detalles ver sección 12.6

### 8.13. Construcción de transformaciones canónicas finitas a partir de transformaciones canónicas infinitesimales

Ya se ha dicho que en una interpretación activa, una transformación canónica que depende de un parámetro, mueve el punto que describe la configuración del sistema formando una trayectoria en el espacio de fase. Una transformación canónica finita que depende de un parámetro continuo  $\alpha$ , puede construirse como una sucesión de TCI's a lo largo de la curva. Formalmente, la TC finita se obtiene integrando las TCI's. Como a cada punto en la trayectoria le corresponde un valor particular del parámetro, comenzamos con la configuración inicial del sistema a la cual por convención le asignamos el parámetro  $\alpha = 0$ . Si  $u$  es alguna función de la configuración del sistema,  $u$  será una función continua de  $\alpha$  a lo largo de la trayectoria, con valor inicial  $u_0 \equiv u(0)$ . Por simplicidad, asumiremos que  $u$  no es función explícita del tiempo. La TCI en una interpretación activa viene dada por (8.49), donde  $\varepsilon$  viene dado por  $d\alpha$

$$\partial u = d\alpha [u, G] \quad (8.63)$$

o a manera de ecuación diferencial

$$\frac{du}{d\alpha} = [u, G] \quad (8.64)$$

Para entender porqué aparece una derivada total (a pesar de la notación engañosa  $\partial u$  en 8.63) recordemos que  $\partial u$  es el cambio en  $u$  asociado a una variación de todas las coordenadas y momentos como se vé en (8.48) y naturalmente, también hay variación en  $\alpha$  ya que es justamente la variación en esta cantidad la que genera el cambio de coordenadas y momentos. En principio, la integración de esta ecuación diferencial nos conduce a  $u(\alpha)$ , es decir al efecto que la TC finita tiene sobre  $u$ . Una solución formal se puede obtener haciendo una expansión de Taylor de  $u(\alpha)$  alrededor de las condiciones iniciales

$$u(\alpha) = u(0) + \alpha \left. \frac{du}{d\alpha} \right|_0 + \frac{\alpha^2}{2!} \left. \frac{d^2u}{d\alpha^2} \right|_0 + \frac{\alpha^3}{3!} \left. \frac{d^3u}{d\alpha^3} \right|_0 + \dots$$

la Ec. (8.64) conduce a

$$\left. \frac{du}{d\alpha} \right|_0 = [u, G]_0$$

donde el subíndice nos indica que el corchete de Poisson debe ser evaluado en el punto inicial  $\alpha = 0$ . Ahora tomando  $[u, G] \equiv u'$  como una función de la configuración del sistema, se tiene

$$\frac{d^2u}{d\alpha^2} = \frac{d}{d\alpha} \left( \frac{du}{d\alpha} \right) = \frac{d}{d\alpha} [u, G] = \frac{du'}{d\alpha}$$

y usando la Ec. (8.64) para la función  $u'$

$$\frac{d^2u}{d\alpha^2} = \frac{du'}{d\alpha} = [u', G] = [[u, G], G]$$

el proceso se repite para las sucesivas derivadas con lo cual la expansión de Taylor queda

$$u(\alpha) = u_0 + \alpha [u, G]_0 + \frac{\alpha^2}{2!} [[u, G], G]_0 + \frac{\alpha^3}{3!} [[[u, G], G], G]_0 + \dots \quad (8.65)$$

en particular, si  $u$  representa una variable canónica  $\zeta_i$ , con  $u_0$  representando el valor inicial de la variable  $\eta_i$ , la ecuación (8.65) nos da la prescripción para calcular la transformación canónica finita generada por  $G$ . Definimos el operador  $\widehat{G}$  como

$$\widehat{G} \equiv [(\dots), G] \quad (8.66)$$

donde  $\widehat{G}$  actúa sobre las funciones  $u$  de la configuración del sistema. Es decir

$$\widehat{G}u \equiv [u, G] \quad (8.67)$$



claramente

$$\widehat{G}^2 u = \widehat{G} (\widehat{G} u) = \widehat{G} ([u, G]) \equiv [[u, G], G]$$

y generalizando a todas las potencias de  $\widehat{G}$ , la expansión (8.65) se escribe

$$\begin{aligned} u(\alpha) &= \mathbf{1}u_0 + \alpha \widehat{G} u \Big|_0 + \frac{1}{2!} (\alpha \widehat{G})^2 u \Big|_0 + \frac{1}{3!} (\alpha \widehat{G})^3 u \Big|_0 + \dots \\ u(\alpha) &= \left[ \left\{ \mathbf{1} + \alpha \widehat{G} + \frac{1}{2!} (\alpha \widehat{G})^2 + \frac{1}{3!} (\alpha \widehat{G})^3 + \dots \right\} u \right]_0 \end{aligned} \quad (8.68)$$

teniendo en cuenta que la sumatoria entre llaves tiene la estructura de la serie correspondiente a la función exponencial, podemos escribir simbólicamente esta transformación de la forma

$$u(\alpha) = e^{\alpha \widehat{G}} u \Big|_0 \quad (8.69)$$

Otra forma interesante de obtener la Ec. (8.69) se basa en la función generatriz para  $e^x$

$$e^x = \lim_{n \rightarrow \infty} \left( 1 + \frac{x}{n} \right)^n$$

si extrapolamos esta expresión para la función de operadores tendremos que

$$e^{\alpha \widehat{G}} = \lim_{n \rightarrow \infty} \left( \widehat{1} + \frac{\alpha \widehat{G}}{n} \right)^n$$

y la ecuación (8.69) se puede reescribir como

$$u(\alpha) = \lim_{n \rightarrow \infty} \left( \widehat{1} + \frac{\alpha \widehat{G}}{n} \right)^n u \Big|_0 \quad (8.70)$$

en (8.70) se puede ver que el valor final para la función  $u(\alpha)$  se puede obtener a partir de su valor inicial, aplicando sucesivamente ( $n$  veces) el operador  $\widehat{1} + \alpha \widehat{G}/n$  con  $n \rightarrow \infty$ . En tal límite, este operador solo difiere infinitesimalmente de la identidad. Es decir la TC completa se está generando como una aplicación sucesiva de TCI's.

Un caso muy importante surge cuando en la Ec. (8.69) tomamos el caso particular en el cual  $G = H$  y el parámetro continuo es el tiempo, se obtiene

$$u(t) = e^{t \widehat{H}} u \Big|_0 \quad (8.71)$$

donde el **operador Hamiltoniano** se define en analogía con (8.66, 8.67)

$$\widehat{H} \equiv [(\dots), H] \quad ; \quad \widehat{H} u \equiv [u, H] \quad (8.72)$$

la ecuación (8.71) nos dice que la evolución temporal de  $u$  se obtiene aplicando el operador  $e^{t \widehat{H}}$  sobre  $u(q_i, p_i)$  y evaluando en  $t = 0$ . Por esta razón, a  $e^{t \widehat{H}}$  se le denomina **operador evolución temporal**.

Nótese que si el corchete de Poisson de  $u$  con el generador  $G$  de la transformación se anula (i.e.  $\widehat{G} u = 0$ ), las Ecs. (8.65, 8.68) nos muestran que la transformación generada por  $G$  dejará invariante a la función  $u$ . Cuando  $G$  es el Hamiltoniano y el parámetro es el tiempo, esto es consistente con la Ec. (8.25), si recordamos que hemos asumido que  $u(q_i, p_i)$  no es función explícita del tiempo.

### 8.13.1. Aplicación del operador evolución temporal en el movimiento uniformemente acelerado

Un ejemplo sencillo de aplicación del operador evolución temporal definido en (8.71), se obtiene cuando estudiamos el problema del movimiento unidimensional con aceleración constante  $a$ , cuyo Hamiltoniano se escribe en la forma

$$H = \frac{p^2}{2m} - max$$

sea  $u = x$  la función que queremos conocer, los corchetes de Poisson nos dan

$$[x, H] = \frac{p}{m} \quad ; \quad [[x, H], H] = \frac{1}{m} [p, H] = a \quad (8.73)$$

y dado que este último corchete es constante, los corchetes de mayor orden se anulan, de modo que la serie definida en (8.65) se trunca y solo posee tres términos. Usando entonces  $G \rightarrow H$ ,  $\alpha \rightarrow t$ , y  $u \rightarrow x$  en (8.65) así como los corchetes (8.73) se obtiene

$$x(t) = x_0 + \frac{p_0}{m}t + \frac{at^2}{2}$$

### 8.13.2. Aplicación del operador evolución temporal en el movimiento armónico simple

Escribiendo explícitamente el operador evolución temporal definido por las Ecs. (8.71, 8.72) tenemos

$$u(t) = \left[ \left\{ \mathbf{1} + t \hat{H} + \frac{1}{2!} (t \hat{H})^2 + \frac{1}{3!} (t \hat{H})^3 + \dots \right\} u \right]_{t=0} \quad (8.74)$$

$$\hat{H} \equiv [(\dots), H] \quad ; \quad \hat{H} u \equiv [u, H] \quad (8.75)$$

El Hamiltoniano del oscilador armónico lineal está dado por

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \quad ; \quad \frac{\partial H}{\partial p} = \frac{p}{m} \quad ; \quad \frac{\partial H}{\partial x} = kx$$

vamos a evaluar la evolución temporal de  $x(t)$ , de modo que haremos  $u \equiv x$ . Evaluemos las sucesivas potencias de la forma  $\hat{H}^n x$

$$\begin{aligned} \hat{H}x &= [x, H] = \frac{\partial x}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial x}{\partial p} \frac{\partial H}{\partial x} = \frac{1}{m}p \\ \hat{H}^2x &= [[x, H], H] = \frac{\partial (\hat{H}x)}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial (\hat{H}x)}{\partial p} \frac{\partial H}{\partial x} = \frac{\partial (\frac{p}{m})}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial (\frac{p}{m})}{\partial p} \frac{\partial H}{\partial x} \\ \hat{H}^2x &= -\frac{k}{m}x \\ \hat{H}^3x &= \frac{\partial (\hat{H}^2x)}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial (\hat{H}^2x)}{\partial p} \frac{\partial H}{\partial x} = \frac{\partial (-\frac{k}{m}x)}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial (-\frac{k}{m}x)}{\partial p} \frac{\partial H}{\partial x} = -\frac{1}{m} \left( \frac{k}{m} \right) p \\ \hat{H}^4x &= \frac{\partial (-\frac{k}{m^2}p)}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial (-\frac{k}{m^2}p)}{\partial p} \frac{\partial H}{\partial x} = \frac{k}{m^2}kx = \left( \frac{k}{m} \right)^2 x \\ \hat{H}^5x &= \frac{\partial \left[ \left( \frac{k}{m} \right)^2 x \right]}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial \left[ \left( \frac{k}{m} \right)^2 x \right]}{\partial p} \frac{\partial H}{\partial x} = \left( \frac{k}{m} \right)^2 \frac{p}{m} = \frac{1}{m} \left( \frac{k}{m} \right)^2 p \\ \hat{H}^6x &= \frac{\partial \left[ \frac{1}{m} \left( \frac{k}{m} \right)^2 p \right]}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial \left[ \frac{1}{m} \left( \frac{k}{m} \right)^2 p \right]}{\partial p} \frac{\partial H}{\partial x} = -\frac{1}{m} \left( \frac{k}{m} \right)^2 kx = -\left( \frac{k}{m} \right)^3 x \end{aligned} \quad (8.76)$$

al colocarlas juntas teniendo en cuenta que  $\widehat{H}^0 \equiv \mathbf{1}$ , tenemos

$$\begin{aligned}\widehat{H}^0 x &= \left(-\frac{k}{m}\right)^0 x ; \widehat{H}^2 x = \left(-\frac{k}{m}\right) x ; \widehat{H}^4 x = \left(-\frac{k}{m}\right)^2 x ; \widehat{H}^6 x = \left(-\frac{k}{m}\right)^3 x \\ \widehat{H} x &= \frac{1}{m} \left(-\frac{k}{m}\right)^0 p ; \widehat{H}^3 x = \frac{1}{m} \left(-\frac{k}{m}\right) p ; \widehat{H}^5 x = \frac{1}{m} \left(-\frac{k}{m}\right)^2 p\end{aligned}$$

lo cual sugiere un patrón para potencias pares y otro para potencias impares en  $\widehat{H}^n x$  de la forma

$$\widehat{H}^{2n} x = \left(-\frac{k}{m}\right)^n x , \widehat{H}^{2n+1} x = \frac{1}{m} \left(-\frac{k}{m}\right)^n p ; n = 0, 1, 2, \dots \quad (8.77)$$

las Ecs. (8.77) pueden demostrarse rigurosamente por inducción usando la Ec. (8.76). Hagámoslo para la primera identidad en (8.77)

$$\begin{aligned}\widehat{H}^{2n+2} x &= \widehat{H}^2 (\widehat{H}^{2n} x) = \widehat{H}^2 \left[ \left(-\frac{k}{m}\right)^n x \right] = \left(-\frac{k}{m}\right)^n (\widehat{H}^2 x) = \left(-\frac{k}{m}\right)^n \left(-\frac{k}{m}\right) x = \left(-\frac{k}{m}\right)^{n+1} x \\ \Rightarrow \widehat{H}^{2(n+1)} x &= \left(-\frac{k}{m}\right)^{n+1} x\end{aligned}$$

y similarmente para la segunda de las Ecs. (8.77). Haciendo  $u \rightarrow x$  en (8.74), y separando potencias pares e impares, resulta

$$\begin{aligned}x(t) &= \left[ \left\{ \mathbf{1} + \frac{1}{2!} (t\widehat{H})^2 + \frac{1}{4!} (t\widehat{H})^4 + \dots \right\} x \right]_0 + \left[ \left\{ t\widehat{H} + \frac{1}{3!} (t\widehat{H})^3 + \frac{1}{5!} (t\widehat{H})^5 + \dots \right\} x \right]_0 \\ x(t) &= \sum_{n=0}^{\infty} \left[ \frac{1}{2n!} (t\widehat{H})^{2n} x \right]_0 + \sum_{n=0}^{\infty} \left[ \frac{1}{(2n+1)!} (t\widehat{H})^{2n+1} x \right]_0\end{aligned} \quad (8.78)$$

e insertando (8.77) en (8.78) y definiendo  $k/m \equiv \omega^2$  se obtiene

$$\begin{aligned}x(t) &= \sum_{n=0}^{\infty} \left[ \frac{t^{2n}}{2n!} \left(-\frac{k}{m}\right)^n x \right]_0 + \sum_{n=0}^{\infty} \left[ \frac{t^{2n+1}}{(2n+1)!} \frac{1}{m} \left(-\frac{k}{m}\right)^n p \right]_0 \\ x(t) &= \sum_{n=0}^{\infty} \left[ \frac{(-1)^n}{2n!} \left(\sqrt{\frac{k}{m}} t\right)^{2n} x \right]_0 + \frac{1}{m} \sum_{n=0}^{\infty} \left[ \frac{(-1)^n}{(2n+1)!} \sqrt{\frac{m}{k}} \left(\sqrt{\frac{k}{m}} t\right)^{2n+1} p \right]_0 \\ x(t) &= x_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} (\omega t)^{2n} + \frac{p_0 \omega}{m} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} (\omega t)^{2n+1}\end{aligned}$$

las series corresponden a las funciones coseno y el seno de modo que

$$x(t) = x_0 \cos \omega t + \frac{p_0 \omega}{m} \sin \omega t \quad (8.79)$$

### 8.13.3. Aplicación del operador evolución paramétrica para la generación de rotaciones

Es importante tener presente que la ecuación general (8.69) nos da la evolución del sistema para un parámetro arbitrario  $\alpha$  que no es necesariamente el tiempo. Para ilustrar este hecho hagamos la asignación  $G \rightarrow L_z$  que corresponde a una transformación canónica que produce una rotación alrededor de  $Z$ . Es natural entonces tomar el parámetro en la forma  $\alpha \rightarrow \theta$  siendo  $\theta$  un ángulo de rotación. La función  $u$  será la coordenada  $x$  de la partícula  $i$ -ésima. La evolución de la coordenada  $x_i$  con  $\theta$  se obtiene haciendo los reemplazos en (8.69) o en (8.65)

$$x_i(\theta) = x_{i0} + \theta [x_i, L_z]_0 + \frac{\theta^2}{2!} [[x_i, L_z], L_z]_0 + \frac{\theta^3}{3!} [[[x_i, L_z], L_z], L_z]_0 + \dots \quad (8.80)$$

Los corchetes de Poisson se pueden evaluar en forma directa o por sus propiedades, usando la expresión (8.60) para  $G = L_z$ , así como los corchetes fundamentales (8.3) para obtener

$$[x_i, L_z] = -y_i \quad ; \quad [y_i, L_z] = x_i$$

y los corchetes sucesivos dan

$$[x_i, L_z] = -y_i \quad ; \quad [[x_i, L_z], L_z] = -[y_i, L_z] = -x_i \quad ; \quad [[[x_i, L_z], L_z], L_z] = -[x_i, L_z] = y_i \quad (8.81)$$

etc. Ahora los corchetes de Poisson se evalúan en el valor inicial  $\theta = 0$  del parámetro de evolución. Reemplazando (8.81) en (8.80) evaluado en  $\theta = 0$  obtenemos

$$\begin{aligned} x_i(\theta) &= x_{i0} - y_{i0}\theta - x_{i0}\frac{\theta^2}{2!} + y_{i0}\frac{\theta^3}{3!} + x_{i0}\frac{\theta^4}{4!} - \dots \\ x_i(\theta) &= x_{i0} \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots \right) - y_{i0} \left( \theta - \frac{\theta^3}{3!} + \dots \right) \\ x_i(\theta) &= x_{i0} \cos \theta - y_{i0} \sin \theta \end{aligned}$$

que corresponde justo a la transformación que genera una rotación alrededor de  $Z$ , se puede ver que las transformaciones de  $y_i$  y  $z_i$  también son las adecuadas (la última no se transforma). Esto corresponde a una prueba alternativa de que  $L_z$  es un generador de rotaciones alrededor de  $z$  y como  $z$  es arbitrario esto indica que el generador  $\mathbf{L} \cdot \mathbf{n}$  produce una rotación alrededor de  $\mathbf{n}$ . Este hecho nos inspira a encontrar las propiedades de los corchetes de Poisson de los momentos angulares en forma más sistemática.

## 8.14. Propiedades de los corchetes de Poisson de los momentos angulares

Las Ecs. (8.55) y (8.57) nos llevan a ver que los momentos  $p_i$  conjugados a una coordenada  $q_i$  pueden actuar como generadores de traslaciones generalizadas en donde solo hay variación de la coordenada  $q_i$  en el espacio de fase. Con base en ello, podemos ver que el momento angular al estar asociado a una variable angular, podría ser el generador de una “traslación angular” i.e. una rotación. Lo cual está respaldado por los resultados de las secciones 5.1 y 8.12.2. El hecho de que el momento angular canónico sea el generador de rotaciones rígidas del sistema conduce a una serie de importantes propiedades de los corchetes de Poisson que involucran a estas cantidades. La Ec. (8.49) que nos determina el cambio de una función del sistema  $u$ , ante una transformación canónica en una visión activa, es también válida si dicha función es la componente de un vector a lo largo de un eje **fijo** en el espacio ordinario. Por lo tanto si  $\mathbf{F}$  es una función vectorial de la configuración del sistema, se tiene que

$$\partial F_i = d\alpha [F_i, G]$$

es importante que la dirección a lo largo de la cual se calcula la componente sea fija en el sentido de que no debe cambiar con la TC. Si la dirección como tal es función de las variables del sistema, la TC no solo cambia el valor de la función sino también su naturaleza, tal como ocurre con el Hamiltoniano. Con esta aclaración el cambio de un vector  $\mathbf{F}$  bajo una rotación del sistema alrededor de un eje fijo  $\mathbf{n}$ , generada por  $\mathbf{L} \cdot \mathbf{n}$  se escribe vectorialmente como

$$\partial \mathbf{F} = d\theta [\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] \quad (8.82)$$

la Ec. (8.82) es válida solo si los vectores unitarios  $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$  que forman la base para escribir  $\mathbf{F}$ , no son rotadas por el generador  $\mathbf{L} \cdot \mathbf{n}$ .

Otra aclaración muy importante sobre el significado de la Ec. (8.82), el generador  $\mathbf{L} \cdot \mathbf{n}$  produce una rotación del sistema bajo la TCI, pero no necesariamente produce una rotación del vector  $\mathbf{F}$ . Dicho generador induce una rotación espacial de las variables del sistema pero no por ejemplo de algún vector externo tal como un campo magnético externo o el vector que describe la aceleración de la gravedad. Surge naturalmente la pregunta de cuales son las condiciones bajo las cuales  $\mathbf{L} \cdot \mathbf{n}$  genera una rotación espacial de  $\mathbf{F}$ . Claramente, este vector es rotado por el generador en cuestión cuando  $\mathbf{F}$  es una función de las variables del sistema  $(q, p)$

únicamente y no involucra ninguna cantidad externa o vector que no sea afectado por la TCI. Bajo estas condiciones una rotación espacial del sistema implica una rotación de  $\mathbf{F}$ . Los vectores que cumplen estas condiciones serán denominados **vectores del sistema**. El cambio en un vector Euclidiano bajo rotaciones infinitesimales alrededor de  $\mathbf{n}$ , fué discutido en la sección 5.1, Ec. (5.9)

$$d\mathbf{F} = \mathbf{n} d\theta \times \mathbf{F}$$

de tal modo que para un vector del sistema, una TCI generada por  $\mathbf{L} \cdot \mathbf{n}$  induce un cambio en dicho vector de la forma<sup>12</sup>

$$\partial\mathbf{F} = d\theta [\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} d\theta \times \mathbf{F} \quad (8.83)$$

la Ec. (8.83) conduce a una importante identidad concerniente al corchete de Poisson entre  $\mathbf{L} \cdot \mathbf{n}$  y un vector del sistema

$$[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{F} \quad (8.84)$$

La condición (8.84) ya no hace ninguna referencia a una TC o incluso a una rotación espacial. Es simplemente una propiedad del corchete de Poisson entre un cierto tipo de vectores (del sistema) y una componente del momento angular canónico en cierta dirección.

La relación (8.84) se puede expresar en componentes usando el tensor de Levi Civitá.

$$[F_i, L_j n_j] = \varepsilon_{ijk} n_j F_k$$

y dado que  $\mathbf{n}$  es arbitrario, podemos elegir en particular que  $n_j = 1$  para un  $j$  específico y cero para las otras componentes, lo cual nos conduce a

$$[F_i, L_j] = \varepsilon_{ijk} F_k \quad (8.85)$$

lo cual se puede escribir también de la forma

$$[F_l, L_m] = F_n \quad l, m, n \text{ en orden cíclico} \quad (8.86)$$

otra relación interesante es ver lo que ocurre con el producto  $\mathbf{F} \cdot \mathbf{G}$  entre dos vectores del sistema. Siendo un escalar, este producto punto debe ser invariante bajo rotaciones<sup>13</sup>. La identidad (8.84) nos muestra que efectivamente, el corchete de Poisson entre este producto y  $\mathbf{L} \cdot \mathbf{n}$  se anula.

$$\begin{aligned} [\mathbf{F} \cdot \mathbf{G}, \mathbf{L} \cdot \mathbf{n}] &= \mathbf{F} \cdot [\mathbf{G}, \mathbf{L} \cdot \mathbf{n}] + [\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] \cdot \mathbf{G} = \mathbf{F} \cdot (\mathbf{n} \times \mathbf{G}) + (\mathbf{n} \times \mathbf{F}) \cdot \mathbf{G} \\ &= \mathbf{F} \cdot (\mathbf{n} \times \mathbf{G}) + \mathbf{F} \cdot (\mathbf{G} \times \mathbf{n}) \Rightarrow \\ [\mathbf{F} \cdot \mathbf{G}, \mathbf{L} \cdot \mathbf{n}] &= 0 \end{aligned} \quad (8.87)$$

recordemos que la anulación del corchete de Poisson de cierta función del sistema con el generador de una transformación, nos conduce a la invarianza de esta función bajo la transformación producida por el generador como se vé en la Ec. (8.65)<sup>14</sup>. Haciendo  $\mathbf{F} = \mathbf{G}$ , se obtiene la propiedad esperada de que la magnitud de un vector se conserva bajo rotaciones, ya que el corchete de Poisson de  $\mathbf{F}^2$  se anularía para cualquier componente de  $\mathbf{L}$ .

Naturalmente, todas las relaciones anteriores son válidas en particular cuando  $\mathbf{F} = \mathbf{L}$  en (8.84). En tal caso resulta

$$[\mathbf{L}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{L} \Leftrightarrow [L_i, L_j] = \varepsilon_{ijk} L_k \quad (8.88)$$

<sup>12</sup>Un vector podría ser mixto i.e. función de  $q$  y  $p$  del sistema pero también función de variables externas. En tal caso es de esperarse que dicho vector se transforme bajo la rotación generada por la TCI pero no transformará de la forma prescrita por la Ec. (8.83).

<sup>13</sup>Esto se puede ver de  $\mathbf{F} \cdot \mathbf{G} = FG \cos \theta$ .  $F$  y  $G$  son invariantes ya que una rotación no cambia la magnitud de los vectores, además el ángulo relativo entre dos vectores también se conserva en una rotación de modo que la invarianza es clara.

<sup>14</sup>Recordemos que esto solo vale si la función del sistema no depende explícitamente del tiempo. En realidad, si tal función depende explícitamente del tiempo, significa que ésta puede cambiar con el tiempo incluso si la configuración del sistema permanece sin alterar. Esto solo es posible si existe algún agente externo que produzca dicho cambio. En conclusión, una función  $u(q_i, p_i, t)$  no sería una función solo del sistema, sino más bien una función mixta que depende del sistema y de algún agente externo. Razonamiento idéntico se da para vectores  $\mathbf{F}$ , del sistema o mixtos.

adicionalmente, haciendo  $\mathbf{F} = \mathbf{G} = \mathbf{L}$  en (8.87) resulta

$$[\mathbf{L}^2, \mathbf{L} \cdot \mathbf{n}] = 0 \quad (8.89)$$

de las relaciones (8.88) y el teorema de Poisson vemos que si  $L_x$  y  $L_y$  son constantes de movimiento entonces  $L_z$  también es una constante de movimiento. Por tanto, la invarianza de dos componentes del momento angular canónico cartesiano nos lleva a la invarianza del vector  $\mathbf{L}$  completo. Adicionalmente, aplicando las relaciones (8.84) y (8.85) con  $\mathbf{F} = \mathbf{p}$ , es decir el momento canónico cartesiano, tenemos que

$$[\mathbf{p}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{p} \quad ; \quad [p_i, L_j] = \varepsilon_{ijk} p_k$$

si ahora asumimos que además de  $L_x, L_y$  la componente  $p_z$  es también una constante de movimiento, entonces además de  $L_z$  obtenemos otras dos constantes de movimiento via teorema de Poisson

$$[p_z, L_x] = p_y \quad , \quad [p_z, L_y] = -p_x$$

de modo que tanto  $\mathbf{L}$  como  $\mathbf{p}$  se conservan. En general si  $L_i, L_j$  y  $p_k$  son constantes de movimiento con  $i, j, k$  diferentes entre sí, obtenemos la invarianza de  $\mathbf{L}$  y  $\mathbf{p}$ . Este es un caso en el cual el teorema de Poisson es útil para encontrar nuevas constantes de movimiento.

En contraste, examinemos un escenario en el cual asumimos que  $p_x, p_y, L_z$  son constantes de movimiento, la aplicación del teorema de Poisson nos da

$$[p_x, p_y] = 0 \quad ; \quad [p_x, L_z] = -p_y \quad ; \quad [p_y, L_z] = p_x$$

y el teorema de Poisson no nos genera en este caso nuevas constantes de movimiento.

Por otro lado, las propiedades (8.3), nos dicen que para cualquier par de momentos canónicos, el corchete de Poisson entre ellos se anula. Pero de las relaciones (8.88) se observa que para ningún par de componentes de  $\mathbf{L}$ , se anula el corchete entre ellas. Por tanto, a pesar de que describimos a  $\mathbf{L}$  como el momento angular canónico total en virtud de su definición como  $\mathbf{r}_i \times \mathbf{p}_i$  (suma sobre índices), no es posible escoger simultáneamente a dos componentes de  $\mathbf{L}$  como momentos canónicos. Nótese que tampoco es posible por ejemplo que  $L_x$  sea una coordenada  $q_i$  y  $L_y$  un momento canónico  $p_k$  ya que en este caso tenemos que

$$[q_i, p_k] = [L_x, L_y] = L_z \neq \delta_{ik}$$

que de nuevo contradice las Ecs. (8.3). Similarmente no es posible que por ejemplo  $L_x$  y  $L_y$  sean ambas coordenadas  $q_i$  y  $q_k$ . En síntesis, no podemos escoger simultáneamente dos componentes de  $\mathbf{L}$  como variables canónicas. Sin embargo, la Ec. (8.89) muestra que es posible escoger simultáneamente a la magnitud al cuadrado de  $\mathbf{L}$ , y a cualquiera de las componentes de  $\mathbf{L}$  (pero solo una), como variables canónicas.

### 8.14.1. Ejemplos de aplicación

#### Corchetes de momentos angulares para partícula libre

La relación (8.84) se puede verificar explícitamente para el caso particular de una partícula libre descrita por coordenadas y momentos cartesianos. El vector momento cartesiano  $\mathbf{p}$ , es claramente un vector del sistema. Tomemos  $\mathbf{n} = \mathbf{u}_z$ , evaluemos directamente el corchete de la izquierda en (8.84), lo haremos componente a componente

$$\begin{aligned} [p_x, L_z] &= [p_x, xp_y - yp_x] = [p_x, xp_y] - [p_x, yp_x] = x[p_x, p_y] + [p_x, x]p_y - y[p_x, p_x] - [p_x, y]p_x \\ &\Rightarrow [p_x, L_z] = -p_y \end{aligned}$$

donde hemos usado las propiedades (8.10, 8.11) y las identidades (8.3). Similarmente

$$[p_y, xp_y - yp_x] = p_x \quad ; \quad [p_z, xp_y - yp_x] = 0$$

de modo que al evaluar la expresión de la izquierda en (8.84) para este caso resulta

$$[\mathbf{p}, \mathbf{L} \cdot \mathbf{u}_z] = -p_y \mathbf{u}_x + p_x \mathbf{u}_y$$

este corchete nos da las componentes de  $\mathbf{u}_z \times \mathbf{p}$ , lo cual confirma la validez de la Ec. (8.84) para este caso particular. Adicionalmente, para este problema el lector puede verificar explícitamente la validez de las siguientes expresiones

$$[\mathbf{p} \cdot \mathbf{L}, \mathbf{L} \cdot \mathbf{n}] = [\mathbf{p}^2, \mathbf{L} \cdot \mathbf{n}] = [\mathbf{L}^2, \mathbf{L} \cdot \mathbf{n}] = 0$$

que nos confirman que estas cantidades son escalares, es decir que no son afectadas por una rotación, en concordancia con (8.87).

### Corchetes de momentos angulares para partícula en campo externo

Supongamos que una partícula se somete a un campo magnético externo uniforme. Tomemos el vector  $\mathbf{F} = \mathbf{A}$  siendo  $\mathbf{A} \equiv \frac{1}{2}(\mathbf{r} \times \mathbf{B})$  donde  $\mathbf{B} = B\mathbf{u}_x$  describe al campo magnético.  $\mathbf{A}$  representa físicamente un potencial vectorial asociado al campo magnético  $\mathbf{B}$ . El vector  $\mathbf{A}$  depende de un vector interno y otro externo al sistema, por tanto es de esperarse que en general se transforme bajo la TCI generada por  $\mathbf{L} \cdot \mathbf{u}_z$  pero no necesariamente de la forma prescrita por (8.84). La evaluación explícita de los corchetes nos da

$$\begin{aligned} \mathbf{A} &= \frac{1}{2}(\mathbf{r} \times \mathbf{B}) = \frac{1}{2}(x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z) \times B\mathbf{u}_x = -\frac{1}{2}yB\mathbf{u}_z + \frac{1}{2}zB\mathbf{u}_y \\ [F_x, L_z] &= [A_x, L_z] = [0, L_z] = 0 \quad ; \quad [F_y, L_z] = [A_y, L_z] = \left[ \frac{1}{2}zB, xp_y - yp_x \right] = 0 \\ [F_z, L_z] &= [A_z, L_z] = \left[ -\frac{1}{2}yB, xp_y - yp_x \right] = -\frac{1}{2}Bx \end{aligned}$$

por tanto el miembro izquierdo de la Ec. (8.84) nos da

$$[\mathbf{A}, \mathbf{L} \cdot \mathbf{u}_z] = -\frac{1}{2}Bx\mathbf{u}_z \quad (8.90)$$

en cuanto al miembro derecho de (8.84), el vector  $\mathbf{u}_z \times \mathbf{A}$  nos da  $-\frac{1}{2}Bz\mathbf{u}_x$ . Efectivamente, vemos que el vector  $\mathbf{A}$  se transforma bajo una rotación del sistema (ya que el corchete de la Ec. 8.90 no es nulo), pero no se transforma con la prescripción dada por (8.84).

## 8.15. Ejercicios

1. Demostrar que las transformaciones de escala, así como las transformaciones canónicas extendidas, no dejan invariantes a los corchetes de Poisson.
2. Demostrar las propiedades (8.9, 8.10) de los corchetes de Poisson.
3. Verifique que se cumple la relación (8.15) entre los corchetes de Poisson y los corchetes de Lagrange.
4. Utilizando corchetes de Poisson, demuestre que las transformaciones definidas en la Ec. (7.94), Pág. 138, son canónicas.
5. Utilizando corchetes de Poisson, demuestre que las transformaciones definidas en la Ec. (7.95), Pág. 139, son canónicas.
6. Demuestre en detalle que la Ec. (8.79) es la solución conocida del movimiento armónico simple en términos de las condiciones iniciales.

## Teoría de Hamilton-Jacobi y variables acción-ángulo

Ya hemos mencionado que las transformaciones canónicas pueden ser usadas como un procedimiento para resolver las ecuaciones de movimiento, para lo cual se sugirieron dos estrategias, (a) Cuando el Hamiltoniano se conserva, se puede obtener la solución transformando a un nuevo conjunto de variables canónicas en donde todas las coordenadas son cíclicas, en cuyo caso las nuevas ecuaciones de movimiento de Hamilton resultan triviales, y (b) encontrar una transformación canónica que partiendo de las coordenadas  $q, p$  que describen la configuración del sistema en un tiempo  $t$ , nos lleve a un nuevo sistema coordinado constituido únicamente por cantidades constantes. En particular, las constantes que constituyen el nuevo sistema de variables canónicas se puede elegir como el conjunto  $q_0, p_0$  que coincide con los valores iniciales de las variables canónicas. Si encontramos esta TC, la correspondiente transformación entre las antiguas y las nuevas coordenadas se puede invertir para obtener

$$q = q(q_0, p_0, t) \quad ; \quad p = p(q_0, p_0, t)$$

lo cual nos da el valor de las coordenadas y momentos en función de los valores iniciales de éstos y el tiempo. Este procedimiento es el más general y es aplicable (al menos formalmente) incluso si el Hamiltoniano es función del tiempo. La teoría de Hamilton Jacobi es un procedimiento sistemático para encontrar la transformación canónica que lleva de los valores  $q, p$  en un tiempo  $t$ , a los valores iniciales o a otros valores constantes que son funciones de las condiciones iniciales.

### 9.1. Ecuación de Hamilton-Jacobi para la función principal de Hamilton

Podemos asegurar que las nuevas variables canónicas son todas constantes si el Hamiltoniano transformado  $K$  es idénticamente nulo, ya que en tal caso las ecuaciones de Hamilton quedan

$$K = 0 \quad \Rightarrow \quad \frac{\partial K}{\partial P_i} = \dot{Q}_i = 0 \quad ; \quad -\frac{\partial K}{\partial Q_i} = \dot{P}_i = 0 \quad (9.1)$$

como ya se ha visto, el antiguo y el nuevo Hamiltoniano están relacionados a través de la función generatriz

$$K = H + \frac{\partial F}{\partial t} \quad (9.2)$$

por tanto el nuevo Hamiltoniano será nulo si  $F$  satisface la ecuación

$$H(q, p, t) + \frac{\partial F}{\partial t} = 0 \quad (9.3)$$

tomaremos a  $F$  como una función de tipo 2 de acuerdo con la clasificación hecha en el capítulo 7<sup>1</sup>. Por tanto,  $F$  es función de  $q_i, P_i$  y el tiempo. Con el fin de escribir el Hamiltoniano como función de las mismas variables

<sup>1</sup>Recordemos que con una función tipo  $F_2$  es fácil contruir una T.C. continúa, dado que una T.C. que difiere infinitesimalmente de la identidad es inmediata para este tipo de función generatriz.



que  $F$ , debemos usar las ecuaciones de transformación (7.19)

$$p_i = \frac{\partial F_2}{\partial q_i} \quad (9.4)$$

de modo que la Ec. (9.3) se escribe

$$H \left( q_1 \dots q_n; \frac{\partial F_2}{\partial q_1}, \dots, \frac{\partial F_2}{\partial q_n}; t \right) + \frac{\partial F_2}{\partial t} = 0 \quad (9.5)$$

la ecuación (9.5) conocida como **ecuación de Hamilton Jacobi**, es una ecuación diferencial parcial en  $n + 1$  variables ( $n$  coordenadas y el tiempo), para la función generatriz que buscamos. Es costumbre denotar las soluciones de  $F_2$  como  $S$  por razones que veremos más adelante, y dicha función  $S$  se denomina **función principal de Hamilton**.

La integración de la Ecuación de Hamilton-Jacobi Ec. (9.5) solo nos provee la dependencia con las coordenadas  $q$  y el tiempo; no aparece ninguna información sobre los nuevos momentos. En realidad, lo único que sabemos es que todos los nuevos momentos deben ser constantes pero no hemos especificado que valores deben tener. Volveremos sobre la determinación de los nuevos momentos más adelante. Matemáticamente, la Ec. (9.5) es una ecuación diferencial parcial de primer orden en  $n + 1$  variables ( $q, t$ ). Un tipo de solución posible de esta ecuación es

$$F_2 \equiv S = S(q_1 \dots q_n; \alpha_1, \dots, \alpha_{n+1}; t) \quad (9.6)$$

donde  $\alpha_i$  son  $n + 1$  constantes de integración independientes. Estas soluciones se conocen como soluciones completas de la ecuación diferencial parcial<sup>2</sup>. No obstante, una de las constantes de movimiento es irrelevante debido a que en todas las ecuaciones en donde aparece  $F_2$ , solo aparece su derivada parcial con respecto a alguna variable (y en particular en la Ec. 9.5). Por tanto si  $F_2$  es una solución,  $F_2 + \alpha$  también es solución de modo que una de las constantes de integración debe ser aditiva en  $S$ . Por tanto, la solución puede remover una constante sin ninguna pérdida de generalidad de modo que podemos escribir

$$S = S(q_1 \dots q_n; \alpha_1, \dots, \alpha_n; t) \quad (9.7)$$

donde ninguna de las constantes independientes  $n$  es solamente aditiva. Obsérvese que (9.7) debe ser de la estructura de una función  $F_2$  i.e.  $S = S(q, P, t)$ . La manera más natural de llegar a esta estructura con (9.7), es asumiendo que

$$P_i = \alpha_i \quad (9.8)$$

veremos enseguida que esta suposición no contradice la aseveración original de que los nuevos momentos están conectados con los valores iniciales  $q(t_0)$  y  $p(t_0)$ .

Describiremos ahora como se obtiene la solución dinámica del problema con base en el método de Hamilton-Jacobi. Una vez que se obtiene una solución  $F_2 \equiv S$ , de la ecuación de Hamilton-Jacobi (9.5), llevamos esta solución a las  $n$  primeras ecuaciones de transformación (7.19) de la forma

$$p_i = \frac{\partial S(q, \alpha, t)}{\partial q_i} \quad (9.9)$$

cuando evaluamos estas  $n$  ecuaciones en  $t = t_0$ , las únicas incógnitas son las  $\alpha$ 's ya que las cantidades  $q, p$  adquieren sus valores iniciales. Por tanto las Ecs. (9.9) evaluadas en  $t = t_0$ , son  $n$  ecuaciones con  $n$  incógnitas  $\alpha$ , que nos permiten al menos formalmente, encontrar las  $\alpha$  en función de los valores iniciales de  $q, p$ . Las otras  $n$  ecuaciones asociadas a  $F_2$  (ver Ecs. 7.19) se escriben

$$Q_i \equiv \beta_i = \frac{\partial S(q, \alpha, t)}{\partial \alpha_i} \quad (9.10)$$

---

<sup>2</sup>Este no es el único tipo de solución posible, la solución mas general incluye una o más funciones arbitrarias en lugar de constantes. Incluso exigiendo que dichas funciones sean constantes hay en general mas de una solución del tipo (9.6). Sin embargo, lo importante es que exista al menos una solución, pues la unicidad de la misma no es necesaria ya que la función principal de Hamilton no representa un observable. De hecho, la falta de unicidad de la solución se debe a la ausencia de condiciones iniciales y/o de frontera sobre las funciones  $F_2 \equiv S$ .

donde hemos tenido en cuenta que los nuevos momentos los hemos elegido como las  $\alpha$ 's y que las nuevas coordenadas también deben ser constantes que denotamos por  $\beta_i$ . Una vez conocidos los valores de las  $\alpha_i$  a través de las Ecs. (9.9), podemos reemplazar estos valores en (9.10) para obtener las  $\beta_i$  haciendo  $t = t_0$  en dichas ecuaciones y apelando a las condiciones iniciales en  $q$ . Ahora bien, en las ecuaciones (9.10) podemos despejar  $q_j$  para obtener las coordenadas en función de  $\alpha, \beta, t$

$$q_j = q_j(\alpha, \beta, t) \quad (9.11)$$

lo cual resuelve la dinámica de las coordenadas generalizadas  $q_j$  en función del tiempo y de las condiciones iniciales. Después de derivar en la Ec. (9.9) se pueden tomar las  $q_i$  encontradas en (9.11) para sustituirlas en (9.9), con lo cual se encuentra el valor de cada  $p_i$  en función de  $\alpha, \beta, t$

$$p_i = p_i(\alpha, \beta, t) \quad (9.12)$$

por tanto, las Ecs. (9.11, 9.12) constituyen las soluciones completas a las ecuaciones de Hamilton.

La función principal de Hamilton es en consecuencia generadora de una transformación canónica que nos lleva a valores constantes de las nuevas coordenadas y los nuevos momentos. Cuando se resuelve la ecuación de Hamilton-Jacobi (9.5), obtenemos al mismo tiempo una solución al problema mecánico. Matemáticamente hemos establecido una equivalencia entre las  $2n$  ecuaciones canónicas de movimiento que son  $2n$  ecuaciones diferenciales ordinarias de primer orden, con la ecuación diferencial parcial de primer orden de Hamilton Jacobi. Esta clase de equivalencia es muy común en la teoría de ecuaciones diferenciales de primer orden. Esencialmente la conexión proviene del hecho de que ambas formulaciones provienen del principio variacional de Hamilton modificado (recordemos que la Ec. 9.2 proviene del principio de Hamilton modificado). Hasta cierto punto, la asignación de las  $\alpha_i$  como los nuevos momentos es arbitraria. Podríamos haber elegido como momentos otro conjunto de constantes  $\gamma_i$ , que son a su vez funciones independientes de las constantes de integración  $\alpha_i$

$$\gamma_i = \gamma_i(\alpha_1, \dots, \alpha_n) \quad (9.13)$$

por medio de estas relaciones, la función principal de Hamilton se puede escribir en función de  $q, \gamma, t$ , el resto de la derivación es análoga teniendo cuidado de reemplazar  $\alpha_i \rightarrow \gamma_i$  en (9.9) y (9.10). Es a veces conveniente definir ciertos  $\gamma_i$  como los nuevos momentos en lugar de las  $\alpha_i$  que aparecen como constantes de integración en (9.5).

Un caso interesante se da cuando el Hamiltoniano no depende explícitamente del tiempo y por tanto se conserva. En este caso la Ec. (9.5) muestra que  $\partial S/\partial t$  no puede ser función explícita del tiempo con lo cual la forma más general de  $S$  se escribe como

$$S(q, \alpha, t) = W(q, \alpha) - kt \quad (9.14)$$

donde  $W(q, \alpha)$  es llamada la **función característica de Hamilton**. Insertando (9.14) en (9.5) se observa que la constante  $k$  debe coincidir con el valor del Hamiltoniano (que es precisamente constante). Recuérdese que aún cuando el Hamiltoniano sea constante, no necesariamente corresponde a la energía del sistema.

### 9.1.1. Interpretación física de las soluciones de Hamilton-Jacobi

Podemos obtener cierta información física adicional sobre la función principal de Hamilton  $S(q, \alpha, t)$ , tomando su derivada total con el tiempo

$$\frac{dS(q, \alpha, t)}{dt} = \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t}$$

dado que  $\dot{\alpha} = 0$ . Usando las ecuaciones (9.9) y (9.5) se obtiene

$$\frac{dS}{dt} = p_i \dot{q}_i - H = L$$

con lo cual se puede escribir

$$S = \int L dt + cte \quad (9.15)$$

recordando que el principio de Hamilton es una aseveración sobre la integral definida de  $L$  en el tiempo, y que a partir de él se obtiene la solución del problema mecánico via ecuaciones de Lagrange, vemos que en este caso la misma integral de acción pero esta vez en forma indefinida, nos provee de otro método de solución. Nótese que la Ec. (9.15) no posee utilidad práctica puesto que para evaluar la integral en el tiempo se debe conocer  $L$  en función de  $t$  lo cual equivale a conocer  $q(t)$  y  $\dot{q}(t)$  es decir hay que saber la solución.

En forma similar, el significado físico de la función característica de Hamilton  $W$ , se puede revelar a través de su derivada total

$$\frac{dW}{dt} = \frac{\partial W}{\partial q_i} \dot{q}_i$$

Usando (9.9) y (9.14) se obtiene

$$\frac{dW}{dt} = p_i \dot{q}_i \Rightarrow W = \int p_i \dot{q}_i dt = \int p_i dq_i \quad (9.16)$$

que es precisamente la acción abreviada definida en (6.92) aunque en forma de integral indefinida.

## 9.2. Solución del oscilador armónico por el método de Hamilton-Jacobi

Una vez más, usaremos el oscilador armónico para ejemplificar la nueva técnica, que en este caso corresponde al método de Hamilton-Jacobi.

### 9.2.1. Oscilador armónico unidimensional con el método de Hamilton Jacobi

Partimos del Hamiltoniano para el oscilador armónico lineal unidimensional

$$H = \frac{1}{2m} (p^2 + m^2 \omega^2 q^2) \equiv E ; \quad \omega = \sqrt{\frac{k}{m}} \quad (9.17)$$

de acuerdo con la ecuación fundamental de Hamilton Jacobi Ec. (9.5), debemos sustituir a  $p$  por  $\partial S / \partial q$  en el Hamiltoniano (recordando que  $F_2$  se redefine como  $S$ ) y plantear la ecuación diferencial (9.5)

$$\frac{1}{2m} \left[ \left( \frac{\partial S}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] + \frac{\partial S}{\partial t} = 0$$

como este Hamiltoniano no depende explícitamente del tiempo, podemos escribir a  $S$  como (9.14) (donde usaremos  $\alpha$  en vez de  $k$ ). De esta forma se elimina la dependencia temporal de la ecuación de Hamilton Jacobi

$$\frac{1}{2m} \left[ \left( \frac{\partial W}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] = \alpha \quad (9.18)$$

la constante  $\alpha$  corresponde a la energía del sistema lo cual se puede ver reemplazando (9.14) en (9.3)

$$H + \frac{\partial S}{\partial t} = 0 \Rightarrow H - \alpha = 0$$

la ecuación (9.18) se puede integrar con lo cual se obtiene

$$W = \sqrt{2m\alpha} \int dq \sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}} \quad (9.19)$$

$$S = -\alpha t + \sqrt{2m\alpha} \int dq \sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}} \quad (9.20)$$

aunque la integral se puede resolver explícitamente, debemos recordar que  $S$  no aparece en las ecuaciones de movimiento sino solo sus derivadas parciales. Resulta más útil en este caso, hacer primero las derivaciones antes de integrar. La solución para  $q$  se obtiene de las ecuaciones (9.10)

$$\beta = \frac{\partial S}{\partial \alpha} = -t + \sqrt{\frac{m}{2\alpha}} \int \frac{dq}{\sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}}}$$

que al ser integrado nos da

$$t + \beta = \frac{1}{\omega} \arcsin \left[ q \sqrt{\frac{m\omega^2}{2\alpha}} \right]$$

la coordenada  $q$  se puede despejar de aquí para obtener a  $q$  en función del tiempo y de las dos constantes  $\alpha$  y  $\delta \equiv \omega\beta$

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega t + \delta) \quad ; \quad \alpha = H, \quad \delta \equiv \omega\beta \quad (9.21)$$

que es la solución que conocemos para el oscilador armónico. La solución para el momento conjugado se obtiene formalmente de las Ecs. (9.9) teniendo en cuenta (9.19) y (9.20)

$$p = \frac{\partial S}{\partial q} = \frac{\partial W}{\partial q} = \sqrt{2m\alpha - m^2\omega^2 q^2} \quad (9.22)$$

y sustituyendo (9.21) en (9.22) se obtiene

$$\begin{aligned} p &= \sqrt{2m\alpha [1 - \sin^2(\omega t + \delta)]} \\ p &= \sqrt{2m\alpha} \cos(\omega t + \delta) \end{aligned} \quad (9.23)$$

este resultado nos dice que  $p = m\dot{q}$  como se esperaba. Finalmente, debemos determinar las constantes  $\alpha$  y  $\delta$  a través de las condiciones iniciales  $q_0, p_0$ . El valor de  $\alpha$  se puede despejar fácilmente evaluando las Ecs. (9.21) y (9.23) en  $t = 0$  y sumando sus cuadrados

$$2m\alpha = p_0^2 + m^2\omega^2 q_0^2$$

esta misma ecuación se puede encontrar recordando que  $\alpha = H = E$  es constante, y usando la Ec. (9.17) en  $t = t_0$ . La fase  $\delta$  se obtiene de nuevo evaluando (9.21) y (9.23) en  $t = 0$  y haciendo el cociente

$$\tan \delta = m\omega \frac{q_0}{p_0}$$

En este caso vemos que la función principal de Hamilton  $S$  asumió el papel de generador de una transformación canónica a una nueva coordenada que mide el ángulo de fase inicial, y a un nuevo momento que se puede identificar como la energía total del sistema.

Ahora bien, si sustituimos la solución (9.21) para  $q$  en la Ec. (9.20) obtenemos

$$S = -\alpha t + 2\alpha \int \cos^2(\omega t + \delta) dt = 2\alpha \int \left[ \cos^2(\omega t + \delta) - \frac{1}{2} \right] dt$$

se puede verificar que el integrando coincide con el Lagrangiano

$$\begin{aligned} L &= \frac{1}{2m} (p^2 - m^2\omega^2 q^2) = \alpha [\cos^2(\omega t + \delta) - \sin^2(\omega t + \delta)] = 2\alpha \left[ \cos^2(\omega t + \delta) - \frac{1}{2} \right] \\ \Rightarrow S &= \int L dt \end{aligned}$$

y vemos que  $S$  efectivamente coincide con la integral indefinida del Lagrangiano con el tiempo, como lo prescribe la Ec. (9.15). Nótese que tal identidad solo se pudo comprobar una vez que se encontró la solución.

### 9.2.2. Oscilador armónico bidimensional anisotrópico

Para ilustrar problemas de más de un grado de libertad, tomemos un oscilador armónico bidimensional con constante de restitución diferente en  $x$  y en  $y$  es decir  $k_x \neq k_y$ , el Hamiltoniano se escribe en la forma

$$H = E = \frac{1}{2m} (p_x^2 + p_y^2 + m^2\omega_x^2 x^2 + m^2\omega_y^2 y^2) \quad ; \quad \omega_x = \sqrt{\frac{k_x}{m}} \quad ; \quad \omega_y = \sqrt{\frac{k_y}{m}} \quad (9.24)$$

como el Hamiltoniano no depende del tiempo la función principal de Hamilton queda

$$S(x, y, \alpha, \alpha_y, t) = W(x, y, \alpha, \alpha_y) - \alpha t$$

dado que las coordenadas y momentos en  $x$  e  $y$  se separan en forma de suma en el Hamiltoniano, es razonable suponer que la función característica de Hamilton  $W$  se separa de forma similar

$$S(x, y, \alpha, \alpha_y, t) = W_x(x, \alpha, \alpha_y) + W_y(y, \alpha, \alpha_y) - \alpha t \quad (9.25)$$

usando (9.24, 9.25) en la Ec. (9.5), obtenemos la ecuación de Hamilton-Jacobi para el oscilador bidimensional anisotrópico

$$\begin{aligned} \frac{1}{2m} \left[ \left( \frac{\partial W_x}{\partial x} \right)^2 + m^2\omega_x^2 x^2 + \left( \frac{\partial W_y}{\partial y} \right)^2 + m^2\omega_y^2 y^2 \right] &= \alpha \\ \frac{1}{2m} \left[ \left( \frac{\partial W_y}{\partial y} \right)^2 + m^2\omega_y^2 y^2 \right] &= \alpha - \frac{1}{2m} \left[ \left( \frac{\partial W_x}{\partial x} \right)^2 + m^2\omega_x^2 x^2 \right] \end{aligned} \quad (9.26)$$

el término de la izquierda solo depende de  $y$ , y el de la derecha solo depende de  $x$ , de modo que ambos deben ser constantes. Igulemos cada término a una constante que denotamos  $\alpha_y$

$$\frac{1}{2m} \left[ \left( \frac{\partial W_y}{\partial y} \right)^2 + m^2\omega_y^2 y^2 \right] = \alpha_y \quad (9.27)$$

$$\alpha - \frac{1}{2m} \left[ \left( \frac{\partial W_x}{\partial x} \right)^2 + m^2\omega_x^2 x^2 \right] = \alpha_y \quad (9.28)$$

la segunda ecuación se puede reescribir como

$$\frac{1}{2m} \left[ \left( \frac{\partial W_x}{\partial x} \right)^2 + m^2\omega_x^2 x^2 \right] = \alpha_x \quad ; \quad \alpha_x \equiv \alpha - \alpha_y \quad (9.29)$$

con lo cual las ecuaciones (9.27, 9.29) quedan más simétricas. Comparando las Ecs. (9.27, 9.29) con la Ec. (9.18), vemos que las soluciones son de la forma dada en las Ecs. (9.21, 9.23)

$$\begin{aligned} x &= \sqrt{\frac{2\alpha_x}{m\omega_x^2}} \sin(\omega_x t + \beta_x) \quad ; \quad p_x = \sqrt{2m\alpha_x} \cos(\omega_x t + \beta_x) \\ y &= \sqrt{\frac{2\alpha_y}{m\omega_y^2}} \sin(\omega_y t + \beta_y) \quad ; \quad p_y = \sqrt{2m\alpha_y} \cos(\omega_y t + \beta_y) \end{aligned} \quad (9.30)$$

vemos que de nuevo las  $\beta'$ s son las fases constantes y la energía es

$$E = \alpha = \alpha_x + \alpha_y$$

### 9.2.3. Oscilador armónico bidimensional isotrópico

Si tomamos el oscilador bidimensional isotrópico es decir con  $k_x = k_y$  podemos usar el resultado anterior y resolverlo como caso particular. No obstante, cuando  $k_x = k_y$  surge una simetría que está ausente en el caso general, ya que la interacción se vuelve central como se puede observar del potencial con  $\omega_x = \omega_y$

$$V(x, y) = \frac{m\omega_x^2}{2}x^2 + \frac{m\omega_y^2}{2}y^2 = \frac{m\omega^2}{2}(x^2 + y^2) = \frac{m\omega^2 r^2}{2} \equiv V(r)$$

en consecuencia, resulta interesante resolver el problema utilizando coordenadas polares, ya que en estas coordenadas el oscilador isotrópico tiene una coordenada cíclica

$$H = E = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} + m^2\omega^2 r^2 \right)$$

vemos que el Hamiltoniano es cíclico en  $\theta$ . Emulando el procedimiento en el caso anisotrópico, asumimos que la función característica  $W$  es separable

$$S(r, \theta, \alpha, \alpha_\theta) = W_r(r, \alpha, \alpha_\theta) + W_\theta(\theta, \alpha, \alpha_\theta) - \alpha t$$

con lo que la ecuación de Hamilton-Jacobi queda

$$\frac{1}{2m} \left[ \left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{1}{r^2} \left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + m^2\omega^2 r^2 \right] = \alpha$$

$$r^2 \left( \frac{\partial W_r}{\partial r} \right)^2 + \left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + m^2\omega^2 r^4 = 2mr^2\alpha$$

$$\left( \frac{\partial W_\theta}{\partial \theta} \right)^2 = 2mr^2\alpha - r^2 \left( \frac{\partial W_r}{\partial r} \right)^2 - m^2\omega^2 r^4$$

el término de la izquierda depende solo de  $\theta$  y el de la derecha solo de  $r$ . Ambos deben ser iguales a una constante  $\alpha_\theta^2$  luego

$$\left( \frac{\partial W_\theta}{\partial \theta} \right) = \alpha_\theta \Rightarrow W_\theta = \alpha_\theta \theta$$

de las ecuaciones básicas para los generadores se tiene que

$$p_\theta = \frac{\partial S}{\partial \theta} = \frac{\partial W_\theta}{\partial \theta} = \alpha_\theta$$

es decir  $\alpha_\theta$  es justamente el momento conjugado (constante) asociado a la variable cíclica. La ecuación para  $W_r$  queda

$$2mr^2\alpha - r^2 \left( \frac{\partial W_r}{\partial r} \right)^2 - m^2\omega^2 r^4 = \alpha_\theta^2 \Rightarrow \left( \frac{\partial W_r}{\partial r} \right)^2 = 2m\alpha - m^2\omega^2 r^2 - \frac{\alpha_\theta^2}{r^2} \Rightarrow$$

$$W_r = \int \sqrt{2mE - m^2\omega^2 r^2 - \frac{p_\theta^2}{r^2}} dr \quad (9.31)$$

aunque se puede resolver directamente para  $W_r$ , es más cómodo usar la solución en coordenadas cartesianas dada en las Ecs. (9.30) para la condición  $k_x = k_y$

$$\begin{aligned} x &= \sqrt{\frac{2\alpha}{m\omega^2}} \sin(\omega t + \beta) & ; & \quad p_x = \sqrt{2m\alpha} \cos(\omega t + \beta) \\ y &= \sqrt{\frac{2\alpha}{m\omega^2}} \sin \omega t & ; & \quad p_y = \sqrt{2m\alpha} \cos \omega t \end{aligned}$$

y las utilizamos para encontrar las contrapartes polares usando

$$r = \sqrt{x^2 + y^2} \quad ; \quad \theta = \arctan \frac{y}{x}$$

con lo cual la solución polar resulta

$$\begin{aligned} r &= \sqrt{\frac{2\alpha}{m\omega^2} \sqrt{\sin^2 \omega t + \sin^2 (\omega t + \beta)}} \quad ; \quad p_r = m\dot{r} \\ \theta &= \arctan \left[ \frac{\sin \omega t}{\sin (\omega t + \beta)} \right] \quad ; \quad p_\theta = mr^2\dot{\theta} \end{aligned}$$

La conservación de  $p_\theta$  (asociado a la variable cíclica  $\theta$ ), nos habla de la conservación del momento angular debido al carácter central que adquiere la interacción con la isotropía. Hay algunos casos límites de interés, el primero cuando  $\beta = 0$ , en cuyo caso

$$r = \sqrt{\frac{4\alpha}{m\omega^2}} \sin \omega t \quad ; \quad p_r = \sqrt{2m\alpha} \cos \omega t \quad ; \quad \theta = \frac{\pi}{4} \quad ; \quad p_\theta = 0$$

el movimiento en el plano  $x - y$  será una oscilación a lo largo de una línea diagonal que pasa por el origen. Otro caso límite ocurre cuando  $\beta = \pi/2$ , en tal caso

$$r = r_0 = \sqrt{\frac{2\alpha}{m\omega^2}} \quad ; \quad p_r = 0 \quad ; \quad \theta = \omega t \quad ; \quad p_\theta = mr_0^2\omega$$

el movimiento en el plano  $x - y$  trazará un círculo centrado en el origen de radio  $r_0$ . Para valores de  $\beta$  en el intervalo  $(0, \pi/2)$ , la órbita en el plano  $x - y$  es una elipse centrada en el origen, aunque los semiejes mayor y menor no van en general a lo largo de los ejes  $x$  e  $y$ . Estos son ejemplos de figuras de Lissajous.

### 9.3. Ecuación de Hamilton-Jacobi para la función característica de Hamilton

La razón principal que permite la fácil integración del oscilador armónico con el método de Hamilton-Jacobi, es el hecho de que la función principal de Hamilton se puede separar de la forma prescrita en (9.14), lo cual a su vez es posible siempre que el Hamiltoniano original no dependa explícitamente del tiempo, es decir cuando éste se conserva. Insertando (9.14) en (9.5) y haciendo  $k \rightarrow \alpha_1$  se obtiene la ecuación de Hamilton-Jacobi restringida

$$H \left( q_i, \frac{\partial W}{\partial q_i} \right) = \alpha_1 \tag{9.32}$$

esta ecuación ya no depende del tiempo, además se observa que una de las constantes de integración (por convención la elegimos como  $\alpha_1$ ) corresponde al valor (constante) del Hamiltoniano. En este caso la función característica de Hamilton aparece como parte de la función generatriz  $S$  cuando  $H$  es constante. Sin embargo, veremos que  $W$  puede ser por sí sola una función generatriz de una transformación canónica con propiedades diferentes a las generadas por  $S$ . Consideremos una transformación canónica en la cual todos los nuevos momentos son constantes de movimiento  $\alpha_i$ , donde en particular  $\alpha_1$  corresponde al valor del Hamiltoniano. Si la función generatriz de esta transformación la postulamos como  $W = W(q, P)$ , entonces esta función es de tipo 2 de tal forma que las ecuaciones de transformación vienen dadas por (7.19)

$$p_i = \frac{\partial W(q, \alpha)}{\partial q_i} \quad ; \quad Q_i = \frac{\partial W(q, \alpha)}{\partial P_i} = \frac{\partial W(q, \alpha)}{\partial \alpha_i} \tag{9.33}$$

estas ecuaciones se asemejan a sus análogos para  $S$ , Ecs. (9.9, 9.10), pero ahora la condición que determina a  $W$  es que  $H$  sea uno de los momentos canónicos nuevos, que por convención denotamos como  $\alpha_1$  <sup>3</sup>

$$H(p_i, q_i) = \alpha_1 . \tag{9.34}$$

---

<sup>3</sup>Nótese que en (9.14)  $k$  es una constante de integración pero no necesariamente es un momento conjugado. Cuando dicha constante se convierte en un momento conjugado, la expresión para  $Q_i$  definida en (9.33) diferirá de la definida en (9.10).

Adicionalmente, aquí solo pedimos que las coordenadas nuevas  $Q_i$  sean cíclicas pero no necesariamente que sean constantes, pues solo exigimos que sean constantes los momentos conjugados  $P_i$ . Usando (9.33) en (9.34) resulta

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1 \quad (9.35)$$

que coincide con la Ec. (9.32). De aquí también se vé que puesto que  $W$  no involucra al tiempo, el antiguo y el nuevo Hamiltoniano coinciden numéricamente i.e.  $K = \alpha_1$ . Esto es una diferencia con respecto al formalismo de la función generatriz  $S$ , la cual cambia el valor del Hamiltoniano y lo vuelve nulo.

La función característica de Hamilton genera una TC que nos lleva a otro conjunto de variables en donde todos los momentos son constantes, usando las Ecs. de Hamilton esto se traduce en

$$\dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0, \quad P_i = \alpha_i \quad (9.36)$$

es decir todas las coordenadas  $Q_i$  son cíclicas, de modo que la solución del problema ya es trivial. Ahora bien, dado que el nuevo Hamiltoniano solo depende de uno de los  $\alpha_i$  puesto que  $K = \alpha_1$ , y como las  $\alpha_i$  son independientes, tenemos entonces que de las otras ecuaciones de Hamilton

$$\dot{Q}_i = \frac{\partial K}{\partial \alpha_i} = \frac{\partial \alpha_1}{\partial \alpha_i} = \delta_{i1} \quad (9.37)$$

se obtienen entonces las soluciones

$$Q_i = \delta_{i1}t + \beta_i \equiv \frac{\partial W(q, \alpha)}{\partial \alpha_i} \quad (9.38)$$

de modo que la única coordenada que no es una constante de movimiento es  $Q_1$ . Dado que  $\alpha_1 = K$  es el momento conjugado a  $Q_1$  y  $Q_1$  es la única coordenada que depende del tiempo, vemos de nuevo un escenario en el cual el Hamiltoniano y el tiempo actúan como si fueran variables canónicamente conjugadas.

La dependencia de  $W$  con las antiguas coordenadas  $q$  se determina con la Ec. Diferencial Parcial (9.32), conocida como ecuación de Hamilton-Jacobi restringida. En este caso hay  $n$  constantes de movimiento (con respecto al caso anterior, no hay constante de integración para el tiempo), pero de nuevo una de ellas es una constante meramente aditiva la cual se identifica de inmediato con  $\alpha_1$ <sup>4</sup>. De nuevo es natural asumir que las  $n$  constantes de integración (incluyendo a  $\alpha_1$ ) sean los nuevos momentos conjugados.

Veamos ahora como se obtiene la solución dinámica en el formalismo de Hamilton-Jacobi restringido (HJR). Una vez que se encuentra una solución para  $W(q, \alpha)$  de la Ec. (9.32), introducimos este  $W$  en la primera mitad de las ecuaciones (9.33) y lo evaluamos en  $t = t_0$ , esto nos relaciona las constantes  $\alpha$  con las condiciones iniciales. Reemplazando  $\alpha_i = \alpha_i(q_0, p_0)$  en las Ecs. (9.38), y evaluando en  $t = t_0$  obtenemos  $\beta_i = \beta_i(q_0, p_0)$ . Una vez conocidos  $\alpha_i, \beta_i$  en términos de las condiciones iniciales, se evalúan de nuevo las Ecs. (9.38) pero esta vez para tiempo arbitrario, lo cual nos permite despejar  $q_i = q_i(\alpha_j, \beta_j, t)$  y por tanto cada  $q_i$  en términos del tiempo y de las condiciones iniciales. Finalmente, estas  $q_i(q_0, p_0, t)$  se reemplazan junto con los  $\alpha_i(q_0, p_0)$  en el primer conjunto de ecuaciones (9.33) para obtener  $p_i$  como función del tiempo y las condiciones iniciales, con lo cual ya tenemos la solución dinámica completa<sup>5</sup>.

Vale la pena enfatizar que solo una de las  $n$  ecuaciones (9.38) involucra al tiempo. Teniendo en cuenta que  $W$  es función de  $q_j$  y  $\alpha_j$  resulta de (9.38) que

$$\begin{aligned} t + \beta_1 &= f_1(q_1, \dots, q_n; \alpha_1, \dots, \alpha_n) \\ \beta_i &= f_i(q_1, \dots, q_n; \alpha_1, \dots, \alpha_n) \quad ; \quad i = 2, 3, \dots, n \end{aligned}$$

vemos que solo la ecuación con  $i = 1$  nos conecta explícitamente al tiempo. Por otro lado, las ecuaciones con  $i > 1$  solo relacionan a las  $q_i$ 's entre sí, es decir son **ecuaciones de trayectoria**. Una de las  $q_i$ 's se puede tomar como independiente y las otras coordenadas se pueden expresar en términos de ella resolviendo solo las

<sup>4</sup>Sin embargo, a diferencia de la constante aditiva que aparecía en (9.14), en este caso la constante tiene un significado físico pues es el valor numérico del Hamiltoniano.

<sup>5</sup>Nótese que la segunda mitad de Ecs. (9.33) no se usó directamente, en virtud de que son equivalentes a las Ecs. (9.38).



ecuaciones que son independientes del tiempo. En consecuencia, el formalismo nos conduce directamente a las ecuaciones de movimiento de la trayectoria (ya que no involucran al tiempo explícitamente). Por ejemplo en el caso de fuerzas centrales veremos que este procedimiento nos genera la ecuación de  $r$  en función de  $\psi$  sin pasar primero por las soluciones de  $r(t)$  y  $\psi(t)$ .

Como ya se mencionó, no es en general necesario elegir las  $\alpha_i$ 's como los nuevos momentos canónicos. En ocasiones puede ser más conveniente elegir otras constantes  $\gamma_i$  que son funciones de las constantes de integración  $\alpha_i$

$$P_i \equiv \gamma_i(\alpha_1, \dots, \alpha_n) \quad ; \quad i = 1, \dots, n \quad (9.39)$$

naturalmente las  $\gamma_i$  deben ser independientes entre sí. En tal caso, la función característica de Hamilton  $W$  tendrá como argumentos a las  $q_i, \gamma_i$ . El nuevo Hamiltoniano  $K = \alpha_1$  (numéricamente igual a  $H$  pero funcionalmente diferente) se obtiene despejando  $\alpha_1$  de (9.39) y será función exclusivamente de las  $\gamma_i$ . Las ecuaciones de Hamilton son entonces

$$\dot{P}_k = -\frac{\partial K}{\partial Q_k} = 0 \quad ; \quad \dot{Q}_k = \frac{\partial K}{\partial \gamma_k} = \frac{\partial \alpha_1(\gamma_i)}{\partial \gamma_k} \equiv v_k \quad (9.40)$$

es obvio que las  $v_k$  son constantes que dependen exclusivamente de las  $\gamma_k$ . Las soluciones para las nuevas variables canónicas son

$$P_k = \gamma_k \quad ; \quad Q_k = v_k t + \beta_k \quad (9.41)$$

la forma de  $W$  no se puede encontrar hasta tener una solución completa de la ecuación de Hamilton Jacobi restringida.

Un comentario final, a priori la expresión  $K = cte$  conduce a las mismas ecuaciones de movimiento que  $K = 0$ . Sin embargo, las ecuaciones (9.1) con  $K = 0$  claramente conducen a soluciones diferentes que las Ecs. (9.36, 9.37) con  $K = \alpha_1 = cte$ . Para comprender la diferencia debemos tener en cuenta que aunque las  $\alpha_i$  son constantes, ellas son tratadas como variables para efectos de obtener las ecuaciones de Hamilton como se vé en las Ecs. (9.37), es decir que aunque  $K$  es constante, dicha constante  $\alpha_1$  es una propiedad del sistema (en muchos casos la energía del sistema), si en cambio  $K$  es igual a una constante independiente del sistema, entonces para todos los efectos será equivalente a escoger  $K = 0$ .

En otros términos, supongamos que un sistema Físico tiene condiciones iniciales  $(q_0, p_0)$ , si para el mismo sistema cambiamos las condiciones iniciales a  $(q'_0, p'_0)$ , entonces cambiará en general el valor numérico del Hamiltoniano

$$\alpha_1(q'_0, p'_0) \neq \alpha_1(q_0, p_0)$$

En tal caso, al usar el formalismo en el cual  $W$  es la función generatriz, el valor numérico del nuevo Hamiltoniano  $K = \alpha_1$  debe cambiar con respecto al caso en el cual usamos las condiciones iniciales  $(q_0, p_0)$ . En contraste, en el formalismo en el cual  $S$  es la función generatriz, dicha función se ajusta de modo que el nuevo Hamiltoniano sea siempre nulo, por tanto en este formalismo  $K = 0$  sin importar las condiciones iniciales.

Esto nos deja una lección muy importante, dos Hamiltonianos que difieren en una constante solo son equivalentes si la constante en cuestión no es una función del sistema, es decir si no depende de las condiciones iniciales de éste. No obstante, en este escenario muy particular (siempre que el Hamiltoniano sea independiente del tiempo) los Hamiltonianos  $K = \alpha_1$  y  $K = 0$  corresponden a la misma Física, ya que ambos están escritos en bases canónicas distintas, y pretenden por métodos diferentes resolver el mismo problema<sup>6</sup>. Pero si dos Hamiltonianos escritos en la misma base canónica difieren en una constante que es función del sistema, las ecuaciones de Hamilton son diferentes y la Física que describen es diferente.

## 9.4. Paralelismo entre el formalismo de Hamilton-Jacobi y el formalismo restringido de Hamilton-Jacobi

Es útil hacer un esquema que nos muestre la estrategia de solución de las ecuaciones de movimiento basados en la función principal de Hamilton  $S$  o en la función característica de Hamilton  $W$ .

<sup>6</sup>Las ecuaciones de Hamilton para los dos Hamiltonianos son diferentes pero no porque estén asociados a Física diferente, sino a bases canónicas distintas.

Estos métodos son aplicables cuando el Hamiltoniano

$$H(q, p, t) \quad \left| \quad \begin{array}{l} \text{se conserva} \\ H(q, p) = cte \equiv \alpha_1 \end{array} \right.$$

La idea del método es encontrar una transformación canónica que nos lleve a nuevas coordenadas de modo que

$$\begin{array}{l} \text{todas las nuevas coordenadas y momentos} \\ Q_i, P_i, \text{ sean constantes de movimiento} \end{array} \quad \left| \quad \begin{array}{l} \text{Todos los momentos } P_i \text{ sean constantes} \end{array} \right.$$

para satisfacer estos requerimientos es condición suficiente que el nuevo Hamiltoniano  $K$  sea idénticamente zero

$$K = 0 \quad \left| \quad \begin{array}{l} \text{Sea cíclico en todas las nuevas coordenadas } Q_i \\ H = K(P_i) = \alpha_1 \end{array} \right.$$

bajo estas condiciones, las ecuaciones de Hamilton para  $Q_i, P_i$  son

$$\begin{array}{l} \dot{Q}_i = \frac{\partial K}{\partial P_i} = 0 \\ \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0 \end{array} \quad \left| \quad \begin{array}{l} \dot{Q}_i = \frac{\partial K}{\partial P_i} = v_i \\ \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0 \end{array} \right.$$

las soluciones son inmediatas

$$\begin{array}{l} Q_i = \beta_i \\ P_i = \gamma_i \end{array} \quad \left| \quad \begin{array}{l} Q_i = v_i t + \beta_i \\ P_i = \gamma_i \end{array} \right.$$

Que satisface los requerimientos inicialmente propuestos. En ambos métodos se han trivializado las ecuaciones de Hamilton de modo que la dificultad para resolver las ecuaciones de movimiento se ha trasladado a la búsqueda de la TC. Para ello se plantea la existencia de una función generatriz de tipo 2 que genera dicha TC, en cada caso la función generatriz es

$$\begin{array}{l} \text{La función principal de Hamilton} \\ F_2(q, P, t) \equiv S(q, P, t) \end{array} \quad \left| \quad \begin{array}{l} \text{Función característica de Hamilton} \\ F_2(q, P) \equiv W(q, P) \end{array} \right.$$

que satisface la ecuación diferencial parcial

$$H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0 \quad \left| \quad H\left(q, \frac{\partial W}{\partial q}\right) - \alpha_1 = 0$$

una solución completa de cada ecuación parcial contiene

$$\begin{array}{l} n \text{ constantes de integración no triviales} \\ \alpha_1, \dots, \alpha_n \end{array} \quad \left| \quad \begin{array}{l} n - 1 \text{ constantes de integración no triviales que junto con } \alpha_1 \\ \text{forman un conjunto de } n \text{ constantes independientes} \end{array} \right.$$

Como los  $P_i$  deben ser constantes, es natural escoger los  $P_i$  como  $n$  funciones independientes de las constantes de integración

$$P_i = \gamma_i(\alpha_1, \dots, \alpha_n) \quad \left| \quad P_i = \gamma_i(\alpha_1, \dots, \alpha_n)$$

las soluciones de la ecuación de Hamilton Jacobi tendrán como argumentos

$$S = S(q_i, \gamma_i, t) \quad \left| \quad W = W(q_i, \gamma_i)$$

En particular podemos escoger  $\gamma_i = \alpha_i$ . Recordando que  $S$  y  $W$  son funciones generatrices tipo 2, sus ecuaciones parciales asociadas son

$$\begin{array}{l} p_i = \frac{\partial S(q_i, \gamma_i, t)}{\partial q_i} \\ Q_i = \frac{\partial S(q_i, \gamma_i, t)}{\partial \gamma_i} = \beta_i \end{array} \quad \left| \quad \begin{array}{l} p_i = \frac{\partial W(q_i, \gamma_i)}{\partial q_i} \\ Q_i = \frac{\partial W(q_i, \gamma_i)}{\partial \gamma_i} = v_i(\gamma_j)t + \beta_i \end{array} \right.$$

La primera mitad de éstas ecuaciones se satisface automáticamente ya que han sido empleadas para construir la ecuación de Hamilton Jacobi. La segunda mitad se puede resolver para  $q_i$  en términos de  $t$  y las  $2n$  constantes  $\alpha_i, \beta_i$ . Para terminar, se deben encontrar los valores específicos de las  $2n$  constantes para lo cual se requieren los dos conjuntos de ecuaciones evaluados en  $t = t_0$  y empleando las condiciones iniciales  $q_0, p_0$ .

Finalmente, cuando el Hamiltoniano no es función explícita del tiempo, i.e.  $H = \alpha_1 = cte$ , ambos métodos son adecuados y nos llevan a una misma ecuación de movimiento Ec. (9.32)

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1 \tag{9.42}$$

sin embargo, la interpretación de la solución  $W(q_i, \gamma_i)$  es muy distinta para el formalismo de Hamilton-Jacobi (HJ) con respecto a la interpretación en el formalismo de Hamilton-Jacobi restringido (HJR). En el formalismo de HJ,  $W(q_i, \gamma_i)$  es solo una parte de la función generatriz de la TC, la función generatriz completa está dada por

$$S(q, \gamma, t) = W(q, \gamma) - \alpha_1 t \tag{9.43}$$

y  $S$  genera una TC que cambia el valor numérico del Hamiltoniano a una valor idénticamente nulo, trayendo como consecuencia que todos los nuevos momentos y coordenadas son constantes de movimiento.

En contraste, para el formalismo HJR, la solución  $W(q_i, \gamma_i)$  representa la función generatriz completa, la cual al ser independiente del tiempo no genera un cambio en el valor numérico del Hamiltoniano de modo que  $K = \alpha_1$ , la TC generada por  $W$  es tal que todas las nuevas coordenadas son cíclicas. Como consecuencia, los nuevos momentos son constantes de movimiento y las nuevas coordenadas son o bien constantes o bien funciones lineales del tiempo.

Debe decirse sin embargo, que aunque  $W$  se interpreta de modo diferente en HJ y en HJR, ambas están resolviendo el mismo problema Físico.

### 9.5. Separación de variables en la ecuación de Hamilton-Jacobi

En esta sección no utilizaremos convención de suma sobre índices repetidos a menos que se especifique lo contrario. Cabe preguntarse en qué casos la ecuación de Hamilton Jacobi es de utilidad práctica teniendo en cuenta que en general las ecuaciones diferenciales parciales con múltiples incógnitas son de difícil solución. Veremos que bajo ciertas condiciones, es posible hacer una separación de variables en la ecuación de Hamilton Jacobi en cuyo caso es posible siempre reducir el problema a cuadraturas. En realidad, salvo casos excepcionales, el formalismo de Hamilton Jacobi solo es efectivo cuando dicha separación es realizable.

Una coordenada  $q_j$  se dice separable en la ecuación de Hamilton Jacobi cuando la función principal de Hamilton (o la función característica), se pueden separar en dos partes aditivas, una que solo depende de la coordenada  $q_j$  (y de los  $\alpha_i$ ) y otra que es totalmente independiente de  $q_j$ . Por simplicidad pensemos que la variable separable es  $q_1$ , la función principal de Hamilton se escribe como

$$S(q_1, \dots, q_n; \alpha_1, \dots, \alpha_n; t) = S_1(q_1; \alpha_1, \dots, \alpha_n; t) + S'(q_2, \dots, q_n; \alpha_1, \dots, \alpha_n; t)$$

y la ecuación de Hamilton Jacobi se puede separar en dos ecuaciones, una separadamente para  $S_1$  y otra para  $S'$ . Nótese que la condición de separabilidad ha sido impuesta a la solución de la ecuación y no al Hamiltoniano. Adicionalmente decimos que la ecuación de Hamilton Jacobi (HJ) es totalmente separable (o simplemente separable) cuando todas las coordenadas en el problema son separables, de modo que la función principal se escribe como

$$S = \sum_{i=1}^n S_i(q_i; \alpha_1, \dots, \alpha_n; t) \tag{9.44}$$

esta hipótesis de separación es consistente siempre que la introducir (9.44) en (9.5), se obtengan  $n$  ecuaciones de la forma

$$H_j \left( q_j; \frac{\partial S_j}{\partial q_j}; \alpha_1, \dots, \alpha_n; t \right) + \frac{\partial S_j}{\partial t} = 0 \tag{9.45}$$

es notable que si el Hamiltoniano no depende explícitamente del tiempo, podemos tratar al tiempo como variable separable y el paso de la función principal a la función característica de Hamilton constituye entonces un ejemplo de la forma en que trabaja la separación de variables. En este caso podemos parametrizar la función principal de Hamilton en la forma

$$S(q, \alpha, t) = S_0(\alpha, t) + W(q, \alpha)$$

al introducir esta solución en la ecuación de HJ, teniendo en cuenta que el Hamiltoniano no depende explícitamente del tiempo resulta

$$H \left( q, \frac{\partial W}{\partial q} \right) + \frac{\partial S_0}{\partial t} = 0$$

el primer término solo depende de las  $q_i$  en tanto que el segundo depende solamente de  $t$ . Por tanto, ambos términos deben ser constantes con valores opuestos

$$\begin{aligned} H \left( q, \frac{\partial W}{\partial q} \right) &= \alpha_1 ; \quad \frac{\partial S_0}{\partial t} = -\alpha_1 \Rightarrow S_0 = -\alpha_1 t \\ S(q, \alpha, t) &= -\alpha_1 t + W(q, \alpha) \end{aligned} \tag{9.46}$$

donde claramente  $\alpha_1$  coincide con el valor numérico del Hamiltoniano. La Ec. (9.46) está en concordancia con (9.43). Naturalmente podemos sumar una constante al valor de  $S_0$  pero esto no tiene ninguna relevancia. Lo importante es encontrar una solución para  $S$ , pues recordemos que la unicidad no se requiere para la función generatriz.

Si además de que el Hamiltoniano es independiente del tiempo ocurre que todas las variables son separables, podemos parametrizar cada  $S_i$  de la forma

$$S_i(q_i; \alpha_1, \dots, \alpha_n; t) = W_i(q_i; \alpha_1, \dots, \alpha_n) - \alpha_i t \quad (9.47)$$

ahora bien, si la ecuación de HJ es realmente separable, la suposición de separabilidad de  $S$  debe conducirnos a ecuaciones de la forma (9.45), que junto con la forma específica de la hipótesis de separación (9.47) nos dará un conjunto de  $n$  ecuaciones restringidas de HJ.

$$H_j \left( q_j; \frac{\partial W_j}{\partial q_j}; \alpha_1, \dots, \alpha_n \right) = \alpha_j \quad (9.48)$$

Es de anotar que las funciones  $H_i$  pueden ser o no Hamiltonianos, esto está relacionado con el hecho de que el concepto de separabilidad no involucra al Hamiltoniano sino a las soluciones de la ecuación de HJ. Similarmente, las constantes de separación  $\alpha_i$  puede ser o no ser una energía, (aunque sus dimensiones son siempre de energía), o alguna otra cantidad dependiendo de la naturaleza de  $q_i$ .

Las constantes  $\alpha_i$  se denominan **constantes de separación**. Cada una de las ecuaciones (9.48) involucra solo una coordenada  $q_j$  y la correspondiente derivada parcial de  $W_j$  con respecto a  $q_j$ . En consecuencia, resultan un conjunto de ecuaciones diferenciales ordinarias de primer orden. Dado que en principio podemos resolver para despejar  $\partial W_j / \partial q_j$  para luego integrar sobre  $q_j$ , se deduce que el problema está reducido a cuadraturas.

En la práctica ocurre con frecuencia que cada  $H_i$  solo contiene uno o unos pocos  $\alpha$ 's. Adicionalmente, existen escenarios en donde solo  $r$  variables son separables y las restantes  $n - r$  variables no pueden separarse. En tal caso, solo podemos reducir a cuadraturas la ecuación de movimiento de las  $r$  variables separables como veremos más adelante.

Finalmente salvo casos excepcionales, casi todas las aplicaciones útiles del formalismo de HJ involucran Hamiltonianos independientes del tiempo en donde  $t$  es una coordenada separable para  $S$  (y "cíclica" para  $W$ ). En consecuencia, la discusión subsecuente sobre separabilidad se limitará a Hamiltonianos constantes y se utilizará solamente la función característica  $W$ .

### 9.5.1. Coordenadas ignorables y separabilidad

Podemos demostrar fácilmente que una coordenada cíclica es separable. Por simplicidad, tomemos a  $q_1$  como la coordenada cíclica, su momento conjugado es una constante de movimiento que denotaremos como  $\gamma_1$ . La ecuación de Hamilton Jacobi restringida queda

$$H \left( q_2, \dots, q_n; \gamma_1; \frac{\partial W}{\partial q_2}, \dots, \frac{\partial W}{\partial q_n} \right) = \alpha_1 \quad (9.49)$$

si proponemos una solución de la forma

$$W = W_1(q_1, \alpha) + W'(q_2, \dots, q_n; \alpha) \quad (9.50)$$

introduciendo (9.50) en (9.49) queda claro que en la Ec. (9.49) solo queda la parte aditiva  $W'$ , en tanto que  $W_1$  es la solución de la ecuación (9.33) con  $i = 1$

$$p_1 = \gamma_1 = \frac{\partial W}{\partial q_1} = \frac{\partial W_1}{\partial q_1} \quad (9.51)$$

$\gamma_1$  es por tanto la constante de separación, la solución obvia para  $W_1$  (excepto por una constante aditiva) es

$$W_1 = \gamma_1 q_1 \quad (9.52)$$

con lo cual  $W$  estará dado por

$$W = W' + \gamma_1 q_1 \quad (9.53)$$

hay una gran semejanza entre la Ec. (9.53) y la forma que  $S$  asume cuando el Hamiltoniano es independiente del tiempo Ec. (9.43). Esto no es sorprendente ya que nuevamente, es como suponer que la variable “cíclica” tiempo asociada al “momento conjugado” constante  $-H$  conduzca a una ecuación del tipo (9.53) con  $\gamma_1 \rightarrow -H = -\alpha_1$  y con  $q_1 \rightarrow t$ . Sin embargo, es importante insistir en que esta es solo una analogía ya que  $t$  no es una coordenada generalizada sino un parámetro. Es cierto que  $H = \alpha_1$  puede ser elegido como un momento conjugado, pero a una coordenada generalizada y no al tiempo.

Si un número  $s$  de las  $n$  coordenadas son no cíclicas y además suponemos que  $S$  es totalmente separable, podemos extender el resultado anterior. El Hamiltoniano viene dado por

$$H = H(q_1, \dots, q_s; \alpha_1, \dots, \alpha_n) \quad (9.54)$$

y la función característica se escribe como

$$W(q_1, \dots, q_s; \alpha_1, \dots, \alpha_n) = \sum_{i=1}^s W_i(q_i; \alpha_1, \dots, \alpha_n) + \sum_{i=s+1}^n q_i \alpha_i \quad (9.55)$$

y hay  $s$  ecuaciones de HJ para solucionar

$$H_i\left(q_i; \frac{\partial W_i}{\partial q_i}; \alpha_1, \dots, \alpha_n\right) = \alpha_i \quad ; \quad i = 1, \dots, s$$

y dado que estas son ecuaciones diferenciales ordinarias de primer orden en la variable independiente  $q_i$ , se pueden reducir inmediatamente a cuadraturas, y la solución completa para  $W$  se puede obtener.

### 9.5.2. Condiciones más generales para la separabilidad

En general, una coordenada  $q_j$  se puede separar si  $q_j$  y su momento conjugado  $p_j$  se pueden aislar en el Hamiltoniano en alguna función  $f(q_j, p_j)$  que no contiene a ninguna otra variable. Postulemos una solución de ensayo de la forma

$$W = W_j(q_j, \alpha) + W'(q_i, \alpha) \quad ; \quad i = 1, \dots, j-1, j+1, \dots, n \quad (9.56)$$

donde  $q_i$  representa el conjunto de todos los  $q$ 's excepto  $q_j$ . Supongamos que el ansatz de separación (9.56) nos conduce a que la ecuación restringida de HJ tenga la forma

$$H\left(q_i; \frac{\partial W'}{\partial q_i}; f\left(q_j, \frac{\partial W_j}{\partial q_j}\right)\right) = \alpha_1 \quad ; \quad i = 1, \dots, j-1, j+1, \dots, n \quad (9.57)$$

esta ecuación se puede invertir (al menos formalmente) para despejar a  $f$

$$f\left(q_j, \frac{\partial W_j}{\partial q_j}\right) = g\left(q_i, \frac{\partial W'}{\partial q_i}, \alpha_1\right) \quad ; \quad i = 1, \dots, j-1, j+1, \dots, n \quad (9.58)$$

ahora bien, el lado izquierdo depende exclusivamente de  $q_j$  en tanto que el lado derecho solo depende de las otras variables  $q$ . Por tanto la Ec. (9.58) solo puede ser cierta si ambos lados son iguales a la misma constante por lo tanto

$$f\left(q_j, \frac{\partial W_j}{\partial q_j}\right) = \alpha_j \quad ; \quad g\left(q_i, \frac{\partial W'}{\partial q_i}\right) = \alpha_j \quad ; \quad i = 1, \dots, j-1, j+1, \dots, n$$

con lo cual se ha logrado la separación de la variable  $q_j$ .

Vale la pena resaltar que la separabilidad depende no solo del problema físico sino también del sistema de coordenadas generalizadas usado. Por ejemplo el problema de un cuerpo bajo una fuerza central es separable en coordenadas polares pero no en cartesianas. En algunos casos la ecuación de HJ no se puede separar

completamente como es el caso en el problema de los tres cuerpos. En contraste, para algunos problemas físicos es posible separar la ecuación de HJ en más de un sistema coordinado. En general solo es posible encontrar una solución cerrada para la ecuación de HJ cuando las variables son completamente separables. En consecuencia, se han estudiado con profundidad diversos métodos para encontrar los sistemas coordinados apropiados para cada problema.

En general no hay un criterio simple para indicar qué sistemas coordinados conducen a soluciones de Hamilton Jacobi separables para un problema específico. Sin embargo, en el caso de sistemas ortogonales de coordenadas, las llamadas **condiciones de Staeckel** resultan muy útiles. Estas son condiciones necesarias y suficientes para la separabilidad bajo ciertas circunstancias. Dichas condiciones son las siguientes:

1. El Hamiltoniano se conserva.
2. El Lagrangiano es solo una función cuadrática de las velocidades generalizadas de tal forma que el Hamiltoniano tiene la estructura dada por la Ec. (6.20) con  $L_0(q, t) = -V(q)$

$$H = \frac{1}{2} (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - V(q)$$

3. El sistema coordinado es un sistema ortogonal, esto trae como consecuencia que la matriz  $\mathbf{T}$  sea diagonal y por lo tanto también lo es su inversa

$$(\mathbf{T}^{-1})_{ij} = \frac{1}{T_{ii}} \delta_{ij} \quad (\text{no suma})$$

4. El vector  $\mathbf{a}$  tiene elementos  $a_i$  que solo dependen de su coordenada correspondiente i.e.  $a_i = a_i(q_i)$
5. La función potencial  $V(q)$  es separable en la forma

$$V(q) = \frac{V_i(q_i)}{T_{ii}} \quad (9.59)$$

6. Existe una matriz  $\Phi$  de dimensión  $n \times n$  con elementos  $\Phi_{ij} = \Phi_{ij}(q_i)$  de tal forma que

$$(\Phi^{-1})_{ij} = \frac{1}{T_{jj}} \quad (\text{no suma})$$

Si se cumplen las condiciones de Staeckel, la función característica de Hamilton será completamente separable.

$$W(q) = \sum_i W_i(q_i)$$

y cada  $W_i$  satisface las ecuaciones

$$\left[ \frac{\partial W_i}{\partial q_i} - a_i \right]^2 = -2V_i(q_i) + 2\phi_{ij}\gamma_j$$

siendo las  $\gamma_j$  constantes de integración y solo hay suma sobre  $j$ . La última condición resulta a priori muy complicada. Sin embargo veremos más adelante que en la práctica ocurre a menudo que solo es necesario probar la existencia de la matriz  $\Phi$ , sin que sea necesario encontrar su forma explícita. Esto facilita enormemente aplicar las condiciones de Staeckel. Para más información sobre las condiciones de Staeckel, remitimos al lector al apéndice D de la segunda edición de la Ref. [1].

## 9.6. Fuerzas centrales en el formalismo de Hamilton-Jacobi

En esta sección nos limitaremos a plantear el problema de una partícula sometida a una fuerza central y reducirlo a cuadraturas empleando el método de HJ. El propósito es solo ilustrar el método ya que los detalles sobre este problema se discutirán en el capítulo 10.

### 9.6.1. Problema bidimensional

Consideremos el problema de una partícula sometida a una fuerza central descrita por el potencial  $V(r)$ . En virtud de la conservación del momento angular, se puede observar que el movimiento debe realizarse en un plano. El Hamiltoniano para este sistema bidimensional se escribe como

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} \right) + V(r) \quad (9.60)$$

dado que  $\theta$  es cíclica, podemos escribir  $W$  en la forma (9.55)

$$W = W_1(r) + \alpha_\theta \theta \quad (9.61)$$

donde

$$\alpha_\theta = p_\theta \equiv l \quad (9.62)$$

es el momento conjugado a  $\theta$  i.e. la magnitud del momento angular. Para el Hamiltoniano (9.60), la ecuación de Hamilton Jacobi restringida (9.32) queda de la forma

$$\begin{aligned} \frac{1}{2m} \left[ \left( \frac{\partial W}{\partial r} \right)^2 + \frac{(\partial W / \partial \theta)^2}{r^2} \right] + V(r) &= \alpha_1 \\ \left( \frac{dW_1}{dr} \right)^2 + \frac{\alpha_\theta^2}{r^2} + 2mV(r) &= 2m\alpha_1 \end{aligned} \quad (9.63)$$

recordemos que  $\alpha_1$  se identifica con el valor numérico del Hamiltoniano y en este caso, de la energía total del sistema. La Ec. (9.63) es una ecuación diferencial ordinaria para  $W_1(r)$  que se puede reducir en forma inmediata a cuadraturas

$$\left( \frac{dW_1}{dr} \right) = \sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\theta^2}{r^2}}$$

y la función  $W$  dada por (9.61) queda en la forma

$$W = \alpha_\theta \theta + \int \left[ \sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\theta^2}{r^2}} \right] dr$$

con esta expresión las Ecs. (9.38) toman la forma

$$t + \beta_1 = \frac{\partial W}{\partial \alpha_1} = \int \frac{m dr}{\sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\theta^2}{r^2}}} \quad (9.64)$$

$$\beta_2 = \frac{\partial W}{\partial \alpha_2} = \frac{\partial W}{\partial \alpha_\theta} = \theta - \int \frac{\alpha_\theta dr}{r^2 \sqrt{2m[\alpha_1 - V(r)] - \frac{\alpha_\theta^2}{r^2}}} \quad (9.65)$$

La ecuación (9.64) nos da  $r = r(t)$  y en general, las constantes  $\alpha_1$  y  $\alpha_\theta$  se pueden relacionar con las cantidades conservadas energía y momento angular (en el problema de una partícula sometida a una fuerza central, es más común conocer la energía y el momento angular como condiciones iniciales, en lugar de las tradicionales condiciones iniciales de posición y momento). Como ya se mencionó, las restantes ecuaciones de  $Q_i$  que no involucran explícitamente al tiempo, proporcionan la ecuación de la órbita. En este caso solo hay una ecuación remanente para la órbita Ec. (9.65). En general resulta útil realizar el cambio de variable  $u = 1/r$  en (9.65) con lo que la ecuación de la órbita queda en la forma

$$\theta = \beta_2 - \int \frac{du}{\sqrt{\frac{2m}{\alpha_\theta^2}(\alpha_1 - V) - u^2}} \quad (9.66)$$

$\beta_2$  estará relacionado con el ángulo inicial medido en un sistema de ejes apropiado.

### 9.6.2. La dinámica de las fuerzas centrales como problema tridimensional

En la anterior sección asumimos desde el principio que el movimiento se ejecutaba en un plano en virtud de la conservación del momento angular. Resulta ilustrativo trabajar el mismo problema pero asumiendo que el movimiento es en general en tres dimensiones para llegar posteriormente a la conclusión de que el movimiento es bidimensional. Comencemos entonces con el Hamiltoniano asociado a una partícula sometida a un potencial central en coordenadas esféricas, Ec. (6.25)

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \theta} \right) + V(r) \quad (9.67)$$

asumiendo separación de variables, la función característica de Hamilton Jacobi queda de la forma

$$W = W_r(r) + W_\theta(\theta) + W_\phi(\phi) \quad (9.68)$$

y dado que la coordenada  $\phi$  es cíclica en el Hamiltoniano (9.67) tenemos que

$$W_\phi = \alpha_\phi \phi \quad (9.69)$$

donde  $\alpha_\phi$  es una constante de integración y coincide con el momento conjugado a  $\phi$  como se puede ver teniendo en cuenta las Ecs. (9.51) y (9.52). Usando la forma de  $W$  dada por las Ecs. (9.68, 9.69) la ecuación de Hamilton Jacobi queda de la forma

$$\left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{1}{r^2} \left[ \left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} \right] + 2mV(r) = 2mE \quad (9.70)$$

donde hemos escrito la constante de integración  $\alpha_1$  como la energía que es el valor numérico de nuestro Hamiltoniano. Nótese que toda dependencia de  $\theta$  ha sido separada en el término entre paréntesis cuadrados, es decir cumple con la estructura de la Ec. (9.57). Esto nos implica que esta función de  $\theta$  se puede igualar a una constante según el razonamiento hecho en la sección 9.5.2, por ejemplo despejando el término mencionado queda

$$\left[ \left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} \right] = 2mr^2 [E - V(r)] - r^2 \left( \frac{\partial W_r}{\partial r} \right)^2$$

el miembro izquierdo depende solo de  $\theta$  y el derecho solo de  $r$  luego ambos deben igualarse a la misma constante

$$\left( \frac{\partial W_\theta}{\partial \theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} = \alpha_\theta^2 \quad (9.71)$$

$$2mr^2 [E - V(r)] - r^2 \left( \frac{\partial W_r}{\partial r} \right)^2 = \alpha_\theta^2 \quad (9.72)$$

y obtenemos dos ecuaciones diferenciales ordinarias para  $\theta$  y  $r$  que ya se reducen el problema a cuadraturas. Nótese que la ecuación diferencial para  $r$  también se puede encontrar reemplazando (9.71) en (9.70) quedando

$$\left( \frac{\partial W_r}{\partial r} \right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m [E - V(r)] \quad (9.73)$$

esta ecuación es claramente equivalente a (9.72). El único aspecto que falta determinar es la asociación de las constantes  $E$ ,  $\alpha_\phi$ ,  $\alpha_\theta$  con cantidades físicas. La cantidad  $E$  es la energía del sistema. Por otro lado,  $\alpha_\phi$  es el momento conjugado a la variable  $\phi$ , y sabemos que el momento conjugado a una variable angular corresponde a la componente del momento angular del sistema a lo largo del eje de rotación que rotaría al sistema como un todo en una cantidad  $d\phi$ , este eje sería claramente el eje  $z$  y por tanto  $\alpha_\phi$  corresponde a la componente polar del momento angular  $L_z$ .



No es tan sencillo identificar el significado Físico de  $\alpha_\theta$  por múltiples razones. En primer lugar, esta cantidad aparece como una constante de separación y no como el momento conjugado a  $\theta$  (de por sí veremos más adelante que no lo es), por otro lado una variación en  $\theta$  no define un único eje de rotación, ya que a diferencia de  $\phi$  cuya variación ocurre siempre sobre el plano  $XY$ , la variación en  $\theta$  se da en el plano definido por el eje  $Z$  y el vector instantáneo de posición, este plano claramente cambia cuando varía la dirección del vector posición. Finalmente,  $\theta$  no puede tomar cualquier valor ( $0 \leq \theta \leq \pi$ ) con lo cual se dificulta su interpretación como ángulo de rotación. Para encontrar el significado físico de  $\alpha_\theta$ , calculemos primero los momentos conjugados a cada variable angular recurriendo al correspondiente Lagrangiano en tres dimensiones para coordenadas esféricas de la partícula sometida al potencial  $V(r)$  i.e.

$$L(r, \theta, \phi) = \frac{1}{2}m \left( \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2 \sin^2 \theta \right) - V(r)$$

los momentos conjugados a las variables angulares quedan

$$p_\theta = \frac{\partial L(r, \theta, \phi)}{\partial \dot{\theta}} = mr^2\dot{\theta} \quad (9.74)$$

$$p_\phi = \frac{\partial L(r, \theta, \phi)}{\partial \dot{\phi}} = mr^2\dot{\phi} \sin^2 \theta \quad (9.75)$$

ahora veremos la relación entre  $\alpha_\theta$ ,  $\alpha_\phi$  y los momentos  $p_\theta$ ,  $p_\phi$ <sup>7</sup>. Ya hemos mencionado que  $\alpha_\phi$  es el momento conjugado a la variable  $\phi$

$$\alpha_\phi = p_\phi = \frac{\partial W_\phi}{\partial \phi} \quad (9.76)$$

Para relacionar  $\alpha_\theta$  con los momentos conjugados, podemos reescribir (9.71) en la forma

$$p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} = \alpha_\theta^2 \quad (9.77)$$

de tal manera que el Hamiltoniano (9.67), se puede reescribir como

$$H = \frac{1}{2m} \left( p_r^2 + \frac{\alpha_\theta^2}{r^2} \right) + V(r) \quad (9.78)$$

y comparando (9.78) con el Hamiltoniano (9.60) y usando (9.62) resulta<sup>8</sup>

$$\alpha_\theta = p_\theta \text{ (asociado a dos dimensiones)} \equiv l \quad (9.79)$$

lo cual ya nos da el significado físico de  $\alpha_\theta$  como el módulo del momento angular del sistema. Nótese que en tres dimensiones la conservación del módulo del momento angular no está asociado a una variable cíclica como sí ocurre en dos dimensiones<sup>9</sup>, pues  $\theta$  no es variable cíclica en tres dimensiones y por tanto  $p_\theta$  no es constante de movimiento como se puede apreciar de (9.77), de esta ecuación también se observa que  $l$  se escribe en términos de los momentos asociados a ambos ángulos.

Vemos entonces que la conservación de las constantes  $E$ ,  $\alpha_\phi$ ,  $\alpha_\theta$  representan físicamente la conservación de la energía, de la componente polar del momento angular y del módulo del momento angular respectivamente. En este ejemplo, vemos que el método de Hamilton Jacobi resulta particularmente poderoso para extraer las constantes de movimiento así como las ecuaciones de  $r = r(t)$  y de la órbita. Adicionalmente, el formalismo

<sup>7</sup>A priori estaríamos tentados a interpretar a  $p_\theta = mr^2\dot{\theta}$  como el módulo del momento angular total, pero en tres dimensiones  $\dot{\theta}$  no es la velocidad angular con que la partícula se mueve sobre el plano. En particular, nótese que  $\theta$  no es cíclica y por tanto  $p_\theta$  no es constante de movimiento.

<sup>8</sup>Los Hamiltonianos (9.78) y (9.60) se pueden comparar apropiadamente, ya que tanto en coordenadas polares planas como en coordenadas esféricas, la coordenada  $r$  se refiere a la distancia al origen.

<sup>9</sup>El modulo del momento angular en tres dimensiones aparece como una constante de separación  $\alpha_\theta$ .

de Hamilton Jacobi para este problema es separable en otras coordenadas como las parabólicas y elípticas, y las constantes aparecen en forma apropiada para cada sistema coordinado.

Este problema también provee un escenario natural para emplear las condiciones de Staeckel de modo que encontremos el potencial más general  $V$  para una partícula que haga que las soluciones de la ecuación sean totalmente separables en coordenadas esféricas. La matriz  $\Phi$  de las condiciones de Staeckel solo depende del sistema coordinado y no del potencial. Como acabamos de demostrar que la ecuación de Hamilton Jacobi es separable para al menos un potencial en coordenadas esféricas, las condiciones de Staeckel me dicen que dicha matriz tiene que existir, y observamos por otro lado que no requerimos de la forma específica de  $\Phi$  para encontrar las condiciones de separabilidad, solo requerimos de su existencia. Adicionalmente, el arreglo rectangular columna  $\mathbf{a}$  es nulo y sabemos que el sistema de coordenadas esféricas es ortogonal, de modo que las condiciones de Staeckel se reducen a aplicar (9.59), para encontrar la forma separable más general del potencial. Para ello encontramos primero los elementos (diagonales) de la matriz de energía cinética (ver Ec. 6.24, Pág. 98)

$$T_{rr} = m \quad ; \quad T_{\theta\theta} = mr^2 \quad ; \quad T_{\phi\phi} = mr^2 \sin^2 \theta \quad (9.80)$$

aplicando (9.59) la estructura más general del potencial separable es de la forma

$$V(q) = \frac{V'_r(r)}{T_{rr}} + \frac{V'_\theta(\theta)}{T_{\theta\theta}} + \frac{V'_\phi(\phi)}{T_{\phi\phi}} = V_r(r) + \frac{V_\theta(\theta)}{r^2} + \frac{V_\phi(\phi)}{r^2 \sin^2 \theta} \quad ; \quad V_q(q) \equiv \frac{V'_q(q)}{m}$$

se puede comprobar la separabilidad de este potencial por sustitución directa en la ecuación de HJ.

## 9.7. Otros problemas de aplicación con el formalismo de H-J

### 9.7.1. Partícula sometida a potencial armónico y campo magnético

Sea una partícula restringida a moverse en un plano, bajo la influencia de un potencial central (no electromagnético)  $V(r) = (k/2)r^2$  y un campo magnético constante  $\mathbf{B}$ , perpendicular al plano. Reduciremos este problema a cuadraturas utilizando HJ. Supondremos que el plano de movimiento pasa por el origen (el foco del potencial central) de modo que las fuerzas armónica y magnética están en dicho plano.

Si bien puede existir una fuerza extra de ligadura que mantenga a la partícula en el plano<sup>10</sup>, ésta no produce trabajo virtual y no contribuye al potencial y por tanto, tampoco contribuye al Lagrangiano ni al Hamiltoniano. Tomando  $XY$  como el plano de movimiento, el Lagrangiano en coordenadas cartesianas se escribe

$$L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2) + q(\dot{\mathbf{r}} \cdot \mathbf{A}) - \frac{k}{2} (x^2 + y^2)$$

un potencial vectorial válido para campo magnético homogéneo y constante es el dado por la Ec. (5.72), Pág. (91)

$$\begin{aligned} \mathbf{A} &= \frac{1}{2} \mathbf{B} \times \mathbf{r} = \frac{1}{2} B \mathbf{u}_z \times (x \mathbf{u}_x + y \mathbf{u}_y + z \mathbf{u}_z) = \frac{1}{2} (x B \mathbf{u}_y - y B \mathbf{u}_x) \\ \dot{\mathbf{r}} \cdot \mathbf{A} &= (\dot{x} \mathbf{u}_x + \dot{y} \mathbf{u}_y + \dot{z} \mathbf{u}_z) \cdot \frac{1}{2} (-y B \mathbf{u}_x + x B \mathbf{u}_y) = \frac{B}{2} (x \dot{y} - y \dot{x}) \end{aligned}$$

y el Lagrangiano queda

$$L = \frac{m}{2} (\dot{x}^2 + \dot{y}^2) + \frac{qB}{2} (x \dot{y} - y \dot{x}) - \frac{k}{2} (x^2 + y^2)$$

introduciendo coordenadas polares

$$x = r \cos \theta \quad , \quad \dot{x} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta \quad ; \quad y = r \sin \theta \quad , \quad \dot{y} = \dot{r} \sin \theta + r \dot{\theta} \cos \theta$$

<sup>10</sup>Aunque las fuerzas aplicadas están en el plano, una componente de la velocidad inicial perpendicular al plano, sacaría a la partícula de éste.

el Lagrangiano queda

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2) + \frac{qB}{2} [(r \cos \theta) (\dot{r} \sin \theta + r \dot{\theta} \cos \theta) - (r \sin \theta) (\dot{r} \cos \theta - r \dot{\theta} \sin \theta)] - \frac{k}{2} r^2$$

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2) + \frac{qB}{2} r^2 \dot{\theta} - \frac{k}{2} r^2 = -\frac{k}{2} r^2 + \begin{pmatrix} \dot{r} & \dot{\theta} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{qB}{2} r^2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \dot{r} & \dot{\theta} \end{pmatrix} \begin{pmatrix} m & 0 \\ 0 & mr^2 \end{pmatrix} \begin{pmatrix} \dot{r} \\ \dot{\theta} \end{pmatrix}$$

comparando con (6.16), tenemos

$$L_0(q, t) = -\frac{k}{2} r^2 ; \quad \mathbf{a} = \begin{pmatrix} 0 \\ \frac{qB}{2} r^2 \end{pmatrix} ; \quad \mathbf{T} = \begin{pmatrix} m & 0 \\ 0 & mr^2 \end{pmatrix} ; \quad \mathbf{T}^{-1} = \begin{pmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{mr^2} \end{pmatrix}$$

y usando (6.20) podemos escribir el Hamiltoniano

$$H = \frac{1}{2} (\tilde{\mathbf{p}} - \tilde{\mathbf{a}}) \mathbf{T}^{-1} (\mathbf{p} - \mathbf{a}) - L_0(q, t) = \frac{1}{2} \begin{pmatrix} p_r & p_\theta - \frac{qB}{2} r^2 \end{pmatrix} \begin{pmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{mr^2} \end{pmatrix} \begin{pmatrix} p_r \\ p_\theta - \frac{qB}{2} r^2 \end{pmatrix} + \frac{k}{2} r^2$$

$$H = \frac{1}{2m} p_r^2 + \frac{1}{2mr^2} \left( p_\theta - \frac{qB}{2} r^2 \right)^2 + \frac{1}{2} k r^2 \quad (9.81)$$

aplicando la ecuación de Hamilton-Jacobi (9.5) al Hamiltoniano (9.81), se obtiene

$$\frac{1}{2m} \left( \frac{\partial S}{\partial r} \right)^2 + \frac{1}{2mr^2} \left( \frac{\partial S}{\partial \theta} - \frac{qB}{2} r^2 \right)^2 + \frac{1}{2} k r^2 + \frac{\partial S}{\partial t} = 0 \quad (9.82)$$

puesto que el Hamiltoniano no depende explícitamente del tiempo y  $\theta$  es cíclica, proponemos una solución de la forma

$$S(r, \theta, E, \alpha, t) = W_r(r, E, \alpha) + p_\theta \theta - Et \quad (9.83)$$

con lo cual la ecuación de HJ (9.82) queda

$$\frac{1}{2m} \left( \frac{dW_r}{dr} \right)^2 + \frac{1}{2mr^2} \left( p_\theta - \frac{qB}{2} r^2 \right)^2 + \frac{1}{2} k r^2 = E$$

cuya solución formal es

$$W_r(r) = \int dr \sqrt{2mE - mkr^2 - \frac{1}{r^2} \left( p_\theta - \frac{qB}{2} r^2 \right)^2} \quad (9.84)$$

definiendo  $k = m\omega_0^2$ , la solución se puede reescribir como

$$W_r(r) = \int dr \sqrt{2m \left( E + \frac{qB p_\theta}{2m} \right) - m^2 \left[ \omega_0^2 + \left( \frac{qB}{2m} \right)^2 \right] r^2 - \frac{p_\theta^2}{r^2}} \quad (9.85)$$

Si comparamos esta solución con la solución (9.31) para el oscilador armónico bidimensional isotrópico, vemos que la Ec. (9.85) equivaldría a un oscilador armónico bidimensional isotrópico con energía  $E'$  y frecuencia angular  $\omega$  dadas por

$$E' = E + \omega_c p_\theta \quad , \quad \omega = \sqrt{\omega_0^2 + \omega_c^2} \quad ; \quad \omega_0 \equiv \sqrt{\frac{k}{m}} \quad , \quad \omega_c \equiv \frac{qB}{2m}$$

nótese que  $\omega_c$  es la mitad de la frecuencia de ciclotrón de la partícula en el campo magnético  $\mathbf{B}$ . Cuando  $B = 0$ , se obtiene  $E' = E$  y  $\omega = \omega_0$  como era de esperarse. Naturalmente, el problema puede formularse directamente con la función característica  $W$  en lugar de la función principal  $S$ .

### 9.7.2. Partícula bajo potencial conservativo en coordenadas elipsoidales

Una partícula en el espacio está sometida a un potencial conservativo  $V(\mathbf{r})$ . Escribiremos las ecuaciones de HJR en coordenadas elipsoidales  $u, v, \phi$ , las cuales se pueden definir con respecto a las coordenadas cilíndricas  $(r, z, \phi)$  de la siguiente manera

$$r = a \sinh v \sin u \quad ; \quad z = a \cosh v \cos u \quad ; \quad \phi = \phi \quad (9.86)$$

donde  $a$  es una distancia arbitraria pero fija ( $a > 0$ ). Encontraremos la ecuación de HJR asociada al problema.

Comenzamos por plantear la energía cinética  $T$  en coordenadas elipsoidales. Para ello partimos de la expresión de  $T$  en coordenadas cilíndricas

$$T = \frac{1}{2}m(\dot{r}^2 + \dot{z}^2) + \frac{1}{2}mr^2\dot{\phi}^2 \quad (9.87)$$

y usando las relaciones (9.86) obtenemos

$$\begin{aligned} \dot{r} &= a\dot{v} \cosh v \sin u + a\dot{u} \sinh v \cos u \quad ; \quad \dot{z} = a\dot{v} \sinh v \cos u - a\dot{u} \cosh v \sin u \\ \dot{r}^2 + \dot{z}^2 &= a^2(\cosh^2 v \sin^2 u + \sinh^2 v \cos^2 u)(\dot{v}^2 + \dot{u}^2) \\ &= a^2[(1 + \sinh^2 v) \sin^2 u + \sinh^2 v (1 - \sin^2 u)](\dot{v}^2 + \dot{u}^2) \\ \dot{r}^2 + \dot{z}^2 &= a^2(\sin^2 u + \sinh^2 v)(\dot{v}^2 + \dot{u}^2) \end{aligned}$$

de modo que la energía cinética (9.87) y el Lagrangiano en coordenadas elipsoidales queda

$$\begin{aligned} T &= \frac{1}{2}ma^2(\sin^2 u + \sinh^2 v)(\dot{v}^2 + \dot{u}^2) + \frac{1}{2}ma^2(\sinh^2 v \sin^2 u)\dot{\phi}^2 \\ L &= \frac{1}{2}ma^2(\sin^2 u + \sinh^2 v)(\dot{v}^2 + \dot{u}^2) + \frac{1}{2}ma^2(\sinh^2 v \sin^2 u)\dot{\phi}^2 - V(u, v, \phi) \\ L &= -V(u, v, \phi) + \\ &+ \frac{1}{2} \begin{pmatrix} \dot{u} & \dot{v} & \dot{\phi} \end{pmatrix} \begin{pmatrix} ma^2(\sin^2 u + \sinh^2 v) & 0 & 0 \\ 0 & ma^2(\sin^2 u + \sinh^2 v) & 0 \\ 0 & 0 & ma^2(\sinh^2 v \sin^2 u) \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{v} \\ \dot{\phi} \end{pmatrix} \end{aligned}$$

comparando con (6.16) tenemos

$$\begin{aligned} L_0 &= -V(u, v, \phi) \quad ; \quad \mathbf{T} = ma^2 \begin{pmatrix} \sin^2 u + \sinh^2 v & 0 & 0 \\ 0 & \sin^2 u + \sinh^2 v & 0 \\ 0 & 0 & \sinh^2 v \sin^2 u \end{pmatrix} \\ \mathbf{a} &= \mathbf{0} \quad ; \quad \mathbf{T}^{-1} = \frac{1}{ma^2} \begin{pmatrix} \frac{1}{\sin^2 u + \sinh^2 v} & 0 & 0 \\ 0 & \frac{1}{\sin^2 u + \sinh^2 v} & 0 \\ 0 & 0 & \frac{1}{\sinh^2 v \sin^2 u} \end{pmatrix} \end{aligned}$$

con lo cual el Hamiltoniano (6.20) queda

$$\begin{aligned} H &= \tilde{\mathbf{p}}\mathbf{T}^{-1}\mathbf{p} - L_0 = \frac{1}{2ma^2} \begin{pmatrix} p_u & p_v & p_\phi \end{pmatrix} \begin{pmatrix} \frac{1}{\sin^2 u + \sinh^2 v} & 0 & 0 \\ 0 & \frac{1}{\sin^2 u + \sinh^2 v} & 0 \\ 0 & 0 & \frac{1}{\sinh^2 v \sin^2 u} \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{v} \\ \dot{\phi} \end{pmatrix} + V(u, v, \phi) \\ H &= \frac{p_u^2 + p_v^2}{2ma^2(\sin^2 u + \sinh^2 v)} + \frac{p_\phi^2}{2ma^2 \sinh^2 v \sin^2 u} + V(u, v, \phi) \quad (9.88) \end{aligned}$$

para este Hamiltoniano, la ecuación de Hamilton-Jacobi restringida (9.32) nos da

$$\frac{\left[ \left( \frac{\partial W}{\partial u} \right)^2 + \left( \frac{\partial W}{\partial v} \right)^2 \right]}{2ma^2(\sin^2 u + \sinh^2 v)} + \frac{\left( \frac{\partial W}{\partial \phi} \right)^2}{2ma^2 \sinh^2 v \sin^2 u} + V(u, v, \phi) = E$$

asumiendo separabilidad aditiva de la función característica de Hamilton, tenemos

$$W = W_u + W_v + W_\phi$$

con lo cual la ecuación de HJR queda

$$\frac{\left[ \left( \frac{\partial W_u}{\partial u} \right)^2 + \left( \frac{\partial W_v}{\partial v} \right)^2 \right]}{2ma^2 (\sin^2 u + \sinh^2 v)} + \frac{\left( \frac{\partial W_\phi}{\partial \phi} \right)^2}{2ma^2 \sinh^2 v \sin^2 u} + V(u, v, \phi) = E \quad (9.89)$$

Simplificaremos un poco el problema asumiendo que el potencial tiene simetría azimutal, de modo que  $V = V(u, v)$ . En tal caso,  $\phi$  será cíclica en el Hamiltoniano (9.88), con lo cual  $W_\phi = p_\phi \phi$ , siendo

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = m\dot{\phi} a^2 \sinh^2 v \sin^2 u = mr^2 \dot{\phi} = cte$$

y la ecuación de HJR se simplifica en la forma

$$\begin{aligned} \frac{\left[ \left( \frac{\partial W_u}{\partial u} \right)^2 + \left( \frac{\partial W_v}{\partial v} \right)^2 \right]}{2ma^2} + \frac{(\sin^2 u + \sinh^2 v) p_\phi^2}{2ma^2 \sinh^2 v \sin^2 u} + (\sin^2 u + \sinh^2 v) V(u, v) &= (\sin^2 u + \sinh^2 v) E \\ \frac{\left[ \left( \frac{\partial W_u}{\partial u} \right)^2 + \left( \frac{\partial W_v}{\partial v} \right)^2 \right]}{2ma^2} + \frac{p_\phi^2}{2ma^2} \left( \frac{1}{\sinh^2 v} + \frac{1}{\sin^2 u} \right) + (\sin^2 u + \sinh^2 v) V(u, v) &= (\sin^2 u + \sinh^2 v) E \end{aligned}$$

esta ecuación se puede reescribir en la forma

$$\left[ \frac{\left( \frac{\partial W_u}{\partial u} \right)^2}{2ma^2} + \frac{p_\phi^2}{2ma^2 \sin^2 u} - E \sin^2 u \right] + (\sin^2 u + \sinh^2 v) V(u, v) = E \sinh^2 v - \frac{\left( \frac{\partial W_v}{\partial v} \right)^2}{2ma^2} - \frac{p_\phi^2}{2ma^2 \sinh^2 v} \quad (9.90)$$

que estructuralmente se escribe como

$$F(u) + (\sin^2 u + \sinh^2 v) V(u, v) = G(v)$$

es claro que esta ecuación admitirá separación de variables si el término  $(\sin^2 u + \sinh^2 v) V(u, v)$  se puede separar, es decir si se cumple

$$\begin{aligned} (\sin^2 u + \sinh^2 v) V(u, v) &= V_u(u) + V_v(v) \\ V(u, v) &= \frac{V_u(u) + V_v(v)}{\sin^2 u + \sinh^2 v} \end{aligned} \quad (9.91)$$

si el potencial tiene la estructura dada en (9.91), la Ec. (9.90) queda

$$\frac{\left( \frac{\partial W_u}{\partial u} \right)^2}{2ma^2} + \frac{p_\phi^2}{2ma^2 \sin^2 u} - E \sin^2 u + V_u(u) = E \sinh^2 v - \frac{\left( \frac{\partial W_v}{\partial v} \right)^2}{2ma^2} - \frac{p_\phi^2}{2ma^2 \sinh^2 v} - V_v(v) \quad (9.92)$$

y dado que el miembro izquierdo solo depende de  $u$  y el derecho solo de  $v$ , ambos deben ser iguales a una constante  $A$ , con lo cual obtenemos dos ecuaciones diferenciales ordinarias que reducen el problema a cuadraturas

$$\frac{1}{2ma^2} \left( \frac{dW_u}{du} \right)^2 + \frac{p_\phi^2}{2ma^2 \sin^2 u} - E \sin^2 u + V_u(u) = A \quad (9.93)$$

$$\frac{1}{2ma^2} \left( \frac{dW_v}{dv} \right)^2 + \frac{p_\phi^2}{2ma^2 \sinh^2 v} - E \sinh^2 v + V_v(v) = -A \quad (9.94)$$

### Partícula puntual bajo el campo gravitacional de dos masas desiguales por HJR

Vamos a ilustrar la separación de variables en la Ec. (9.92), reduciendo a cuadraturas el problema de una masa puntual  $m$ , que se mueve en el campo gravitacional generado por dos masas desiguales fijas  $M_1$  y  $M_2$ . Dado que  $a$  en las Ecs. (9.86) es una cantidad positiva arbitraria pero fija, la definiremos de modo que sea la mitad de la distancia entre  $M_1$  y  $M_2$ . Sin pérdida de generalidad, podemos asumir que  $M_1$  y  $M_2$  yacen en el eje  $Z$  en las posiciones  $\mathbf{r}_1 = a\mathbf{u}_z$ ,  $\mathbf{r}_2 = -a\mathbf{u}_z$ . Si  $\mathbf{r}$  es la posición de la masa  $m$ , el potencial generado por  $M_1$  y  $M_2$  en la posición  $\mathbf{r}$  de  $m$  será

$$V(\mathbf{r}) = -\frac{GmM_1}{|\mathbf{r} - \mathbf{r}_1|} - \frac{GmM_2}{|\mathbf{r} - \mathbf{r}_2|} = -\frac{GmM_1}{|\mathbf{r} - a\mathbf{u}_z|} - \frac{GmM_2}{|\mathbf{r} + a\mathbf{u}_z|} \quad (9.95)$$

una ventaja de poner a  $M_1$  y  $M_2$  sobre el eje  $Z$ , es que el potencial tendrá automáticamente simetría azimutal. Resta entonces demostrar que el potencial (9.95) posee en coordenadas elipsoidales la estructura dada en la Ec. (9.91), para reducir este problema a cuadraturas.

Denotaremos a la posición de la masa  $m$  como  $\mathbf{r} = x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z$  de modo que

$$\begin{aligned} |\mathbf{r} \pm a\mathbf{u}_z|^2 &= |x\mathbf{u}_x + y\mathbf{u}_y + (z \pm a)\mathbf{u}_z|^2 = x^2 + y^2 + (z \pm a)^2 \\ |\mathbf{r} \pm a\mathbf{u}_z|^2 &= r^2 + (z \pm a)^2 \end{aligned} \quad (9.96)$$

donde  $r$ ,  $z$  y  $\phi$  son las coordenadas cilíndricas de  $\mathbf{r}$ . Ahora escribimos la expresión (9.96) en coordenadas elipsoidales, haciendo uso de las Ecs. (9.86)

$$\begin{aligned} |\mathbf{r} \pm a\mathbf{u}_z|^2 &= a^2 \sinh^2 v \sin^2 u + (a \cosh v \cos u \pm a)^2 \\ &= a^2 (\sinh^2 v \sin^2 u + \cosh^2 v \cos^2 u + 1 \pm 2 \cosh v \cos u) \end{aligned}$$

nótese que esta factorización fue posible, debido a que la constante  $a$  de la Ec. (9.86) se hizo coincidir con la mitad de la distancia entre  $M_1$  y  $M_2$  de modo que  $\mathbf{r}_{1,2} = \pm a\mathbf{u}_z$ . Usando  $\cosh^2 v = 1 + \sinh^2 v$ , tenemos

$$\begin{aligned} |\mathbf{r} \pm a\mathbf{u}_z|^2 &= a^2 [\sinh^2 v \sin^2 u + (1 + \sinh^2 v) \cos^2 u + 1 \pm 2 \cosh v \cos u] \\ &= a^2 (\sinh^2 v + \cos^2 u + 1 \pm 2 \cosh v \cos u) \\ &= a^2 (\cosh^2 v + \cos^2 u \pm 2 \cosh v \cos u) \end{aligned}$$

de lo cual se obtiene finalmente

$$|\mathbf{r} \pm a\mathbf{u}_z|^2 = a^2 (\cosh v \pm \cos u)^2$$

que al reemplazar en el potencial (9.95) nos da

$$V = -\frac{GmM_1}{a(\cosh v - \cos u)} - \frac{GmM_2}{a(\cosh v + \cos u)} = -\frac{1}{a} \frac{GmM_1 (\cosh v + \cos u) + GmM_2 (\cosh v - \cos u)}{\cosh^2 v - \cos^2 u}$$

el denominador se puede reescribir como

$$\cosh^2 v - \cos^2 u = \sinh^2 v + 1 - \cos^2 u = \sinh^2 v + \sin^2 u$$

reorganizando numerador y denominador, el potencial queda

$$\begin{aligned} V(u, v) &= -\frac{1}{a} \frac{Gm(M_1 - M_2) \cos u + Gm(M_1 + M_2) \cosh v}{\sin^2 u + \sinh^2 v} \\ V(u, v) &= \frac{[\frac{1}{a} Gm(M_2 - M_1) \cos u] + [-\frac{1}{a} Gm(M_1 + M_2) \cosh v]}{\sin^2 u + \sinh^2 v} \equiv \frac{V_u(u) + V_v(v)}{\sin^2 u + \sinh^2 v} \end{aligned}$$

por tanto este potencial posee la estructura (9.91) con

$$V_u(u) = \frac{1}{a}Gm(M_2 - M_1)\cos u \ ; \ V_v(v) = -\frac{1}{a}Gm(M_1 + M_2)\cosh v \quad (9.97)$$

y reemplazando (9.97) en las Ecs. (9.93, 9.94) el problema queda reducido a cuadraturas

$$\frac{1}{2ma^2} \left( \frac{dW_u}{du} \right)^2 + \frac{p_\phi^2}{2ma^2 \sin^2 u} - E \sin^2 u + \frac{1}{a}Gm(M_2 - M_1)\cos u = A \quad (9.98)$$

$$\frac{1}{2ma^2} \left( \frac{dW_v}{dv} \right)^2 + \frac{p_\phi^2}{2ma^2 \sinh^2 v} - E \sinh^2 v - \frac{1}{a}Gm(M_1 + M_2)\cosh v = -A \quad (9.99)$$

### 9.8. Variables acción-ángulo para sistemas con un grado de libertad

En múltiples ramas de la Física es de interés el estudio de movimientos periódicos. En muchos casos no estamos interesados en los detalles de la órbita sino en la determinación de las frecuencias del movimiento. Una variante del formalismo de Hamilton-Jacobi nos permite hallar el periodo de estos movimientos sin resolver completamente la ecuación de movimiento. Ya hemos enfatizado que en el método de HJ, no estamos obligados a tomar como nuevos momentos conjugados a las  $n$  constantes de integración  $\alpha_i$ , sino que podemos tomar un conjunto de  $n$  funciones  $\gamma_k(\alpha)$  independientes entre sí. Aprovechando esta arbitrariedad, no definiremos los nuevos momentos conjugados como las constantes de integración  $\alpha_i$  sino como ciertas funciones de las  $\alpha_i$  que contienen información sobre un ciclo completo de movimiento, estos nuevos momentos se denotan por  $J_i$  y se denominan **variables de acción**.

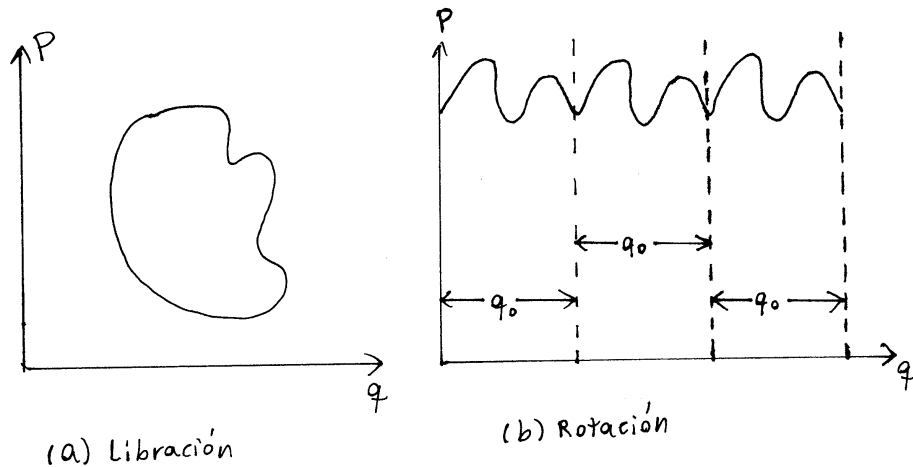


Figura 9.1: Trayectoria en el espacio de fase de un sistema unidimensional (a) para periodicidad tipo libración y (b) para periodicidad tipo rotación.

Consideraremos en primera instancia el caso de un solo grado de libertad. Asumiremos además que el Hamiltoniano es constante

$$H(q, p) = \alpha_1$$

podemos resolver para el momento y obtener

$$p = p(q, \alpha_1) \quad (9.100)$$

la Ecuación (9.100) se puede considerar como una ecuación de trayectoria en el espacio de fase. Cuando el movimiento es periódico se habla de una órbita en el espacio de fase. Definiremos dos tipos de periodicidad de acuerdo con las características de la órbita en el espacio de fase:

1. En el primer tipo la órbita es **cerrada en el espacio de fase** como se ve en la figura 9.1(a). De modo que el punto que define la dinámica en el espacio de fase, regresa sobre la misma trayectoria en forma periódica. Tanto  $q$  como  $p$  son funciones periódicas en el tiempo con la misma frecuencia. Un movimiento periódico de esta naturaleza se puede encontrar cuando la posición inicial yace entre dos ceros de la energía cinética, este movimiento se conoce con el término astronómico de **libración** y tiene como principal exponente al oscilador armónico unidimensional.
2. En el segundo tipo de órbita, la trayectoria en el espacio de fase es tal que  $p$  es una función periódica de  $q$ , con periodo  $q_0$ , como se ilustra en la figura 9.1(b). Equivalentemente este tipo de movimiento implica que cuando la coordenada se incrementa en  $q_0$ , el sistema permanece básicamente inalterado. El ejemplo más familiar es el cuerpo rígido que rota sobre un eje fijo, siendo  $q$  el ángulo de rotación. Cuando se incrementa  $q$  en  $2\pi$ , no se produce ningún cambio esencial en el estado del cuerpo rígido. En realidad para este tipo de periodicidad la coordenada de posición está siempre asociada a un ángulo de rotación de tal modo que el movimiento periódico asociado se denomina simplemente **rotación**, en contraste con la libración. En este caso los valores de  $q$  ya no están acotados sino que pueden crecer indefinidamente.

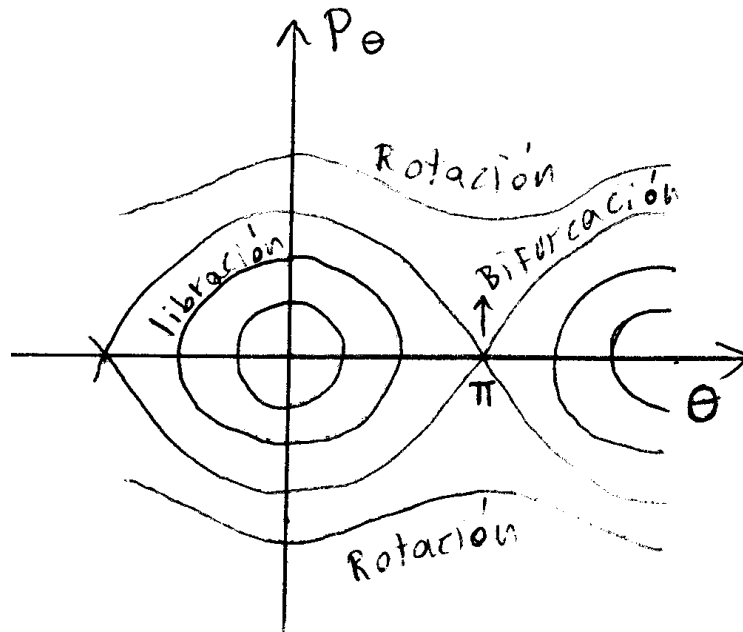


Figura 9.2: Trayectorias en el espacio de fase de un péndulo simple, para diferentes condiciones iniciales. Cuando  $E < mgl$ , obtenemos una órbita cerrada i.e periodicidad tipo libración. Para el caso  $E = mgl$  tenemos bifurcación. Finalmente, cuando  $E > mgl$ , tenemos periodicidad de rotación.

Hay sistemas físicos que pueden exhibir cualquiera de estos dos tipos de movimiento periódico. Un ejemplo sencillo es el péndulo simple consistente en una lenteja sostenida por una varilla fija a un punto de suspensión, donde el movimiento de la lenteja es en un plano<sup>11</sup>. En este caso,  $q$  es el ángulo de deflexión respecto a la vertical. Tomando el cero de potencial en el punto de suspensión de la varilla sin masa, la energía es constante y coincide con el Hamiltoniano quedando

$$E = \frac{p_\theta^2}{2ml^2} - mgl \cos \theta$$

<sup>11</sup> Asumimos una varilla en lugar de una cuerda para que la distancia al punto de suspensión sea siempre la misma, sin importar la amplitud angular.



siendo  $l$ , la longitud de la varilla. Despejando  $p_\theta$ , obtenemos la ecuación para el camino que describe el punto del sistema en el espacio de fase en la forma descrita por (9.100)

$$p_\theta = \pm \sqrt{2ml^2 (E + mgl \cos \theta)} \quad (9.101)$$

si  $E$  es menor que  $mgl$  sabemos por consideraciones de energía que el movimiento es acotado en  $\theta$ , de modo que  $|\theta| < \theta_{\text{máx}}$  el valor de la cota superior se obtiene cuando  $p_\theta = 0$  y corresponde al valor

$$\cos \theta_{\text{máx}} = -\frac{E}{mgl}$$

en este caso la energía es **siempre negativa**. Claramente el péndulo oscila entre  $-\theta_{\text{máx}}$  y  $\theta_{\text{máx}}$  lo cual nos da un movimiento periódico tipo libración (ver figura 9.2). Nótese además que el doble signo en (9.101) es necesario para que la órbita sea cerrada, es decir no define una función. Por otro lado, si  $E > mgl$ , la energía es suficiente para que el péndulo pueda girar completamente de modo que  $\theta$  puede tomar cualquier valor y no está acotado, este sería un movimiento periódico tipo rotación (ver figura 9.2). El límite  $E = mgl$  se ilustra también en la Fig. 9.2, y corresponde a un péndulo que llega a  $\theta = \pi$  con energía cinética cero es decir con  $p_\theta = 0$ . Este es un punto de equilibrio inestable en el cual se puede quedar indefinidamente, pero si hay la menor perturbación el péndulo toma uno de dos caminos drásticamente diferentes (giro en cualquiera de los sentidos). El punto  $\theta = \pi$ ,  $p_\theta = 0$  es un punto de silla para el Hamiltoniano  $H = E(p_\theta, \theta)$ , y en dicho punto se intersectan dos caminos del espacio de fase con energía constante. Este fenómeno se conoce como **bifurcación**.

### 9.8.1. Formulación general de la variables acción-ángulo en una dimensión

Ignorando por el momento la posibilidad de la bifurcación, asumiremos que nuestro sistema unidimensional tiene movimiento de libración o de rotación dependiendo de las condiciones iniciales. Para cualquiera de estos movimientos periódicos es útil introducir una nueva variable  $J$ , designada para reemplazar a  $\alpha_1$  como el nuevo momento conjugado del sistema. Es decir el nuevo momento conjugado ya no será el Hamiltoniano sino la variable acción  $J$  definida por

$$J = \oint p dq \quad (9.102)$$

la integración se realiza en un ciclo completo de libración o rotación según el caso. Obsérvese que este término se asemeja a la acción abreviada definida en (6.92) lo cual justifica su nombre de variable de acción. Las dimensiones de  $J$  siempre serán en consecuencia unidades de momento angular. Teniendo en cuenta (9.100) y (9.102) se observa que  $J$  es función exclusivamente de  $\alpha_1$  i.e. del Hamiltoniano

$$J = J(\alpha_1) \Rightarrow \alpha_1 \equiv H = H(J)$$

es decir cumple con la estructura dada por (9.39). Vemos además que la coordenada conjugada a  $J$  es cíclica, como se espera con el método de HJR. La función característica de Hamilton puede escribirse como

$$W = W(q, J) \quad (9.103)$$

La coordenada generalizada conjugada a  $J$  se denomina **variable angular**  $w$ , y se define usando el segundo conjunto de ecuaciones de transformación (9.33)

$$w = \frac{\partial W(q, J)}{\partial J} \quad (9.104)$$

las ecuaciones de Hamilton para el conjunto canónico  $w, J$  nos dan

$$\dot{w} = \frac{\partial H(J)}{\partial J} = v(J) \quad (9.105)$$

$$\dot{J} = \frac{\partial H(J)}{\partial w} = 0 \Rightarrow J = cte \quad (9.106)$$

$J$  es constante ya que está asociado a una coordenada cíclica. A su vez, puesto que  $v(J)$  solo depende de  $J$ , también es constante. Por tanto, la Ec. (9.105) tiene solución inmediata

$$w = vt + \beta \quad (9.107)$$

y  $w$  es función lineal del tiempo lo cual es lógico ya que este es un caso particular de la Ec. (9.41) con  $w$  haciendo el papel de la nueva coordenada  $Q$ .

Una vez obtenida  $W(q, J)$ , la solución formal del problema se puede obtener despejando  $q$  en términos de  $w, J$  a partir de (9.104), y luego reemplazando  $w$  por la expresión en (9.107). Esto nos da  $q$  como función del tiempo y de las constantes  $J, v, \beta$  que a su vez se pueden obtener de las condiciones iniciales. Sin embargo, este procedimiento no posee ninguna ventaja significativa con respecto a otras elecciones de los nuevos momentos. El verdadero poder de la formulación en variables acción ángulo radica en la interpretación Física de las constantes  $v(J)$ . Para ver su significado primero consideremos el cambio en  $w$  cuando  $q$  hace un ciclo completo de libración o rotación, este cambio está dado por

$$\Delta w = \oint \frac{\partial w}{\partial q} dq \quad (9.108)$$

usando (9.104) queda<sup>12</sup>

$$\Delta w = \oint \frac{\partial^2 W}{\partial q \partial J} dq$$

dado que  $J$  es constante a lo largo de todo el ciclo, la derivada con respecto a  $J$  se puede sacar fuera del signo integral

$$\Delta w = \frac{d}{dJ} \oint \frac{\partial W}{\partial q} dq = \frac{d}{dJ} \oint p dq = 1$$

donde hemos usado (9.33) y la definición de  $J$  (9.102). Esta ecuación establece que  $w$  cambia en la unidad ( $w$  es adimensional) cuando  $q$  se mueve a lo largo de un ciclo. Si ahora asumimos que el movimiento es periódico en el tiempo<sup>13</sup>, entonces este cambio también se puede evaluar de (9.107)

$$\Delta w = w(\tau) - w(0) = v\tau = 1$$

donde  $\tau$  denota el periodo para un ciclo completo de  $q$  en cualquiera de los dos tipos de periodicidad. De aquí resulta entonces que  $v$  es el inverso del periodo, es decir la frecuencia asociada al movimiento periódico de  $q$ . Por lo tanto el formalismo de las variables acción ángulo nos permite evaluar la frecuencia del movimiento periódico **sin resolver completamente el movimiento del sistema**. Si sabemos a priori que cierto sistema de un grado de libertad es periódico en cualquiera de las dos formas, la frecuencia se puede determinar una vez que  $H$  se escribe en términos de  $J$  y aplicando (9.105). La identificación de  $v(J)$  como una frecuencia y la Ec. (9.107) nos refuerza el hecho de que a la variable  $w$  se le denomine variable angular. De la misma forma se puede ver que si  $w$  tiene dimensiones de ángulo (adimensional) el momento conjugado  $J$  debe tener dimensiones de momento angular.

<sup>12</sup>Nótese que en (9.108) no aparece dependencia explícita del tiempo como a priori se vé en (9.107). La razón es que cuando  $w$  se escribe en términos de  $q$  queda en la forma  $w = w(q, J)$  según se puede ver de las Ecs. (9.103, 9.104). Además  $J$  es constante a lo largo del ciclo por construcción.

<sup>13</sup>Nótese que hasta este punto no se ha utilizado la periodicidad en el tiempo, solo el hecho de que la trayectoria en el espacio de fase sea cerrada o que  $p = p(q)$  sea periódica en  $q$ , pero podría ocurrir que cada ciclo en el espacio de fase tomara un tiempo distinto en ejecutarse.

## 9.9. Problemas de aplicación de variables acción-ángulo con un solo grado de libertad

### 9.9.1. El oscilador armónico unidimensional en variables acción-ángulo

Como ejemplo de aplicación de este formalismo veamos el oscilador armónico unidimensional. Retomando el Hamiltoniano del oscilador armónico denotándolo por  $\alpha$

$$\alpha = \frac{1}{2m} (p^2 + m^2 \omega^2 q^2) \quad (9.109)$$

Para calcular la acción debemos despejar el momento  $p$  en función de la coordenada  $q$ , lo cual nos da

$$p = \pm \sqrt{2m\alpha - m^2 \omega^2 q^2}$$

que coincide con la Ec. (9.22). Podemos ahora calcular la variable de acción en la forma

$$J = \oint p dq = \oint \left( \pm \sqrt{2m\alpha - m^2 \omega^2 q^2} \right) dq$$

siendo  $\alpha$  la energía total (conservada) y  $\omega^2 = k/m$ . Nótese que el signo positivo vale para la mitad del ciclo (en el que se incrementa  $q$ ) y el negativo para la otra mitad (en el que decrece la variable  $q$ ). Por simetría esto se puede escribir como cuatro veces la integral que comprende al movimiento desde el origen hasta el punto de máxima elongación en dirección positiva, en este cuarto de ciclo  $p > 0$  y por tanto el radical es positivo

$$J = 4 \int_0^{q_{\text{máx}}} \sqrt{2m\alpha - m^2 \omega^2 q^2} dq$$

el cambio de variable

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \sin \theta$$

nos define claramente los límites de integración requeridos para el cuarto de ciclo de la franja superior del plano de fase, y la integral se convierte en

$$\begin{aligned} J &= 4 \int_0^{\pi/2} \sqrt{2m\alpha - m^2 \omega^2 \left( \frac{2\alpha}{m\omega^2} \sin^2 \theta \right)} \cos \theta d\theta = 4 \int_0^{\pi/2} \sqrt{\left( \frac{2\alpha}{\omega} \right)^2 [1 - \sin^2 \theta]} \cos \theta d\theta \\ &= \frac{8\alpha}{\omega} \int_0^{\pi/2} |\cos \theta| \cos \theta d\theta \end{aligned}$$

pero en el intervalo  $[0, \pi/2]$  tenemos que  $|\cos \theta| = \cos \theta$ , por tanto

$$J = \frac{8\alpha}{\omega} \int_0^{\pi/2} \cos^2 \theta d\theta = \frac{2\pi\alpha}{\omega} \quad (9.110)$$

despejando  $\alpha$  i.e. el Hamiltoniano resulta

$$\alpha \equiv H(J) = \frac{J\omega}{2\pi} \quad (9.111)$$

con lo cual la frecuencia de oscilación se obtiene aplicando (9.105)

$$v = \frac{\partial H}{\partial J} = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (9.112)$$

en concordancia con la frecuencia obtenida por otros métodos. Es interesante escribir las soluciones para las Ecs. (9.21, 9.23) en función de  $J$  y  $w$ , a pesar de que no se requieren para encontrar las frecuencias. Teniendo en cuenta las Ecs. (9.112, 9.107) vemos que

$$w = vt + \beta = \frac{\omega}{2\pi} t + \beta \Rightarrow \quad (9.113)$$

$$2\pi w = \omega t + \delta ; \quad \delta \equiv 2\pi\beta \quad (9.114)$$

donde hemos definido adecuadamente la constante arbitraria de integración  $\beta$ . Con esta relación, las soluciones de (9.21, 9.23) en función de  $J$  y  $w$  quedan de la forma

$$q = \sqrt{\frac{J}{\pi m \omega}} \sin 2\pi w \quad ; \quad p = \sqrt{\frac{m J \omega}{\pi}} \cos 2\pi w \quad (9.115)$$

Las Ecs. (9.115) definen la transformación canónica que nos lleva del sistema canónico  $(q, p)$  al sistema canónico  $(w, J)$ . Vale la pena enfatizar que la forma explícita de la TC no fué necesaria para calcular la frecuencia del movimiento.

### 9.9.2. Partícula en movimiento periódico en una dimensión bajo un potencial $V(x) = F|x|$

Una partícula posee movimiento periódico en una dimensión bajo la influencia de un potencial  $V(x) = F|x|$ , donde  $F$  es constante positiva. Encontraremos el periodo de movimiento utilizando variables acción-ángulo. El Hamiltoniano del sistema es simplemente

$$H \equiv E = \frac{p^2}{2m} + F|q|$$

despejando el momento conjugado en términos de la coordenada y las constantes de movimiento nos da

$$p = \sqrt{2m(E - F|q|)}$$

con lo cual tenemos una ecuación del tipo (9.100) que describe la trayectoria en el espacio de fase. La variable de acción viene dada por la Ec. (9.102)

$$J = \oint p \, dq = \oint \sqrt{2m(E - F|q|)} \, dq \quad (9.116)$$

Puesto que  $F > 0$ , el movimiento está acotado dentro del intervalo  $[-E/F, E/F]$ , como se puede ver por consideraciones de energía. En un ciclo completo se recorre este intervalo de ida y vuelta. Este movimiento es del tipo libración ya que la coordenada no crece indefinidamente, sino que retorna a los mismos valores. Por simetría, podemos recorrer solo el primer cuadrante desde  $q = 0$  hasta  $q = E/F$  (donde  $p = 0$ ), y multiplicamos por el factor 4 para tener en cuenta el recorrido completo en el espacio de fase. En el recorrido desde 0 hasta  $E/F$  tanto  $p$  como  $q$  son positivos. La integral (9.116) queda entonces

$$J = 4\sqrt{2m} \int_0^{E/F} \sqrt{E - Fq} \, dq$$

con la sustitución

$$u = E - Fq \quad ; \quad du = -F \, dq$$

la integral queda

$$\begin{aligned} J &= 4\sqrt{2m} \int_E^0 u^{1/2} \frac{du}{(-F)} = \frac{4\sqrt{2m}}{F} \int_0^E u^{1/2} \, du \\ J &= \frac{8\sqrt{2m}}{3F} E^{3/2} \end{aligned} \quad (9.117)$$

despejando  $E = H$  en términos de  $J$  encontramos la frecuencia de movimiento

$$\begin{aligned} E &= H = \left( \frac{3FJ}{8\sqrt{2m}} \right)^{2/3} = \frac{(3F)^{2/3}}{4 \sqrt[3]{2m}} J^{2/3} \\ \nu &= \frac{\partial H}{\partial J} = \frac{(3F)^{2/3}}{6 \sqrt[3]{2m}} J^{-1/3} \end{aligned}$$

y utilizando de nuevo la Ec. (9.117) tenemos

$$\begin{aligned}\nu &= \frac{\partial H}{\partial J} = \frac{(3F)^{2/3}}{6 \sqrt[3]{2m}} \left( \frac{3F}{8\sqrt{2m} E^{3/2}} \right)^{1/3} = \frac{3F}{12\sqrt{2mE}} \\ \nu &= \frac{F}{4\sqrt{2mE}} ; \quad \tau = \frac{1}{\nu} = \frac{4\sqrt{2mE}}{F}\end{aligned}\quad (9.118)$$

que nos da la frecuencia y periodo del movimiento en términos de los parámetros del sistema ( $F$  y  $m$ ) y las condiciones iniciales (la energía).

## 9.10. Variables acción-ángulo para sistemas completamente separables<sup>14</sup>

Cuando tenemos un número arbitrario de grados de libertad, el formalismo de Hamilton Jacobi es particularmente relevante cuando existe uno o mas conjuntos coordinados en los cuales la ecuación de Hamilton Jacobi es completamente separable. Si además existe algún tipo de periodicidad análoga a la discutida en la sección anterior, entonces tendremos que la variante del formalismo de Hamilton Jacobi conocida como el formalismo de las variables acción ángulo, puede ser particularmente ventajosa. Tomaremos además la suposición simplificada de que  $H$  es constante, de modo que podemos usar el formalismo de HJR. La separabilidad completa nos lleva a que las ecuaciones de la transformación canónica tengan la forma

$$p_i = \frac{\partial W_i(q_i; \alpha_1, \dots, \alpha_n)}{\partial q_i} \quad (9.119)$$

lo cual nos da cada  $p_i$  en función de su coordenada conjugada  $q_i$  y las  $n$  constantes de integración

$$p_i = p_i(q_i; \alpha_1, \dots, \alpha_n) \quad (9.120)$$

la Ec. (9.120) es el equivalente de (9.100), pero aplicado ahora a varios grados de libertad. La Ec. (9.120) representa la ecuación de órbita de la proyección del punto del sistema sobre el plano  $(q_i, p_i)$  en el espacio de fase ( $i$  es un índice fijo en este caso). Es útil definir variables acción ángulo cuando las ecuaciones de órbita para todos los planos  $(q_i, p_i)$ , describen órbitas cerradas (libraciones) o funciones periódicas de  $q_i$  (rotaciones).

Es de anotar que las condiciones anteriores no significan que necesariamente las variables  $q_i, p_i$  sean funciones periódicas del tiempo, es decir que repitan su valor para intervalos de tiempo regulares. Incluso cuando cada conjunto  $(q_i, p_i)$  es periódico en este sentido, el movimiento completo no es necesariamente periódico. Para tomar un ejemplo concreto, el oscilador armónico con tres grados de libertad puede tener una frecuencia diferente sobre cada coordenada cartesiana, el movimiento solo será periódico si las tres frecuencias son conmensurables entre sí, es decir si los cocientes entre las frecuencias son racionales; de no ser así, se describen figuras de Lissajous que no son curvas cerradas<sup>15</sup>. Tal movimiento se denomina movimiento **periódico múltiple**. Una de las ventajas de la formulación de Hamilton Jacobi en variables acción ángulo, es que permite evaluar todas las frecuencias de un movimiento periódico múltiple sin resolver completamente el problema del movimiento.

En analogía con (9.102) las variables de acción  $J_i$  se definen en términos de integrales de línea sobre periodos completos de la órbita en el plano  $(q_i, p_i)$ .

$$J_i = \oint p_i dq_i \quad (9.121)$$

nótese que la separabilidad manifestada en la ecuación (9.120) es indispensable para que cada  $J_i$  sea una constante, ya que si cada  $p_i$  depende de varios  $q$ 's, esta integral dependerá de todos los  $q$ 's excepto de  $q_i$ . Si una coordenada es cíclica, su momento conjugado es constante de modo que la trayectoria u órbita en el espacio de fase  $(q_i, p_i)$  es una línea recta horizontal, que no parece ser la naturaleza de un movimiento periódico. En realidad el movimiento se puede considerar un caso límite de movimiento periódico de rotación, en el cual

<sup>14</sup>En lo que sigue del capítulo no adoptaremos la convención de suma de índices repetidos a menos que se indique lo contrario.

<sup>15</sup>Vale decir que este ejemplo es en el espacio real y no en el espacio de fase. Pero el razonamiento para el movimiento en el espacio de fase es idéntico.

se le puede asignar un periodo arbitrario a  $q_i$ . Dado que una coordenada de rotación es invariablemente un ángulo, si la coordenada cíclica es angular, entonces tendrá un periodo natural de  $2\pi$ . En consecuencia, es natural (aunque no obligatorio) evaluar la integral en la definición de la variable de acción correspondiente a una coordenada angular cíclica entre 0 y  $2\pi$ . Por tanto, para toda variable cíclica  $q_i$  su variable de acción asociada se definirá como

$$J_i = 2\pi p_i \quad (9.122)$$

usando la Ec. (9.119) en la definición de  $J_i$  Ec. (9.121) se tiene

$$J_i = \oint \frac{\partial W_i(q_i; \alpha_1, \dots, \alpha_n)}{\partial q_i} dq_i \quad (9.123)$$

como ya mencionamos, gracias a la completa separabilidad en las coordenadas, cada  $J_i$  es solo función de las constantes de integración  $\alpha_i$  que aparecen al integrar la ecuación de Hamilton Jacobi y es por tanto constante. Por otro lado, debido a la independencia de los pares  $(q_i, p_i)$ , se tiene que además las  $J$ 's son  $n$  funciones independientes de las  $\alpha$ 's de modo que forman un conjunto adecuado de constantes para definir los nuevos momentos. Al expresar los  $\alpha$ 's en términos de los  $J$ 's, podemos redefinir los argumentos en la función característica<sup>17</sup>

$$W = W(q_1, \dots, q_n; J_1, \dots, J_n) = \sum_{j=1}^n W_j(q_j; J_1, \dots, J_n) \quad (9.124)$$

en tanto que el Hamiltoniano sería función exclusiva de las  $J$

$$H = \alpha_1 = H(J_1, \dots, J_n) \quad (9.125)$$

al igual que en el caso de un grado de libertad, podemos definir variables angulares conjugadas  $w_i$  a través de la función característica  $W$  teniendo en cuenta que esta última es una función generatriz de tipo 2

$$w_i = \frac{\partial W}{\partial J_i} = \sum_{j=1}^n \frac{\partial W_j(q_j; J_1, \dots, J_n)}{\partial J_i} \quad (9.126)$$

esta ecuación muestra que las  $w_i$  son en general funciones de todas las  $q_j$  y todas las  $J_j$ .

$$w_i = w_i(q_1, \dots, q_n; J_1, \dots, J_n) \quad (9.127)$$

De nuevo, la mitad de las ecuaciones de Hamilton conduce a

$$\dot{w}_i = \frac{\partial H(J_1, \dots, J_n)}{\partial J_i} = v_i(J_1, \dots, J_n) \quad (9.128)$$

y de nuevo la otra mitad de las ecuaciones de Hamilton nos conduce a que las  $J$ 's son constantes. Adicionalmente, las  $v_i$  son todas constantes puesto que son funciones exclusivas de las  $J_i$ . Por lo tanto, las ecuaciones de Hamilton anteriores se pueden integrar para obtener  $w_i$  como funciones lineales del tiempo

$$w_i = v_i t + \beta_i \quad (9.129)$$

en general cada  $w_i$  se incrementa de forma diferente a los otros.

En este caso se puede ver que las  $v_i$  son las frecuencias asociadas al movimiento periódico múltiple, pero el argumento para llegar a esta aseveración es más sutil que en el caso de un grado de libertad. Las ecuaciones de transformación que conducen al conjunto canónico  $(w, J)$  implican que cada  $q_j$  y  $p_j$  es una función de las constantes  $J_i$  y las variables  $w_i$ . El objetivo ahora es encontrar la función matemática de los  $q_i$  en términos de los  $w_i$ , para lo cual examinamos el cambio de cada  $w_i$  cuando cada una de las variables  $q_j$  ha ejecutado un número entero de ciclos  $m_j$  de libración o rotación. Este es un procedimiento puramente matemático ya

<sup>17</sup>Dado que  $W$  es de tipo  $F_2$ , debe ser escrita en términos de las antiguas coordenadas y los nuevos momentos  $W = W(q, J)$ , como en la Ec (9.124).

que en este caso no estamos ejecutando un movimiento del sistema en el tiempo. Es como si congeláramos el tiempo y cada  $q_i$  fuera movido independientemente a través de un cierto número de ciclos de su movimiento. Esto es análogo al concepto de desplazamiento virtual desarrollado para el principio de D'Alembert, de modo que los cambios infinitesimales en las  $w_i$  cuando cambian las  $q_i$  infinitesimalmente, son también de naturaleza virtual por lo cual usamos la notación  $\delta w_i$

$$\delta w_i = \sum_{j=1}^n \frac{\partial w_i}{\partial q_j} dq_j = \sum_{j=1}^n \frac{\partial^2 W}{\partial J_i \partial q_j} dq_j = \sum_{j=1}^n \left[ \frac{\partial^2}{\partial J_i \partial q_j} \sum_{k=1}^n W_k(q_k; J_1, \dots, J_n) \right] dq_j$$

donde hemos usado (9.126). La derivada con respecto a  $q_j$  se anula excepto para  $W_j$  y usando (9.119) y (9.120) se obtiene

$$\delta w_i = \frac{\partial}{\partial J_i} \sum_{j=1}^n p_j(q_j; J_1, \dots, J_n) dq_j \tag{9.130}$$

la Ec. (9.130) representa a  $\delta w_i$  como una suma de contribuciones independientes cada una asociada al movimiento virtual de un  $q_j$ . El cambio total en  $w_i$  se puede escribir entonces

$$\Delta w_i = \sum_{j=1}^n \frac{\partial}{\partial J_i} \oint_{m_j} p_j(q_j, J) dq_j \tag{9.131}$$

el operador diferencial con respecto a  $J_i$  se puede mantener por fuera de la integral ya que en el proceso de integración i.e. de variación de las  $q_i$  el conjunto de las  $J_i$  permanece inalterado. El símbolo  $m_j$  indica que cada una de las  $n$  integrales se realiza sobre un cierto número de ciclos, siendo  $m_j$  el número de ciclos que ejecuta la coordenada  $q_j$ . Por otro lado, en virtud de la definición de variables acción ángulo Ec. (9.121) cada una de estas integrales es  $m_j J_j$ . Adicionalmente, dado que cada  $J_i$  es independiente, se sigue que

$$\Delta w_i = m_i \tag{9.132}$$

si algún  $q_j$  no se barre en un número entero de ciclos en la correspondiente integración, habrá un remanente debido a la fracción del ciclo que no se completó de modo que  $\Delta w_i$  no tendrá un valor entero. Si tratamos a los  $w_i$  y los  $m_i$  como arreglos vectoriales podemos escribir la relación anterior de la forma

$$\Delta \mathbf{w} = \mathbf{m} \tag{9.133}$$

en estas instancias conviene reafirmar la importancia de la naturaleza virtual del movimiento de rotación o libración en cada plano  $q_i, p_i$ . Nótese que en (9.130)  $\delta w_i$  se representa como una suma de contribuciones independientes gracias a la virtualidad del movimiento de cada  $q_j$ ; pues de lo contrario, el parámetro tiempo hace que estos ciclos se recorran en forma correlacionada. Más importante aún, en el movimiento real si las frecuencias en cada plano  $q_i, p_i$  no son conmesurables, no existirá un valor finito del tiempo para el cual se hayan ejecutado ciclos completos en cada plano. Es decir para un movimiento real con frecuencias no conmesurables, no hay un valor del tiempo para el cual las Ecs. (9.132, 9.133) sean válidas.

Finalmente, es importante insistir en que hasta el momento no se ha utilizado periodicidad en el tiempo, lo cual está enfatizado por el carácter virtual de los  $\delta w_i$ .

### 9.10.1. Movimientos periódicos múltiples de libración

Supongamos que cada movimiento separado es de tipo libración, de modo que cada  $q_j, p_j$  vuelve a su valor inicial cuando se completa un ciclo. El resultado dado por la Ec. (9.133) puede ser expresado de la siguiente forma: sea  $\boldsymbol{\eta}$  el arreglo vectorial de dimensión  $2n$  de las  $q_i, p_i$ , que es en general función de  $\mathbf{w}$  de tal manera que un cambio no trivial en el cual  $\Delta \boldsymbol{\eta} = 0$ , debe corresponder a un cambio  $\Delta \mathbf{w} = \mathbf{m}$ , que es un arreglo vectorial de componentes enteras<sup>18</sup>. Dado que el número de ciclos para cada coordenada se puede elegir arbitrariamente,

<sup>18</sup>Estrictamente  $\boldsymbol{\eta}$  es función de  $\mathbf{w}$  y  $\mathbf{J}$ , pero  $\mathbf{J}$  es constante durante el proceso de completar uno o más ciclos.

tomemos  $m_k = \delta_{ki}$ , de modo que completamos un ciclo en las variables  $q_i, p_i$  dejando a las otras variables sin transformar, es decir que todas las componentes  $q_k, p_k$  con  $k \neq i$  permanecen sin cambiar en tanto que las componentes  $q_i, p_i$  vuelven a sus valores originales luego de completar un ciclo, esto nos da como resultado que  $\Delta \boldsymbol{\eta} = 0$ . Por tanto, en el caso más general las componentes de  $\boldsymbol{\eta}$  tienen que ser funciones periódicas de **cada**  $w_i$  con periodo unidad; esto es, las  $q$ 's y las  $p$ 's son funciones periódicas múltiples de las  $w$ 's con periodos unidad<sup>19</sup>. Una función periódica múltiple como esta puede ser siempre representada por una expansión de Fourier, la cual para un cierto  $q_k$  se escribe como

$$q_k = \sum_{j_1=-\infty}^{\infty} \dots \sum_{j_n=-\infty}^{\infty} a_{j_1, \dots, j_n}^{(k)} \exp [2\pi i (j_1 w_1 + \dots + j_n w_n)] \quad (9.134)$$

donde los  $j_m$  son enteros. Naturalmente podemos tratar al conjunto de las  $j$ 's y las  $w$ 's como arreglos vectoriales para escribir esta relación en forma mucho más compacta

$$q_k = \sum_{\mathbf{j}} a_{\mathbf{j}}^{(k)} \exp [2\pi i (\mathbf{j} \cdot \mathbf{w})] \quad (9.135)$$

análogamente, la Ec. (9.129) también se puede escribir en forma vectorial

$$\mathbf{w} = \mathbf{v}t + \boldsymbol{\beta} \quad (9.136)$$

y la dependencia temporal de  $q_k$  se escribe de la forma

$$q_k(t) = \sum_{\mathbf{j}} a_{\mathbf{j}}^{(k)} \exp [2\pi i \mathbf{j} \cdot (\mathbf{v}t + \boldsymbol{\beta})] \quad (9.137)$$

obsérvese que en general  $q_k(t)$  no es necesariamente una función periódica en el tiempo; pues la expansión de Fourier solo ha requerido la periodicidad en cada coordenada  $w_i$  por aparte y cada una de ellas crece a diferente ritmo, no es necesario ni siquiera que haya periodicidad temporal para completar cada ciclo en un  $q_i, p_i$  dado. Incluso si asumimos periodicidad temporal para cada  $q_i, p_i$ , la periodicidad en  $q_k(t)$  ocurrirá solo si los  $v_i$  son todos conmensurables entre sí, es decir múltiplos racionales unos de otros. Por lo tanto, en el caso más general  $q_k$  se considera una función **cuasi periódica del tiempo**. Finalmente, debe recordarse que los coeficientes  $a_{\mathbf{j}}^{(k)}$  pueden encontrarse usando el procedimiento estándar para encontrar coeficientes de Fourier. Ellos están dados por integrales múltiples sobre la celda unitaria en el espacio  $\mathbf{w}$

$$a_{\mathbf{j}}^{(k)} = \int_0^1 \dots \int_0^1 q_k(\mathbf{w}) \exp [-2\pi i (\mathbf{j} \cdot \mathbf{w})] d\mathbf{w}$$

$d\mathbf{w}$  define el elemento de volumen en el espacio  $n$  dimensional de las  $w_i$ . Un procedimiento análogo se puede hacer para  $p_k(t)$ .

### 9.10.2. Movimientos cuasi periódicos múltiples de rotación

Cuando el movimiento periódico es del tipo rotación, la coordenada  $q_k$  no retorna a su valor original cuando se realiza un ciclo completo del par de variables separables  $(q_i, p_i)$ , sino que se incrementa en una cantidad  $q_{0k}$ . Tal coordenada de rotación no es en sí múltiplemente periódica. Sin embargo, cuando se realiza un número entero de ciclos  $m_k$  vemos que la variable  $w_k$  ha aumentado en  $m_k$  unidades. En consecuencia, la función  $q_k - w_k q_{0k}$  sí retorna a su valor inicial para cualquier número entero de ciclos y a semejanza del caso de la libración, es una función periódica múltiple de todos los  $w$ 's con periodo unidad. Por tanto, podemos expandir

<sup>19</sup>Nótese que para llegar a esta conclusión, fué necesario que cada plano  $q_i, p_i$  se barriera independientemente, es decir fué importante que el movimiento fuera virtual.



la función en una serie de Fourier múltiple análoga a (9.134), y encontrar la dependencia temporal usando (9.129)

$$q_k - w_k q_{0k} = \sum_{\mathbf{j}} a_{\mathbf{j}}^{(k)} \exp [2\pi i (\mathbf{j} \cdot \mathbf{w})] \Rightarrow$$

$$q_k = q_{0k} (v_k t + \beta_k) + \sum_{\mathbf{j}} a_{\mathbf{j}}^{(k)} \exp [2\pi i \mathbf{j} \cdot (\mathbf{v}t + \boldsymbol{\beta})]$$

por lo tanto, es siempre posible generar una función periódica múltiple a partir de una coordenada de rotación, la cual puede ser manipulada exactamente de la misma forma que una coordenada de libración. En consecuencia, para simplificar la discusión nos restringiremos a trabajar con el movimiento periódico tipo libración.

### 9.10.3. Movimientos periódicos simples y múltiples tipo libración

En el movimiento periódico múltiple tipo libración, los momentos separables  $p_k$  son funciones periódicas múltiples de las  $w$ 's, y pueden ser expandidos en una serie de Fourier múltiple similar a la presentada en (9.134). De esto se deduce que cualquier función de varios de los pares de variables  $(q_i, p_i)$ , también serán funciones periódicas múltiples de los  $w$ 's. De modo que se pueden escribir como

$$f(q, p) = \sum_{\mathbf{j}} b_{\mathbf{j}} \exp [2\pi i \mathbf{j} \cdot \mathbf{w}] = \sum_{\mathbf{j}} b_{\mathbf{j}} \exp [2\pi i \mathbf{j} \cdot (\mathbf{v}t + \boldsymbol{\beta})] \quad (9.138)$$

por ejemplo, cuando las coordenadas cartesianas de las partículas en el sistema no son en sí las coordenadas de separación, ellas aún se pueden escribir como funciones del tiempo en la forma de las Ecs. (9.138), ya que  $\mathbf{r}_i = \mathbf{r}_i(q)$ .

Cuando asumimos periodicidad temporal en cada plano  $q_i, p_i$ ; las Ecs. (9.134, 9.137), representan el tipo más general de movimiento periódico múltiple del tipo libración. Sin embargo, no todos los sistemas con este tipo de movimiento poseen todas las características mostradas en (9.134, 9.137). Por ejemplo, en una amplia gama de problemas de aplicación de las variables acción ángulo las Ecs. (9.126, 9.127) se simplifican a

$$w_i = \frac{\partial W_i}{\partial J_i}(q_i; J_1, \dots, J_n) \quad ; \quad w_i = w_i(q_i; J_1, \dots, J_n) \quad (9.139)$$

de modo que cada coordenada  $q_i$  de separación es función únicamente de su  $w_i$  asociado. Cuando esto ocurre,  $q_k$  es una función periódica solo de  $w_k$ , y la serie de Fourier múltiple se reduce a una serie de Fourier sencilla

$$q_k = \sum_{j=-\infty}^{\infty} a_j^{(k)} \exp [2\pi i (j w_k)] = \sum_{j=-\infty}^{\infty} a_j^{(k)} \exp [(2\pi i) j (v_k t + \beta_k)] \quad (9.140)$$

En el lenguaje de los sistemas de osciladores acoplados, se puede decir que estas  $q$ 's son las coordenadas normales del sistema. Sin embargo, incluso cuando el movimiento de las  $q$ 's se pueda simplificar así, ocurre con frecuencia que las funciones  $f(q)$  de todas las  $q$ 's, tales como las coordenadas cartesianas, continúan siendo funciones periódicas múltiples de las  $w$ 's y deben ser representadas como en la Ec. (9.138). Aunque haya periodicidad temporal en cada plano  $(q_k, p_k)$ , solo habrá periodicidad temporal de tales funciones si las varias frecuencias  $v_k$  son conmensurables. Una vez más, el movimiento de un oscilador armónico bidimensional anisotrópico es un buen ejemplo para ilustrar estas consideraciones.

Supongamos que en un conjunto particular de coordenadas cartesianas, el Hamiltoniano viene dado por

$$H = \frac{1}{2m} [(p_x^2 + 4\pi^2 m^2 v_x^2 x^2) + (p_y^2 + 4\pi^2 m^2 v_y^2 y^2)]$$

estas coordenadas cartesianas son entonces coordenadas de separación adecuadas y cada una exhibe movimiento armónico simple con frecuencias  $v_x$  y  $v_y$  respectivamente. Luego, las soluciones para  $x, y$  son formas

particularmente simples de expansiones sencillas de Fourier de la forma (9.140). Supongamos ahora que las coordenadas están rotadas  $\pi/4$  alrededor del eje  $z$ . Las componentes del movimiento a lo largo de los ejes  $x', y'$  serán<sup>21</sup>

$$\begin{aligned} x' &= \frac{1}{\sqrt{2}} [x_0 \cos 2\pi (v_x t + \beta_x) + y_0 \cos 2\pi (v_y t + \beta_y)] \\ y' &= \frac{1}{\sqrt{2}} [y_0 \cos 2\pi (v_y t + \beta_y) - x_0 \cos 2\pi (v_x t + \beta_x)] \end{aligned} \quad (9.141)$$

si el cociente  $v_x/v_y$  es racional, estas dos expresiones serán conmensurables y corresponderán a figuras de Lissajous cerradas. Pero si el cociente no es racional, la figura es tal que el punto en el espacio de fase nunca vuelve exactamente sobre el mismo trazo y las Ecs. (9.141) nos dan un ejemplo sencillo de expansiones de Fourier de periodicidad múltiple de la forma (9.138).

Incluso cuando  $q_k$  es una función periódica múltiple de todas las  $w$ 's, intuitivamente parece existir una relación entre  $q_k$  y su  $w_k$  asociado (y por tanto con la frecuencia  $v_k$ ). Después de todo es de anotar que el argumento que nos llevó a la Ec. (9.132), nos dice que cuando  $q_k$  completa un ciclo (moviendo solo a  $q_k$ ),  $w_k$  se incrementa en la unidad en tanto que las otras  $w$ 's retornan a sus valores iniciales. Un argumento riguroso para esta relación fué desarrollado por J. Vinti en 1961.

Supongamos que un cierto intervalo de tiempo  $T$  contiene a  $m_k$  ciclos completos de  $q_k$  mas una fracción de un ciclo. En general el tiempo requerido para cada ciclo sucesivo puede ser diferente, puesto que  $q_k$  no es necesariamente una función periódica en el tiempo. Vinti demostró, sobre la base de un teorema de la teoría de números, que cuando  $T$  crece indefinidamente

$$\lim_{T \rightarrow \infty} \frac{m_k}{T} = v_k$$

de modo que la frecuencia promedio del movimiento de  $q_k$  está siempre dada por  $v_k$ , incluso cuando el movimiento completo es más complicado que una función periódica con frecuencia  $v_k$ .

Es notable el hecho de que la ecuación (9.123) nos dice que cuando  $q_i$  hace un recorrido virtual de un ciclo completo (es decir cuando  $w_i$  se incrementa en la unidad) la función característica se incrementa en  $J_i$ . Con una estrategia similar a la que se sigue para el movimiento periódico de rotación, se tiene que la función

$$W' = W - \sum_k w_k J_k \quad (9.142)$$

permanece invariante cuando cada  $w_k$  se incrementa en la unidad, con las otras variables de acción permaneciendo constantes. La ecuación (9.142) representa entonces una función periódica múltiple que se puede expandir en términos de las  $w_i$  (o las frecuencias  $v_i$  y el tiempo), por una serie de la forma (9.138). Puesto que las ecuaciones de transformación para las variables angulares vienen dadas por

$$w_k = \frac{\partial W}{\partial J_k}$$

se puede reconocer que la ecuación (9.142) define una transformación de Legendre desde la base  $(q, J)$  hacia la base  $(q, w)$ . Recordando que  $W$  es una función generatriz tipo 2 y usando las Ecs. (7.14, 7.18) tenemos que si  $W(q, J)$  es de tipo  $F_2$  entonces  $W'(q, w)$  es la función correspondiente de tipo  $F_1$ . Por tanto, ambas transforman de las variables  $(q, p)$  a las variables  $(w, J)$ . Sin embargo aunque  $W'$  genera la misma transformación que  $W$ , **no** es una solución de la ecuación de Hamilton Jacobi, ya que el formalismo de Hamilton Jacobi está diseñado para funciones tipo 2.

Hemos visto que asumiendo periodicidad temporal para cada plano  $(q_i, p_i)$ , la conmensurabilidad es un criterio esencial para distinguir entre un sistema periódico múltiple o un sistema periódico simple. Cuando las frecuencias todas son conmensurables entre sí, la configuración se repite después de un tiempo suficientemente

<sup>21</sup>La clave es que  $x = x(v_x)$ ,  $y = y(v_y)$  en tanto que  $x' = x'(v_x, v_y)$  y  $y' = y'(v_x, v_y)$ .

largo y será periódico simple. Matemáticamente la conmesurabilidad entre pares de frecuencias se manifiesta en las siguientes ecuaciones

$$j_i v_i = j_k v_k \quad (\text{no suma}) \tag{9.143}$$

siendo  $j_i$  y  $j_k$  enteros positivos. Basta con probar que un cierto  $v_i$  es conmesurable con los demás para demostrar que todos son conmesurables entre sí, en cuyo caso se habla de una sistema **completamente conmesurable**. Pero si solo  $m$  de las  $n$  frecuencias satisfacen (9.143), el sistema se dice  $m$ -conmesurable. Por ejemplo si tenemos las frecuencias

$$\begin{aligned} v_1 &= 3MHz, & v_2 &= 5MHz, & v_3 &= 7MHz, & v_4 &= 2\sqrt{2}MHz \\ v_5 &= 3\sqrt{2}MHz, & v_6 &= \sqrt{3}MHz, & v_7 &= \sqrt{7}MHz \end{aligned}$$

Los tres primeros son 3-conmesurados y los dos siguientes son doble conmesurados. Hay una relación interesante entre conmesurabilidad y las coordenadas en las cuales las Ecuaciones de Hamilton Jacobi son separables. Puede demostrarse que el camino que recorre el punto en el espacio de configuraciones o de fase para un sistema no conmesurable llena completamente una región limitada del correspondiente espacio (es decir para cualquier punto de esta región, la curva pasa por dicho punto en algún instante de tiempo). Esto se vé en las figuras de Lissajous para frecuencias inconmesurables.

Supongamos que el sistema es tal que en cualquiera de las coordenadas de separación el movimiento es simplemente periódico y que por tanto es independiente del movimiento en las otras coordenadas. En consecuencia, el camino que traza el punto del sistema como un todo tiene que estar limitado por las superficies de  $q_i$  y  $p_i$  constantes que marcan los límites del movimiento oscilatorio de las variables de separación (esto es extensible al movimiento periódico rotacional si restringimos el valor del ángulo entre 0 y  $2\pi$ ). Estas superficies por tanto definen un volumen en el espacio que está densamente lleno por la órbita del punto del sistema. De esto se sigue que las variables de separación en un sistema no conmesurado tienen que ser únicas; la ecuación de Hamilton Jacobi no puede ser separable en dos sistemas coordenados diferentes (excepto por variaciones triviales tales como cambios de escala). La posibilidad de separar el movimiento de un sistema en más de un sistema coordenado es usualmente una evidencia de conmesurabilidad.

#### 9.10.4. Variables acción-ángulo para sistemas degenerados

Un ejemplo particularmente simple de conmesurabilidad lo da el caso de la degeneración, que ocurre cuando dos o más de las frecuencias son iguales. Si dos de las constantes elásticas en el oscilador armónico tridimensional son iguales, sus frecuencias asociadas son iguales y hay una degeneración simple. Si el oscilador es isotrópico, el sistema es completamente degenerado.

Cuando hay degeneración presente, las frecuencias no son todas independientes. Si ordenamos las  $n$  frecuencias de modo que las  $m + 1$  primeras sean iguales, tendremos  $m$  ecuaciones de degeneración (degeneración de orden  $m$ )

$$\nu_1 - \nu_2 = 0 \quad , \quad \nu_2 - \nu_3 = 0 \quad , \quad \dots \quad , \quad \nu_m - \nu_{m+1} = 0 \tag{9.144}$$

de modo que solo una de estas frecuencias (digamos  $\nu_{m+1}$ ) es independiente, es claro entonces que tendremos  $n - m$  frecuencias independientes

$$\{\nu_{m+1}, \nu_{m+2}, \dots, \nu_n\}$$

Veamos un método sistemático de reducir las frecuencias usando una transformación canónica de las variables acción-ángulo  $(w, J)$ , hacia otras variables también del tipo acción-ángulo  $(w', J')$ . Si tenemos  $m$  condiciones de degeneración sobre las frecuencias  $\nu_i$ , podemos escribir las ecuaciones (9.144) en la forma

$$\sum_{i=1}^n j_{ki} \nu_i = 0, \quad k = 1, \dots, m. \tag{9.145}$$

donde las  $j_{ki}$  toman los valores 0 y  $\pm 1$ . Haremos ahora una transformación canónica desde  $(w, J)$  hasta  $(w', J')$  definida por una función generatriz del tipo descrito en (7.23), con  $g = 0$ . En nuestro caso tenemos las asignaciones

$$q_i \rightarrow w_i \quad ; \quad p_i \rightarrow J_i \quad ; \quad Q_i \rightarrow w'_i \quad ; \quad P_i \rightarrow J'_i \tag{9.146}$$

definiendo la función

$$f_k(w_1, \dots, w_n) = \begin{cases} \sum_{i=1}^n j_{ki} w_i & \text{si } k = 1, \dots, m \\ w_k & \text{si } k = m+1, \dots, n \end{cases} \quad (9.147)$$

reemplazando (9.147) en (7.23) teniendo en cuenta las asignaciones (9.146) y el hecho de que en (7.23) hay convención de suma sobre índices repetidos, se obtiene una función tipo  $F_2$  de la forma

$$F_2(w, J') = \sum_{k=1}^n f_k(w_1, \dots, w_n) J'_k = \sum_{k=1}^m \left( \sum_{i=1}^n j_{ki} w_i \right) J'_k + \sum_{k=m+1}^n J'_k w_k$$

de modo que

$$F_2(w, J') = \sum_{k=1}^m \sum_{i=1}^n J'_k j_{ki} w_i + \sum_{k=m+1}^n J'_k w_k \quad (9.148)$$

las coordenadas nuevas se encuentran con las ecuaciones de transformación (7.19) y las asignaciones (9.146)

$$w'_k = \frac{\partial F_2(w, J')}{\partial J'_k} = \frac{\partial}{\partial J'_k} \left[ \sum_{p=1}^m \sum_{i=1}^n J'_p j_{pi} w_i + \sum_{p=m+1}^n J'_p w_p \right]$$

con lo cual resulta

$$\begin{aligned} w'_k &= \sum_{i=1}^n j_{ki} w_i ; & k = 1, \dots, m \\ w'_k &= w_k ; & k = m+1, \dots, n \end{aligned} \quad (9.149)$$

La forma funcional de las  $J$  con respecto a las  $J'$  se obtiene utilizando la ecuación (7.26) con las asignaciones (9.146)

$$\begin{aligned} \mathbf{J} &= \left( \frac{\partial \mathbf{f}}{\partial \mathbf{w}} \right) \mathbf{J}' \Rightarrow J_i = \sum_{k=1}^n \frac{\partial f_k}{\partial w_i} J'_k \Rightarrow \\ J_i &= \left\{ \sum_{k=1}^m J'_k \frac{\partial}{\partial w_i} \left[ \sum_{l=1}^n j_{kl} w_l \right] \right\} + \left\{ \sum_{k=m+1}^n J'_k \frac{\partial w_k}{\partial w_i} \right\} \end{aligned}$$

donde hemos usado (9.147). Las correspondientes constantes asociadas a las variables de acción  $J_i$  son las soluciones de las  $n$  ecuaciones de transformación<sup>24</sup>

$$J_i = \sum_{k=1}^m J'_k j_{ki} + \sum_{k=m+1}^n J'_k \delta_{ki} ; \quad i = 1, \dots, n \quad (9.150)$$

Nuestra hipótesis es que el sistema es periódico en las variables canónicas  $(q_i, p_i)$  y ello nos lleva a que las variables del sistema son periódicas múltiples en las coordenadas  $w_k$ . Cuando se ejecutan  $m_i$  ciclos completos en cada plano  $(q_i, p_i)$  el cambio en las variables  $w_i$  viene dado por

$$\Delta w_i = m_i$$

ahora bien, de las Ecs. (9.149) vemos que para  $k > m$  tenemos  $\Delta w'_i = \Delta w_i$ . Por otro lado, las Ecs. (9.149) para  $k \leq m$ , se pueden escribir

$$k \leq m \Rightarrow w'_k = w_k - w_{k+1} \Rightarrow \Delta w'_k = \Delta w_k - \Delta w_{k+1} = m_k - m_{k+1} = \text{entero}$$

<sup>24</sup>Nótese que las Ecs. (9.149, 9.150), nos dicen que esta es una transformación canónica en donde las nuevas coordenadas solo son funciones de las antiguas coordenadas y los nuevos momentos solo son funciones de los antiguos momentos  $w'_k = w'_k(w)$ ,  $J'_k = J'_k(J)$ .

Si hacemos una solo ciclo en el plano  $(q_k, p_k)$  y ningún ciclo en los otros planos tenemos  $\Delta w'_k = 1$  (para cualquier valor de  $k$ ). Estas consideraciones nos llevan a concluir que el sistema también es periódico múltiple en las  $w'_k$  coordenadas con periodo unidad. Por tanto, estas nuevas variables también son del tipo acción-ángulo, y las nuevas frecuencias se pueden calcular como en las ecuaciones (9.128, 9.129)

$$\begin{aligned} v'_k &= \dot{w}'_k = \sum_{i=1}^n j_{ki} v_i = 0 \quad ; \quad k = 1, \dots, m \\ v'_k &= \dot{w}'_k = v_k \quad ; \quad k = m + 1, \dots, n \end{aligned} \quad (9.151)$$

donde hemos tenido en cuenta (9.145). En el nuevo sistema canónico tenemos entonces  $n - m$  frecuencias independientes (cuando  $k \geq m + 1$ ) que además coinciden con las frecuencias originales del sistema, y las  $m$  frecuencias espúreas asociadas con la  $m$ -degeneración se han traducido en  $m$  frecuencias nulas<sup>25</sup>. Puesto que  $w'_k = v'_k t + \beta'_k$ , las frecuencias nulas corresponden a  $w'_k = \beta'_k$  y por tanto a factores constantes en la expansión de Fourier, como se puede ver por ejemplo en la Ec. (9.134). Estos factores constantes también aparecen en la expansión original, siempre que los índices  $j_i$  satisfagan una condición de degeneración. Por ejemplo, supongamos que tenemos tres frecuencias  $\nu_1, \nu_2, \nu_3$  y que  $\nu_1 = \nu_2$ , la expansión (9.137) en las bases  $(w, J)$  y  $(w', J')$  se escribe

$$\begin{aligned} q_k(t) &= \sum_{j_1} \sum_{j_2} \sum_{j_3} a_{j_1, j_2, j_3}^{(k)} \exp \{2\pi i [j_1 (\nu_2 t + \beta_1) + j_2 (\nu_2 t + \beta_2) + j_3 (\nu_3 t + \beta_3)]\} \\ q_k(t) &= \sum_{m_1} \sum_{m_2} \sum_{m_3} a_{m_1, m_2, m_3}^{(k)} \exp \{2\pi i [m_1 (\beta_1 - \beta_2) + m_2 (\nu_2 t + \beta_2) + m_3 (\nu_3 t + \beta_3)]\} \end{aligned}$$

Volviendo al caso general, dado que las frecuencias  $\nu'_i$  se calculan como

$$v'_i = \frac{\partial H}{\partial J'_i} \quad (9.152)$$

el Hamiltoniano tiene que ser independiente de las variables de acción  $J'_i$  asociadas a las frecuencias nulas, que por construcción son las  $m$  primeras frecuencias. Por tanto

$$H = H(J'_{m+1}, J'_{m+2}, \dots, J'_n) \quad (9.153)$$

En particular, si el sistema es completamente degenerado (i.e. degeneración de orden  $n - 1$ ), existirán  $m = n - 1$  condiciones de degeneración del tipo (9.145), y el Hamiltoniano dependerá de solo una de las  $J'_i$ . Vemos entonces que el paso del sistema  $(w, J)$  al sistema  $(w', J')$  conduce a una simplificación en la estructura del Hamiltoniano y en las expansiones de Fourier de las variables del sistema.

Dado que las  $J'$  son  $n$  cantidades constantes independientes, las constantes originales de integración pueden ser expresadas en términos de los  $J'$ , y por tanto  $W$  puede expresarse como  $W = W(q, J')$ . Con estos argumentos  $W$  genera una transformación a un nuevo conjunto canónico en donde las  $J'$  son los nuevos momentos canónicos<sup>26</sup>. Pero en virtud de la transformación puntual generada por  $F_2$  en (9.148), sabemos que  $w'$  es el conjugado de  $J'$ . Se sigue entonces que las nuevas coordenadas generadas por  $W(q, J')$  tienen que ser el conjunto de variables angulares  $w'$ , con ecuaciones de transformación dadas por

$$w'_i = \frac{\partial W}{\partial J'_i} \quad (9.154)$$

en conclusión, la función característica  $W$  sirve también como generatriz de la transformación desde  $(q, p)$  hasta  $(w', J')$ .

<sup>25</sup>Es obvio que las  $m$  frecuencias nulas no son frecuencias físicas, sino artificios matemáticos para simplificar el problema. En todo caso, el número de frecuencias no-nulas independientes permanece constante.

<sup>26</sup>El hecho de que  $W(q, J)$  se pueda escribir como  $W(q, J')$ , también se puede ver teniendo en cuenta que la transformación  $(w, J) \rightarrow (w', J')$  es tal que  $J' = J'(J)$ .

### 9.11. Comentarios finales sobre las variables acción-ángulo

Hemos visto que para sistemas que poseen movimientos tipo libración o rotación en el espacio de fase, existe un formalismo adecuado para estudiar las frecuencias del movimiento sin resolver completamente las ecuaciones de movimiento. El formalismo de las variables acción ángulo es una variante de la técnica de Hamilton Jacobi en la cual los nuevos momentos no se eligen como las constantes de integración sino como ciertas cantidades constantes que contienen información sobre la órbita de libración o rotación, y que son en todo caso función exclusiva de las constantes de integración. Cuando asumimos que el movimiento es periódico en el tiempo este formalismo nos permite extraer de forma sencilla las frecuencias múltiples sin resolver completamente las ecuaciones de movimiento. No obstante, debe tenerse claro que en el formalismo no hay ningún criterio que nos asegure que el movimiento es de tipo libración o rotación, ni mucho menos que sea periódico en el tiempo. Estas son hipótesis de trabajo que deben estar sustentadas por argumentos alternativos teóricos y/o experimentales. Vale destacar sin embargo, que si los movimientos son de tipo rotación o libración pero no son periódicos en el tiempo, el formalismo permite extraer la frecuencia promedio tomada sobre muchos ciclos.

De otra parte, cuando existe degeneración en ciertas frecuencias del sistema, es posible pasar de las variables canónicas acción-ángulo  $(w, J)$  a otras variables canónicas de acción-ángulo  $(w', J')$  que me conducen a las mismas frecuencias independientes, y en donde las frecuencias no independientes se convierten en frecuencias nulas. Estas frecuencias nulas no son físicas pero permiten simplificar considerablemente el problema. En algunos casos, la hipótesis de periodicidad puede conducir automáticamente a la presencia de degeneración debido a algunas simetrías del sistema (ver por ejemplo la sección 9.12). No obstante, en problemas generales pueden existir degeneraciones accidentales (por ejemplo debidas a condiciones iniciales) que deben incorporarse en este formalismo como hipótesis de trabajo.

### 9.12. El problema de Kepler en variables acción-ángulo

Para exhibir todas las propiedades de la solución usaremos las tres dimensiones en coordenadas esféricas tal como en la sección 9.6.2. Lo primero que debemos determinar es el tipo de periodicidad en los planos  $(\theta, p_\theta)$ ,  $(\phi, p_\phi)$  y  $(r, p_r)$ . Para ello despejamos a cada momento generalizado en función de su coordenada de las ecuaciones (9.73, 9.76, 9.77, 9.79) con lo cual obtenemos

$$p_\phi = \alpha_\phi \quad (9.155)$$

$$p_\theta^2 = \alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta} \quad (9.156)$$

$$p_r^2 = 2m [E - V(r)] - \frac{\alpha_\theta^2}{r^2} \quad (9.157)$$

como  $\phi$  es variable angular cíclica se le puede asociar la acción  $J_\phi = 2\pi\alpha_\phi$  de acuerdo con (9.122), y asumir que es una rotación eligiendo por conveniencia el periodo  $2\pi$  (recuérdese que para variable cíclica el periodo de rotación es arbitrario). A primera vista, la Ec. (9.156) parece mostrar que la trayectoria en el plano de fase  $(\theta, p_\theta)$  es periódica. No obstante, es necesario tener cierta información del movimiento para saber si  $\theta$  está en un intervalo acotado que se recorre de ida y vuelta en un intervalo finito de tiempo (o si tiene periodo tipo rotación donde el periodo en  $\theta$  se recorra en un intervalo finito de tiempo). Algo similar ocurre con el plano  $(\phi, p_\phi)$  y con el plano  $(r, p_r)$ . En el movimiento real estos planos definirán trayectorias periódicas solo si la energía del sistema es negativa, pues de lo contrario la órbita no será cerrada ni acotada (para más detalles, ver Cap. 10). Por ejemplo cuando la trayectoria es abierta,  $\phi$  no barre el intervalo completo entre 0 y  $2\pi$  y el intervalo acotado solo se recorre de ida, no hay en consecuencia movimiento de rotación ni de libración en el plano  $(\phi, p_\phi)$ . Por otro lado, cuando la trayectoria es cerrada y acotada se puede ver que  $\theta$  y  $r$  son acotados (van y vuelven dentro de un intervalo finito de tiempo) de modo que su movimiento es de libración<sup>28</sup>. En contraste  $\phi$  aumenta indefinidamente y su movimiento es de rotación. La discusión anterior refuerza el hecho de que el

<sup>28</sup>Si suponemos que el plano de movimiento pasa por el origen, y que el eje  $z$  no yace en el plano de movimiento, podemos asegurar que  $\theta \neq 0$  (y en general  $\theta \neq n\pi$ ) en todo punto de la trayectoria, lo cual nos evita una singularidad en la Ec. (9.156).

carácter periódico en el espacio de fase y/o en el tiempo son una hipótesis de trabajo para iniciar el tratamiento con variables acción ángulo, pero no hay nada en este formalismo que nos garantice tal periodicidad, la cual debe ser extraída por otros argumentos.

Asumiendo entonces periodicidad en cada plano  $(q_i, p_i)$ , vamos a construir las variables de acción asociadas a cada coordenada. Para ello usamos los valores de los momentos conjugados a  $\phi, \theta, r$  de las Ecs. (9.155, 9.156, 9.157), y usando un potencial de la forma  $V = -k/r$

$$J_\phi = \oint \frac{\partial W}{\partial \phi} d\phi = \oint \alpha_\phi d\phi = 2\pi\alpha_\phi \quad (9.158)$$

$$J_\theta = \oint \frac{\partial W}{\partial \theta} d\theta = \oint \left[ \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta}} \right] d\theta \quad (9.159)$$

$$J_r = \oint \frac{\partial W}{\partial r} dr = \oint \left[ \sqrt{2mE + \frac{2mk}{r} - \frac{\alpha_\theta^2}{r^2}} \right] dr \quad (9.160)$$

la primera integral es trivial y está asociada con el hecho de que la variable asociada es cíclica. Para calcular la segunda integral es conveniente introducir el ángulo polar del momento angular total que denotaremos por  $\theta_i$ . Recordando que  $\alpha_\phi$  representa la componente polar del momento angular (ver Sec. 9.6.2), el ángulo  $\theta_i$  viene dado claramente por

$$\cos \theta_i = \frac{L_z}{L} = \frac{\alpha_\phi}{\alpha_\theta} \quad (9.161)$$

en términos de  $\cos \theta_i$ , la integral (9.159) queda

$$\begin{aligned} J_\theta &= \alpha_\theta \oint \left[ \sqrt{1 - \frac{\alpha_\phi^2}{\alpha_\theta^2} \frac{1}{\sin^2 \theta}} \right] d\theta \\ J_\theta &= \alpha_\theta \oint \sqrt{1 - \frac{\cos^2 \theta_i}{\sin^2 \theta}} d\theta \end{aligned} \quad (9.162)$$

podemos asumir sin pérdida de generalidad, que el plano del movimiento pasa por el origen de coordenadas. Para analizar los límites de integración debemos calcular los ángulos mínimo y máximo  $\theta_0$  y  $\theta_1$  entre un vector de posición de la partícula y el eje  $Z$ .

La Fig. 9.3 nos ayuda a encontrar estos ángulos extremos. En esta figura, el plano  $YZ$  se define de modo que sea perpendicular al plano de movimiento. Por simplicidad, el plano  $YZ$  se hace coincidir con el plano del papel y el eje  $X$  sale del papel. La línea punteada que pasa por el origen define la intersección entre el plano de movimiento y el plano  $YZ$ . Puesto que  $\mathbf{L}$  es perpendicular al plano de movimiento, si lo trasladamos al origen claramente yacerá en el plano  $YZ$ . Teniendo en cuenta que  $\theta_i$  es el ángulo entre el momento angular  $\mathbf{L}$  y el eje  $Z$ , y que  $\mathbf{L}$  es perpendicular a cualquier vector de posición dentro del plano de movimiento, es fácil ver que el menor y el mayor valor de  $\theta$  se obtienen cuando el momento angular, el eje  $Z$  y el vector de posición son coplanarios. Hay dos vectores unitarios de posición que cumplen esta condición: los dos vectores unitarios que van a lo largo de la línea punteada denotados por  $\mathbf{r}_A$  y  $\mathbf{r}_B$ , los cuales definen los ángulos  $\theta_0$  y  $\theta_1$  respectivamente. Adicionalmente, la Fig. 9.3 muestra que

$$\theta_0 = \frac{\pi}{2} - \theta_i \quad , \quad \theta_1 = \pi - \theta_0 = \frac{\pi}{2} + \theta_i.$$

De lo anterior, se deduce que el circuito completo del ángulo  $\theta$  consiste en ir desde  $(\pi/2) - \theta_i$ , hasta  $(\pi/2) + \theta_i$  y volver, donde  $\sin \theta_0 = \cos \theta_i$ . La integral circuital (9.162) se puede escribir como 4 veces la integral entre  $\pi/2$  y  $(\pi/2) + \theta_i$  quedando<sup>29</sup>

$$J_\theta = 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \sqrt{1 - \frac{\cos^2 \theta_i}{\sin^2 \theta}} d\theta = 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \sqrt{\frac{\sin^2 \theta - \cos^2 \theta_i}{\sin^2 \theta}} d\theta \quad (9.163)$$

<sup>29</sup> $\theta_0$  es el ángulo mínimo pero no necesariamente el ángulo inicial.

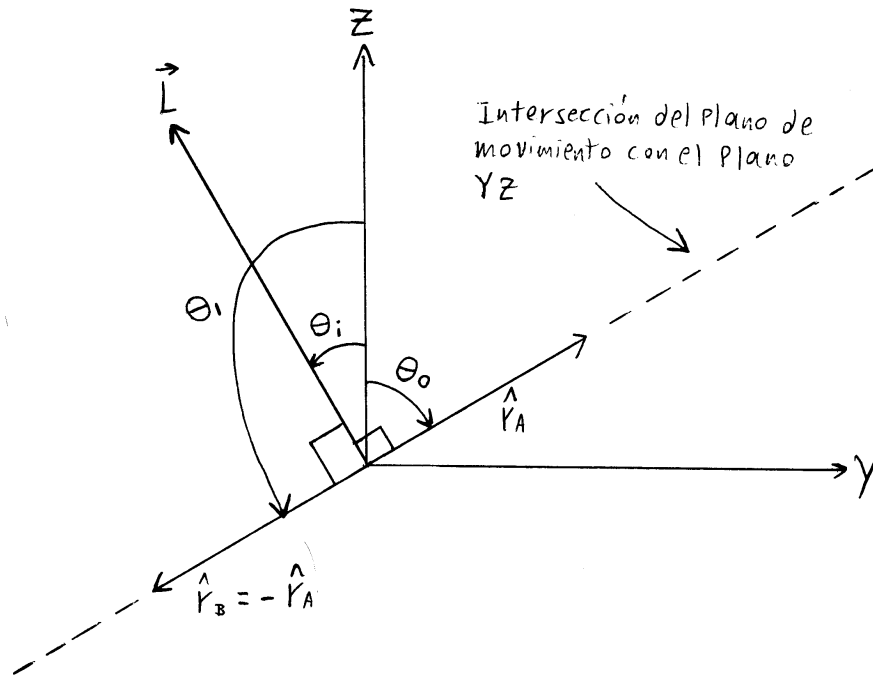


Figura 9.3: Movimiento bajo una fuerza central. El plano  $YZ$  es perpendicular al plano de movimiento. Cuando la partícula alcanza la posición de mínimo o máximo valor de  $\theta$ , el vector posición de la partícula está en el mismo plano que el eje  $Z$  y el momento angular.

ahora bien, en nuestra convención  $\mathbf{L}$  está en el plano  $YZ$ . Por tanto, siempre es posible escoger el sentido del eje  $Z$  de tal forma que el ángulo  $\theta_i$  entre  $Z$  y  $\mathbf{L}$  esté en el intervalo  $[0, \pi/2)$ . Esto equivale a escoger  $\cos \theta_i$  y  $\sin \theta_i$  como positivos<sup>30</sup>, y dado que la integral anterior se evalúa entre  $\pi/2$  y  $(\pi/2) + \theta_i$ , tendremos que  $\sin \theta$  también será positivo. Por tanto, la integral (9.163) queda

$$\begin{aligned}
 J_\theta &= 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \frac{\sqrt{\sin^2 \theta - \cos^2 \theta_i}}{\sin \theta} d\theta = 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \frac{\sqrt{-\cos^2 \theta + 1 - \cos^2 \theta_i}}{\sin \theta} d\theta \\
 J_\theta &= 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \frac{\sqrt{\sin^2 \theta_i - \cos^2 \theta}}{\sin \theta} d\theta
 \end{aligned} \tag{9.164}$$

La sustitución

$$\cos \theta = -\sin \theta_i \sin \psi \quad \Rightarrow \quad \sin \theta d\theta = \sin \theta_i \cos \psi d\psi$$

convierte la integral (9.164) en

$$\begin{aligned}
 J_\theta &= 4\alpha_\theta \int_{\pi/2}^{(\pi/2)+\theta_i} \frac{\sin \theta d\theta}{\sin^2 \theta} \sqrt{\sin^2 \theta_i - \cos^2 \theta} = 4\alpha_\theta \int_0^{\pi/2} \frac{\sin \theta_i \cos \psi d\psi}{1 - (\sin \theta_i \sin \psi)^2} \sqrt{\sin^2 \theta_i - (\sin \theta_i \sin \psi)^2} \\
 &= 4\alpha_\theta \int_0^{\pi/2} \frac{\sin^2 \theta_i \cos \psi d\psi}{1 - \sin^2 \theta_i \sin^2 \psi} \sqrt{1 - \sin^2 \psi}
 \end{aligned}$$

quedando

$$J_\theta = 4\alpha_\theta \sin^2 \theta_i \int_0^{\pi/2} \frac{\cos^2 \psi d\psi}{1 - \sin^2 \theta_i \sin^2 \psi} \tag{9.165}$$

<sup>30</sup>Elegimos el intervalo  $[0, \pi/2)$ , ya que  $\theta_i = \pi/2$ , equivale a que el eje  $Z$  esté sobre el plano de movimiento, y esto permite que  $\theta = 0$ , en los puntos de la trayectoria en los cuales la órbita cruza al eje  $Z$ . Esto a su vez nos lleva a una singularidad en la expresión (9.156).



la sustitución adicional

$$u = \tan \psi \quad ; \quad du = (1 + \tan^2 \psi) d\psi = (1 + u^2) d\psi$$

nos da

$$\begin{aligned} J_\theta &= 4\alpha_\theta \sin^2 \theta_i \int_0^{\pi/2} \frac{d\psi}{\frac{1}{\cos^2 \psi} - \sin^2 \theta_i \tan^2 \psi} = 4\alpha_\theta \sin^2 \theta_i \int_0^{\pi/2} \frac{d\psi}{(1 + \tan^2 \psi) - \sin^2 \theta_i \tan^2 \psi} \\ &= 4\alpha_\theta \sin^2 \theta_i \int_0^\infty \frac{du/(1+u^2)}{(1+u^2) - \sin^2 \theta_i u^2} = 4\alpha_\theta \sin^2 \theta_i \int_0^\infty \frac{du}{(1+u^2)[1+u^2(1-\sin^2 \theta_i)]} \\ J_\theta &= 4\alpha_\theta \int_0^\infty \frac{\sin^2 \theta_i}{(1+u^2)[1+u^2 \cos^2 \theta_i]} du \end{aligned}$$

al utilizar fracciones parciales esta integral queda

$$J_\theta = 4\alpha_\theta \int_0^\infty du \left[ \frac{1}{1+u^2} - \frac{\cos^2 \theta_i}{1+u^2 \cos^2 \theta_i} \right] = 4\alpha_\theta \int_0^\infty \frac{du}{1+u^2} - 4\alpha_\theta \cos^2 \theta_i \int_0^\infty \frac{du}{1+u^2 \cos^2 \theta_i}$$

haciendo el cambio de variable  $u' = u \cos \theta_i$ ,  $du' = \cos \theta_i du$  en la segunda integral queda

$$\begin{aligned} J_\theta &= 4\alpha_\theta \int_0^\infty \frac{du}{1+u^2} - 4\alpha_\theta \cos^2 \theta_i \int_0^\infty \frac{du'/\cos \theta_i}{1+u'^2} = 4\alpha_\theta \left[ \int_0^\infty \frac{du}{1+u^2} - \cos \theta_i \int_0^\infty \frac{du'}{1+u'^2} \right] \\ J_\theta &= 4\alpha_\theta [1 - \cos \theta_i] \int_0^\infty \frac{du}{1+u^2} = 4\alpha_\theta (1 - \cos \theta_i) [\arctan u]_0^\infty = 4\alpha_\theta (1 - \cos \theta_i) \frac{\pi}{2} = 2\pi\alpha_\theta (1 - \cos \theta_i) \end{aligned}$$

con lo cual se obtiene finalmente

$$J_\theta = 2\pi\alpha_\theta \left( 1 - \frac{\alpha_\phi}{\alpha_\theta} \right) = 2\pi (\alpha_\theta - \alpha_\phi) \quad (9.166)$$

puesto que la idea es escribir todo en términos de las  $J$ 's, invertimos las Ecs. (9.158, 9.166) para obtener  $\alpha_\theta$ ,  $\alpha_\phi$  en términos de  $J_\theta$ ,  $J_\phi$

$$\alpha_\phi = \frac{J_\phi}{2\pi} \quad ; \quad \alpha_\theta = \frac{J_\theta + J_\phi}{2\pi} \quad (9.167)$$

sustituyendo (9.167) en la expresión (9.160) para  $J_r$  resulta

$$J_r = \oint \frac{\partial W}{\partial r} dr = \oint \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2}} dr \quad (9.168)$$

después de realizar esta integración, obtenemos una función cuyos argumentos son

$$J_r = J_r(E, J_\theta + J_\phi) \quad (9.169)$$

y si despejamos la energía  $E = H$  en esta ecuación obtenemos

$$\begin{aligned} E &= H = H(J_r, J_\theta + J_\phi) \Rightarrow \frac{\partial H}{\partial J_\theta} = \frac{\partial H}{\partial J_\phi} \\ &\Rightarrow \nu_\theta = \nu_\phi \end{aligned} \quad (9.170)$$

de modo que las frecuencias en los ángulos están degeneradas. **Este resultado no depende de la ley del inverso cuadrado sino solo del hecho de que la fuerza sea central**, y de que la órbita sea acotada, en cuyo caso el movimiento será al menos simplemente degenerado. Esta degeneración es consecuencia del hecho de que el movimiento se hace en un plano perpendicular al momento angular. El movimiento en este plano indica que  $\theta$  y  $\phi$  están relacionados de tal manera que cuando  $\phi$  completa un periodo entre 0 y  $2\pi$ ,  $\theta$  recorre

un ciclo completo entre los límites  $\theta_1$  y  $\theta_2$  ida y vuelta siendo  $\theta_{1,2} = (\pi/2) \pm \theta_i$ , de modo que las frecuencias en  $\theta$  y  $\phi$  son necesariamente iguales.

Para realizar la integración en  $J_r$  primero tengamos en cuenta que el movimiento es acotado solo para energías negativas y dado que el integrando es igual a  $p_r = m\dot{r}$ , los límites del movimiento están comprendidos entre las raíces  $r_1$  y  $r_2$  de la expresión que está dentro del radical en (9.168) y que forman los puntos de retorno en  $r$ . Los extremos  $r_1$  y  $r_2$  son entonces las raíces de

$$\begin{aligned} 2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2} &= 0 \\ 2mEr^2 + 2mkr - \frac{(J_\theta + J_\phi)^2}{4\pi^2} &= 0 \end{aligned} \quad (9.171)$$

Si  $r_1, r_2$  son los límites inferior y superior respectivamente (las dos raíces de la cuadrática 9.171), un ciclo completo incluye ir desde  $r_1$  hasta  $r_2$  y volver. En la ida (vuelta)  $p_r$  es positivo (negativo) y por tanto la raíz cuadrada en (9.168) debe ser positiva (negativa). Con lo cual (9.168) se escribe como

$$J_r = 2 \int_{r_1}^{r_2} \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2}} dr$$

La integración por variable compleja se puede ver en la Pág 469 de la Ref. [1], el resultado es

$$J_r = -(J_\theta + J_\phi) + \pi k \sqrt{\frac{2m}{-E}}$$

el cual tiene la estructura dada por (9.169) como se anticipó. Nótese que  $J_r$  solo es real para energías negativas, en concordancia con el hecho de que las trayectorias deben tener energía negativa para ser acotadas. Despejando la energía i.e. el Hamiltoniano se tiene

$$E = H = -\frac{2\pi^2 m k^2}{(J_r + J_\theta + J_\phi)^2} \quad (9.172)$$

se vé que para fuerzas atractivas inversas al cuadrado de la distancia, la degeneración es aún mayor de lo predicho para fuerzas centrales en general. En este caso los tres periodos coinciden y el movimiento es completamente degenerado, lo cual nos dice que la órbita es cerrada cuando la energía es negativa. Con una órbita cerrada el movimiento es simplemente periódico y en este caso, completamente degenerado. Si la fuerza central contiene un término proporcional a  $r^{-3}$  (por ejemplo una corrección relativista de primer orden), la órbita ya no es cerrada sino que tiene la forma de una elipse que precesa. Una de las degeneraciones se remueve en este caso, pero el movimiento es aún simplemente degenerado ya que la fuerza aún es central.

Volviendo al caso de Kepler, la frecuencia y el periodo vienen dados por

$$v = \frac{\partial H}{\partial J_r} = \frac{\partial H}{\partial J_\theta} = \frac{\partial H}{\partial J_\phi} = \frac{4\pi^2 m k^2}{(J_r + J_\theta + J_\phi)^3} ; \quad \tau = \frac{1}{\nu} = \frac{(J_r + J_\theta + J_\phi)^3}{4\pi^2 m k^2} \quad (9.173)$$

la suma de los  $J$ 's se puede escribir en términos de la energía a partir de (9.172), y el periodo resulta

$$\tau = \pi k \sqrt{\frac{m}{-2E^3}} \quad (9.174)$$

esta fórmula para el periodo está de acuerdo con la tercera ley de Kepler si tenemos en cuenta que el semieje mayor  $a$  es igual a  $-k/2E$ , como se verá en el Cap. 10, Ecs(10.94, 10.110).

Recalcamos finalmente, que para encontrar el periodo del movimiento no utilizamos la ecuación de la trayectoria ni la dependencia de las variables con el tiempo. Para integrar los  $J_i$ , solo fué necesario hacer algunas suposiciones generales tales como: (a) El movimiento es periódico y acotado en todos los planos de

fase, solo si la energía es negativa<sup>31</sup>. (b) El momento angular es una constante de movimiento y esto implica que el movimiento es en un plano.

Vale decir que en este caso fué posible predecir con base en estas hipótesis, la presencia de degeneración simple en el caso de fuerzas centrales ( $\nu_\theta = \nu_\phi$ ), y de degeneración total para interacción kepleriana atractiva i.e.  $V(r) = -k/r$ .

### 9.12.1. Variables acción-ángulo teniendo en cuenta la degeneración

La degeneración completa nos indica que existe un conjunto de variables acción-ángulo ( $w', J'$ ) en las cuales  $H$  solo depende de un  $J'_i$  y solo una de las frecuencias es no nula. Vamos a encontrar la transformación canónica caracterizada por

$$(w_\theta, w_\phi, w_r, J_\theta, J_\phi, J_r) \rightarrow (w'_1, w'_2, w'_3, J'_1, J'_2, J'_3)$$

para lo cual usamos los resultados de la sección (9.10.4). Expresemos las condiciones de degeneración en la forma

$$v_\phi - v_\theta = 0 \quad ; \quad v_\theta - v_r = 0 \quad (9.175)$$

Ordenaremos las variables acción-ángulo originales en la forma

$$(w_\theta, w_\phi, w_r, J_\theta, J_\phi, J_r) \equiv (w_1, w_2, w_3, J_1, J_2, J_3)$$

con este criterio de orden, conviene reescribir las  $m = 2$  condiciones de degeneración (9.175), para las  $n = 3$  frecuencias, en la forma

$$-v_\theta + v_\phi + 0 \cdot v_r = 0 \quad ; \quad v_\theta + 0 \cdot v_\phi - v_r = 0 \quad (9.176)$$

comparando con (9.145) se tiene que

$$j_{11} = -1, \quad j_{12} = 1, \quad j_{13} = 0 \quad ; \quad j_{21} = 1, \quad j_{22} = 0, \quad j_{23} = -1$$

y reemplazando los elementos de  $j_{ki}$  en (9.148)

$$\begin{aligned} F_2 &= \sum_{k=1}^2 \sum_{i=1}^3 J'_k j_{ki} w_i + \sum_{k=3}^3 J'_k w_k = \left[ \sum_{k=1}^2 J'_k (j_{k1} w_1 + j_{k2} w_2 + j_{k3} w_3) \right] + J'_3 w_3 \\ F_2 &= J'_1 (j_{11} w_1 + j_{12} w_2 + j_{13} w_3) + J'_2 (j_{21} w_1 + j_{22} w_2 + j_{23} w_3) + J'_3 w_3 \\ F_2 &= J'_1 (-w_\theta + w_\phi) + J'_2 (w_\theta - w_r) + J'_3 w_r \end{aligned} \quad (9.177)$$

las nuevas variables angulares se obtienen de (9.149)

$$\begin{aligned} w'_1 &= \sum_{i=1}^3 j_{1i} w_i = j_{11} w_1 + j_{12} w_2 + j_{13} w_3 = -w_\theta + w_\phi \\ w'_2 &= \sum_{i=1}^3 j_{2i} w_i = j_{21} w_1 + j_{22} w_2 + j_{23} w_3 = w_\theta - w_r \quad ; \quad w'_3 = w_r \end{aligned} \quad (9.178)$$

para obtener la transformación de los momentos, aplicamos (9.150), de lo cual resulta

$$\begin{aligned} J_\theta &\equiv J_1 = \sum_{k=1}^2 J'_k j_{k1} + \sum_{k=3}^3 J'_k \delta_{k1} = J'_1 j_{11} + J'_2 j_{21} + J'_3 \delta_{31} = -J'_1 + J'_2 \\ J_\phi &\equiv J_2 = \sum_{k=1}^2 J'_k j_{k2} + \sum_{k=3}^3 J'_k \delta_{k2} = J'_1 j_{12} + J'_2 j_{22} + J'_3 \delta_{32} = J'_1 \\ J_r &\equiv J_3 = \sum_{k=1}^2 J'_k j_{k3} + \sum_{k=3}^3 J'_k \delta_{k3} = J'_1 j_{13} + J'_2 j_{23} + J'_3 \delta_{33} = -J'_2 + J'_3 \end{aligned}$$

<sup>31</sup>En la sección 10.4.1, Pág. 221, veremos que se puede concluir que la curva es acotada solo para energías negativas, a través del análisis de curvas de energía potencial efectiva. En tal análisis, no es necesario recurrir a las ecuaciones de movimiento o de trayectoria.

quedando

$$J_\theta = J'_2 - J'_1 ; \quad J_\phi = J'_1 ; \quad J_r = J'_3 - J'_2 \quad (9.179)$$

las Ecs. (9.178) y el inverso de las ecuaciones (9.179), me caracterizan a  $(w', J')$  en términos de los  $(w, J)$  originales

$$w'_1 = w_\phi - w_\theta ; \quad w'_2 = w_\theta - w_r ; \quad w'_3 = w_r \quad (9.180)$$

$$J'_1 = J_\phi ; \quad J'_2 = J_\theta + J_\phi ; \quad J'_3 = J_\theta + J_\phi + J_r \quad (9.181)$$

y reemplazando estas expresiones en (9.172) resulta

$$H(J'_3) = E = -\frac{2\pi^2 m k^2}{J_3'^2} \quad (9.182)$$

como se predijo, el Hamiltoniano es función de una sola  $J'_i$  en esta base de variables, debido a la completa degeneración del sistema. La única frecuencia característica (no nula) del problema es

$$\nu'_3 = \frac{\partial H}{\partial J'_3} = \frac{4\pi^2 m k^2}{J_3'^3}$$

para escribir la frecuencia y el periodo en términos de constantes más Físicas, podemos despejar  $J'_3$  de la Ec. (9.182) de modo que queda en términos de la energía, esencialmente en la misma forma en que pasamos de la Ec. (9.173) a la Ec. (9.174).

Si realizamos un análisis detallado de las variables angulares  $w'$  veremos que éstas también conducen a constantes de movimiento. Esto se puede ver más fácilmente si trabajamos el problema bidimensionalmente desde el principio (ver sección 10.16).

### 9.13. Ejercicios

1. Un sistema de un grado de libertad está descrito por el Hamiltoniano

$$H = \frac{p^2}{2m} - mAtx$$

siendo  $A$  una constante. Resuelva el problema dinámico utilizando la función principal de Hamilton bajo las condiciones iniciales  $x(0) = 0$  y  $p(0) = mv_0$ . Nótese que este es un ejemplo de Hamiltoniano dependiente del tiempo que se puede resolver por HJ. Es claro que para este Hamiltoniano no se puede emplear la ecuación de HJR para la función característica de Hamilton.

2. Una partícula de masa  $m$  está restringida a moverse sobre una curva en el plano vertical definida por las ecuaciones paramétricas

$$y = l(1 - \cos 2\phi) ; \quad x = l(2\phi + \sin 2\phi)$$

sobre la partícula actúa la fuerza gravitacional en la dirección vertical  $y$ . Encuentre las frecuencias del movimiento empleando variables acción-ángulo, empleando todas las condiciones iniciales que nos lleven a que el máximo de  $\phi$  sea menor o igual a  $\pi/4$ .

3. Consideremos una partícula de carga  $q$  que se mueve en el plano  $XY$  sujeta a un campo magnético constante y uniforme  $\mathbf{B}$ , perpendicular al plano  $XY$ . Escogeremos el vector potencial  $\mathbf{A}$  de modo que solo tenga componente  $y$  (esta escogencia se conoce como gauge de Landau), de modo que  $A_y = Bx$  y el Hamiltoniano del sistema será

$$H(x, y, p_x, p_y) = \frac{p_x^2}{2m} + \frac{(p_y + bx)^2}{2m} ; \quad b \equiv -qB \quad (9.183)$$

asuma separación de variables para la función principal de Hamilton

$$S = W_x(x) + W_y(y) - \alpha t$$

y reduzca el problema a cuadraturas.

4. Con respecto al Hamiltoniano (9.183), consideremos la transformación gauge  $\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Psi$ , con  $\Psi = -bxy/2$ . (a) Demuestre que el nuevo Hamiltoniano queda

$$H(x, y, p_x, p_y) = \frac{1}{2m} \left( p_x - \frac{by}{2} \right)^2 + \frac{1}{2m} \left( p_y + \frac{bx}{2} \right)^2 \quad (9.184)$$

si bien este Hamiltoniano es claramente equivalente al Hamiltoniano (9.183), la ecuación de Hamilton-Jacobi para el Hamiltoniano (9.184), es considerablemente más difícil de separar. (b) Pruebe el ansatz de separación

$$S = \frac{1}{2}Kxy + W_x(x) + W_y(y) - \alpha t$$

y reduzca el problema a cuadraturas. Este problema ilustra el hecho de que la separabilidad de la ecuación de HJ no solo depende del sistema coordenado elegido, sino también del gauge elegido.

5. Una partícula de masa  $m$  está restringida a moverse sobre el eje  $X$  sujeta al potencial  $V = a \sec^2(x/l)$ . (a) Resuelva la ecuación de HJ y a partir de la función generatriz, encuentre  $x(t)$ . (b) Encuentre las variables acción-ángulo y la frecuencia  $\nu$  asociada al sistema. Obtenga la dependencia de la frecuencia con la amplitud y encuentre el límite de pequeñas amplitudes para  $\nu$ .
6. Reduzca a cuadraturas la ecuación de HJ para el Hamiltoniano

$$H(q, p, t) = f(t) \left[ \frac{p^2}{2m} + \frac{kq^2}{2} \right]$$

donde  $m$  y  $k$  son constantes y  $f(t)$  una función integrable. Encuentre  $q(t)$  y  $p(t)$  así como la trayectoria en el espacio de fase, para los tres casos siguientes

$$f(t) = e^{\alpha t} \quad ; \quad f(t) = e^{-\alpha t} \quad ; \quad f(t) = \cos \Omega t \quad ; \quad \alpha > 0$$

donde  $\alpha$  y  $\Omega$  son constantes.

# Capítulo 10

## Fuerzas centrales

Discutiremos a continuación, el problema de la interacción entre dos masas puntuales que se mueven bajo la influencia de una fuerza que va a lo largo de la línea que las une. Este es un problema que posee muchas aplicaciones tanto en Física Clásica como en Física Moderna. Siguiendo el espíritu de las formulaciones aquí presentadas, primeros nos concentraremos en las primeras integrales que se pueden hallar sin resolver el problema completo, para luego analizar algunos potenciales específicos.

### 10.1. Reducción al problema equivalente de dos partículas desacopladas

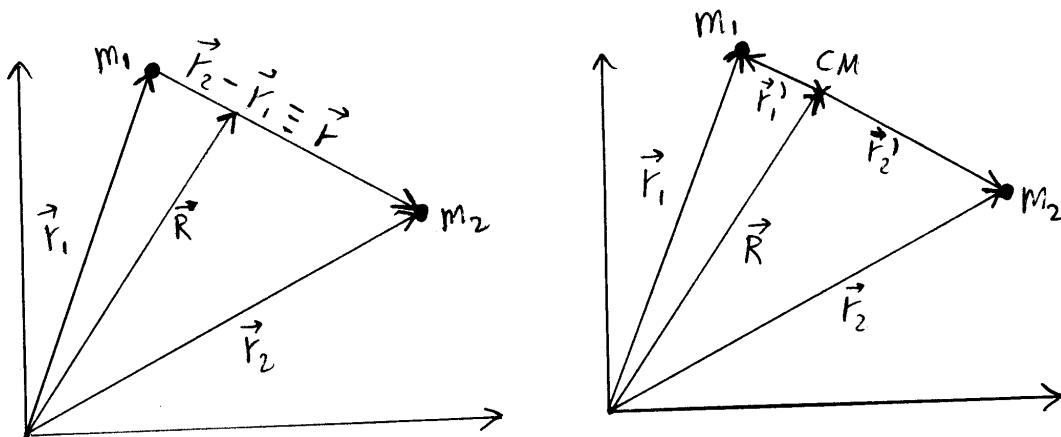


Figura 10.1: Variables de posición fundamentales en el problema de dos cuerpos.

Consideremos un sistema monogénico de dos masas puntuales  $m_1$  y  $m_2$  como lo indica la Fig. 10.1, donde las únicas fuerzas que actúan sobre ellas son las debidas al potencial mutuo  $U$ . La isotropía del espacio nos sugiere que si las masas no poseen alguna propiedad vectorial, la interacción entre ellas debe ir dirigida a lo largo de la línea que las une, esto indica que el potencial debe ser función del valor absoluto de la coordenada relativa  $\mathbf{r}_2 - \mathbf{r}_1 \equiv \mathbf{r}$ . Este sistema tiene seis grados de libertad y por tanto requiere de seis coordenadas generalizadas. Quizás el sistema de coordenadas generalizadas más conveniente lo constituye las coordenadas de posición del centro de masa  $\mathbf{R}$ , y las coordenadas que determinan al vector relativo  $\mathbf{r}$ . Estas coordenadas se pueden escribir en términos de las coordenadas de posición de las partículas  $\mathbf{r}_1$  y  $\mathbf{r}_2$

$$\mathbf{r} \equiv \mathbf{r}_2 - \mathbf{r}_1 ; \quad \mathbf{R} \equiv \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} \quad (10.1)$$

estas ecuaciones se pueden invertir para obtener

$$\mathbf{r}_1 = \mathbf{R} - \frac{m_2}{m_1 + m_2} \mathbf{r} \quad ; \quad \mathbf{r}_2 = \mathbf{R} + \frac{m_1}{m_1 + m_2} \mathbf{r} \quad (10.2)$$

también son útiles las coordenadas de posición de las partículas relativas al centro de masa  $\mathbf{r}'_1$  y  $\mathbf{r}'_2$

$$\mathbf{r}_1 = \mathbf{R} + \mathbf{r}'_1 \quad ; \quad \mathbf{r}_2 = \mathbf{R} + \mathbf{r}'_2 \quad (10.3)$$

con lo cual

$$\mathbf{r}'_1 = -\frac{m_2}{m_1 + m_2} \mathbf{r} \quad ; \quad \mathbf{r}'_2 = \frac{m_1}{m_1 + m_2} \mathbf{r} \quad (10.4)$$

En esta sección consideraremos una situación algo más general en donde el potencial puede depender también de las derivadas temporales del vector relativo  $\mathbf{r}$ . El Lagrangiano del sistema se puede escribir como

$$L = T(\dot{\mathbf{R}}, \dot{\mathbf{r}}) - U(\mathbf{r}, \dot{\mathbf{r}}, \dots)$$

es bien sabido que la energía cinética de un sistema de partículas se puede escribir como la energía cinética del centro de masa mas la energía cinética con respecto al centro de masa (ver Ec. 1.32, Pág. 13)

$$T = \frac{1}{2} m_1 \dot{\mathbf{r}}_1^2 + \frac{1}{2} m_2 \dot{\mathbf{r}}_2^2 = \frac{1}{2} m_1 \dot{\mathbf{r}}_1'^2 + \frac{1}{2} m_2 \dot{\mathbf{r}}_2'^2 + \frac{1}{2} M \dot{\mathbf{R}}^2 \quad (10.5)$$

donde  $M \equiv m_1 + m_2$ . Usando (10.4) se puede escribir la energía cinética en términos de las coordenadas generalizadas elegidas i.e. las componentes de  $\dot{\mathbf{R}}$  y  $\dot{\mathbf{r}}$

$$T(\dot{\mathbf{R}}, \dot{\mathbf{r}}) = \frac{1}{2} \frac{m_1 m_2}{M} \dot{\mathbf{r}}^2 + \frac{1}{2} M \dot{\mathbf{R}}^2 \quad (10.6)$$

el Lagrangiano queda de la forma

$$L = \frac{1}{2} M \dot{\mathbf{R}}^2 + \frac{1}{2} \frac{m_1 m_2}{M} \dot{\mathbf{r}}^2 - U(\mathbf{r}, \dot{\mathbf{r}}, \dots) \quad (10.7)$$

se puede ver que las 3 coordenadas de  $\mathbf{R}$  son cíclicas. Si elegimos como coordenadas generalizadas las tres componentes cartesianas de  $\mathbf{R}$ , vemos que los tres momentos lineales (que serían los momentos canónicos) son constantes y por tanto,  $\dot{\mathbf{R}} = \mathbf{cte}$ , de modo que el centro de masa está en reposo o movimiento rectilíneo uniforme<sup>1</sup>

$$\mathbf{R} = \mathbf{R}_0 + \dot{\mathbf{R}} t \quad (10.8)$$

si nuestro sistema original de referencia es inercial, entonces el sistema con origen en el centro de masa también lo es. Podemos entonces ver el movimiento a partir del centro de masa, en cuyo caso el Lagrangiano (10.7) queda

$$L = \frac{1}{2} \mu \dot{\mathbf{r}}^2 - U(\mathbf{r}, \dot{\mathbf{r}}, \dots) \quad (10.9)$$

donde hemos definido la **masa reducida** del sistema como

$$\mu \equiv \frac{m_1 m_2}{M} \quad (10.10)$$

El Lagrangiano (10.9) es el equivalente al Lagrangiano que se obtendría si tuviéramos una partícula de masa  $\mu$  (que llamaremos la  $\mu$ -partícula) sometida a una fuerza que apunta siempre hacia un punto fijo (fuerza central), y a una distancia  $r$  del centro de fuerza. Por otro lado, el Lagrangiano (10.7) que se escribe desde el sistema de referencia del laboratorio, es equivalente al Lagrangiano de dos partículas desacopladas, una de ellas es la  $\mu$ -partícula ya mencionada y la otra es una partícula libre de masa  $M = m_1 + m_2$ , que se mueve

<sup>1</sup>Desde el punto de vista Newtoniano esto se puede ver por el hecho de que el sistema está aislado, de modo que el centro de masa no puede estar acelerado. En términos de simetrías, se dice que el sistema tiene invarianza traslacional que conduce a la conservación del momento lineal.

con velocidad constante como se vé en la Ec. (10.8) y que llamaremos la  $M$ -partícula. Sin embargo, dado que la dinámica de la  $M$ -partícula es trivial, solo necesitaremos resolver la dinámica de la  $\mu$ -partícula. Por esta razón suele decirse que el problema de dos cuerpos sometidos a fuerzas centrales mutuas, se puede reducir a un problema de una sola partícula que interactúa con un centro de fuerzas<sup>2</sup>.

Debemos recordar sin embargo, que tanto la  $\mu$ -partícula como la  $M$ -partícula **SON IMAGINARIAS**, no hay ninguna partícula en el sistema Físico con masa  $\mu$  o con masa  $M$ . Las trayectorias que encontraremos son las trayectorias de estas partículas imaginarias. Para encontrar la trayectoria de las partículas reales con respecto al sistema inercial original (laboratorio), es necesario devolverse tomando las Ecs. (10.2, 10.8) junto con las soluciones que encontremos para  $\mathbf{r}$ .

No obstante, si ocurre que  $m_1 \ll m_2$  entonces tanto la trayectoria como la masa de la  $\mu$ -partícula, van a ser muy semejantes a la trayectoria y masa real de  $m_1$ . En el mismo límite, la trayectoria y la masa de la  $M$ -partícula son muy similares a la trayectoria y la masa de la partícula real  $m_2$ . De hecho, en esta aproximación la posición del centro de masa es casi igual a la posición de la partícula  $m_2$ .

Podemos decir entonces que el paso de las coordenadas  $\mathbf{r}_1, \mathbf{r}_2$  a las coordenadas  $\mathbf{R}, \mathbf{r}$  nos lleva de un sistema de dos partículas (reales) acopladas o interactuantes, a otro sistema de dos partículas (imaginarias) desacopladas o no interactuantes entre sí, donde una de ellas es libre y la otra interactúa con el potencial central.

## 10.2. Ecuaciones de movimiento y primeras integrales

Nos concentraremos de aquí en adelante en el problema de **una sola partícula** sometida a un potencial que solo depende de  $r$  y no de sus derivadas temporales, de modo que la fuerza va en la dirección de  $\mathbf{r}$ . Es natural tomar el origen en el punto de convergencia de las fuerzas. Dado que en tal sistema de referencia la energía potencial solo depende de la distancia al origen, el problema exhibe simetría esférica, por lo tanto cualquier rotación sobre cualquier eje fijo que pase por el origen no tiene efecto en la solución. Por lo tanto debe haber algún ángulo que representa la rotación sobre el eje en cuestión que debe ser cíclico. Estas propiedades de simetría llevan a una simplificación significativa del problema.

La simetría esférica del Lagrangiano (invarianza ante rotaciones arbitrarias) nos lleva a la conservación del momento angular  $\mathbf{L}$  como vimos en la sección 5.4.3 aplicando el teorema de Noether<sup>3</sup>

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \text{cte}$$

Naturalmente los vectores  $\mathbf{r}$  y  $\mathbf{p}$  forman el plano instantáneo de movimiento. Si el vector  $\mathbf{L}$  es constante, esto significa que todos los planos instantáneos de movimiento pertenecen a un mismo plano. En síntesis, el movimiento es bidimensional, además el “sentido de giro” dictaminado por las condiciones iniciales no se puede invertir ya que esto implicaría que  $\mathbf{L}$  invirtiera su sentido. Un caso particular resulta cuando  $\mathbf{L}$  es nulo, esto solo se logra si  $\mathbf{r}$  es paralelo a  $\mathbf{p}$ , en cuyo caso el movimiento es lineal y la recta que contiene a la trayectoria pasa por el centro de fuerzas (aunque no necesariamente la partícula pasará por el centro de fuerzas)<sup>4</sup>.

Por comodidad podemos asumir que el movimiento yace en el plano  $XY$  de modo que  $\mathbf{L}$  es paralelo al eje  $Z$ . La conservación de  $\mathbf{L}$  nos provee de tres constantes de movimiento, al restringir el movimiento a un plano hemos usado dos de ellas (los dos grados de libertad necesarios para definir el vector unitario en la dirección de  $\mathbf{L}$ ), la tercera deberá aparecer como la magnitud de  $\mathbf{L}$ . El momento angular se puede calcular usando los

<sup>2</sup>Es importante tener en cuenta que en el sistema de referencia del laboratorio, el problema equivalente es de dos cuerpos desacoplados en donde uno de ellos tiene dinámica trivial. Por ejemplo, el número de grados de libertad se tiene que conservar en una transformación de coordenadas, y los seis grados de libertad originales no se pueden manifestar en una sola partícula. Cuando pasamos al sistema de referencia centro de masa (que también es inercial) absorbemos tres grados de libertad triviales y solo nos quedamos con los tres grados de libertad no triviales de la  $\mu$ -partícula.

<sup>3</sup>Recuérdese que desde el punto de vista Newtoniano, esto se puede ver teniendo en cuenta que para fuerzas centrales  $\mathbf{r}$  es paralelo a  $\mathbf{F}$  de modo que el torque  $\mathbf{r} \times \mathbf{F}$  es nulo y por tanto se conserva el vector momento angular (ver sección 1.4 Pág. 4).

<sup>4</sup> $\mathbf{L}$  también es nulo si  $\mathbf{r} = \mathbf{0}$  y/o  $\mathbf{p} = \mathbf{0}$  para todo tiempo. Pero esto implicaría que la partícula es libre.



vectores unitarios radial y transversal de las coordenadas polares

$$\begin{aligned}\mathbf{L} &= \mathbf{r} \times \mathbf{p} = m\mathbf{r} \times \mathbf{v} = mr\mathbf{u}_r \times (v_r\mathbf{u}_r + v_\theta\mathbf{u}_\theta) = mrv_\theta\mathbf{u}_r \times \mathbf{u}_\theta \\ \mathbf{L} &= mr^2\dot{\theta}\mathbf{u}_z\end{aligned}\quad (10.11)$$

Trabajaremos entonces en coordenadas polares planas para las cuales el Lagrangiano se escribe

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r)$$

como se previó, existe una coordenada angular cíclica  $\theta$ . El momento conjugado a esta variable es constante y corresponde a la magnitud del momento angular

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} = cte$$

nótese que esta cantidad contiene la información sobre el sentido de giro (dada por el signo de  $\dot{\theta}$ ) y por tanto del sentido de  $\mathbf{L}$ , como se observa en la Ec. (10.11), es decir  $\dot{\theta}$  nos dice si  $\mathbf{L}$  apunta en forma paralela o antiparalela al eje  $Z$ . Una de las primeras integrales de movimiento es entonces

$$\dot{p}_\theta = \frac{d}{dt}(mr^2\dot{\theta}) = 0 \quad (10.12)$$

y se puede integrar en forma inmediata

$$mr^2\dot{\theta} = l \quad (10.13)$$

siendo  $l$  la magnitud del momento angular. Podemos extraer mas información útil a partir de (10.12), de la cual se deduce

$$\frac{d}{dt}\left(\frac{1}{2}r^2\dot{\theta}\right) = 0 \quad (10.14)$$

veamos la interpretación del término entre paréntesis. Sea  $ds$  la longitud de arco recorrida por la partícula en un intervalo infinitesimal de tiempo  $dt$ , el diferencial de área que barre el radio vector de posición en un tiempo  $dt$  es

$$dA = \frac{r ds}{2} = \frac{1}{2}r(rd\theta)$$

y la velocidad de área, es decir el área que barre el radio vector de posición por unidad de tiempo es

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}r^2\dot{\theta} \quad (10.15)$$

pero de acuerdo con (10.14) se llega a que

$$\frac{d}{dt}\left(\frac{dA}{dt}\right) = 0 \Rightarrow \frac{dA}{dt} = cte$$

de modo que la velocidad de área es constante y se obtiene la segunda ley de Kepler, **el radio vector de posición barre áreas iguales en tiempos iguales**. El valor de la velocidad de área se puede encontrar fácilmente a partir del valor del momento angular  $l$ . Kepler obtuvo esta ley con base en la interacción gravitacional en donde  $V(r) = -k/r$ . Sin embargo, en nuestra presente derivación solo se empleó la conservación del momento angular, mostrando que esta ley es válida para cualquier fuerza central.

La ecuación de Lagrange para la coordenada no cíclica  $r$  es

$$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r} = 0$$

recordando que la fuerza viene dada por  $-\nabla V$  que en este caso es  $-(\partial V/\partial r)\mathbf{u}_r$  se tiene que

$$m\ddot{r} - mr\dot{\theta}^2 = f(r)$$

teniendo en cuenta la primera integral dada por (10.13) podemos eliminar  $\dot{\theta}$  de la anterior ecuación

$$m\ddot{r} - \frac{l^2}{mr^3} = f(r) \quad (10.16)$$

con lo cual se obtiene una ecuación diferencial ordinaria solo en la variable  $r$ . Aunque esto reduce formalmente el problema a cuadraturas, resulta más ventajoso obtener otra primera integral de movimiento teniendo en cuenta que la función energía (o el Hamiltoniano) reúne las condiciones para ser la energía del sistema y además no es función explícita del tiempo lo cual nos lleva a la conservación de la energía del sistema (esto también se puede ver por el hecho de que las fuerzas centrales son conservativas). Esta ley de conservación se escribe

$$E = \frac{1}{2}m(v_r^2 + v_\theta^2) + V(r) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) = \frac{1}{2}m\left(\dot{r}^2 + \frac{r^2l^2}{m^2r^4}\right) + V(r) \quad (10.17)$$

$$E = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}\frac{l^2}{mr^2} + V(r) \quad (10.18)$$

donde hemos tenido en cuenta de nuevo (10.13). Nótese que la Ec. (10.18) es de primer orden en tanto que (10.16) es de segundo orden, es decir formalmente hemos hecho un primer proceso de integración. Este primer proceso de integración se puede ver de manera más transparente usando las Ecs. (10.12, 10.16). La Ec. (10.16) se puede reescribir como

$$m\ddot{r} = -\frac{d}{dr}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right)$$

multiplicando a ambos lados por  $\dot{r}$

$$m\dot{r}\ddot{r} = -\dot{r}\frac{d}{dr}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right) \quad (10.19)$$

el miembro de la izquierda se escribe como

$$m\dot{r}\ddot{r} = \frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2\right)$$

para el miembro de la derecha se tiene en cuenta que la derivada total respecto al tiempo de una función  $g$  que solo depende de  $r$ , se escribe como  $dg(r)/dt = (dg/dr)\dot{r}$ . De modo que

$$-\frac{d}{dt}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right) = -\left[\dot{r}\frac{d}{dr}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right)\right]$$

con lo cual la Ec. (10.19) queda

$$\begin{aligned} \frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2\right) &= -\frac{d}{dt}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right) \Rightarrow \\ \frac{1}{2}m\dot{r}^2 + V + \frac{1}{2}\frac{l^2}{mr^2} &= cte \equiv E \end{aligned} \quad (10.20)$$

y la Ec. (10.20) coincide con (10.18). Se observa que el último paso es un proceso de integración que me redujo la ecuación diferencial de segundo orden a una de primer orden. El procedimiento anterior nos muestra también la ventaja de conocer constantes de movimiento por argumentos de simetría, ya que permite obtener primeras integrales de movimiento sin tener que realizar explícitamente el primer proceso de integración.

Dado que originalmente tenemos dos variables  $r$  y  $\theta$  y ecuaciones diferenciales de segundo orden, se requieren cuatro integraciones para resolver las ecuaciones de movimiento de las cuales hemos obtenido dos. Como se puede ver las primeras integrales obtenidas han convertido las ecuaciones de Lagrange en dos ecuaciones de primer orden que requieren de dos integraciones Ecs. (10.13, 10.18). Nótese que la Ec. (10.18) tiene una apariencia muy similar a una ecuación unidimensional con potencial conservativo

$$E = \frac{1}{2}m\dot{x}^2 + V(x)$$

excepto por el término  $l^2/(2mr^2)$ . Esto nos induce a definir un **potencial efectivo** de la forma

$$V_{eff} \equiv V(r) + \frac{1}{2} \frac{l^2}{mr^2} \quad (10.21)$$

de modo que la Ec. (10.18) se reescribe como

$$E = \frac{1}{2} m \dot{r}^2 + V_{eff}(r) \quad (10.22)$$

con la Ec. (10.22) el análogo unidimensional se hace perfecto. Esto nos permitirá realizar curvas de energía potencial efectiva versus  $r$  e interpretar el movimiento de manera análoga al caso unidimensional, aunque con algunas diferencias que ya indicaremos más adelante.

Ya hemos visto varios métodos para obtener las integrales de movimiento, el más sencillo a partir del procedimiento que hemos seguido es el de resolver para  $\dot{r}$  en la Ec. (10.22)

$$\dot{r} = \sqrt{\frac{2}{m} [E - V_{eff}(r)]} \quad (10.23)$$

con lo cual

$$dt = \frac{dr}{\sqrt{\frac{2}{m} [E - V_{eff}(r)]}}$$

integrando y teniendo en cuenta los valores iniciales  $r(t_0) = r_0$  se escribe

$$t - t_0 = \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{m} [E - V_{eff}(r)]}} \quad (10.24)$$

la Ec. (10.24) nos da  $t$  en función de  $r$  y de las constantes de integración  $E, l, r_0$ , recordemos que  $l$  está contenido en el potencial efectivo. Este puede invertirse al menos formalmente para obtener  $r$  en función del tiempo y las constantes. Una vez encontrada la solución para  $r(t)$  se puede sustituir en (10.13) para obtener la solución en  $\theta$

$$\dot{\theta} = \frac{l}{mr^2(t)} \quad (10.25)$$

$$d\theta = \frac{l dt}{mr^2(t)} \quad (10.26)$$

tomando a  $\theta_0$  como el valor inicial de  $\theta$ , la integración nos da

$$\theta - \theta_0 = \frac{l}{m} \int_{t_0}^t \frac{dt}{r^2(t)} \quad (10.27)$$

finalmente, si estamos interesados en la ecuación de la trayectoria, esta se puede obtener haciendo el cociente de (10.25) sobre (10.23)

$$\frac{d\theta/dt}{dr/dt} = \frac{d\theta}{dr} = \frac{l}{mr^2 \sqrt{\frac{2}{m} (E - V_{eff})}} \quad (10.28)$$

$$\theta - \theta_0 = \int_{r_0}^r \frac{l dr}{mr^2 \sqrt{\frac{2}{m} (E - V_{eff})}} \quad (10.29)$$

formalmente las Ecs. (10.24, 10.27) nos proveen de las dos integrales restantes necesarias para determinar las ecuaciones de movimiento<sup>5</sup>. Dado que se requieren cuatro integrales también se requieren cuatro constantes

<sup>5</sup>La integral (10.28) no es independiente y se obtiene a partir de las anteriores.

de movimiento, las cuales pueden ser  $r_0, \theta_0, \dot{r}_0, \dot{\theta}_0$ , sin embargo un conjunto mas conveniente para la mayoría de aplicaciones es  $E, l, r_0, \theta_0$ . En todo caso, un conjunto se puede convertir en el otro. En los problemas que analizaremos será mas conveniente usar  $E$  y  $l$ . En mecánica cuántica, los valores iniciales de las coordenadas y sus derivadas pierden su significado, pero  $E, l$  son todavía útiles en este formalismo y de por sí muchas de las diferencias entre los comportamientos clásicos y cuánticos estriban en el comportamiento de estas dos cantidades. Por este motivo, para cuantizar una teoría es conveniente que esté escrita en términos de la energía y el momento angular del sistema.

### 10.3. El problema unidimensional equivalente y la clasificación de órbitas

Ya hemos visto que la definición de un nuevo potencial efectivo (10.21) conduce a que una de las primeras integrales asociada a la conservación de la energía, tenga un análogo unidimensional ya que queda escrito de la forma

$$E = \frac{1}{2}m\dot{r}^2 + V_{eff}(r) \quad (10.30)$$

Donde el potencial efectivo  $V_{eff}$  corresponde al potencial real más un término adicional en la forma

$$V_{eff} \equiv V(r) + V_{cent} \quad ; \quad V_{cent} \equiv \frac{1}{2} \frac{l^2}{mr^2} = \frac{1}{2}mv_{\theta}^2 \quad (10.31)$$

$V_{cent}$  se conoce como **potencial centrífugo**. Este nombre se debe a que si este término correspondiera a un potencial real, correspondería a una fuerza central repulsiva (ya que es positivo). No obstante, debe tenerse claro que este término no da cuenta de una interacción real, es decir no es un verdadero potencial, realmente corresponde a una porción de la energía cinética de la partícula (la correspondiente a la velocidad transversal) como puede verse en la sección anterior Ecs. (10.17, 10.18). Este término centrífugo depende de  $l$  y por tanto depende en general de las condiciones iniciales, un cambio en el momento angular corresponderá a un cambio en el potencial centrífugo, obsérvese que el potencial real unidimensional es independiente de las condiciones iniciales lo cual constituye una de las principales diferencias con el potencial efectivo. Sin embargo, mientras el sistema no interactúe con otros cuerpos, el momento angular será una constante y por lo tanto el potencial efectivo no cambiará.

Veamos ahora qué observables se pueden determinar con base en las primeras integrales, la velocidad por ejemplo se puede obtener en magnitud y dirección, su magnitud se obtiene de la conservación de la energía

$$E = \frac{1}{2}mv^2 + V(r) \quad \Rightarrow \quad v = \sqrt{\frac{2}{m}[E - V(r)]} \quad (10.32)$$

en este caso solo interviene el potencial real ya que el efectivo lo que hace es tomar una porción de la energía cinética la cual aquí se escribe completa. La componente radial de la velocidad se puede obtener de (10.23)<sup>6</sup>. Con la rapidez y la velocidad radial es suficiente para obtener el vector velocidad<sup>7</sup>. Otra manera de hacerlo es obtener  $\dot{\theta}$  a través de  $l$  en la Ec. (10.13), y  $\dot{r}$  con la ecuación (10.23).

Las siguientes relaciones se siguen de las Ecs. (10.30, 10.31)

$$E - V_{eff} = \frac{1}{2}m\dot{r}^2 = \frac{1}{2}mv_r^2 \quad (10.33)$$

$$E - V(r) = \frac{1}{2}mv^2 = \frac{1}{2}m(v_r^2 + v_{\theta}^2) \quad (10.34)$$

$$V_{eff} - V = \frac{1}{2}mr^2\dot{\theta}^2 = \frac{1}{2}mv_{\theta}^2 \quad (10.35)$$

<sup>6</sup>Nótese que esto es equivalente a tomar (10.32), descomponiendo la velocidad en sus términos radial y transversal y absorbiendo el término transversal en el potencial para definir el potencial efectivo que aparece en (10.23).

<sup>7</sup>La conservación del momento angular evita una posible ambigüedad en el sentido de la velocidad transversal.

Las Ecs. (10.33, 10.34, 10.35) nos inspiran a hacer gráficas de la energía total, el potencial efectivo, y el potencial real en función de  $r$ . Naturalmente, la curva de  $E$  corresponde a una recta horizontal en virtud de su conservación. De las mencionadas ecuaciones vemos que las regiones accesibles, es decir los valores permitidos para  $r$  deben cumplir que  $E \geq V_{eff}(r)$  puesto  $E - V_{eff}$  es la cantidad no negativa  $\frac{1}{2}m\dot{r}^2$ , esta cantidad a lo más puede ser cero lo cual ocurre en los **puntos de retorno**<sup>8</sup>  $r_i$ , en donde se anula la velocidad radial. De la misma manera, la región accesible debe cumplir que  $E \geq V(r)$  ya que  $E - V(r)$  es la energía cinética completa para un valor dado de  $r$ . Finalmente, en la región permitida se debe cumplir que  $V_{eff}(r) \geq V(r)$  puesto que  $V_{eff} - V$  es la energía cinética transversal.

Las Ecs. (10.33, 10.34, 10.35), muestran que la brecha entre la energía total y la curva de potencial efectivo es la energía cinética radial, la brecha entre  $E$  y la curva  $V(r)$  nos da la energía cinética total y la brecha entre las curvas  $V_{eff}$  y  $V(r)$  nos da la energía cinética transversal (término centrífugo). Por tanto, estas curvas proveen la rapidez de la partícula así como sus componentes radial y transversal para una distancia dada  $r$ , con una energía y momento angular dados.

Por otro lado, vemos que la definición (10.21) conduce automáticamente a la condición  $V_{eff}(r) \geq V(r)$ , y que el cumplimiento de la condición  $E \geq V_{eff}$ , nos lleva automáticamente a  $E \geq V(r)$ . En consecuencia, solo la condición  $E \geq V_{eff}$  nos da información física sobre regiones permitidas o excluidas para la partícula.

El análisis cualitativo aquí descrito permite pintar la órbita en forma aproximada.

## 10.4. Análisis de curvas de potencial efectivo

Vamos a ilustrar el análisis de curvas de potencial efectivo a través de algunos ejemplos

### 10.4.1. Potencial efectivo para interacción kepleriana

Tomemos primero el ejemplo de una fuerza atractiva proporcional al inverso al cuadrado de la distancia al centro de fuerzas

$$\mathbf{f}(r) = -\frac{k}{r^2}\mathbf{u}_r \quad ; \quad V(r) = -\frac{k}{r} \quad ; \quad V_{eff}(r) = -\frac{k}{r} + \frac{l^2}{2mr^2}$$

este potencial efectivo se ilustra en la figura 10.2a, y con líneas punteadas se grafican el término centrífugo y el potencial real  $V(r)$ . El término centrífugo domina para  $r \rightarrow 0$ , en tanto que el potencial real domina cuando  $r \rightarrow \infty$ . Existe un mínimo local que nos permitirá tener equilibrio estable (órbitas acotadas) en ciertas circunstancias.

Si la partícula tiene una energía positiva  $E_1$ , ésta tendrá una distancia mínima de aproximación  $r_1$  al centro de fuerza de acuerdo con el diagrama 10.2b (valores menores de  $r$  implican  $E_1 < V_{eff}$  que de acuerdo con la Ec. 10.33 nos dice que la energía cinética radial es negativa, lo cual no es posible), pero no posee un valor máximo para  $r$  de modo que la órbita no está acotada. Una partícula que viene desde el infinito se tropieza con la “barrera efectiva repulsiva” para ser “repelida” y vuelve hacia el infinito. Esta situación se ilustra en la Fig. 10.4a. Nótese que esta **barrera repulsiva efectiva** no es causada solamente por la interacción real sino que es generada por la conjugación entre la interacción real y las condiciones iniciales. Efectivamente, si redirigimos la partícula (cambiamos la dirección de  $\mathbf{v}_0$ ) de manera que “apunte” desde el infinito directamente hacia el centro atractor, la trayectoria sería una línea recta que en principio puede pasar por el centro atractor. Este cambio en las condiciones iniciales (en  $\mathbf{v}_0$ ) corresponde a un cambio en el potencial centrífugo, mostrando que esta barrera efectiva está relacionada con las condiciones iniciales y no solamente con la interacción<sup>9</sup>. Por otro lado, en el análisis anterior con la condición inicial  $r_0 > r_1$ , hemos supuesto que  $r$  disminuirá hasta rebotar

<sup>8</sup>Hay que tener cuidado con la interpretación de los puntos de retorno en problemas bidimensionales como este. En el caso unidimensional significaba un verdadero “volver sobre los pasos” de la partícula. En el caso bidimensional se devuelven los valores de  $r$ , pero  $\theta$  sigue avanzando en el mismo sentido de giro. De modo que corresponde a un retorno en el valor de la coordenada pero no en el movimiento.

<sup>9</sup>Naturalmente, puede verse que la interacción juega un papel en la creación de la barrera efectiva, ya que si la interacción estuviera ausente, la trayectoria sería una línea recta para cualquier dirección inicial  $\mathbf{v}_0$  y la distancia mínima de acercamiento sería diferente al caso en que la interacción está presente y la trayectoria se curva. Esto también se puede ver teniendo en cuenta que el potencial efectivo es la suma del potencial real (interacción) y la energía cinética transversal (condiciones iniciales).

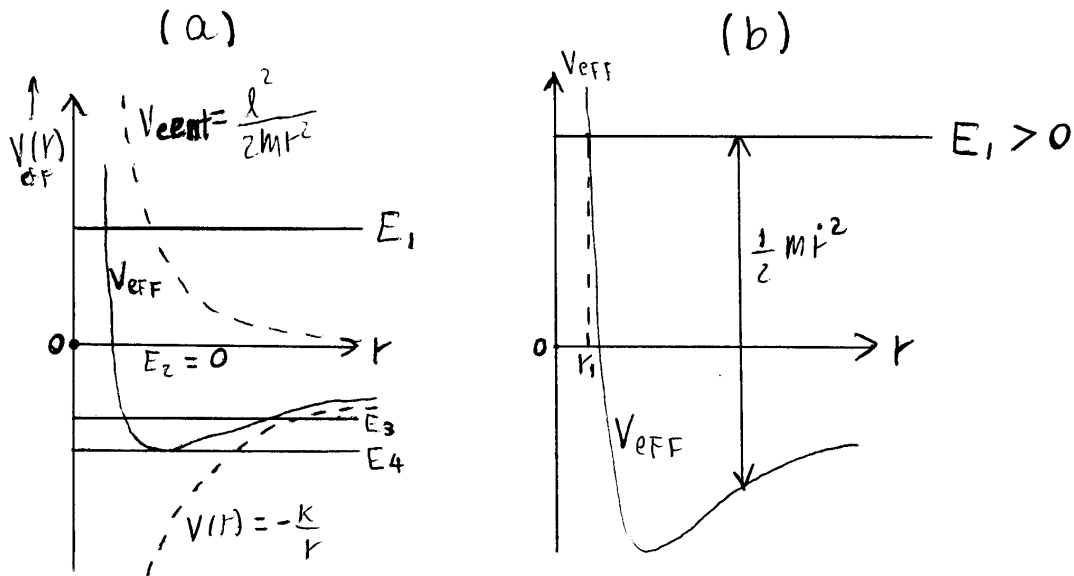


Figura 10.2: (a) Gráficas del potencial centrífugo (línea punteada superior), potencial real (gráfica punteada inferior) y potencial efectivo (gráfica continua), como funciones de  $r$ , para un potencial Kepleriano. Las líneas horizontales representan diversos valores de la energía total de la partícula. (b) Gráfica del potencial efectivo y de la energía total para el caso de energía total positiva.

en la barrera efectiva para volver a aumentar ahora indefinidamente. No obstante, el perfil de la curva no nos prohíbe que la partícula empiece a aumentar su coordenada  $r$  desde el principio aumentando indefinidamente sin acercarse nunca a la barrera de potencial, efectivamente esta sería la situación si  $\dot{r}_0 > 0$  como se puede ver de la Fig. 10.4a, si invertimos el sentido de la velocidad inicial. Nuevamente las condiciones iniciales son las que nos permiten discernir cual es la situación que está ocurriendo<sup>10</sup>. De lo anterior, vemos que dependiendo de las condiciones iniciales es posible que la partícula no acceda a toda la región permitida por la curva de potencial efectivo.

Ahora bien, para el caso de una energía  $E_2 = 0$ , la historia es muy parecida al caso de energía positiva excepto que en el infinito la partícula no poseerá energía cinética (la energía cinética en el infinito es precisamente la energía total ya que todos los potenciales se van para cero). En contraste, cuando la energía es positiva la partícula posee energía cinética incluso en el infinito.

La Fig. 10.3a, muestra que para cualquier valor  $E_3 < 0$  que sea mayor que el mínimo de  $V_{eff}$ <sup>11</sup>, la coordenada  $r$  está acotada entre dos puntos de retorno  $r_1$  y  $r_2$  que serán los valores máximo y mínimo de la distancia al centro atractor, conocidas como **distancias apsidales**. Esto no significa que las órbitas sean necesariamente cerradas, solo nos demuestra que están acotadas por la región definida entre los círculos de radios  $r_1$  y  $r_2$ , donde los puntos de retorno siempre están sobre una de las circunferencias. La forma genérica de la órbita esperada se muestra en la figura 10.4b.

La Fig. 10.3b ilustra la situación en la cual la energía  $E_4$  es igual al mínimo de  $V_{eff}$  de modo que solo hay un valor accesible de  $r$  donde  $\dot{r} = 0$ , en cuyo caso la trayectoria es circular. Definiendo una **fuerza efectiva** de la forma

$$f_{eff}(r) = -\frac{\partial V_{eff}}{\partial r} = f + \frac{l^2}{mr^3} \tag{10.36}$$

el requerimiento de órbita circular  $\partial_r V_{eff} = 0$ , corresponde claramente a la anulación de la fuerza efectiva. Es

<sup>10</sup>Recordemos que en principio se requieren cuatro condiciones iniciales para determinar la dinámica del sistema, y en la curva de potencial hay solo dos, la energía y el momento angular. Usualmente las dos condiciones restantes que permiten un mejor análisis son los valores iniciales de  $r$  y de  $\dot{r}$ .

<sup>11</sup>Si  $E < V_{eff}(r)$  para todo  $r$ , no hay regiones accesibles ya que conduce a energía cinética radial negativa.

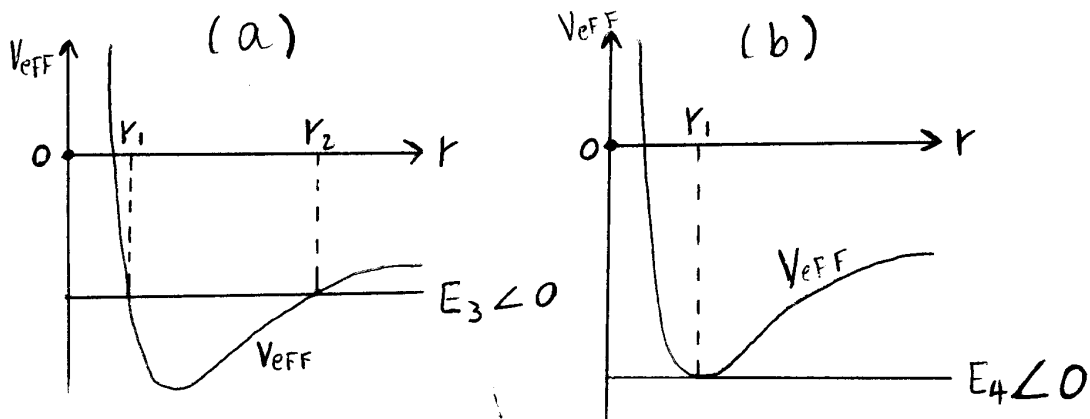


Figura 10.3: (a) Gráfica del potencial efectivo kepleriano y de la energía total para el caso de energía total negativa donde  $E_3$  es mayor que el valor del mínimo del potencial efectivo. (b) Gráfica del potencial efectivo kepleriano y la energía total para el caso de energía total negativa donde  $E_4$  coincide con el valor del mínimo del potencial.

el equivalente a pararse en el sistema no inercial atado a la partícula para ver la anulación entre la fuerza real y la fuerza centrífuga de carácter ficticio.

Por otro lado, recordemos que el cambio en  $l$  cambia el perfil de  $V_{eff}$  aunque esto no cambia la clasificación general de los tipos de órbitas (a menos que  $l$  se vuelva nulo). Es decir para valores negativos de la energía seguimos teniendo órbitas acotadas, y para energías no negativas las órbitas continúan siendo no acotadas sin importar el valor del momento angular (siempre que sea no nulo), aunque por supuesto el perfil específico del potencial efectivo y la forma específica de la órbita cambian con el momento angular. Sin embargo, cuando el momento angular se vuelve nulo la clasificación de las órbitas sí cambia, ya que el potencial centrífugo se anula y la trayectoria es una línea recta.

Más adelante veremos que en el caso de la ley de inverso cuadrado atractivo, las energías positivas conducen a órbitas hiperbólicas, energía cero conduce a órbitas parabólicas y energía negativa a elipses. Estos resultados están en concordancia con el análisis cualitativo que mostramos aquí. Sin embargo, además de los detalles de la órbita, veremos que toda órbita acotada para potencial kepleriano corresponde a una trayectoria cerrada, lo cual no se puede garantizar con el análisis del potencial efectivo. De hecho, son muy pocos los potenciales para los cuales toda órbita acotada es cerrada (ver sección 10.9, Pág. 237).

#### 10.4.2. Potencial efectivo equivalente para dos cuerpos no interactuantes

La gran simplicidad de este problema nos permitirá reforzar el significado físico del potencial centrífugo. Sean dos cuerpos no interactuantes que se mueven con velocidades paralelas y opuestas como se vé en la Fig. 10.5a. Las líneas paralelas que forman las trayectorias de las partículas están separadas una distancia  $b$  usualmente denominada **parámetro de impacto**.

Antes de realizar el análisis del potencial efectivo, realizaremos un análisis puramente cinemático. El gráfico 10.5a muestra que inicialmente  $r$  es muy grande y cuando las partículas se acercan,  $r$  disminuye hasta alcanzar su valor mínimo en  $r = b$  (ver Fig. 10.5b). Posteriormente,  $r$  vuelve a aumentar y crece indefinidamente. Esto nos muestra que  $r = b$  es un punto de retorno para la coordenada  $r$ , la cual disminuye hasta llegar a  $b$  para luego volver a aumentar. Nótese que  $r = b$  es un punto de retorno para  $r$ , pero no para el movimiento como tal. Por otro lado, es fácil ver que si en la Fig. 10.5a invertimos el sentido de ambas velocidades, y  $r(0) > b$ , la coordenada  $r$  aumenta a partir de su valor inicial  $r(0)$  y crece indefinidamente, de modo que  $r$  nunca pasa por el punto de retorno.

Ahora para construir el potencial efectivo, veremos como son las ecuaciones para la  $\mu$ -partícula equivalente

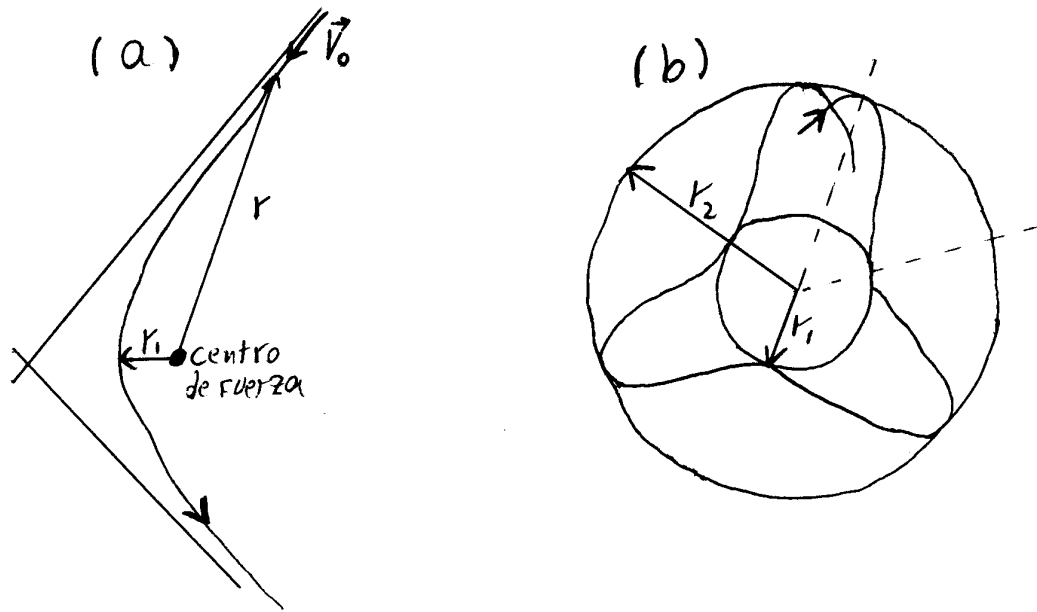


Figura 10.4: (a) Ilustración del tipo de órbita no acotada que corresponde a una energía total no negativa en un potencial kepleriano.  $r_1$  es la menor distancia de aproximación al centro de fuerzas. (b) Ilustración del tipo de órbita acotada entre dos valores de  $r$  que corresponde a una energía total negativa. De momento no podemos garantizar que esta órbita sea cerrada.

a este problema de dos cuerpos. Denotaremos por  $\mathbf{v}_0$  la velocidad relativa la cual viene dada por

$$\mathbf{v}_0 = \dot{\mathbf{r}} = \dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1 = \mathbf{v}_2 - \mathbf{v}_1$$

claramente  $\mathbf{v}_0$  es constante ya que  $\mathbf{v}_1$  y  $\mathbf{v}_2$  lo son. La energía total del sistema relativa al centro de masa es

$$E = \frac{1}{2}\mu v_0^2 + V(r) = \frac{1}{2}\mu v_0^2$$

puesto que  $V(r) = 0$ . El potencial efectivo está dado por

$$V_{eff} = \frac{l^2}{2\mu r^2} + V(r) = \frac{l^2}{2\mu r^2}$$

podemos evaluar  $l$  en forma directa pero es más fácil usando la relación

$$E = \frac{1}{2}\mu \dot{r}^2 + \frac{l^2}{2\mu r^2} = \frac{1}{2}\mu v_0^2 \quad (10.37)$$

es claro que cuando los dos cuerpos alcanza su distancia mínima de aproximación  $r = b$  y  $\dot{r} = 0$  puesto que  $r$  es un mínimo local en este punto, esto se ilustra en la Fig. 10.5b. Por tanto

$$\frac{l^2}{2\mu b^2} = \frac{1}{2}\mu v_0^2 \Rightarrow l = \mu b v_0 \quad (10.38)$$

con lo cual el potencial efectivo tendrá la forma

$$V_{eff} = \frac{1}{2}\mu v_0^2 \frac{b^2}{r^2}$$



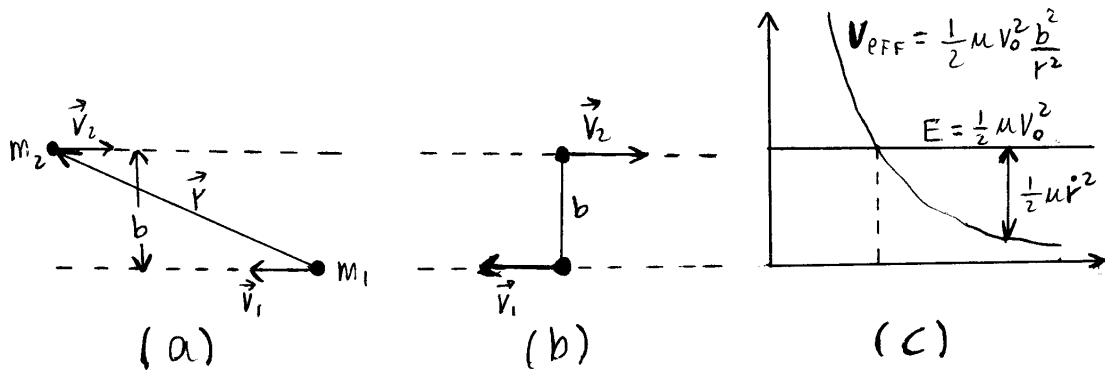


Figura 10.5: Ilustración del comportamiento de la coordenada  $r$  para el caso de dos partículas no interactuantes. (b) Ilustración de la “barrera de potencial efectivo” que nos indica el valor de la mínima distancia de aproximación. (c) Gráfica de potencial efectivo para la partícula equivalente al problema de dos cuerpos no interactuantes.

de modo que el gráfico del potencial efectivo es el indicado en la Fig. 10.5c. La región permitida es aquella en la cual  $E \geq V_{eff}$ . Haremos un análisis del movimiento de la  $\mu$ -partícula equivalente. Denotaremos el punto de retorno como  $r_t$ . Si las condiciones iniciales son tales que  $r(0) > r_t$  y  $\dot{r}(0) < 0$ , entonces la partícula inicialmente se aproxima al punto de retorno y al llegar a él “rebota” en la “barrera de potencial efectivo” luego de lo cual se invierte el sentido radial de movimiento y la coordenada  $r$  crece indefinidamente. Debe tenerse presente que el retorno es en la coordenada radial de la  $\mu$ -partícula, pero no en la coordenada  $\theta$ , de modo que no hay retorno del movimiento como tal. Si cambiamos las condiciones iniciales de modo que  $r(0) > r_t$  y  $\dot{r}(0) > 0$ , la partícula inicialmente se aleja del punto de retorno y la coordenada  $r$  aumenta indefinidamente puesto que no hay puntos de retorno a la derecha de  $r(0)$ , en cuyo caso la  $\mu$ -partícula nunca pasa por el punto de retorno. Compárese este análisis con el estudio puramente cinemático de las dos partículas reales al principio de este sección.

La Fig. 10.5c muestra además que cuando la  $\mu$ -partícula se dirige hacia el punto de retorno, la energía cinética radial va disminuyendo<sup>12</sup>, y cuando se aleja del punto de retorno, la energía cinética radial está aumentando hasta llegar al valor  $E$  en el infinito.

A manera de consistencia, encontraremos el punto de retorno con la condición  $E = V_{eff}(r_t)$  lo cual nos da

$$\frac{1}{2}\mu v_0^2 = \frac{1}{2}\mu v_0^2 \frac{b^2}{r_t^2} \Rightarrow r_t = b$$

en concordancia con nuestro análisis cinemático. Esto se vé de las gráficas 10.5a,b ya que  $r$  disminuye hasta llegar a  $r = b$ , y a partir de allí vuelve a aumentar. En nuestra gráfica unidimensional 10.5c, esto se interpreta diciendo que la  $\mu$ -partícula equivalente “rebota” en la “barrera de potencial efectivo”. Insistiendo en que estos “rebotes” o “retornos” son en la coordenada  $r$  y no en el movimiento del sistema.

Este ejemplo es muy enfático en su mensaje, el potencial efectivo (que aquí coincide con el potencial centrífugo) no está de ninguna manera relacionado con interacción, ya que estas partículas no interactúan entre sí. En este caso, la barrera de potencial efectiva (centrífuga) es debida exclusivamente a las condiciones iniciales del problema. Efectivamente, si el choque fuera frontal i.e.  $b = 0$ , el potencial efectivo se anularía y se desvanecería la barrera de potencial efectiva, lo cual se refleja en el hecho de que dos partículas no interactuantes que se aproximan frontalmente, puede acercarse una a otra en forma indefinida.

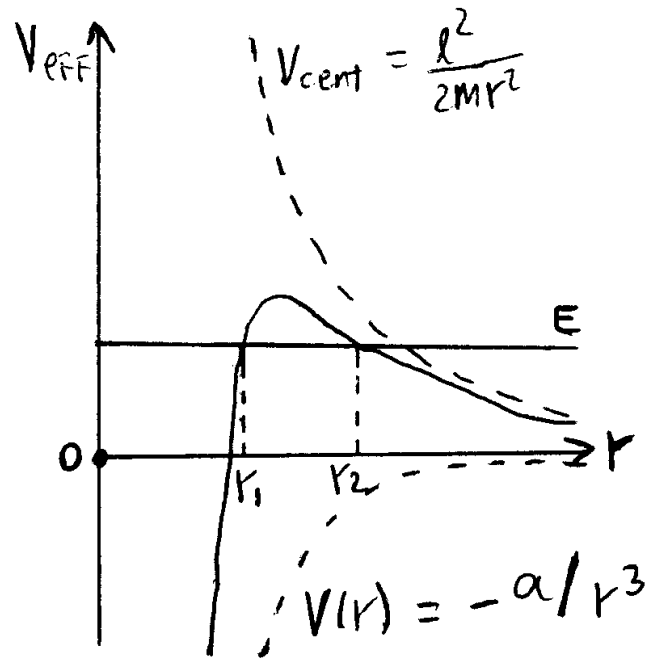


Figura 10.6: Gráficas del potencial centrífugo (línea punteada superior), potencial real (gráfica punteada inferior) y potencial efectivo (gráfica continua), como funciones de  $r$ , para un potencial de la forma  $V = -a/r^3$ .

### 10.4.3. Potencial atractivo proporcional al inverso del cubo de la distancia

Tomemos como ejemplo un potencial atractivo de la forma

$$V(r) = -\frac{a}{r^3} \Rightarrow f = -\frac{3a}{r^4} \quad (10.39)$$

$$V_{eff} = -\frac{a}{r^3} + \frac{l^2}{2mr^2} \quad (10.40)$$

el potencial efectivo junto con el potencial real y el término centrífugo, se grafican en la figura 10.6. Este potencial efectivo solo tiene un máximo local y tiende a cero por la derecha cuando  $r \rightarrow \infty$ , también tiende a  $-\infty$  cuando  $r \rightarrow 0$ . Para una energía  $E$  positiva menor que el máximo local, hay dos tipos de movimiento dependiendo del valor inicial de  $r$ . Si  $r_0 \leq r_1$  el movimiento será acotado entre 0 y  $r_1$ ; además la energía cinética tenderá a infinito a medida que se acerca al centro atractor. Si  $r_0 \geq r_2$  el movimiento es no acotado y su distancia mínima de acercamiento es  $r_2$ , la partícula no podrá nunca acceder al **pozo de potencial**, debido a que existe entre  $r_1$  y  $r_2$  una **barrera de potencial**. El intervalo  $r_1 < r < r_2$  es claramente inaccesible.

Para el caso  $E \leq 0$  la órbita estará acotada entre  $r = 0$  y un punto de retorno, la energía tiende a infinito cuando  $r \rightarrow 0$ . Cuando  $E$  es mayor que el máximo del potencial todas las regiones son permitidas.

Resulta interesante el caso en el cual  $E$  coincide con el valor del máximo del potencial, llamemos  $r_m$  al valor de  $r$  en el cual ocurre el máximo. Si  $r_0 > r_m$  y  $\dot{r}_0 < 0$  la coordenada  $r$  disminuirá acercándose a  $r_m$ , en este proceso disminuye la energía cinética radial hasta anularse cuando  $r = r_m$ , y la partícula queda atrapada en una trayectoria circular de radio  $r_m$ . Nótese sin embargo que si la energía es ligeramente mayor o ligeramente menor a este valor del máximo de potencial, la naturaleza del movimiento cambia drásticamente, de modo que

<sup>12</sup>Naturalmente, la energía cinética total es constante y está dada por  $E - V(r) = E$ . De modo que la energía cinética transversal debe estar aumentando. Efectivamente, esta energía está dada por  $V_{eff}(r) - V(r) = V_{eff}(r)$ , y el potencial efectivo aumenta cuando disminuye  $r$ .

tenemos una órbita circular inestable. De otra parte, si  $r_0 > r_m$  y  $\dot{r}_0 > 0$  la partícula se aleja indefinidamente del valor  $r_m$  y nunca retorna. Un análisis similar se puede hacer para la condición inicial  $r_0 < r_m$ .

10.4.4. Potencial efectivo para fuerza restauradora lineal

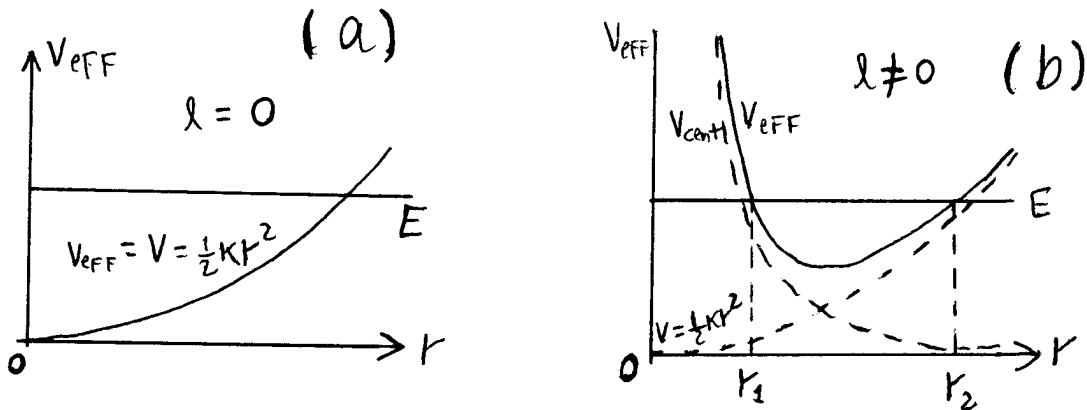


Figura 10.7: (a) Gráfica del potencial efectivo correspondiente a una fuerza restauradora lineal para momento angular nulo. (b) Gráficas de potencial centrífugo, potencial real y potencial efectivo para una fuerza restauradora lineal con momento angular no nulo.

Otro caso interesante es el de una fuerza restauradora lineal (oscilador armónico isotrópico)

$$f = -kr \ ; \ V = \frac{1}{2}kr^2$$

para momento angular cero, correspondiente a movimiento a lo largo de una línea recta,  $V_{eff} = V$  y la situación es como la que se ilustra en la Fig. 10.7a. Para cualquier valor positivo de la energía el movimiento está acotado y como se sabe, es armónico simple. Si  $l \neq 0$ , surge un potencial centrífugo y las características del movimiento se ilustran en la Fig. 10.7b. El movimiento es siempre acotado para todas las energías físicamente posibles y no pasa por el centro de fuerzas. En este caso particular es fácil ver que la órbita es elíptica, ya que si  $\mathbf{f} = -kr$ , las componentes  $x, y$  de las fuerzas son

$$f_x = -kx \ ; \ f_y = -ky$$

el movimiento consta de la composición de dos movimientos armónicos simples de la misma frecuencia cada uno perpendicular al otro. Esto conduce en general a órbitas elípticas.

Un ejemplo bien conocido es el del péndulo esférico de pequeñas oscilaciones. Las famosas figuras de Lissajous se obtienen como la composición de dos movimientos armónicos simples perpendiculares entre sí, y son cerradas cuando los cocientes entre las frecuencias son números racionales. Para dos oscilaciones con la misma frecuencia, la figura es una línea recta cuando las oscilaciones están en fase, un círculo cuando su diferencia de fase es  $\pi/2$ , y una forma elíptica en los demás casos. En consecuencia, el movimiento bajo una fuerza central restauradora lineal nos provee las figuras de Lissajous más sencillas.

10.4.5. Consideraciones generales sobre curvas de potencial efectivo

Para un potencial central (real) dado, la forma detallada de las órbitas puede ser muy compleja y depender fuertemente de las condiciones iniciales. Sin embargo, de los análisis anteriores podemos deducir la siguiente división cualitativa de las órbitas para una partícula sometida a una fuerza central (1) Movimiento acotado, (2) Movimiento no acotado, (3) Movimiento circular, (4) Movimiento rectilíneo.

Cuando el potencial efectivo posee un mínimo local, existen valores de la energía para los cuales la órbita está acotada en un intervalo  $[r_a, r_b]$  siendo  $r_a$  y  $r_b$  puntos de retorno de  $r$  en donde la energía cinética radial se anula. Sin embargo, el movimiento estará acotado en este intervalo solo si la posición inicial es tal que  $r_0 \in [r_a, r_b]$ . Aunque la presencia de un mínimo local en el potencial efectivo es una condición suficiente para que el movimiento acotado sea posible, el ejemplo de la sección 10.4.3 muestra que no es una condición necesaria. Adicionalmente, el ejemplo del potencial tipo Hooke de la sección 10.4.4, nos muestra que existen potenciales (reales) para los cuales el movimiento es siempre acotado sin importar las condiciones iniciales<sup>13</sup>.

Los análisis anteriores también muestran que para muchos potenciales (reales) el movimiento no acotado es posible. Para un par de valores fijos de la energía y el momento angular, esto ocurre cuando existe un punto  $r_b$  para el cual no hay puntos de retorno a la derecha de  $r_b$ . Si en este caso las condiciones iniciales son tales que  $r(0) > r_b$  y  $\dot{r}(0) > 0$ , la coordenada  $r$  crece indefinidamente.

Cuando el potencial efectivo posee un mínimo o un máximo local en un punto  $r_1$ , el movimiento circular es posible si se cumple la condición  $E = V_{eff}(r_1)$ . Para el caso de un mínimo local se requiere además la condición<sup>14</sup>  $r(0) = r_1$ . El movimiento circular tendrá radio  $r_1$  y será estable (inestable) si  $V_{eff}(r_1)$  corresponde a un mínimo (máximo) local. Para más detalles, ver sección 10.7.

Una condición especial interesante ocurre cuando el momento angular es nulo, lo cual cinemáticamente implica que la velocidad inicial de la partícula es tal que su prolongación pasa por el centro de fuerzas. En otras palabras, ocurre cuando la partícula “apunta” directamente al centro de fuerzas. En este caso la simetría esférica del potencial implica que la trayectoria debe ser una línea recta (no hay velocidad inicial transversal ni aceleración transversal). En realidad, la velocidad inicial  $\mathbf{v}_0$  rompe la simetría esférica del potencial y la reduce a una simetría cilíndrica donde el eje de simetría es aquél paralelo a  $\mathbf{v}_0$  que pasa por el centro de fuerzas. No obstante, hay un remanente de la simetría esférica si tenemos en cuenta que la dirección de este eje es arbitraria. Es decir no importa cual sea la dirección de  $\mathbf{v}_0$  (siempre y cuando la partícula “apunte” hacia el centro de fuerzas) vamos a obtener una simetría cilíndrica. Podemos expresar esto diciendo que hay una simetría esférica que nos permite escoger el eje de simetría en dirección arbitraria, pero una vez elegido un eje específico, la simetría esférica se “rompe” reduciéndose a una simetría cilíndrica alrededor de dicho eje. En todo caso, los argumentos de simetría nos dicen que el movimiento de la partícula para  $\mathbf{L} = \mathbf{0}$  debe realizarse a lo largo del eje de simetría determinado por  $\mathbf{v}_0$  y el centro de fuerzas, de modo que la trayectoria debe ser recta.

No todos los potenciales pueden exhibir los cuatro tipos de órbita. Por ejemplo, el potencial real de Hooke no puede conducir a órbitas no acotadas para ningún conjunto de condiciones iniciales. Existe un conjunto de condiciones suficientes (pero no necesarias), para que los cuatro tipos de movimiento sean posibles para un potencial dado, y son las siguientes: (a) el potencial real decae más lentamente que  $1/r^2$  cuando  $r \rightarrow \infty$  y (b) diverge más lentamente que  $1/r^2$  cuando  $r \rightarrow 0$ . La primera condición asegura que el potencial real predomina sobre el término centrífugo para valores grandes de  $r$ , en tanto que la segunda condición asegura que para pequeños valores de  $r$  predomina el término centrífugo. En síntesis, estas condiciones nos garantizan que el potencial en cuestión tendrá un comportamiento asintótico similar al del potencial de Kepler  $V(r) = -k/r$ .

## 10.5. El teorema del virial

Algunas propiedades de las fuerzas centrales se pueden derivar como un caso particular del **teorema del virial**, el cual es de carácter estadístico ya que se refiere a los promedios temporales de varias cantidades mecánicas. Consideremos un sistema de masas puntuales con vectores de posición  $\mathbf{r}_i$  y fuerzas resultantes  $\mathbf{F}_i$ . Las ecuaciones fundamentales de movimiento son

$$\dot{\mathbf{p}}_i = \mathbf{F}_i$$

<sup>13</sup>En el caso del potencial de Hooke esto es de esperarse, debido a que se trata de una interacción que aumenta con la distancia.

<sup>14</sup>Puede pensarse que para mínimos locales con  $E = V_{eff}(r_1)$  la condición  $r(0) = r_1$  es redundante, ya que en cierta vecindad de  $r_1$ , el único valor permitido para  $r$  es precisamente  $r_1$ . No obstante, si el mínimo local no es un mínimo absoluto del potencial efectivo, pueden existir otras regiones permitidas para  $r$ .

ahora nos concentraremos en la cantidad

$$G = \sum_{i=1}^N \mathbf{p}_i \cdot \mathbf{r}_i$$

donde la suma es sobre todas las partículas del sistema.  $\mathbf{F}_i$  incluye a las fuerzas internas y externas sobre la partícula  $i$ . La derivada total en el tiempo de esta cantidad es

$$\begin{aligned} \frac{dG}{dt} &= \sum_{i=1}^N \dot{\mathbf{r}}_i \cdot \mathbf{p}_i + \sum_{i=1}^N \dot{\mathbf{p}}_i \cdot \mathbf{r}_i = \sum_{i=1}^N m_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i + \sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{r}_i = 2 \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 + \sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{r}_i \\ \frac{dG}{dt} &= 2T + \sum_{i=1}^N \mathbf{F}_i \cdot \mathbf{r}_i \end{aligned}$$

calculemos ahora el promedio temporal de esta cantidad tomado sobre un intervalo  $[0, \tau]$ , lo cual se obtiene integrando en dicho intervalo y dividiendo por  $\tau$

$$\begin{aligned} \frac{1}{\tau} \int_0^\tau \left( \frac{dG}{dt} \right) dt &\equiv \overline{\left( \frac{dG}{dt} \right)} = \overline{2T} + \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i} \\ \overline{2T} + \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i} &= \frac{1}{\tau} [G(\tau) - G(0)] \end{aligned}$$

hay dos situaciones interesantes para las cuales el término de la derecha se anula. (a) Cuando el movimiento es periódico de tal modo que todas las coordenadas y velocidades se repiten después de cierto tiempo, en tal caso el término de la derecha se anula si elegimos a  $\tau$  como el periodo del movimiento. (b) Las coordenadas y velocidades de todas las partículas permanecen finitas para todo tiempo, el sistema está entonces acotado y el valor de  $G$  también. En este caso el término de la derecha tiende a cero para tiempos suficientemente largos. En cualquiera de estas situaciones se tiene

$$\overline{T} = -\frac{1}{2} \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i} \quad (10.41)$$

La ecuación (10.41) se conoce como el teorema del virial, y el lado derecho se conoce como el virial de Clausius. Con frecuencia es conveniente separar las fuerzas externas  $\mathbf{F}_i^{(e)}$  e internas  $\mathbf{F}_{ij}$  de cada partícula para escribir cada contribución en forma separada. Si asumimos que se cumple el principio de acción y reacción el teorema queda

$$\overline{T} = -\frac{1}{2} \left\{ \overline{\sum_i \mathbf{F}_i^{(e)} \cdot \mathbf{r}_i} + \overline{\sum_{\text{pares}} \mathbf{F}_{ij} \cdot \mathbf{r}_{ij}} \right\} \quad (10.42)$$

Si las fuerzas son derivables de un potencial, el teorema del virial (10.41) queda de la forma

$$\overline{T} = \frac{1}{2} \overline{\sum_i \nabla V \cdot \mathbf{r}_i} \quad (10.43)$$

En el contexto de las fuerzas centrales, examinaremos la información que nos da el teorema del virial aplicado a una sola partícula sujeta a un potencial central. Si usamos una ley de potencial de la forma

$$V(r) = ar^{n+1}$$

de manera que la fuerza va como  $r^n$ , se tiene que

$$\nabla V \cdot \mathbf{r} = \left( \frac{\partial V}{\partial r} \mathbf{u}_r \right) \cdot (r \mathbf{u}_r) = \frac{\partial V}{\partial r} r = (n+1) V$$

de modo que la Ec. (10.43) para una partícula queda

$$\bar{T} = \frac{n+1}{2} \bar{V} \quad (10.44)$$

este mismo resultado se puede obtener aplicando el teorema de Euler Ec. (5.16) para un potencial homogéneo en  $r$  de grado  $n+1$  (lo cual es más general). Para el caso particular de una fuerza con ley de inverso al cuadrado, usamos  $n = -2$  y resulta

$$\bar{T} = -\frac{1}{2} \bar{V} \quad (10.45)$$

nótese que aunque el teorema del virial está relacionado con promedios temporales, se puede emplear para una sola partícula o para muchas. Ya vimos que el promedio temporal debe tomarse sobre un periodo si el movimiento es periódico, o para tiempos muy grandes si el movimiento no es periódico pero es acotado en el espacio y las velocidades. En el caso de fuerzas centrales, el promedio dado por la Ec. (10.44) solo será válido si la órbita es acotada. Si la órbita no es periódica, el promedio se debe tomar sobre un intervalo muy grande de tiempo, si es periódica se debe tomar sobre el periodo.

### 10.5.1. Otras aplicaciones del teorema del virial

Una de las aplicaciones más interesantes del teorema del virial es la derivación de la ecuación de estado de un gas. Aplicando el teorema del virial junto con ciertas consideraciones estadísticas para un gas en un contenedor, se obtiene

$$pV = NkT + \frac{1}{3} \overline{\sum_{\text{pares}} \mathbf{F}_{ij} \cdot \mathbf{r}_{ij}}$$

El lector interesado puede consultar los detalles por ejemplo en la Ref. [2]. Nótese que la contribución del término de fuerzas internas puede ser negativo (positivo) para fuerzas atractivas (repulsivas). En el caso de un gas ideal, se desprecia la interacción entre moléculas y solo se considera la interacción con las paredes, de modo que todas las fuerzas internas se anulan y la ecuación de estado queda en la forma

$$pV = NkT \quad ; \quad \text{gas ideal}$$

Adicionalmente, si la fuerza resultante (para una o más partículas) es la suma entre fuerzas no friccionales y fuerzas friccionales proporcionales a la velocidad, entonces el virial solo depende de las primeras, no hay contribución de las fuerzas friccionales viscosas. No obstante, es necesario que se le inyecte energía al sistema para mantener el movimiento ya que si las fuerzas viscosas detienen el movimiento todos los promedios temporales tienden a cero para tiempos suficientemente grandes.

## 10.6. Ecuación de la órbita y potenciales integrables

Ya hemos escrito la ecuación de la órbita Ec. (10.29). Una forma alternativa que nos provee de cierta información interesante, se obtiene reescribiendo la Ec. (10.13) de la siguiente forma

$$dt = \frac{mr^2}{l} d\theta$$

esta ecuación permitirá relacionar la derivada temporal de una función arbitraria  $F$  con su derivada con respecto a  $\theta$ . Para nuestros propósitos nos interesa también la segunda derivada

$$\frac{dF}{dt} = \frac{dF}{\left(\frac{mr^2}{l}\right) d\theta} = \frac{l}{mr^2} \frac{dF}{d\theta}$$

$$\begin{aligned} \frac{d^2 F}{dt^2} &= \frac{d}{dt} \left( \frac{dF}{dt} \right) = \frac{d}{dt} \left( \frac{l}{mr^2} \frac{dF}{d\theta} \right) = \frac{d}{\left( \frac{mr^2}{l} \right) d\theta} \left( \frac{l}{mr^2} \frac{dF}{d\theta} \right) \\ \frac{d^2 F}{dt^2} &= \frac{l}{mr^2} \frac{d}{d\theta} \left( \frac{l}{mr^2} \frac{dF}{d\theta} \right) \end{aligned}$$

y como  $F$  es arbitrario, hemos encontrado una relación entre las derivadas en  $\theta$  y en  $t$

$$\frac{d(\dots)}{dt} = \frac{l}{mr^2} \frac{d(\dots)}{d\theta} \quad ; \quad \frac{d^2(\dots)}{dt^2} = \frac{l}{mr^2} \frac{d}{d\theta} \left[ \frac{l}{mr^2} \frac{d(\dots)}{d\theta} \right] \quad (10.46)$$

esta ecuación se puede sustituir bien sea en la Ec. (10.16) o en (10.23). La sustitución en (10.16) nos da una ecuación de segundo orden, en tanto que la sustitución en (10.23) nos da una ecuación de primer orden en el tiempo. Aunque las ecuaciones de segundo orden son en general más difíciles de resolver, aplicaremos primero la Ec. (10.46) a la Ec. (10.16) ya que esta ecuación diferencial también nos dará un información útil. Aplicando las relaciones (10.46) en la Ec. (10.16) se obtiene

$$m \frac{l}{mr^2} \frac{d}{d\theta} \left[ \frac{l}{mr^2} \frac{dr}{d\theta} \right] - \frac{l^2}{mr^3} = f(r) \quad (10.47)$$

usando la identidad

$$\frac{1}{r^2} \frac{dr}{d\theta} = - \frac{d(1/r)}{d\theta}$$

en la Ec. (10.47) se obtiene

$$\frac{l}{r^2} \frac{d}{d\theta} \left[ - \frac{l}{m} \frac{d(1/r)}{d\theta} \right] - \frac{l^2}{m} \left( \frac{1}{r} \right)^3 = f(r) \Rightarrow \frac{l^2}{m} \left( \frac{1}{r} \right)^2 \frac{d^2(1/r)}{d\theta^2} + \frac{l^2}{m} \left( \frac{1}{r} \right)^3 = -f(r)$$

de lo cual se vé conveniente el cambio de variable  $u = 1/r$ , resultando

$$\frac{l^2 u^2}{m} \left[ \frac{d^2 u}{d\theta^2} + u \right] = -f \left( \frac{1}{u} \right) \quad (10.48)$$

$$\left[ \frac{d^2 u}{d\theta^2} + u \right] = \frac{m}{l^2} \frac{1}{u^2} \frac{d}{dr} V \left( \frac{1}{u} \right) \quad (10.49)$$

y teniendo en cuenta

$$\frac{d}{du} = \frac{dr}{du} \frac{d}{dr} = - \frac{1}{u^2} \frac{d}{dr}$$

la Ec. (10.49) se puede escribir en función del potencial

$$\left[ \frac{d^2 u}{d\theta^2} + u \right] = - \frac{m}{l^2} \frac{d}{du} V \left( \frac{1}{u} \right) \quad (10.50)$$

empleando cualquiera de las Ecs. (10.48, 10.50) podemos hallar la órbita con base en la fuerza o el potencial de interacción. Es también interesante el caso inverso, es decir dada la órbita (determinada experimentalmente) encontrar el potencial de interacción. Por el momento, deseamos demostrar a partir de (10.50) que la órbita es simétrica respecto a los puntos de retorno del movimiento. Notemos que si la órbita es simétrica será posible reflejarla respecto a la dirección del ángulo de retorno sin producir ninguna variación. Si se eligen las coordenadas de tal modo que el punto de retorno corresponda a  $\theta = 0$ , la reflexión podrá hacerse matemáticamente sustituyendo  $\theta$  por  $-\theta$ . La ecuación diferencial (10.50) que describe la órbita es evidentemente invariante ante dicha sustitución, ya que aparece la segunda derivada en  $\theta$  pero no la primera derivada. Veamos ahora si las condiciones iniciales son también invariantes<sup>15</sup>, para verlo será más útil escribir las condiciones iniciales en

<sup>15</sup>Hemos visto en la sección 10.4.5, que las condiciones iniciales pueden romper las simetrías del potencial. Similarmente, las condiciones iniciales pueden romper la simetría de las ecuaciones de movimiento. Podemos expresarlo diciendo que las simetrías dinámicas pueden ser rotas por la cinemática.

términos de derivadas en  $\theta$  en lugar de derivadas temporales

$$\left. \frac{du}{d\theta} \right|_{t=0} = \left. \frac{d(1/r)}{d\theta} \right|_{t=0} = - \left. \frac{1}{r^2} \frac{dr}{d\theta} \right|_{t=0} = - \left. \frac{1}{r^2} \frac{dr}{dt} \frac{dt}{d\theta} \right|_{t=0} = - \left. \frac{\dot{r}}{\dot{\theta} r^2} \right|_{t=0} = 0$$

donde hemos tenido en cuenta que  $\dot{r}_0 = 0$  ya que empezamos a una distancia apsidal. Por otro lado, también por nuestras hipótesis de trabajo  $t = 0$  corresponde a  $\theta = 0$  por tanto las condiciones iniciales quedan

$$u = u(\theta = 0) \quad , \quad \left( \frac{du}{d\theta} \right)_{t=0} = \left( \frac{du}{d\theta} \right)_{\theta=0} = 0$$

y estas condiciones iniciales tampoco se ven afectadas por la transformación de inversión angular. Por lo tanto, la ecuación de la órbita es la misma tanto si la expresamos en términos de  $\theta$ , como si la expresamos en función de  $-\theta$ , siempre que  $\theta$  se mida con respecto a una línea apsidal. Esto nos lleva a la conclusión de que **la órbita es invariante ante una reflexión respecto a los vectores apsidales**. Esto implica que se puede construir la órbita completa si se conoce la porción de la órbita comprendida entre dos puntos de retorno cualesquiera. La reflexión de la porción dada respecto a uno de los vectores apsidales produce un trozo adicional de la órbita y se puede repetir este proceso indefinidamente hasta completar el resto de la órbita como se ilustra en la Fig. 10.8.

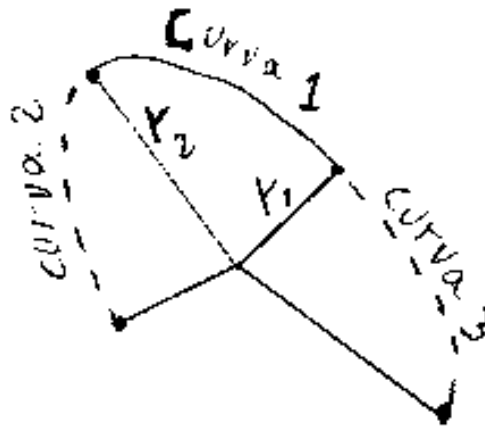


Figura 10.8: Construcción de la órbita a partir de una sección orbital entre dos líneas apsidales (curva 1 continua). La curva 2 se formó por imagen especular de la curva 1 con respecto al ápside  $r_2$  en tanto que la curva 3 se formó por reflexión de la curva 1 con respecto al ápside  $r_1$ .

Retornaremos ahora a la ecuación de órbita en la forma (10.29) pero escribiéndola de nuevo en términos del potencial real  $V(r)$

$$\theta - \theta_0 = \int_{r_0}^r \frac{l \, dr}{mr^2 \sqrt{\frac{2}{m} \left[ E - V(r) - \frac{l^2}{2mr^2} \right]}}$$

la cual se puede reescribir apropiadamente como

$$\theta = \int_{r_0}^r \frac{dr}{r^2 \sqrt{\frac{2mE}{l^2} - \frac{2mV}{l^2} - \frac{1}{r^2}}} + \theta_0 \quad (10.51a)$$

de nuevo haciendo un cambio de variable  $u = 1/r$ ,  $du = -(1/r^2) dr = -u^2 dr$  resulta

$$\theta = \theta_0 - \int_{u_0}^u \frac{du}{\sqrt{\frac{2mE}{l^2} - \frac{2mV}{l^2} - u^2}} \quad (10.52)$$



la integración detallada de esta expresión no es en general sencilla. En realidad solo ciertos tipos de potenciales han sido estudiados en forma detallada. Los más importantes son los potenciales de la forma

$$V = kr^{n+1} \quad (10.53)$$

para los cuales la ecuación (10.52) queda

$$\theta = \theta_0 - \int_{u_0}^u \frac{du}{\sqrt{\frac{2mE}{l^2} - \frac{2mk}{l^2}u^{-n-1} - u^2}} \quad (10.54)$$

la cual es integrable en términos de funciones sencillas solo en algunos casos. Las soluciones se pueden escribir en términos de funciones trigonométricas en los casos en que

$$n = 1, -2, -3$$

en tanto que para las potencias

$$n = 5, 3, 0, -4, -5, -7$$

las soluciones se pueden expresar en términos de funciones elípticas. Estos son todos los casos para potencias enteras en los cuales las soluciones se pueden escribir en términos de funciones sencillas. Algunos exponentes fraccionarios también se pueden escribir en términos de funciones elípticas y muchos otros en términos de funciones hipergeométricas. Las funciones trigonométricas y elípticas son casos especiales de las funciones integrales hipergeométricas. Un análisis más detallado se puede ver en la segunda edición de la Ref. [1].

## 10.7. Condición para órbitas circulares estables e inestables

Aún es posible extraer información adicional del problema unidimensional equivalente así como de la ecuación de la trayectoria. En particular, se puede deducir un teorema relacionado con los tipos de fuerzas centrales atractivas que nos llevan a órbitas cerradas, es decir órbitas en las cuales la partícula trace de nuevo los mismos pasos.

Ya hemos descrito un tipo de órbita cerrada: la órbita circular centrada en el centro de fuerzas. Para un valor dado de  $l$  (es decir para un perfil dado del potencial efectivo), este movimiento se da si la energía total  $E$  de la partícula coincide con un mínimo o máximo local del potencial efectivo, y el radio  $r_0$  del círculo estará dado por el valor de  $r$  para el cual se encuentra dicho mínimo o máximo. Recordemos que el requisito de que  $V_{eff}$  tenga un extremo coincide con el requisito de que la **fuerza efectiva** dada por (10.36) se anule para el valor  $r_0$  donde se ubica el extremo, por lo tanto

$$f(r_0) = -\frac{l^2}{mr_0^3} \quad (10.55)$$

esto nos dice que la fuerza (real) debe ser atractiva al menos a la distancia  $r_0$ , con el fin de lograr una órbita circular. Adicionalmente, la energía de la partícula se obtiene de las Ecs. (10.30, 10.31) simplemente teniendo en cuenta que la energía cinética radial es cero ya que  $\dot{r}_0$  debe ser nulo.

$$E = V_{eff}(r_0) = V(r_0) + \frac{l^2}{2mr_0^2} \quad (10.56)$$

las Ecs. (10.55) y (10.56) implican que para toda fuerza central atractiva se puede obtener una órbita circular de radio arbitrario  $r_0$  dado, si el momento angular viene dado por (10.55) y la energía viene dada por (10.56).

Si el potencial efectivo presenta un mínimo local y elevamos la energía ligeramente, la órbita ya no será circular pero estará acotada entre dos círculos de radios cercanos al de la órbita original, de modo que la trayectoria no se desvía significativamente de la original (aunque podría dejar de ser cerrada). Tomando la terminología del caso unidimensional decimos que esta órbita circular es **estable**. Por el contrario si estamos en un máximo

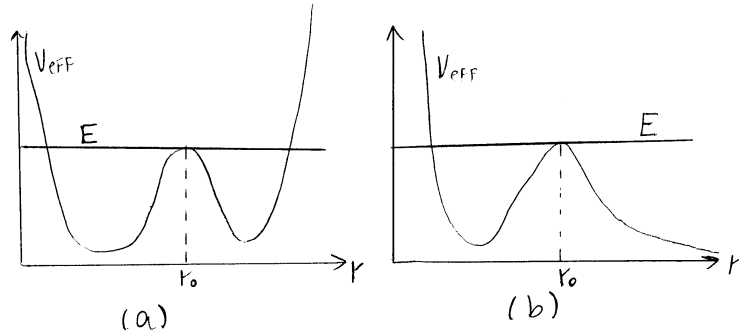


Figura 10.9: (a) Para este potencial efectivo el movimiento es acotado para energías alrededor de  $E$ . (b) Para este potencial un ligero aumento en la energía conduce a movimiento no acotado, una ligera disminución de la energía conduce a movimiento no acotado si  $r(t=0) > r_0$ .

local, el más leve aumento (o disminución) de la energía puede llevar a órbitas totalmente diferentes aunque podrían todavía ser acotadas como se vé en la Fig. 10.9a, pero en algunos casos como el de la figura 10.9b el movimiento se vuelve no acotado. El hecho importante es que un ligero cambio en la energía (y por tanto en las condiciones iniciales) conduce a trayectorias totalmente distintas por lo cual decimos que la órbita circular es **inestable**. Dado que la estabilidad (inestabilidad) está dictaminada por la condición de mínimo (máximo), del potencial efectivo, podemos en consecuencia traducirlo algebraicamente en segunda derivada positiva (negativa) que corresponde a un perfil cóncavo hacia arriba (abajo) en el punto donde se encuentra el extremo. El criterio de estabilidad se escribe entonces como

$$\begin{aligned} \left. \frac{\partial^2 V_{eff}}{\partial r^2} \right|_{r=r_0} &= \left. \frac{\partial^2 V}{\partial r^2} \right|_{r=r_0} + \left. \frac{3l^2}{mr^4} \right|_{r=r_0} = - \left. \frac{\partial f}{\partial r} \right|_{r=r_0} + \frac{3l^2}{mr_0^4} > 0 \Rightarrow - \left. \frac{df}{dr} \right|_{r=r_0} > - \frac{3l^2}{mr_0^4} \\ \left. \frac{df}{dr} \right|_{r=r_0} &< \frac{3l^2}{mr_0^4} \end{aligned}$$

usando (10.55) resulta:

$$\left. \frac{df}{dr} \right|_{r=r_0} < - \frac{3f(r_0)}{r_0} \quad (10.57)$$

y teniendo en cuenta que  $r_0/f(r_0)$  es negativo lo cual se puede verificar de (10.55), obtenemos

$$\frac{r_0}{f(r_0)} \left. \frac{df}{dr} \right|_{r=r_0} > -3 \Rightarrow \left. \frac{\left(\frac{df}{f}\right)}{\left(\frac{dr}{r}\right)} \right|_{r=r_0} > -3 \Rightarrow \left. \frac{d \ln f}{d \ln r} \right|_{r=r_0} > -3$$

la condición de estabilidad se puede escribir entonces de dos maneras equivalentes

$$\frac{r_0}{f(r_0)} \left. \frac{df}{dr} \right|_{r=r_0} = \left. \frac{d \ln f}{d \ln r} \right|_{r=r_0} > -3 \quad (10.58)$$

como caso particular, si la fuerza está gobernada por una ley de potencias de  $r$  de la forma

$$f = -kr^n ; \quad k > 0 \quad (10.59)$$

sacando logaritmo resulta

$$\begin{aligned} \ln f &= \ln(-k) + \ln r^n = \ln(-k) + n \ln r \Rightarrow d \ln f = n d \ln r \\ \frac{d \ln f}{d \ln r} &= n \end{aligned} \quad (10.60)$$

combinando (10.60) con (10.58), la condición de estabilidad (10.58) resulta

$$n > -3$$

Nótese que esta condición de estabilidad no depende del radio. Por tanto, una ley de fuerza atractiva del tipo (10.59) que varíe mas lentamente que  $1/r^3$  puede sostener órbitas circulares estables para todos los valores de  $r_0$ .

## 10.8. Órbitas circulares perturbadas a primer orden

Ya hemos visto que si la órbita circular es estable, un ligero aumento en la energía conduce a una pequeña variación de  $r$  alrededor de  $r_0$ . Para  $u = 1/r$ , usaremos la identidad

$$\frac{d}{du} V\left(\frac{1}{u}\right) = \frac{dr}{du} \frac{d}{dr} V(r) = -\frac{dr}{du} f(r) = \frac{1}{u^2} f\left(\frac{1}{u}\right)$$

Escribamos la Ec. (10.50) en la forma

$$\left[ \frac{d^2 u}{d\theta^2} + u \right] = J(u) \quad ; \quad J(u) \equiv -\frac{m}{l^2} \frac{d}{du} V\left(\frac{1}{u}\right) = -\frac{m}{l^2 u^2} f\left(\frac{1}{u}\right) \quad (10.61)$$

La condición de órbita circular dada por (10.55), para un radio  $r_0 = u_0^{-1}$  queda en la forma

$$f\left(\frac{1}{u_0}\right) = -\frac{l^2}{m} u_0^3 \quad \Rightarrow \quad -\frac{m}{l^2 u_0^2} f\left(\frac{1}{u_0}\right) = u_0 \quad (10.62)$$

sustituyendo (10.62) en la segunda de las Ecs. (10.61), con  $u = u_0$  resulta

$$J(u_0) = u_0 \quad (10.63)$$

y la energía debe cumplir la condición (10.56). Ahora bien, bajo la condición de estabilidad, con una energía ligeramente superior a la dada por (10.56) obtendremos un movimiento acotado en el que para todo tiempo la variable  $u$  solo difiere ligeramente de  $u_0$ . Utilizando (10.63), la expansión de Taylor de  $J(u)$  alrededor de  $u_0$ , es

$$J(u) = u_0 + (u - u_0) \frac{dJ}{du_0} + O\left[(u - u_0)^2\right] \quad (10.64)$$

y para  $u \approx u_0$ , esta expansión se podrá tomar a primer orden. Denotando  $x \equiv u - u_0$  y reemplazando esta expansión a primer orden en (10.61) resulta

$$\begin{aligned} \left[ \frac{d^2 u}{d\theta^2} + u \right] &= u_0 + (u - u_0) \frac{dJ}{du_0} \Rightarrow \left[ \frac{d^2 (u - u_0)}{d\theta^2} + (u - u_0) \right] = (u - u_0) \frac{dJ}{du_0} \Rightarrow \\ \left[ \frac{d^2 x}{d\theta^2} + x \right] &= x \frac{dJ}{du_0} \Rightarrow \frac{d^2 x}{d\theta^2} + x \left( 1 - \frac{dJ}{du_0} \right) = 0 \end{aligned}$$

la ecuación se puede escribir en la forma

$$\frac{d^2 x}{d\theta^2} + \beta^2 x = 0 \quad ; \quad \beta^2 \equiv \left( 1 - \frac{dJ}{du_0} \right) \quad (10.65)$$

es claro que  $\beta^2$  es una cantidad real, adicionalmente la condición de estabilidad para  $x$  exige que  $\beta^2$  sea definido positivo<sup>16</sup>. Para encontrar la relación de  $\beta^2$  con la interacción y las condiciones iniciales debemos evaluar  $dJ/du_0$ . De la definición de  $J(u)$  en (10.61) resulta

$$\frac{dJ}{du} = \frac{2m}{l^2 u^3} f\left(\frac{1}{u}\right) - \frac{m}{l^2 u^2} \frac{d}{du} f\left(\frac{1}{u}\right) = -\frac{2J}{u} - \frac{m}{l^2 u^2} \frac{d}{du} f\left(\frac{1}{u}\right)$$

<sup>16</sup>Se podría pensar que una solución estable con  $\beta^2$  negativo es posible ya que en este caso  $x = Ae^{|\beta|\theta} + Be^{-|\beta|\theta}$  y si  $A = 0$ , entonces  $x$  permanece acotado. No obstante podemos ver que cuando  $\theta \rightarrow \infty$  se tiene que  $x \rightarrow 0$ . Físicamente, esto significa que la energía  $E$  inicial ha disminuído al valor  $E_0$  de la energía que corresponde a la condición de circularidad. Pero las fuerzas centrales son conservativas de modo que esto solo es posible si se introduce una fuerza disipativa adicional.

y aplicando las condiciones de circularidad (10.55, 10.63), se tiene que

$$\begin{aligned}\frac{dJ}{du_0} &= -\frac{2J(u_0)}{u_0} - \frac{m}{l^2 u_0^2} \frac{df}{du_0} = -2 + u_0 \left( -\frac{m}{l^2 u_0^3} \right) \frac{df}{du_0} \\ \frac{dJ}{du_0} &= -2 + \frac{u_0}{f_0} \frac{df}{du_0}\end{aligned}\quad (10.66)$$

reemplazando (10.66) en la segunda de las Ecs. (10.65),  $\beta^2$  queda en la forma

$$\begin{aligned}\beta^2 &= 3 - \frac{u_0}{f_0} \frac{df}{du_0} = 3 - \frac{u_0}{f_0} \left( \frac{dr}{du} \frac{df}{dr} \right) \Big|_{r=r_0} = 3 + \frac{u_0}{f_0} \frac{1}{u_0^2} \frac{df}{dr} \Big|_{r=r_0} = 3 + \frac{1}{f_0 u_0} \frac{df}{dr} \Big|_{r=r_0} \\ \beta^2 &= 3 + \frac{r_0}{f_0} \frac{df}{dr} \Big|_{r=r_0} = 3 + \frac{d \ln f}{d \ln r} \Big|_{r=r_0}\end{aligned}\quad (10.67)$$

nótese que la condición de estabilidad (i.e. de positividad de  $\beta^2$ ) coincide con la condición de estabilidad dada por la Ec. (10.58). Con  $\beta^2$  definida positiva y tomando un origen adecuado para  $\theta$ , la ecuación diferencial (10.65) tendrá como solución

$$x \equiv u - u_0 = a \cos \beta \theta \quad ; \quad \beta^2 \equiv 3 + \frac{r_0}{f_0} \frac{df}{dr} \Big|_{r=r_0} = 3 + \frac{d \ln f}{d \ln r} \Big|_{r=r_0}\quad (10.68)$$

En resumen, hemos demostrado que para pequeñas variaciones con respecto a la condición de circularidad estable, la partícula ejecuta un movimiento armónico simple en  $u$  ( $\equiv 1/r$ ) alrededor de  $u_0$

$$u = u_0 + a \cos \beta \theta \quad (10.69)$$

donde  $a$  es la amplitud del movimiento, la cual depende de la desviación de la energía con respecto al valor de ésta cuando la órbita es circular,  $\beta$  es un valor que surge de la expansión de Taylor de  $J(u)$  alrededor de la órbita circular de radio  $r_0 = u_0^{-1}$  y está dado por la Ec. (10.67). Por otro lado, la ecuación (10.69) muestra que cuando el radio vector de la partícula ha barrido completamente el plano (i.e.  $\theta$  ha barrido un intervalo entre 0 y  $2\pi$ ),  $u$  ha ejecutado  $\beta$  oscilaciones. Si  $\beta$  es un número racional, de tal forma que  $\beta = n/m$  con  $n$  y  $m$  enteros, entonces después de  $m$  revoluciones del radio vector la órbita comenzará a repetirse<sup>17</sup>, es decir obtenemos una **órbita cerrada a primer orden**<sup>18</sup>.

Para cada valor de  $r_0$  que cumpla la desigualdad (10.57) o (10.58), es posible construir una órbita circular estable de radio  $r_0$  si el momento angular y la energía adquieren los valores prescritos por las Ecs. (10.55, 10.56). La pregunta natural es ¿para qué formas funcionales de la fuerza, las órbitas ligeramente perturbadas con respecto a la circular son cerradas a primer orden?. Es claro que la condición de que  $\beta$  sea racional es necesaria, pero se requiere un ingrediente adicional: el valor de  $\beta$  debe ser el mismo para todos los valores de  $r_0$  para los cuales se pueden construir órbitas circulares estables. De no ser así, puesto que  $\beta$  solo puede tomar valores discretos (por ser racional), el número de periodos de oscilación cambiaría discontinuamente con  $r_0$ , y las órbitas no podrían ser cerradas en la discontinuidad. Con  $\beta^2$  constante para todo el rango de  $r_0$ , podemos sin ambigüedad quitar la evaluación en  $r_0$  de la expresión de la derecha en la Ec. (10.67)

$$\beta^2 = 3 + \frac{r}{f} \frac{df}{dr} = 3 + \frac{d \ln f}{d \ln r} \quad (10.70)$$

con lo cual resulta una ecuación diferencial para  $f$  en términos de la variable  $r$ , siempre que tengamos en cuenta que solo es válida en el rango de  $r$  en donde las órbitas circulares estables son posibles. Tenemos entonces

$$\frac{d \ln f}{d \ln r} = \beta^2 - 3 \equiv \frac{dF'}{dr'}$$

<sup>17</sup>Naturalmente, para obtener el mínimo de revoluciones necesarias para repetir la órbita, se requiere que  $n$  y  $m$  sean primos entre sí.

<sup>18</sup>Esta órbita tal vez no es exactamente cerrada ya que aquí solo estamos en aproximación de primer orden, en seguida veremos un criterio para tener órbitas exactamente cerradas.

la solución para  $F'$  en términos de  $r'$  es inmediata

$$F' = (\beta^2 - 3) r' + C' \quad \Rightarrow \quad \ln f = (\beta^2 - 3) \ln r + C'$$

ahora tendremos en cuenta que  $C'$  es una constante que en general puede ser compleja. Definiremos entonces  $C' \equiv \ln(-k)$  con lo cual queda

$$\begin{aligned} \ln f &= \ln r^{\beta^2-3} + \ln(-k) \\ \ln f &= \ln \left[ -kr^{\beta^2-3} \right] \end{aligned}$$

de aquí resulta

$$f(r) = -kr^{\beta^2-3} \quad ; \quad k > 0 \quad (10.71)$$

donde la condición  $k > 0$  proviene del hecho ya discutido de que la condición de circularidad requiere de una fuerza atractiva como se vé en la Ec. (10.55). Todas las leyes de fuerza de la forma (10.71) con  $\beta$  racional, conducen a ecuaciones de órbitas que son cerradas a primer orden, es decir cuando las condiciones iniciales solo difieren ligeramente de aquellas que conducen a una órbita circular. Aparecen dentro del espectro de fuerzas permitidas las leyes de fuerza mas familiares: ley del inverso cuadrado ( $\beta = 1$ ) y ley de Hooke ( $\beta = 2$ ). También aparecen por supuesto un infinito espectro de otras leyes de fuerza.

## 10.9. Órbitas circulares perturbadas a orden superior al primero y condiciones para órbitas cerradas (teorema de Bertrand)

Ahora supongamos que las condiciones iniciales se mueven en una forma no necesariamente ligera, pero de tal forma que la órbita sigue estando acotada. El problema se puede resolver tomando un término adicional en la expansión de Taylor y resolviendo la ecuación de órbita resultante (primer término anarmónico) J. Bertrand resolvió el problema en 1873 y encontró que para desviaciones a segundo orden en la circularidad, las órbitas son cerradas para todo movimiento acotado solo si  $\beta^2 = 1$  ó  $4$ . Estos valores corresponden a la ley de inverso cuadrado y la ley de Hooke respectivamente. Por supuesto, aún es posible que a orden más alto estas órbitas puedan ser abiertas. Por fortuna, dado que solo quedan dos leyes de interacción, éstas se pueden examinar en todo detalle y se puede demostrar que ambas conducen a órbitas cerradas de manera exacta (de hecho lo demostraremos para la ley del inverso al cuadrado), siempre que la órbita sea acotada. Estas son entonces las únicas leyes de fuerza que conducen a órbitas cerradas para cualquier órbita acotada, es decir para condiciones iniciales arbitrarias en  $E$  y  $l$  salvo por la condición de que estos valores conduzcan a órbita acotada. De aquí resulta el **teorema de Bertrand**

**Theorem 6** *Las únicas fuerzas centrales (atractivas) que producen órbitas cerradas para toda trayectoria acotada de una partícula, son la ley del inverso cuadrado y la ley de Hooke.*

Este resultado es muy notable ya que las apreciaciones astronómicas muestran que muchos cuerpos celestes se mueven en órbitas cerradas al menos a primer orden. Esto nos conduce a leyes de la forma (10.71). Sin embargo, si pensamos que la órbita debe ser cerrada cuando solo interactúan dos cuerpos y las pequeñas desviaciones se atribuyen a la interacción con otros cuerpos quedamos con solo dos leyes posibles. La ley de Hooke es descartable ya que implicaría que la interacción aumenta con la distancia haciendo imposible despreciar la interacción con muchos cuerpos. Nos queda entonces que la ley de gravitación debe ser de la forma  $1/r^2$ .

En el formalismo de Hamilton Jacobi vimos una forma alterna de ver el movimiento orbital cerrado: El movimiento orbital en el plano se puede ver como la composición de dos movimientos oscilatorios periódicos uno en  $r$  y el otro en  $\theta$ . En el caso de la ley inverso cuadrado y de Hooke, ambos movimientos tienen el mismo periodo y tenemos entonces una degeneración que como vimos, tiene una fuerte relación con la naturaleza del potencial.

Un comentario final, el teorema de Bertrand no prohíbe la existencia de órbitas cerradas para otras leyes de fuerzas. Lo que el teorema prohíbe para otras leyes de fuerzas es que toda trayectoria acotada sea cerrada, pero

es posible que ciertas trayectorias acotadas con condiciones iniciales muy específicas nos lleven a trayectorias cerradas. De hecho hemos demostrado que toda fuerza central atractiva puede generar órbitas circulares para cualquier valor del radio del círculo, siempre que los valores de  $E$  y  $l$  se ajusten de una manera muy específica.

### 10.10. Órbitas circulares perturbadas usando variables acción-ángulo (Opcional)

Describiremos las pequeñas oscilaciones radiales alrededor de un movimiento circular estable perturbado, utilizando el formalismo del potencial efectivo y las variables acción-ángulo. Expandiendo el potencial efectivo alrededor de algún valor  $r = r_0$ , tenemos

$$V_{eff}(r) = V_{eff}(r_0) + (r - r_0) \frac{dV_{eff}(r_0)}{dr} + \frac{1}{2}(r - r_0)^2 \frac{d^2V_{eff}(r_0)}{dr^2} + \dots$$

el Hamiltoniano en coordenadas polares asociado al potencial central real  $V(r)$ , está dado por

$$H = \frac{1}{2m} \left( p_r^2 + \frac{l^2}{r^2} \right) + V(r) \equiv \frac{1}{2m} p_r^2 + V_{eff}(r)$$

una órbita circular estable existe si  $r_0$  corresponde a un mínimo local del potencial efectivo, en cuyo caso

$$\begin{aligned} \frac{dV_{eff}(r_0)}{dr} &= -\frac{l^2}{mr_0^3} + \frac{dV(r_0)}{dr} = 0 \\ \frac{d^2V_{eff}(r_0)}{dr^2} &= \frac{3l^2}{mr_0^4} + \frac{d^2V(r_0)}{dr^2} \equiv k > 0 \end{aligned} \quad (10.72)$$

dado que las oscilaciones alrededor del movimiento circular son pequeñas, podemos definir

$$r = r_0 + \lambda \quad ; \quad \lambda \ll r_0$$

y el Hamiltoniano queda

$$\begin{aligned} H &= \frac{1}{2m} p_r^2 + V_{eff}(r_0 + \lambda) \simeq \frac{1}{2m} p_r^2 + V_{eff}(r_0) + \frac{1}{2}(r - r_0)^2 \frac{d^2V_{eff}(r_0)}{dr^2} \\ H &= E \simeq \frac{1}{2m} p_r^2 + V_{eff}(r_0) + \frac{1}{2} \lambda^2 k \end{aligned} \quad (10.73)$$

las pequeñas oscilaciones se dan cuando la energía es apenas un poco mayor al valor del mínimo del potencial efectivo. Podemos entonces definir el pequeño valor de energía dado por

$$\epsilon \equiv E - V_{eff}(r_0) \quad (10.74)$$

y teniendo en cuenta que  $\dot{r} = \dot{\lambda}$  se tiene que

$$p_r = \frac{\partial L}{\partial \dot{r}} = \frac{\partial L}{\partial \dot{\lambda}} = p_\lambda \quad (10.75)$$

utilizando las Ecs. (10.74, 10.75), la Ec. (10.73) queda

$$\epsilon \simeq \frac{1}{2m} p_\lambda^2 + \frac{1}{2} \lambda^2 k = \frac{1}{2m} \left[ p_\lambda^2 + m^2 \left( \frac{k}{m} \right) \lambda^2 \right]$$

lo cual es equivalente a redefinir el Hamiltoniano con un corrimiento constante, que claramente no afecta a la Física del problema. Este es el Hamiltoniano de un oscilador armónico para la variable  $\lambda$ . Comparándolo con el Hamiltoniano (9.109), con  $q \rightarrow \lambda$  resulta

$$\epsilon \simeq \frac{1}{2m} (p_\lambda^2 + m^2 \omega^2 \lambda^2) \quad ; \quad \omega^2 \equiv \frac{k}{m}$$

Por tanto, podemos utilizar la Ec. (9.111) para escribir

$$\epsilon = \frac{J\omega}{2\pi}$$

con lo cual la frecuencia está dada por (9.112)

$$\nu = \frac{\partial H}{\partial J} = \frac{\partial \epsilon}{\partial J} = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (10.76)$$

y el movimiento está descrito por las Ecs. (9.115) con la asignación  $q \rightarrow \lambda$ .

$$r = r_0 + \sqrt{\frac{J}{\pi m \omega}} \sin 2\pi w \quad ; \quad p_r = \sqrt{\frac{mJ\omega}{\pi}} \cos 2\pi w$$

se deja como ejercicio al lector escribir la frecuencia de pequeñas oscilaciones radiales (10.76) en términos de las condiciones iniciales, y comprobar que coincide con la obtenida en las Ecs. (10.68).

## 10.11. El problema de Kepler: Ley del inverso al cuadrado (atractiva)

Examinemos brevemente el problema inverso<sup>19</sup>: dada la trayectoria, encontrar la ley de fuerzas que gobierna la dinámica de la partícula. Vale recordar que Newton empleó las leyes de Kepler para deducir la ley de gravitación universal, es decir abordó el problema de manera inversa. Es lógico pensar que esta interacción sea de naturaleza central, en virtud de la isotropía del espacio. Emplearemos las leyes de Kepler que describen el movimiento planetario como postulados observacionales:

1. Todo planeta se mueve con una trayectoria elíptica, en uno de cuyos focos se encuentra el Sol
2. El radio vector que une el Sol y el planeta recorre áreas iguales en tiempos iguales
3. El cociente  $k$  entre el cubo de la distancia  $r$  del sol al planeta y el cuadrado del periodo de revolución  $T$  del planeta

$$k = \frac{r^3}{T^2} \quad (10.77)$$

es idéntico para todos los planetas.

Tomaremos por simplicidad una trayectoria circular (elipse con distancia focal cero) con rapidez constante. El movimiento lunar es un buen ejemplo de esta trayectoria, la fuerza centrípeta es de la siguiente forma

$$F = ma_c = m \frac{v^2}{r} \quad (10.78)$$

como la velocidad tangencial constante es de la forma

$$v = \frac{2\pi r}{T} \quad (10.79)$$

entonces

$$F = \frac{m(4\pi^2 r^2)}{rT^2} = \frac{4\pi^2 m r}{T^2} \quad (10.80)$$

<sup>19</sup>Este problema inverso se conoce usualmente como problema de dispersión o scattering como veremos más adelante.

Aplicando la tercera ley de Kepler Ec. (10.77) a la Ec. (10.80), resulta

$$F = \frac{4\pi^2 m r^3}{r^2 T^2} = \frac{4\pi^2 k m}{r^2}$$

lo cual nos lleva a una ley proporcional al inverso cuadrado de la distancia que además involucra como constante de proporcionalidad a la masa (inercial) de la partícula. Es natural además asumir que esta fuerza depende de dos constantes de acople  $m_1$  y  $m_2$  asociadas con cada cuerpo y que llamaremos masas gravitacionales, llegando a la llamada ley de gravitación universal

$$F = \frac{G m_1 m_2}{r^2} \quad (10.81)$$

Donde  $G$  es una constante que se ajusta experimentalmente y que se denomina constante de gravitación universal. Por otro lado, la segunda ley de Kepler fortalece la idea de que la fuerza es de naturaleza central, ya que vimos que esta ley se cumple para cualquier fuerza central (ver sección 10.2).

Newton calculó la relación entre la aceleración de la gravedad en la tierra  $g$  y la aceleración centrípeta  $a_c$  de la Luna. Aplicando la ley de gravitación universal (asumiendo que la masa inercial y la masa gravitacional son idénticas) podemos calcular la fuerza  $F_R$  que la tierra hace sobre una masa  $m_1$  en su superficie, así como la fuerza centrípeta  $F_r$  que la tierra hace sobre la luna.

$$F_R = \frac{G m_1 m_T}{R^2} = m_1 g \quad ; \quad F_r = \frac{G m_L m_T}{r^2} = m_L a_c$$

siendo  $R$  el radio de la tierra y  $r$  la distancia de la tierra a la luna;  $m_T$  y  $m_L$  son las masas de la tierra y la luna. De esta relación se obtiene

$$\frac{g}{a_c} = \frac{r^2}{R^2} \quad (10.82)$$

Cavendish realizó la corroboración de esta igualdad. Veamos algunos datos típicos

$$R = 6,37 \times 10^6 m \quad ; \quad r = 3,84 \times 10^8 m \quad ; \quad T = 2,36 \times 10^6 s \quad ; \quad g = 9,8 m/s^2$$

siendo  $T$  un periodo lunar. Aplicando estos valores a las Ecs. (10.78, 10.79) se tiene

$$a_c = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} \simeq 2,72 \times 10^{-3} m/s^2 \quad \Rightarrow \quad \frac{g}{a_c} \approx (60)^2$$

en tanto que los cocientes entre  $r^2$  y  $R^2$  son

$$\frac{(3,84 \times 10^8)^2}{(6,37 \times 10^6)^2} = 3634 \approx (60)^2$$

de modo que la igualdad (10.82) está sustentada por los valores experimentales. Este análisis le da consistencia a la ley del inverso al cuadrado.

## 10.12. Solución para la órbita en el problema de Kepler

Habiendo establecido que la atracción gravitacional es inversamente proporcional al cuadrado de la distancia en el caso de trayectoria circular, tomaremos esta ley de fuerza como hipótesis de trabajo para obtener las posibles trayectorias para diversas condiciones iniciales y contrastar con las observaciones astronómicas. La fuerza y el potencial vienen dados por

$$f = -\frac{k}{r^2} \quad ; \quad V = -\frac{k}{r} \quad (10.83)$$

como ya vimos, existen varias maneras de integrar la ecuación de movimiento. La mas simple, consiste en sustituir (10.83) en (10.48). Con lo cual se obtiene

$$\frac{d^2 u}{d\theta^2} + u = -\frac{mf}{l^2 u^2} = \frac{mk}{l^2}$$



usando el cambio de variable  $y = u - \frac{mk}{l^2}$ , la ecuación diferencial queda

$$\frac{d^2y}{d\theta^2} + y = 0 \quad (10.84)$$

cuya solución más general se puede escribir de la forma

$$\begin{aligned} y &= B \cos(\theta - \theta') \Rightarrow u - \frac{mk}{l^2} = B \cos(\theta - \theta') \\ u &= \frac{mk}{l^2} + B \cos(\theta - \theta') \end{aligned}$$

donde  $B$  y  $\theta'$  son las constantes de integración. Escribiendo la solución en función de  $r$

$$\frac{1}{r} = \frac{mk}{l^2} [1 + \varepsilon \cos(\theta - \theta')] \quad ; \quad \varepsilon \equiv B \frac{l^2}{mk} \quad (10.85)$$

Otra alternativa consiste en hacer las asignaciones  $n \rightarrow -2$  y  $k \rightarrow -k$  en las ecuaciones (10.53, 10.54)<sup>20</sup>

$$\theta = \theta' - \int \frac{du}{\sqrt{\frac{2mE}{l^2} + \frac{2mku}{l^2} - u^2}} \quad (10.86)$$

tomaremos la integral como indefinida y a cambio  $\theta'$  absorbe las constantes de integración que surjan, de modo que  $\theta'$  no necesariamente coincide con el valor inicial de  $\theta$ . Puesto que  $\theta'$  absorbe a la constante  $\theta_0$  y a los límites de integración en la ecuación (10.54) se sigue que  $\theta'$  dependerá tanto de  $\theta_0$  como de  $r_0$ .

La integral indefinida tiene la estructura siguiente

$$\int \frac{du}{\sqrt{\alpha + \beta u + \gamma u^2}} = \frac{1}{\sqrt{-\gamma}} \arccos \left[ -\frac{\beta + 2\gamma u}{\sqrt{q}} \right] \quad ; \quad q \equiv \beta^2 - 4\alpha\gamma \quad (10.87)$$

la integral (10.86) se resuelve a través de (10.87) con las siguientes sustituciones

$$\alpha = \frac{2mE}{l^2} \quad , \quad \beta = \frac{2mk}{l^2} \quad ; \quad \gamma = -1 \quad ; \quad q = \left( \frac{2mk}{l^2} \right)^2 \left( 1 + \frac{2El^2}{mk^2} \right) \quad (10.88)$$

donde el valor de  $q$  se obtuvo de

$$\begin{aligned} q &\equiv \beta^2 - 4\alpha\gamma = \left( \frac{2mk}{l^2} \right)^2 - 4 \left( \frac{2mE}{l^2} \right) (-1) = \left( \frac{2mk}{l^2} \right)^2 + 2 \left( \frac{2^2 m^2 k^2 E l^2}{l^4 m k^2} \right) \\ q &= \left( \frac{2mk}{l^2} \right)^2 + \left( \frac{2mk}{l^2} \right)^2 \left( \frac{2El^2}{mk^2} \right) = \left( \frac{2mk}{l^2} \right)^2 \left( 1 + \frac{2El^2}{mk^2} \right) \end{aligned}$$

sustituyendo (10.87, 10.88) en la integral (10.86) se tiene

$$\begin{aligned} \theta &= \theta' - \arccos \left[ -\frac{\left( \frac{2mk}{l^2} \right) - 2u}{\sqrt{\left( \frac{2mk}{l^2} \right)^2 \left( 1 + \frac{2El^2}{mk^2} \right)}} \right] = \theta' - \arccos \left[ -\frac{\left( \frac{2mk}{l^2} \right) \left[ 1 - \frac{2ul^2}{2mk} \right]}{\left( \frac{2mk}{l^2} \right) \sqrt{\left( 1 + \frac{2El^2}{mk^2} \right)}} \right] \\ \theta &= \theta' - \arccos \left[ \frac{\frac{l^2 u}{mk} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}} \right] \end{aligned}$$

<sup>20</sup>Aunque este camino es más largo, tiene la ventaja de que la constante  $\varepsilon$  queda directamente en términos de la energía y el momento angular

$\varepsilon > 1$	$E > 0$	$l \neq 0$	hipérbola
$\varepsilon = 1$	$E = 0$	$l \neq 0$	parábola
$\varepsilon < 1$	$E < 0$	$l \neq 0$	elipse
$\varepsilon = 0$	$E = -\frac{mk^2}{2l^2}$	$l \neq 0$	circunferencia

Cuadro 10.1: Tipos de cónicas que se obtienen con diferentes valores de la excentricidad. La segunda columna nos da el valor de la energía para el cual se obtienen estas cónicas. El momento angular se asume diferente de cero.

y resolviendo para  $u \equiv 1/r$  la ecuación de la trayectoria queda

$$\begin{aligned} \cos(\theta' - \theta) &= \frac{\frac{l^2 u}{mk} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}} \Rightarrow \sqrt{1 + \frac{2El^2}{mk^2}} \cos(\theta - \theta') = \frac{l^2 u}{mk} - 1 \\ \frac{l^2 u}{mk} &= 1 + \sqrt{1 + \frac{2El^2}{mk^2}} \cos(\theta - \theta') \Rightarrow u = \frac{mk}{l^2} \left[ 1 + \sqrt{1 + \frac{2El^2}{mk^2}} \cos(\theta - \theta') \right] \\ \frac{1}{r} &= \frac{mk}{l^2} \left[ 1 + \sqrt{1 + \frac{2El^2}{mk^2}} \cos(\theta - \theta') \right] \end{aligned} \quad (10.89)$$

se puede ver que esta ecuación coincide con (10.85) excepto que  $\varepsilon$  aparece en (10.89) como función de  $E$  y  $l$  tal como se anticipó. La Ec. (10.89) nos permite identificar a  $\theta'$  como uno de los ángulos correspondientes a la inversión del movimiento de la órbita, ya que si tomamos  $\theta' = 0$  vemos que esta ecuación queda invariante ante la transformación especular  $\theta \rightarrow -\theta$ . Esto nos indica según la discusión de la sección 10.6 que  $\theta = \theta'$  corresponde al lugar geométrico de un ápside, de hecho la Ec. (10.89) nos indica que es un periápside, puesto que para  $\theta = \theta'$  la coordenada  $r$  adquiere su valor mínimo. Notemos que de las cuatro constantes de integración solo aparecen tres ( $E, l, \theta'$ ), esto se debe a que la cuarta constante ( $\theta_0$  ó  $r_0$ ) ubica la posición inicial de la partícula, la cual es claramente irrelevante en la ecuación de la órbita<sup>21</sup>. Sin embargo, esta cuarta constante deberá aparecer cuando se solucione el problema de  $r$  y  $\theta$  en función del tiempo. En particular, si queremos resolver la primera integral de movimiento asociada a la conservación del momento angular Ec. (10.13), tenemos que conocer el valor inicial de  $\theta$  i.e.  $\theta_0$ .

### 10.12.1. Clasificación de las órbitas según los valores de $E$ y $l$

Teniendo en cuenta que en coordenadas polares, la ecuación general de una cónica con foco en el origen es de la forma

$$\frac{1}{r} = C [1 + \varepsilon \cos(\theta - \theta')]$$

siendo  $\varepsilon$  la excentricidad de la sección cónica, se tiene por comparación con (10.89) que la órbita siempre corresponde a una sección cónica con foco en el centro de fuerzas y excentricidad

$$\varepsilon = \sqrt{1 + \frac{2El^2}{mk^2}} \quad (10.90)$$

como se sabe de la teoría de secciones cónicas, el tipo de cónica depende del valor de la excentricidad según la tabla 10.1

Los valores de la tabla 10.1 corresponden a todos los casos posibles con  $l \neq 0$ . Cuando el momento angular es nulo la Ec. (10.89) ya no es válida, y recalculando se observa fácilmente que la trayectoria consiste en una

<sup>21</sup>No obstante, si la órbita es abierta, esta información será necesaria para encontrar la porción de la órbita que la partícula recorre, la cual depende del punto de partida. Para trayectoria cerrada el punto de partida no es relevante para trazar la trayectoria completa.

línea recta, que corresponde al caso en el cual la velocidad inicial es tal que la partícula “apunta” directamente al centro de fuerzas. La clasificación anterior coincide con la clasificación general de trayectorias acotadas y no acotadas que se hizo en la sección 10.4.1, basados en el análisis de la curva de energía potencial efectiva. Nótese que siempre que garanticemos que  $l \neq 0$ , el tipo de cónica solo depende de la energía, aunque la trayectoria detallada depende de  $E$  y  $l$ . Adicionalmente, si la trayectoria es abierta y la posición inicial de la partícula no está en el infinito<sup>22</sup>, la partícula solo recorrerá una porción de la sección cónica (parábola o hipérbola).

### 10.12.2. Condición de circularidad

Veamos como se relaciona la condición de circularidad con la estudiada en la sección 10.7. Aplicando el teorema del virial (10.45) y teniendo en cuenta que en la órbita circular  $E$ ,  $T$  y  $V$  son constantes, tenemos

$$E = \overline{T} + \overline{V} = -\frac{1}{2}\overline{V} + \overline{V} = \frac{\overline{V}}{2} = \frac{V}{2}$$

por lo tanto

$$E = -\frac{k}{2r_0} \quad (10.91)$$

pero recordando la condición de circularidad (10.55), que equivale a la anulación de la fuerza efectiva podemos escribir

$$f(r_0) = -\frac{k}{r_0^2} = -\frac{l^2}{mr_0^3} \Rightarrow r_0 = \frac{l^2}{mk} \quad (10.92)$$

que al reemplazar en (10.91) resulta

$$E = -\frac{mk^2}{2l^2} \quad (10.93)$$

esta expresión coincide con la que se obtendría haciendo  $\varepsilon = 0$  (condición de circularidad) en la Ec. (10.90). Un breve cálculo nos muestra que esta expresión también coincide con la que se obtuvo en (10.56) para la condición de circularidad con un potencial central arbitrario. Partiendo de (10.93) y usando (10.92) se tiene que

$$E = -\frac{mk^2}{l^2} + \frac{mk^2}{2l^2} = -k\frac{mk}{l^2} + \frac{l^2}{2m} \left(\frac{mk}{l^2}\right)^2 = -\frac{k}{r_0} + \frac{l^2}{2mr_0^2} = V(r_0) + \frac{l^2}{2mr_0^2}$$

que coincide con (10.56).

### 10.12.3. Órbitas elípticas

Para el caso de órbitas elípticas ( $\varepsilon < 1$ ,  $E < 0$ ,  $l \neq 0$ ), se puede demostrar que el eje mayor solo depende de la energía, hecho fundamental en la teoría atómica de Bohr. De las propiedades de la elipse, la suma de las distancias apsidales  $r_1 + r_2$  es la longitud del eje mayor. Como en los puntos  $r_i$  que determinan las ápsides la velocidad radial es por definición nula, la conservación de la energía nos dice que

$$E = V(r_i) + \frac{1}{2}mv_\theta^2 = -\frac{k}{r_i} + \frac{l^2}{2mr_i^2} \quad ; \quad i = 1, 2$$

por tanto, las distancias apsidales corresponden a las raíces de la siguiente ecuación

$$\begin{aligned} E - \frac{l^2}{2mr_i^2} + \frac{k}{r_i} &= 0 \\ r_i^2 + \frac{k}{E}r_i - \frac{l^2}{2mE} &= 0 \end{aligned}$$

<sup>22</sup>Si la partícula está inicialmente en el infinito (que en la práctica significa “muy lejos”), aún es posible que la velocidad inicial lo aleje aún más del centro de fuerza.

en una ecuación cuadrática, la suma de las raíces es el cociente entre el coeficiente del término lineal sobre el coeficiente del término cuadrático cambiado de signo. Por lo tanto el semieje mayor viene dado por

$$a = \frac{r_1 + r_2}{2} = -\frac{k}{2E} \quad (10.94)$$

y el semieje mayor solo depende de la energía como se anticipó. En el límite de circularidad la Ec. (10.94) coincide con (10.91). Despejando la energía en (10.94) y reemplazando en (10.90), se obtiene la excentricidad de la elipse en función del semieje mayor

$$\varepsilon = \sqrt{1 - \frac{l^2}{mka}} \quad (10.95)$$

la cual se puede despejar en la forma

$$\frac{l^2}{mk} = a(1 - \varepsilon^2) \quad (10.96)$$

con lo cual la ecuación de la órbita (10.85) cuando ésta es elíptica se puede reescribir en la forma

$$r = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon \cos(\theta - \theta')} \quad (10.97)$$

de la Ec. (10.97) se encuentra que las dos distancias apsidales (correspondientes a  $\theta - \theta'$  igual a cero y  $\pi$ ) son iguales a

$$r_1 = r_{\min} = a(1 - \varepsilon) \quad ; \quad r_2 = r_{\max} = a(1 + \varepsilon)$$

lo cual es consistente con las propiedades de la elipse.

### 10.13. Movimiento en el tiempo en el problema de Kepler

La relación entre las coordenadas y el tiempo es mucho más complicada de obtener que la ecuación de la órbita. Las Ecs. (10.24, 10.27) nos dan en principio la relación entre la distancia radial y el tiempo y la posición angular y el tiempo, que para el potencial (10.83) toman la forma

$$t - t_0 = \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{m}(E - V_{eff})}} = \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{m}\left(E + \frac{k}{r} - \frac{l^2}{2mr^2}\right)}} \quad (10.98)$$

$$\theta = \theta_0 + \frac{l}{m} \int_{t_0}^t \frac{dt}{r^2(t)} \quad (10.99)$$

sin embargo, la expresión (10.99) tiene la desventaja de requerir la integración previa de (10.98) y la inversión de la relación obtenida para encontrar  $r(t)$ . Por esta razón, es más fácil encontrar la relación entre  $\theta$  y el tiempo usando la Ec. (10.13) de conservación del momento angular.

$$mr^2 \frac{d\theta}{dt} = l \quad \Rightarrow \quad dt = \frac{mr^2}{l} d\theta \quad (10.100)$$

combinando esta ecuación con la ecuación de órbita (10.85) resulta

$$dt = \frac{m \left[ \frac{l^2}{mk} \frac{1}{1 + \varepsilon \cos(\theta - \theta')} \right]^2}{l} d\theta \quad \Rightarrow \quad dt = \frac{ml^4 d\theta}{lm^2 k^2 [1 + \varepsilon \cos(\theta - \theta')]^2}$$

$$t - t_0 = \frac{l^3}{mk^2} \int_{\theta_0}^{\theta} \frac{d\theta}{[1 + \varepsilon \cos(\theta - \theta')]^2} \quad (10.101)$$

aunque esta integral se puede escribir en forma cerrada, la forma funcional es muy compleja y su inversión para obtener las coordenadas en función del tiempo es un problema colosal.

### 10.13.1. Dependencia temporal en el caso parabólico

Ilustraremos entonces un caso más simple reduciendo nuestra atención al movimiento parabólico ( $\varepsilon = 1$ ). Es usual medir el ángulo polar a partir del radio vector en el punto de mayor acercamiento conocido como **perihelio** o **periapsis**. Este convenio implica hacer  $\theta'$  nulo en la ecuación de órbita (10.85). Correspondientemente el tiempo se comienza a medir a partir del paso por el periapsis. Usando la identidad

$$1 + \cos \theta = 2 \cos^2 \frac{\theta}{2}$$

la ecuación (10.101) aplicada a una sección parabólica ( $\varepsilon = 1$ ) con  $\theta' = 0$  nos da

$$\begin{aligned} t &= \frac{l^3}{mk^2} \int_{\theta_0}^{\theta} \frac{d\theta}{[1 + \cos \theta]^2} = \frac{l^3}{mk^2} \int_{\theta_0}^{\theta} \frac{d\theta}{[2 \cos^2 \frac{\theta}{2}]^2} = \frac{l^3}{4mk^2} \int_0^{\theta} \sec^4 \frac{\theta}{2} d\theta \\ t &= \frac{l^3}{4mk^2} \int_0^{\theta} \left(1 + \tan^2 \frac{\theta}{2}\right) \sec^2 \frac{\theta}{2} d\theta \end{aligned}$$

usando el cambio de variable  $x = \tan(\theta/2)$ ,  $dx = \frac{1}{2} \sec^2 \frac{\theta}{2} d\theta$

$$\begin{aligned} t &= \frac{l^3}{2mk^2} \int_0^x (1 + x^2) dx = \frac{l^3}{2mk^2} \left[ x + \frac{x^3}{3} \right] \\ t &= \frac{l^3}{2mk^2} \left[ \tan \frac{\theta}{2} + \frac{1}{3} \tan^3 \frac{\theta}{2} \right] \end{aligned} \quad (10.102)$$

aún la inversión de esta ecuación implica resolver una cúbica para  $\tan(\theta/2)$ , para luego hallar el arcotangente respectivo. La distancia radial en ese instante se encuentra mediante la ecuación orbital.

### 10.13.2. Dependencia temporal para el movimiento elíptico

Para el movimiento elíptico la integración se realiza de manera más conveniente mediante una variable auxiliar  $\psi$  conocida como **anomalía excéntrica**, definida por la relación

$$r = a(1 - \varepsilon \cos \psi) \quad (10.103)$$

al comparar con la Ec. (10.97) para la órbita, queda claro que  $\psi$  también cubre el intervalo entre 0 y  $2\pi$  cuando  $\theta$  recorre una revolución completa, además  $\psi = 0$  en el perihelio al igual que  $\theta$  (por convenio), así mismo en el afelio  $\psi = \theta = \pi$ . A partir de las Ecs. (10.94, 10.95) podemos expresar a  $E$  y  $l$  en función de  $a$ ,  $\varepsilon$ , y  $k$

$$E = -\frac{k}{2a} ; \quad \varepsilon^2 = 1 - \frac{l^2}{mka} \quad \Rightarrow \quad l^2 = mka(1 - \varepsilon^2)$$

estos reemplazos tienen la ventaja de que  $a$  y  $\varepsilon$  son variables más geométricas y por tanto más observables astronómicamente. Reemplazando estas expresiones en la Ec. (10.98) para movimiento elíptico resulta

$$\begin{aligned} t &= \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{m} \left( -\frac{k}{2a} + \frac{k}{r} - \frac{mka(1-\varepsilon^2)}{2mr^2} \right)}} = \int_{r_0}^r \frac{dr}{\sqrt{\frac{2k}{mr^2} \left( -\frac{r^2}{2a} + r - \frac{a(1-\varepsilon^2)}{2} \right)}} \\ t &= \sqrt{\frac{m}{2k}} \int_{r_0}^r \frac{r dr}{\sqrt{r - \frac{r^2}{2a} - \frac{a(1-\varepsilon^2)}{2}}} \end{aligned} \quad (10.104)$$

por convención  $r(t=0) = r_0$  es la distancia al perihelio, para la cual  $\psi = 0$ . La introducción de la anomalía excéntrica (10.103)

$$r = a(1 - \varepsilon \cos \psi) ; \quad dr = a\varepsilon \sin \psi d\psi$$

reduce la integral (10.104) a una forma más simple:

$$\begin{aligned}
 t &= \sqrt{\frac{m}{2k}} \int_0^\psi \frac{a(1-\varepsilon \cos \psi) (a\varepsilon \sin \psi d\psi)}{\sqrt{a(1-\varepsilon \cos \psi) - \frac{[a(1-\varepsilon \cos \psi)]^2}{2a} - \frac{a(1-\varepsilon^2)}{2}}} \\
 t &= \sqrt{\frac{m}{2k}} \int_0^\psi \frac{a^2\varepsilon(1-\varepsilon \cos \psi) \sin \psi d\psi}{\sqrt{a(1-\varepsilon \cos \psi) - \frac{a^2(1-2\varepsilon \cos \psi + \varepsilon^2 \cos^2 \psi)}{2a} - \frac{a(1-\varepsilon^2)}{2}}} \\
 t &= \sqrt{\frac{ma^4}{2k}} \int_0^\psi \frac{\varepsilon(1-\varepsilon \cos \psi) \sin \psi d\psi}{\sqrt{\frac{a}{2}\sqrt{2}(1-\varepsilon \cos \psi) - (1-2\varepsilon \cos \psi + \varepsilon^2 \cos^2 \psi) - (1-\varepsilon^2)}} \\
 t &= \sqrt{\frac{2ma^4}{2ka}} \int_0^\psi \frac{\varepsilon(1-\varepsilon \cos \psi) \sin \psi d\psi}{\sqrt{2-2\varepsilon \cos \psi - 1 + 2\varepsilon \cos \psi - \varepsilon^2 \cos^2 \psi - 1 + \varepsilon^2}} \\
 t &= \sqrt{\frac{ma^3}{k}} \int_0^\psi \frac{\varepsilon(1-\varepsilon \cos \psi) \sin \psi d\psi}{\sqrt{\varepsilon^2 - \varepsilon^2 \cos^2 \psi}} = \sqrt{\frac{ma^3}{k}} \int_0^\psi \frac{(1-\varepsilon \cos \psi) \sin \psi d\psi}{\sqrt{\sin^2 \psi}} \\
 t &= \sqrt{\frac{ma^3}{k}} \int_0^\psi (1-\varepsilon \cos \psi) d\psi \tag{10.105}
 \end{aligned}$$

primero podemos observar que esta ecuación nos permite obtener una expresión para el periodo  $\tau$  del movimiento elíptico, si extendemos la integral a todo el dominio de  $\psi$  entre 0 y  $2\pi$

$$\tau = 2\pi a^{3/2} \sqrt{\frac{m}{k}} \tag{10.106}$$

Ahora escribamos (10.105) introduciendo la frecuencia de revolución  $\omega$  en la forma

$$\omega = \frac{2\pi}{\tau} = \sqrt{\frac{k}{ma^3}}$$

donde hemos usado (10.106). La integral de (10.105) se calcula fácilmente y se obtiene

$$\omega t = \psi - \varepsilon \sin \psi \quad ; \quad \psi(t=0) = 0 \tag{10.107}$$

esta relación es conocida como ecuación de Kepler. La cantidad  $\omega t$  barre el dominio entre 0 y  $2\pi$  junto con  $\psi$  y  $\theta$ , en el curso de una revolución completa, razón por la cual también se le llama una anomalía, más exactamente **anomalía media**.

Para hallar la posición en la órbita en un tiempo  $t$ , se debería en primer lugar invertir la ecuación de Kepler (10.107) de tal manera que se obtiene la anomalía excéntrica  $\psi$  en función del tiempo. Esta última se reemplaza entonces en (10.103) que es la ecuación que define la anomalía excéntrica, con ello obtenemos  $r(t)$ . Por otro lado, el ángulo polar  $\theta$  se puede expresar en función de  $\psi$  por comparación de (10.103) con la ecuación de la órbita elíptica (10.97), donde nuevamente por convención tomaremos  $\theta' = 0$ , que equivale a que  $\theta = 0$  en el perihelio y por tanto en  $t = 0$ .

$$\begin{aligned}
 \frac{a(1-\varepsilon^2)}{1+\varepsilon \cos \theta} &= a(1-\varepsilon \cos \psi) \Rightarrow 1+\varepsilon \cos \theta = \frac{1-\varepsilon^2}{1-\varepsilon \cos \psi} \\
 \cos \theta &= \frac{1-\varepsilon^2}{\varepsilon(1-\varepsilon \cos \psi)} - \frac{1}{\varepsilon} = \frac{1-\varepsilon^2 - (1-\varepsilon \cos \psi)}{\varepsilon(1-\varepsilon \cos \psi)} = \frac{\varepsilon(-\varepsilon + \cos \psi)}{\varepsilon(1-\varepsilon \cos \psi)}
 \end{aligned}$$

de modo que esta ecuación se puede escribir como

$$\cos \theta = \frac{\cos \psi - \varepsilon}{1 - \varepsilon \cos \psi} \tag{10.108}$$

tenemos a partir de (10.108) que

$$\begin{aligned} 1 - \cos \theta &= 1 - \frac{\cos \psi - \varepsilon}{1 - \varepsilon \cos \psi} = \frac{1 - \varepsilon \cos \psi - \cos \psi + \varepsilon}{1 - \varepsilon \cos \psi} = \frac{(1 - \cos \psi) + \varepsilon(1 - \cos \psi)}{1 - \varepsilon \cos \psi} \\ 1 - \cos \theta &= \frac{(1 - \cos \psi)(1 + \varepsilon)}{1 - \varepsilon \cos \psi} \end{aligned}$$

similarmente

$$\begin{aligned} 1 + \cos \theta &= 1 + \frac{\cos \psi - \varepsilon}{1 - \varepsilon \cos \psi} = \frac{1 - \varepsilon \cos \psi + \cos \psi - \varepsilon}{1 - \varepsilon \cos \psi} = \frac{(1 + \cos \psi) - \varepsilon(1 + \cos \psi)}{1 - \varepsilon \cos \psi} \\ 1 + \cos \theta &= \frac{(1 + \cos \psi)(1 - \varepsilon)}{1 - \varepsilon \cos \psi} \end{aligned}$$

por tanto

$$\frac{1 - \cos \theta}{1 + \cos \theta} = \frac{(1 - \cos \psi)(1 + \varepsilon)}{(1 + \cos \psi)(1 - \varepsilon)}$$

y usando la identidad

$$\frac{1 - \cos \theta}{1 + \cos \theta} = \tan^2 \frac{\theta}{2}$$

tenemos que

$$\tan^2 \frac{\theta}{2} = \frac{(1 + \varepsilon)}{(1 - \varepsilon)} \tan^2 \frac{\psi}{2}$$

teniendo en cuenta que  $\tan \theta/2$ , tiene el mismo signo que  $\tan \psi/2$  en todo el rango de ambos ángulos, nos queda finalmente

$$\tan \frac{\theta}{2} = \sqrt{\frac{1 + \varepsilon}{1 - \varepsilon}} \tan \frac{\psi}{2} \quad (10.109)$$

la ecuación (10.109) nos da  $\theta$  como función de  $\psi$ . Para encontrar  $\theta$  en función del tiempo, es necesario invertir la ecuación de Kepler (10.107). La solución de esta ecuación ha sido tema de amplio estudio (ver Ref. [1] segunda edición).

### 10.13.3. Tercera ley de Kepler

Elevando al cuadrado la Ec. (10.106), obtenemos

$$\tau^2 = \frac{4\pi^2 m}{k} a^3 \quad (10.110)$$

este resultado también se puede obtener a través de las propiedades de la elipse. Ya hemos visto que la conservación del momento angular conduce a que la velocidad aerolar es constante y viene dada por la Ec. (10.15)

$$\frac{dA}{dt} = \frac{1}{2} r^2 \dot{\theta} = \frac{l}{2m}$$

si integramos sobre un periodo completo, se obtiene el área total, la cual para una elipse viene dada por  $\pi ab$ , siendo  $b$  el semieje menor.

$$\int_0^\tau \frac{dA}{dt} dt = A_T = \frac{l\tau}{2m} = \pi ab \quad (10.111)$$

el semieje menor viene dado por

$$b = a\sqrt{1 - \varepsilon^2} \quad (10.112)$$

y combinando (10.112) con la relación (10.95) nos permite escribir el semieje menor en la forma

$$b = l\sqrt{\frac{a}{mk}} \quad (10.113)$$

y sustituyendo (10.113) en (10.111) resulta

$$\frac{l\tau}{2m} = \pi a l \sqrt{\frac{a}{mk}} \Rightarrow \tau = 2\pi a^{3/2} \sqrt{\frac{m}{k}}$$

en consistencia con (10.110). Estas expresiones coinciden con la tercera ley de Kepler, la cual se puede enunciar de la siguiente manera: los cuadrados de los periodos de los distintos planetas son proporcionales a los cubos de sus ejes (o semi-ejes) mayores.

Kepler enunció además que la constante de proporcionalidad es la misma para todos los planetas. No obstante, esta última afirmación es solo aproximadamente cierta. Para verlo, recordemos que el movimiento de un planeta alrededor del sol, es un problema de dos cuerpos y la masa que aparece aquí equivale a la masa reducida del sistema

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

por convención tomaremos a  $m_1$  como la masa de un planeta y a  $m_2$  como la masa del sol. La fuerza gravitatoria de atracción se escribe

$$f = -\frac{Gm_1 m_2}{r^2} = -\frac{k}{r^2} ; \quad k \equiv Gm_1 m_2$$

y la Ec. (10.110) queda de la forma

$$\begin{aligned} \tau^2 &= \frac{4\pi^2 \mu}{k} a^3 = \frac{4\pi^2 \frac{m_1 m_2}{m_1 + m_2}}{Gm_1 m_2} a^3 \\ \tau^2 &= \frac{4\pi^2 a^3}{G(m_1 + m_2)} \cong \frac{4\pi^2 a^3}{Gm_2} \end{aligned} \quad (10.114)$$

nótese que la constante de proporcionalidad es diferente para cada planeta, pero si despreciamos la masa  $m_1$  del planeta con respecto a la masa  $m_2$  del sol, la constante será la misma para todos los planetas. En las órbitas del átomo de Bohr esta constante de proporcionalidad es la misma, ya que todos los electrones poseen la misma masa y carga (es decir igual valor de la masa reducida y de  $k$ ).

## 10.14. Vector de Laplace-Runge-Lenz

A continuación veremos que en el problema de Kepler existe un vector conservado además del momento angular. La segunda ley de Newton para fuerzas centrales se puede escribir como

$$\dot{\mathbf{p}} = f(r) \mathbf{u}_r = f(r) \frac{\mathbf{r}}{r} \quad (10.115)$$

teniendo en cuenta (10.115), desarrollemos la cantidad

$$\begin{aligned} \dot{\mathbf{p}} \times \mathbf{L} &= \left[ f(r) \frac{\mathbf{r}}{r} \right] \times (\mathbf{r} \times \mathbf{p}) = \frac{f(r)}{r} \mathbf{r} \times (\mathbf{r} \times m\dot{\mathbf{r}}) = \frac{mf(r)}{r} [\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}})] \\ \dot{\mathbf{p}} \times \mathbf{L} &= \frac{mf(r)}{r} [\mathbf{r} (\mathbf{r} \cdot \dot{\mathbf{r}}) - r^2 \dot{\mathbf{r}}] \end{aligned} \quad (10.116)$$

donde hemos usado la identidad

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$$

podemos simplificar la ecuación (10.116) teniendo en cuenta la identidad

$$\mathbf{r} \cdot \dot{\mathbf{r}} = \frac{1}{2} \frac{d}{dt} (\mathbf{r} \cdot \mathbf{r}) = \frac{1}{2} \frac{d}{dt} (rr) = r\dot{r}$$

o visto de otra forma

$$\mathbf{r} \cdot \dot{\mathbf{r}} = r\mathbf{u}_r \cdot \dot{\mathbf{r}} = r\mathbf{u}_r \cdot (v_r \mathbf{u}_r + v_\theta \mathbf{u}_\theta) = rv_r = r\dot{r}$$



de modo que

$$\dot{\mathbf{p}} \times \mathbf{L} = \frac{mf(r)}{r} [\mathbf{r}(r\dot{r}) - r^2\dot{\mathbf{r}}] = mf(r) [\mathbf{r}\dot{r} - r\dot{\mathbf{r}}] = mf(r) r^2 \left[ \frac{\mathbf{r}\dot{r}}{r^2} - \frac{\dot{\mathbf{r}}}{r} \right]$$

Teniendo en cuenta esta identidad y el hecho de que  $\mathbf{L}$  es constante, se tiene que

$$\begin{aligned} \frac{d}{dt}(\mathbf{p} \times \mathbf{L}) &= \dot{\mathbf{p}} \times \mathbf{L} = -mf(r) r^2 \left[ \frac{\dot{\mathbf{r}}}{r} - \frac{\mathbf{r}\dot{r}}{r^2} \right] = -mf(r) r^2 \frac{d}{dt} \left( \frac{\mathbf{r}}{r} \right) \\ \frac{d}{dt}(\mathbf{p} \times \mathbf{L}) &= -mf(r) r^2 \frac{d\mathbf{u}_r}{dt} \end{aligned} \quad (10.117)$$

hasta aquí podemos llegar para una expresión general de la fuerza central. Ahora tomando una fuerza atractiva proporcional al inverso al cuadrado de la distancia  $f(r) = -k/r^2$ , la integración resulta inmediata, ya que la Ec. (10.117) se convierte en

$$\frac{d}{dt}(\mathbf{p} \times \mathbf{L}) = m \frac{k}{r^2} r^2 \frac{d\mathbf{u}_r}{dt} = \frac{d}{dt}(mk\mathbf{u}_r) \Rightarrow \frac{d}{dt}(\mathbf{p} \times \mathbf{L} - mk\mathbf{u}_r) = 0$$

que nos dice que en el problema de Kepler hay un vector conservado  $\mathbf{A}$  definido por

$$\mathbf{A} \equiv \mathbf{p} \times \mathbf{L} - mk\mathbf{u}_r \quad (10.118)$$

conocido como vector de Laplace-Runge-Lenz (L-R-L). Dado que  $\mathbf{L} = l\mathbf{u}_z$ , la definición de  $\mathbf{A}$  nos muestra que

$$\mathbf{A} \cdot \mathbf{L} = 0 \quad (10.119)$$

esta ortogonalidad implica que  $\mathbf{A}$  es un vector fijo que yace en el plano de movimiento. Si llamamos  $\theta$  el ángulo que hace  $\mathbf{r}$  con el vector fijo  $\mathbf{A}$ , el producto escalar de estos dos vectores será

$$\mathbf{A} \cdot \mathbf{r} = Ar \cos \theta = (\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} - mkr \quad (10.120)$$

permutando términos en el producto mixto

$$\mathbf{r} \cdot (\mathbf{p} \times \mathbf{L}) = \mathbf{L} \cdot (\mathbf{r} \times \mathbf{p}) = l^2$$

con lo que la Ec. (10.120) queda en la forma

$$Ar \cos \theta = l^2 - mkr \Rightarrow (A \cos \theta + mk)r = l^2$$

o alternativamente

$$\frac{1}{r} = \frac{mk}{l^2} \left[ 1 + \frac{A}{mk} \cos \theta \right] \quad (10.121)$$

vemos entonces que el vector de L-R-L proporciona otra manera de llegar a la ecuación de la órbita en el problema de Kepler. Comparando (10.121) con la ecuación de la órbita en la forma (10.85), se ve que el módulo de  $\mathbf{A}$  viene dado por

$$A = mk\varepsilon \quad (10.122)$$

de la Ec. (10.121) es claro que  $\theta = 0$  corresponde a la dirección del periápside  $\mathbf{r}(\theta = 0) = \mathbf{r}_{\text{mín}} \equiv \mathbf{r}_m$ . Recordando además que  $\theta$  por definición es el ángulo entre  $\mathbf{A}$  y  $\mathbf{r}$  (según la Ec. 10.120), vemos que  $\theta = 0$  indica que  $\mathbf{A}$  es paralelo a  $\mathbf{r}_m$ . Por tanto,  $\mathbf{A}$  tiene la dirección del radio vector que corresponde al periapsis o perihelio de la órbita.

Hemos identificado en consecuencia dos vectores y un escalar que se conservan ( $\mathbf{L}$ ,  $\mathbf{A}$ ,  $E$ ) para un total de siete cantidades conservadas. Por otro lado, el sistema requiere de seis constantes de movimiento que corresponden por ejemplo a las tres componentes de la posición inicial y a las tres componentes de la velocidad inicial de la partícula. Nótese por otro lado, que todas las constantes que aparecen en el conjunto ( $\mathbf{L}$ ,  $\mathbf{A}$ ,  $E$ ) son funciones algebraicas de  $\mathbf{r}$  y  $\mathbf{p}$  que describen la órbita en su conjunto (orientación en el espacio, excentricidad

etc.), ninguna de las siete cantidades se refiere a la localización de la partícula en el instante inicial. Como una de las constantes de movimiento debe referirse a esta información, por ejemplo en forma del tiempo de paso  $T$  de la partícula por el perihelio, solo podrá haber cinco cantidades de movimiento independientes en el conjunto  $(\mathbf{L}, \mathbf{A}, E)$  que describan el tamaño, forma y orientación de la órbita. Es necesario entonces que existan dos relaciones entre las cantidades  $\mathbf{L}$ ,  $\mathbf{A}$  y  $E$ . La primera viene dada por la relación de ortogonalidad entre  $\mathbf{L}$  y  $\mathbf{A}$ , Ec. (10.119). La otra se deduce de la Ec. (10.122) cuando se escribe la excentricidad en términos de  $E$  y  $l$ , Ec. (10.90)

$$A = mk\sqrt{1 + \frac{2El^2}{mk^2}} \Rightarrow A^2 = m^2k^2 + 2mEl^2 \quad (10.123)$$

confirmando que solo hay cinco cantidades independientes entre las siete descritas por el conjunto  $(\mathbf{L}, \mathbf{A}, E)$ .

Dado que el momento angular y la energía proveen cuatro cantidades independientes, el vector de L-R-L nos da una más. Es entonces natural preguntarse si para una fuerza central arbitraria existe una cantidad conservada análoga al vector de L-R-L que junto con  $\mathbf{L}$  y  $E$  nos sirvan para definir la órbita. Parece ser que sí se pueden construir cantidades de este tipo pero en general no son funciones muy simples del movimiento (D. M. Fradkin, *Progress of Theoretical Physics* **37**, 798, mayo 1967). Esto se debe a que las constantes de movimiento relacionadas con la órbita definen la forma funcional  $r(\theta)$ , pero hemos visto que en general las fuerzas centrales conducen a órbitas no cerradas, como se vé en el teorema de Bertrand. Es una característica general de las órbitas no cerradas el hecho de que la curva llegará a pasar por un punto  $(r, \theta)$  arbitrario que esté entre las cotas de los puntos de retorno de  $r$ . Podemos verlo heurísticamente teniendo en cuenta que si la órbita no es cerrada, al variar  $\theta$  a lo largo de todo un ciclo la partícula no deberá recorrer de nuevo sus pasos sobre ninguna órbita anterior. En consecuencia la órbita es tal que  $r$  es función multiforme de  $\theta$  (módulo  $2\pi$ ), en realidad es una función infinitiforme de  $\theta$ . En consecuencia, la cantidad conservada adicional a  $\mathbf{L}$  y a  $E$  que define la órbita, deberá contener una función infinitiforme del movimiento de la partícula. Solo en el caso de órbitas cerradas o más generalmente cuando el movimiento sea degenerado, como en el problema de Kepler, se puede esperar que la cantidad conservada adicional sea una función sencilla de  $\mathbf{r}$  y  $\mathbf{p}$  tal como el vector de L-R-L.

Es interesante ver el aporte geométrico de cada constante de movimiento, dos grados de libertad de  $\mathbf{L}$  determinan el plano de movimiento, los valores de  $E$  y  $l$  determinan el tamaño y forma de la elipse<sup>23</sup>, finalmente el vector de L-R-L determina la orientación de la elipse en el plano de movimiento. Nótese que una elipse solo requiere un parámetro angular para ser orientada dentro de un plano, lo cual coincide con el hecho de que el vector L-R-L solo contribuye con un parámetro<sup>24</sup>.

A la luz del teorema de Bertrand, es de esperarse que en el caso de una ley de fuerza central tipo Hooke, encontremos una cantidad conservada semejante al vector de L-R-L, puesto que en este caso las órbitas también son degeneradas. Tal cantidad existe pero la manera más natural de definirla es con un tensor de segundo rango. De esta forma la existencia de una constante de movimiento adicional a  $E$ ,  $\mathbf{L}$  y relacionada con la órbita que sea función algebraica simple del movimiento, es suficiente para indicar que el movimiento es degenerado y que las órbitas acotadas son cerradas.

## 10.15. Parametrización de las órbitas keplerianas en el espacio

Aunque matemáticamente es posible determinar la órbita en el espacio usando cantidades como  $E$ ,  $\mathbf{L}$  y  $\mathbf{A}$ . Es más conveniente parametrizar la órbita con cantidades más geométricas. Hemos visto que en el espacio tridimensional que es un problema de tres grados de libertad, se requieren seis condiciones iniciales (por ejemplo

<sup>23</sup>El tamaño y forma de la elipse se determinan geoméricamente con la excentricidad  $\varepsilon$  y el semi-eje mayor  $a$ . Las Ecs. (10.90, 10.94) nos dicen que  $E$  y  $l$  nos determinan las cantidades  $a$  y  $\varepsilon$ .

<sup>24</sup>Estrictamente un plano requiere para su determinación un vector normal a él y un punto por donde pasa, igualmente la ubicación de la elipse dentro del plano requiere además de su forma, tamaño y orientación, conocer dos puntos por donde pasa o un punto por donde pasa con su velocidad. Aquí son indispensables la posición y velocidad inicial. Nótese sin embargo, que cambiar la posición inicial con las mismas cinco constantes de movimiento, implica que la elipse sufre una traslación paralela con lo cual es obvio que las características del movimiento son idénticas. No obstante, el análisis anterior muestra que la determinación completa de la posición de la trayectoria en el espacio requiere de un grado de libertad adicional asociado a la posición inicial.

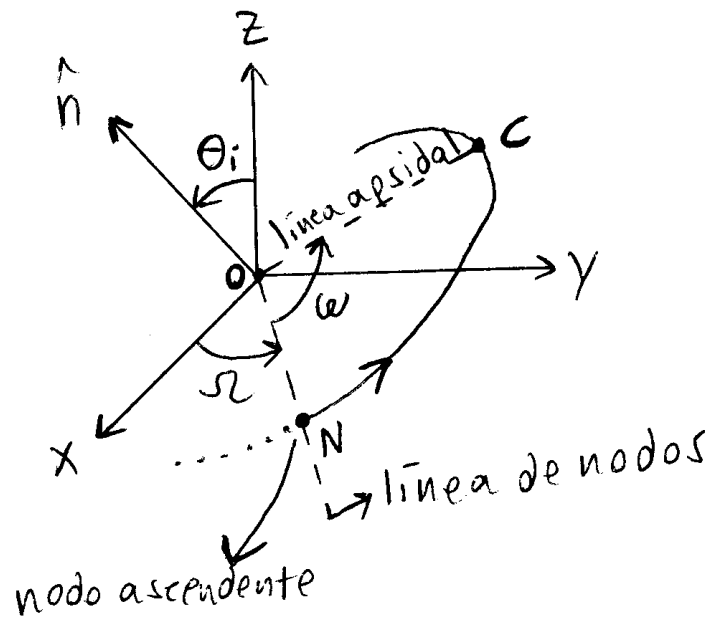


Figura 10.10: Parametrización de la órbita elíptica en el espacio para el problema de Kepler.

$q_i, \dot{q}_i$  para las tres coordenadas generalizadas), para resolver completamente la dinámica. Ya hemos observado que para el problema de fuerzas centrales es más conveniente tomar las seis constantes de modo que cinco de ellas sean funciones de las coordenadas y momentos que describen la órbita en el espacio y solo la última se refiere a la posición de la partícula en la órbita en un instante dado (ver sección 10.14). Para el caso de la órbita elíptica, las cinco constantes que determinan la órbita se pueden tomar de la siguiente forma: dos constantes típicamente angulares determinan la orientación del plano de movimiento, un parámetro para la escala de la elipse (por ejemplo el semieje mayor), otro para la forma de la elipse (digamos la excentricidad o el eje menor), y finalmente un parámetro que especifique la orientación de la elipse dentro del plano de la órbita.

Veamos de forma específica la manera de determinar la órbita con cinco parámetros de esta naturaleza y un sexto parámetro que me introduce la dependencia temporal. La Fig. 10.10, muestra un sistema  $xyz$  donde el origen se eligió en el centro de fuerzas, de modo que el plano de movimiento pasa por dicho origen. En el proceso de determinación de la órbita seguiremos los siguientes pasos: (a) determinación del plano de la órbita, (b) determinación de la escala y forma de la elipse, (c) determinación de la orientación de la elipse en el plano de movimiento y el sentido en que la partícula recorre la elipse. Estos pasos nos determinan completamente la órbita, de modo que para determinar la dinámica completa solo se necesita un paso adicional (d) determinación de la dependencia de la posición de la partícula en la órbita con el parámetro tiempo.

Para realizar el paso (a), observemos que el vector unitario  $\mathbf{n}$  especifica el plano de movimiento, podríamos en consecuencia pensar en utilizar los ángulos que definen a este vector unitario como los parámetros que definen la orientación de este plano. Resulta no obstante, más práctico definir la dirección de este plano en la forma siguiente: La intersección del plano  $xy$  con el plano orbital se llama **línea de nodos**. Si determinamos la dirección de la línea de nodos, lo que nos falta para determinar el plano orbital es el ángulo diedro entre el plano  $xy$  y el plano de movimiento. Para determinar la dirección de la línea de nodos, notemos que en dicha línea hay dos puntos que corresponden a la intersección de la órbita elíptica con el plano  $xy$ . El punto en el cual la partícula pasa del hemisferio inferior o hemisferio sur ( $z < 0$ ) al hemisferio superior o hemisferio norte ( $z > 0$ ) se denomina **nodo ascendente**. La parte punteada de la órbita en la Fig. 10.10 corresponde a la porción que está en el hemisferio sur. La recta  $ON$  es un segmento de la línea de nodos que va del origen al nodo ascendente. La dirección de  $ON$  en el plano  $xy$  (y por tanto la dirección de la línea de nodos) puede determinarse con el ángulo que este segmento hace con el eje  $x$  conocido como **longitud del nodo ascendente**

y denotado por  $\Omega$ . Una vez determinada la dirección de la línea nodal a través de  $\Omega$ , vemos fácilmente que el ángulo diedro entre el plano de movimiento y el plano  $xy$  corresponde justamente a  $\theta_i$  i.e. el ángulo entre  $\mathbf{n}$  y el eje  $z$  que usualmente se denomina **inclinación de la órbita**.

El paso (b) se realiza de manera inmediata si conocemos dos de los siguientes tres parámetros: semieje menor, semieje mayor y excentricidad. Usualmente se toman los dos últimos, de modo que típicamente obtenemos la escala de la elipse con el semieje mayor y la forma de ésta con la excentricidad.

Ahora para el paso (c) determinamos la orientación de la elipse en el plano de movimiento. Es natural tomar como referente la orientación del periápside, la cual se puede obtener a través del ángulo  $NOC$  siendo  $C$  el punto de la órbita que corresponde al periápside, este ángulo se denota por  $\omega$  y se denomina **argumento del perihelio** (incluso para órbitas no relacionadas con el sol). Este es el ángulo entre la línea nodal (con su sentido determinado por el nodo ascendente), y la línea del periápside. Nótese que la determinación del nodo ascendente ya nos determina el sentido en el cual la partícula recorre la elipse.

Finalmente, para el paso (d) de determinar la dependencia de la posición de la partícula con el parámetro tiempo, podemos determinar por ejemplo la posición angular inicial de la partícula, digamos con respecto al periápside. Sin embargo, un parámetro más utilizado es el tiempo  $T$  para el cual la partícula pasa por el periápside. Es decir el tiempo que le toma a la partícula en ir desde su posición inicial en  $t = 0$  hasta el periápside.

En síntesis, la dinámica completa de movimiento se determina con los siguientes parámetros

$$\theta_i, \Omega, a, \varepsilon, \omega, T \quad (10.124)$$

los dos primeros parámetros definen la orientación del plano por medio de la orientación de la línea de nodos ( $\Omega$ ), y el ángulo diedro  $\theta_i$  entre el plano orbital y el plano  $xy$ . El semieje mayor  $a$  define la escala de la elipse, y la excentricidad  $\varepsilon$  define la forma de ésta. El ángulo  $\omega$  determina la dirección de la elipse en el plano de movimiento tomando como referente la línea del periápside. El sentido del movimiento lo determina la localización del nodo ascendente (que a su vez se usa para definir  $\omega$ ). Finalmente, el parámetro  $T$  define la dependencia de la posición de la partícula en la órbita en función del tiempo.

## 10.16. Problema de Kepler en variables acción-ángulo revisado (opcional)

En la sección 9.12 vimos la reducción a cuadraturas del problema de Kepler con el formalismo de variables acción-ángulo, comenzando con un problema tridimensional en coordenadas esféricas. Para entonces solo nos interesaba indicar la solución formal. En esta sección nos ocuparemos de mostrar algunos detalles adicionales que se pueden observar para el problema de Kepler en variables acción-ángulo, aprovechando además el conocimiento adicional que hemos obtenido en este capítulo. En particular usaremos desde el principio el hecho de que el movimiento corresponde a una órbita bidimensional cerrada, y que la frecuencia es completamente degenerada.

En primer lugar, cuando utilizamos coordenadas esféricas o polares planas como coordenadas generalizadas, la transformación de coordenadas cartesianas a las generalizadas no depende explícitamente del tiempo, por consiguiente la energía cinética en estas coordenadas solo depende cuadráticamente de  $\dot{q}_i$  (ver sección 2.3.1) i.e.  $T = T_2$ . Por otro lado, para una partícula sometida a una fuerza central, el Lagrangiano y el Hamiltoniano en coordenadas esféricas o en coordenadas polares planas poseen las siguientes propiedades estructurales

$$\begin{aligned} L &= T - V = L_2 + L_0 \quad ; \quad L_2 = T \quad , \quad L_0 = -V \\ H &\equiv p_i \dot{q}_i - L = T + V \quad \Rightarrow \quad p_i \dot{q}_i - (T - V) = T + V \\ &\Rightarrow p_i \dot{q}_i = 2T = 2L_2 \end{aligned}$$

por tanto, la energía cinética se puede expresar en coordenadas esféricas  $(r, \theta, \phi)$  o en coordenadas polares

planas  $(r, \psi)$ , si tenemos en cuenta la naturaleza bidimensional del movimiento<sup>25</sup>

$$2T = p_r \dot{r} + p_\theta \dot{\theta} + p_\phi \dot{\phi} = p_r \dot{r} + p_\psi \dot{\psi} \quad (10.125)$$

con lo cual se obtiene

$$p_\theta \dot{\theta} + p_\phi \dot{\phi} = p_\psi \dot{\psi} \Rightarrow p_\theta d\theta = p_\psi d\psi - p_\phi d\phi$$

esta expresión nos permite calcular la integral de  $J_\theta$ , teniendo en cuenta además que  $p_\psi = l$  es el módulo del momento angular, y que  $p_\phi$  es constante ya que  $\phi$  es cíclica. La definición de  $J_\theta$  queda

$$J_\theta \equiv \oint p_\theta d\theta = \oint p_\psi d\psi - \oint p_\phi d\phi = l \oint d\psi - p_\phi \oint d\phi$$

y dado que las frecuencias de  $\theta$  y  $\phi$  son iguales,  $\phi$  y  $\psi$  varían en  $2\pi$  cuando  $\theta$  realiza un ciclo completo de libración, de modo que las integrales se simplifican en la forma

$$J_\theta = 2\pi(l - p_\phi) = 2\pi(\alpha_\theta - \alpha_\phi) \quad (10.126)$$

donde hemos tenido en cuenta que  $\alpha_\phi$  coincide con el momento conjugado a  $\phi$ , en tanto que  $\alpha_\theta$  corresponde al módulo del momento angular, según las Ecs. (9.76, 9.79)<sup>26</sup>. La Ec. (10.126) concuerda con la Ec. (9.166). La variable  $J_\phi$  es inmediata dado el carácter cíclico de  $\phi$

$$J_\phi = 2\pi\alpha_\phi = 2\pi p_\phi \quad (10.127)$$

### 10.16.1. Determinación del Hamiltoniano en términos de variables de acción

Es posible extraer información de manera más directa con las variables acción-ángulo, utilizando la completa degeneración entre las frecuencias en  $r$ ,  $\theta$  y  $\phi$ . Dicha degeneración nos permite utilizar los resultados de la sección 9.12.1, en la cual obtenemos la transformación canónica

$$(w_\theta, w_\phi, w_r, J_\theta, J_\phi, J_r) \rightarrow (w'_1, w'_2, w'_3, J'_1, J'_2, J'_3)$$

la cual es independiente de la forma explícita de las variables acción-ángulo originales. La completa degeneración de las frecuencias de  $r, \theta, \phi$  implica que el Hamiltoniano es función de un solo  $J'_i$  como se vé en las Ecs. (9.181, 9.182), aquí omitiremos la notación primada de estas ecuaciones. La dependencia de  $H$  con este  $J$  único se puede obtener de (10.125). Para verlo, aplicamos el teorema del virial al problema de Kepler Ec. (10.45)

$$\overline{V} = -2\overline{T} \quad (10.128)$$

de modo que el Hamiltoniano se escribe como

$$H = E = \overline{T} + \overline{V} = -\overline{T} \quad (10.129)$$

integrando la Ec. (10.125) en el tiempo sobre un periodo (degenerado y denotado por  $\tau_3$ ) se tiene que

$$\begin{aligned} 2\frac{\tau_3}{\tau_3} \int_0^{\tau_3} T dt &= \oint p_r \dot{r} dt + \oint p_\theta \dot{\theta} dt + \oint p_\phi \dot{\phi} dt \\ 2\tau_3 \overline{T} &= \oint p_r dr + \oint p_\theta d\theta + \oint p_\phi d\phi \end{aligned}$$

utilizando la definición de  $J_i$  y la Ec. (9.181), obtenemos

$$\frac{2\overline{T}}{v_3} = J_r + J_\theta + J_\phi = J_3 \quad (10.130)$$

<sup>25</sup>Nótese que en tanto el movimiento esté confinado a un plano y el origen pase por dicho plano, la coordenada  $r$  de distancia al origen es igual en coordenadas esféricas, que en las coordenadas polares asociadas a dicho plano.

<sup>26</sup>Recordemos que  $\alpha_\theta$  y  $\alpha_\phi$  surgen como constantes de integración (ver Sec. 9.6.2, Pág. 184).

donde  $v_3$  es la frecuencia degenerada del movimiento. Combinando las Ecs. (10.129, 10.130) y teniendo en cuenta (9.128) resulta

$$-2H = J_3 v_3 = J_3 \frac{dH}{dJ_3} \Rightarrow -\frac{2}{J_3} = \frac{1}{H} \frac{dH}{dJ_3}$$

esta ecuación diferencial tiene solución inmediata

$$H = \frac{D}{J_3^2} \quad (10.131)$$

donde  $D$  es una constante que no puede contener ningún  $J$  y por tanto solo depende de  $m$  y  $k$ . Podemos calcular  $D$  usando una trayectoria circular de radio  $r_0$  para la cual  $J_r = 0$  (ya que  $p_r = m\dot{r} = 0$ ), y por tanto

$$\begin{aligned} J_3 &= J_r + J_\theta + J_\phi = J_\theta + J_\phi = 2\pi(l - \alpha_\phi) + 2\pi\alpha_\phi \\ J_3 &= 2\pi l \end{aligned} \quad (10.132)$$

el teorema del virial Ec. (10.128) nos dice que

$$H = \bar{T} + \bar{V} = -\frac{\bar{V}}{2} + \bar{V} = \frac{\bar{V}}{2}$$

y dado que en el movimiento circular el potencial es constante, tenemos

$$H = E = -\frac{k}{2r_0} \quad (10.133)$$

Para potenciales keplerianos la condición de circularidad descrita por la Ec. (10.55) se puede escribir en la forma

$$\frac{k}{r_0^2} = \frac{l^2}{mr_0^3} = \frac{J_3^2}{4\pi^2 mr_0^3} \quad (10.134)$$

donde hemos tenido en cuenta la Ec. (10.132). Despejando  $r_0$  en esta ecuación tenemos

$$r_0 = \frac{J_3^2}{4\pi^2 mk} \quad (10.135)$$

sustituyendo (10.135) en (10.133) obtenemos  $H$  en función de  $J_3$

$$\begin{aligned} H &= -\frac{k}{2r_0} = -\frac{k}{2} \left( \frac{4\pi^2 mk}{J_3^2} \right) \\ H &= E = -\frac{2\pi^2 mk^2}{J_3^2} \end{aligned} \quad (10.136)$$

aunque esta relación la hemos encontrado para órbitas circulares, la Ec. (10.131) nos dice que también debe ser válida para cualquier órbita cerrada kepleriana, ya que el Hamiltoniano es una cantidad dinámica (no cinemática), y por tanto es independiente de las condiciones iniciales. Esta relación coincide además con (9.182). Vemos que hemos evaluado  $H(J)$  sin calcular las integrales circuitales, usando la degeneración completa de las frecuencias de movimiento. Esto significa una simplificación importante, ya que la realización de las integrales circuitales es considerablemente extensa como vimos en la sección 9.12.

### 10.16.2. Relación entre variables acción-ángulo y variables orbitales

El tratamiento del problema de Kepler en variables acción-ángulo para degeneración completa, nos permitirá escribir las cinco constantes que determinan la órbita en una parametrización diferente a la establecida en la sección 10.15, Ec. (10.124). Para nuestro caso, tres constantes están asociadas a las variables de acción

$J_1, J_2, J_3$ , en tanto que las otras dos constantes estarán asociadas a  $w_1$  y  $w_2$  ya que sus frecuencias correspondientes son nulas<sup>27</sup>. En consecuencia, debe ser posible relacionar las cinco primeras constantes en (10.124) con las cinco constantes  $J_1, J_2, J_3, w_1, w_2$  generadas por las variables acción-ángulo (la variable temporal se deja intacta). Combinando las Ecs. (9.181, 10.126, 10.127) se observa que<sup>28</sup>

$$\begin{aligned} J_2 &= J_\theta + J_\phi = 2\pi(\alpha_\theta - \alpha_\phi) + 2\pi\alpha_\phi \\ J_2 &= 2\pi\alpha_\theta \equiv 2\pi l \end{aligned} \quad (10.137)$$

y usando la Ec. (9.161) se tiene que

$$\begin{aligned} \cos \theta_i &= \frac{\alpha_\phi}{\alpha_\theta} = \frac{2\pi\alpha_\phi}{2\pi\alpha_\theta} = \frac{J_\phi}{J_2} \\ \cos \theta_i &= \frac{J_1}{J_2} \end{aligned} \quad (10.138)$$

donde hemos tenido en cuenta las Ecs. (9.181, 10.137). Ahora bien, puesto que el semieje mayor  $a$  solo es función de la energía según (10.94), la Ec. (10.136) nos indica que  $a$  solo es función de  $J_3$

$$a = -\frac{k}{2E} = \frac{J_3^2}{4\pi^2 mk} \quad (10.139)$$

la ecuación (10.95) para la excentricidad se puede escribir en función de  $J_2$  y  $J_3$  usando (10.137) y (10.139)

$$\begin{aligned} \varepsilon &= \sqrt{1 - \frac{l^2}{mka}} = \sqrt{1 - \frac{J_2^2}{4\pi^2 mka}} = \sqrt{1 - \frac{J_2^2}{4\pi^2 mk} \left( \frac{4\pi^2 mk}{J_3^2} \right)} \\ \varepsilon &= \sqrt{1 - \left( \frac{J_2}{J_3} \right)^2} \end{aligned} \quad (10.140)$$

Las Ecs. (10.138, 10.139, 10.140) conectan las variables orbitales  $(\theta_i, a, \varepsilon)$  con las variables de acción  $(J_1, J_2, J_3)$ .

Falta la identificación de  $w_1$  y  $w_2$  con los elementos clásicos de órbita (10.124). Es claro que  $w_1$  y  $w_2$  deben contener a  $\Omega$  y  $\omega$  las cuales no están contenidas en los  $J$ 's. Veremos que para una elección adecuada de las constantes de integración aditivas  $w_1$  y  $w_2$  serán proporcionales a  $\Omega$  y  $\omega$ .

Veamos el caso de  $w_1$ , la ecuación (9.154) define a  $w_1$  en términos de la función generatriz  $W$

$$w_1 = \frac{\partial W}{\partial J_1} \quad (10.141)$$

podemos utilizar la Ec. (9.16) para escribir

$$W = \int p_\phi d\phi + \int p_\theta d\theta + \int p_r dr \quad (10.142)$$

debemos recordar que en (9.16) las integrales eran indefinidas. Para poder usar (10.141), debemos escribir  $W$  en términos de  $(J_1, J_2, J_3)$ . De acuerdo con (10.142), esto equivale a encontrar  $p_\phi, p_\theta, p_r$  en términos de los  $J_i$ . La Ec. (9.168), Pág. 209 nos da una expresión para  $p_r$

$$p_r = \frac{\partial W}{\partial r} = \sqrt{2mE + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2}} \quad (10.143)$$

<sup>27</sup>Esto se puede ver derivando temporalmente las ecuaciones (9.180), y teniendo en cuenta la degeneración en las frecuencias de  $r, \theta, \phi$ .

<sup>28</sup>Debe tenerse en cuenta que la relación (10.132) solo es válida para movimiento circular en tanto que la relación (10.137) es válida para todas las trayectorias elípticas (en particular las circulares). En general,  $J_2$  y  $J_3$  son momentos canónicos independientes. Alternativamente, se puede ver que para el movimiento circular  $J_r = 0$ , con lo cual  $J_2 = J_\theta + J_\phi = J_\theta + J_\phi + J_r = J_3$ .

podemos obtener  $p_r$  en función de los  $J_i$ , sustituyendo (10.136, 9.181) en (10.143)

$$p_r(J_1, J_2, J_3) = \sqrt{-\frac{4\pi^2 m^2 k^2}{J_3^2} + \frac{2mk}{r} - \frac{J_2^2}{4\pi^2 r^2}} \quad (10.144)$$

por tanto  $p_r$  es independiente de  $J_1$  de manera que

$$w_1 = \frac{\partial W}{\partial J_1} = \frac{\partial}{\partial J_1} \int p_\phi d\phi + \frac{\partial}{\partial J_1} \int p_\theta d\theta \quad (10.145)$$

Usando (10.127) y (9.181) se tiene

$$p_\phi = \alpha_\phi = \frac{J_1}{2\pi} \quad (10.146)$$

finalmente, el momento  $p_\theta$  se puede obtener de la Ec. (9.71) Pág. 184

$$\left(\frac{\partial W_\theta}{\partial \theta}\right)^2 + \frac{\alpha_\phi^2}{\sin^2 \theta} = \alpha_\theta^2 \Rightarrow p_\theta^2 = \alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta} \quad (10.147)$$

para obtener  $p_\theta$  en términos de los  $J_i$ , sustituimos las Ecs. (10.137, 10.146) en (10.147)

$$\begin{aligned} p_\theta &= \pm \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2 \theta}} = \pm \sqrt{\left(\frac{J_2}{2\pi}\right)^2 - \frac{1}{\sin^2 \theta} \left(\frac{J_1}{2\pi}\right)^2} \\ p_\theta &= \pm \frac{1}{2\pi} \sqrt{J_2^2 - J_1^2 \csc^2 \theta} \end{aligned} \quad (10.148)$$

para relacionar  $w_1$  con el nodo ascendente debe tomarse el signo negativo de la raíz cuadrada, ya que cuando la partícula pasa por el nodo ascendente, la variable  $\theta$  está disminuyendo como se aprecia en la Fig. 10.10. En consecuencia,  $\dot{\theta} < 0$  al pasar por el nodo ascendente y por tanto  $p_\theta = mr^2 \dot{\theta} < 0$ . Obsérvese que cuando calculamos  $J_\theta$  no tuvimos que preocuparnos por la elección del signo, ya que al efectuar un ciclo completo los dos signos aparecen cada uno en un semiciclo. La variable angular  $w_1$  está dada por la Ec. (10.145)

$$w_1 = \int \frac{\partial p_\phi}{\partial J_1} d\phi + \int \frac{\partial p_\theta}{\partial J_1} d\theta = \int \frac{\partial}{\partial J_1} \left(\frac{J_1}{2\pi}\right) d\phi - \int \frac{\partial}{\partial J_1} \left(\frac{1}{2\pi} \sqrt{J_2^2 - J_1^2 \csc^2 \theta}\right) d\theta$$

donde hemos usado las Ecs. (10.146, 10.148). Recordando de nuevo que estas integrales son indefinidas, se obtiene

$$\begin{aligned} w_1 &= \frac{\phi}{2\pi} + \frac{J_1}{2\pi} \int \frac{d\theta}{\sin^2 \theta \sqrt{J_2^2 - J_1^2 \csc^2 \theta}} = \frac{\phi}{2\pi} + \frac{1}{2\pi} \frac{J_1}{J_2} \int \frac{d\theta}{\sin^2 \theta \sqrt{1 - \left(\frac{J_1}{J_2}\right)^2 \csc^2 \theta}} \\ w_1 &= \frac{\phi}{2\pi} + \frac{\cos \theta_i}{2\pi} \int \frac{d\theta}{\sin^2 \theta \sqrt{1 - \cos^2 \theta_i \csc^2 \theta}} \end{aligned}$$

donde hemos usado (10.138), tenemos entonces

$$\begin{aligned} 2\pi w_1 &= \phi + \cos \theta_i \int \frac{d\theta}{\sin^2 \theta \sqrt{1 - \cos^2 \theta_i \csc^2 \theta}} = \phi + \frac{\cos \theta_i}{\sin \theta_i} \int \frac{d\theta}{\frac{\sin^2 \theta}{\sin \theta_i} \sqrt{1 - \cos^2 \theta_i \csc^2 \theta}} \\ &= \phi + \cot \theta_i \int \frac{\csc^2 \theta d\theta}{\sqrt{\frac{1}{\sin^2 \theta_i} - \frac{\cos^2 \theta_i \csc^2 \theta}{\sin^2 \theta_i}}} \end{aligned}$$

desarrollando el radical tenemos

$$\begin{aligned} \frac{1}{\sin^2 \theta_i} - \frac{\cos^2 \theta_i \csc^2 \theta}{\sin^2 \theta_i} &= \frac{1 - \cos^2 \theta_i \csc^2 \theta}{\sin^2 \theta_i} = \frac{\sin^2 \theta_i + \cos^2 \theta_i - \cos^2 \theta_i \csc^2 \theta}{\sin^2 \theta_i} \\ &= 1 + \frac{\cos^2 \theta_i (1 - \csc^2 \theta)}{\sin^2 \theta_i} = 1 - \cot^2 \theta_i \cot^2 \theta \end{aligned}$$



con lo cual resulta

$$2\pi w_1 = \phi + \int \frac{\cot \theta_i \csc^2 \theta \, d\theta}{\sqrt{1 - \cot^2 \theta_i \cot^2 \theta}}$$

realizamos el cambio de variable

$$\sin u = \cot \theta_i \cot \theta \quad ; \quad \cos u \, du = -\cot \theta_i \csc^2 \theta \, d\theta \quad (10.149)$$

con lo cual obtenemos

$$2\pi w_1 = \phi - \int \frac{\cos u \, du}{\sqrt{1 - \sin^2 u}} = \phi - \int du$$

resultando finalmente

$$2\pi w_1 = \phi - u \quad (10.150)$$

La ecuación (10.149), nos da la relación entre  $u$  y el ángulo  $\theta$ . Es posible encontrar la interpretación de

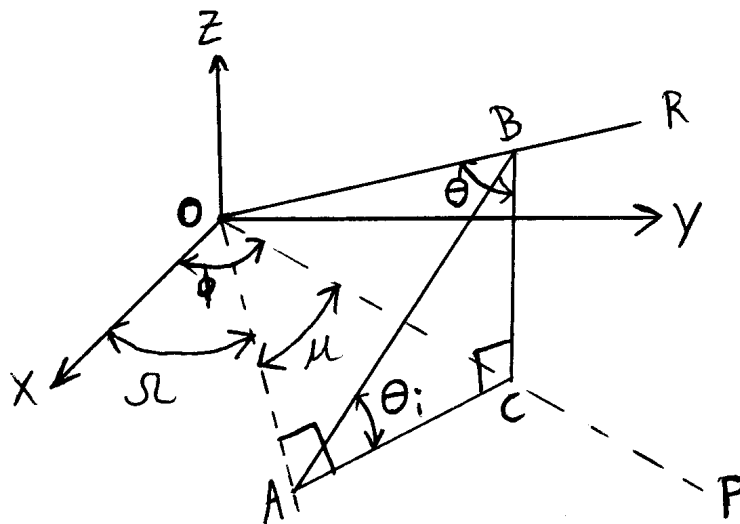


Figura 10.11: Variables angulares para el posicionamiento de la órbita que aparecen en el problema de Kepler, cuando se usa el formalismo de variables acción ángulo.

$u$  aplicando trigonometría esférica al triángulo esférico definido por la línea de nodos, el radio vector y la proyección del radio vector sobre el plano  $xy$ . Resulta más simple sin embargo, realizar algunas manipulaciones con la trigonometría plana. En la Fig. 10.11 la recta que contiene al segmento  $OA$  es la misma recta que contiene al segmento  $ON$  de la Fig. 10.10, y por tanto define la línea de nodos,  $OR$  es colineal con el radio vector de posición en un instante dado y  $OP$  es paralelo a la proyección del radio vector sobre el plano  $xy$ . Definiremos a  $\phi$  como el ángulo azimutal i.e. entre el eje  $x$  y la proyección  $OP$ . Vamos a demostrar que  $u$  es el ángulo entre la proyección del radio vector y la línea de nodos i.e. entre  $OP$  y  $OA$ . Para ello imaginemos un plano normal al plano  $xy$  y a la línea de nodos y que corte al radio vector a una distancia del origen  $O$  igual a la unidad. Este plano corta a las rectas  $OA$ ,  $OR$  y  $OP$  en los puntos  $A$ ,  $B$  y  $C$  respectivamente. Los puntos  $A$ ,  $B$  y  $C$  junto con el origen  $O$ , definen cuatro triángulos rectángulos  $OAC$ ,  $OCB$ ,  $OAB$  y  $ACB$  donde el vértice en ángulo recto está definido por la letra de la mitad<sup>29</sup>. De los triángulos  $OCB$  y  $ACB$  vemos que

$$\begin{aligned} \cos \theta &= \frac{BC}{OB} = BC \quad ; \quad \cot \theta_i = \frac{AC}{BC} \quad \Rightarrow \\ AC &= \cos \theta \cot \theta_i \end{aligned} \quad (10.151)$$

<sup>29</sup>Aunque los segmentos  $ON$  y  $OA$  de las figuras 10.10, 10.11 son paralelos, vale decir que el punto  $N$  (definido por el nodo ascendente), es en general diferente al punto  $A$  (definido por la intersección de una recta con un plano), de modo que la longitud de estos segmentos es en general distinta.

donde hemos tenido en cuenta que  $OB$  tiene longitud unidad por construcción. Por otro lado, los triángulos  $OCB$  y  $OAC$  nos dicen que

$$\begin{aligned} \sin \theta &= \frac{OC}{OB} = OC ; \quad \sin u = \frac{AC}{OC} = \frac{AC}{\sin \theta} \Rightarrow \\ AC &= \sin \theta \sin u \end{aligned} \quad (10.152)$$

igualando la ecuaciones (10.151, 10.152), se obtiene

$$\sin u = \cot \theta_i \cot \theta$$

que coincide con la Ec. (10.149) con lo cual vemos que  $u$  en dicha ecuación es el ángulo indicado en la Fig. 10.11. Es claro también de dicha figura y de la definición de  $\Omega$  en la Fig. 10.10, que  $\phi - u = \Omega$  de lo cual se obtiene

$$2\pi w_1 = \Omega \quad (10.153)$$

un procedimiento similar se puede realizar para encontrar la interpretación física de  $w_2$ . De las integrales que aparecen en (10.142), las correspondientes a  $\theta$  y  $r$  contienen a  $J_2$  y por tanto intervienen en el cálculo de  $w_2$ . A partir de  $w_2 = \partial W / \partial J_2$  se puede calcular la integral asociada a  $\theta$  con una sustitución trigonométrica similar a la utilizada para calcular  $w_1$ . La integral asociada a  $r$  se puede calcular de varias formas, la más directa es usando la ecuación de la órbita que da  $r$  en términos del ángulo polar en el plano orbital. Eligiendo adecuadamente el límite inferior de integración que es arbitrario se puede encontrar que  $2\pi w_2$  está dado por la diferencia entre dos ángulos que están en el plano de la órbita, uno de ellos es el ángulo entre el radio vector y la línea de nodos y el otro es el ángulo entre el radio vector y la línea del periápside (ver Figs. 10.10 y 10.11), con lo cual se obtiene

$$2\pi w_2 = \omega \quad (10.154)$$

donde  $\omega$  es el argumento del perihelio como se aprecia en la Fig. 10.10.

Si bien el método de las variables acción-ángulo no es el más ventajoso para trabajar el problema de Kepler, veremos que las variables acción-ángulo constituyen una parametrización muy adecuada para el tratamiento de la teoría canónica de perturbaciones en el problema de Kepler. En el lenguaje de la mecánica celeste, a las variables acción-ángulo en el problema de Kepler se les conoce como **variables de Delauney**, si bien en algunas convenciones las variables de Delauney difieren de las acción-ángulo por constantes multiplicativas.

## 10.17. Ejercicios

1. A partir de la expresión (10.41) para el teorema del virial obtenga la expresión (10.42), teniendo en cuenta la ley de acción y reacción y el hecho de que la fuerza resultante sobre la  $i$ -ésima partícula se escribe como

$$\mathbf{F}_i = \mathbf{F}_i^{(e)} + \sum_{j \neq i} \mathbf{F}_{ij}$$

siendo  $\mathbf{F}_i^{(e)}$  la fuerza externa neta y siendo  $\mathbf{F}_{ij}$  la fuerza interna que la partícula  $j$  hace sobre la partícula  $i$  del sistema de partículas.

2. Sea una fuerza atractiva central de la forma

$$f(r) = -ke^{-ar} ; \quad k > 0$$

encuentre los valores de  $r$  para los cuales es posible una órbita circular estable. Repita el ejercicio para  $f(r) = -kr^3 e^{-ar^2}$ .

3. Escriba la frecuencia de pequeñas oscilaciones radiales (10.76) en términos de las condiciones iniciales, y compruebe que coincide con la obtenida en las Ecs. (10.68). Sugerencia: Utilice la Ec. (10.72).

4. Con base en la Ec. (10.114), evalúe el valor de la constante de proporcionalidad que relaciona el cuadrado del periodo con el cubo del semi-eje mayor, para cada planeta del sistema solar. Evalúe la desviación porcentual máxima que se presenta entre estas constantes.
5. Un cometa se aproxima a la tierra y en un instante dado, se mide su posición y su velocidad inicial tomando a la tierra como origen de coordenadas. Asuma que la masa de la tierra es mucho mayor a la del cometa, y que el sistema de referencia fijo a la tierra es inercial. Prescindiendo del tamaño de ambos astros, describa bajo que condiciones el cometa retornará en una órbita cerrada, no retornará o colisionará con la tierra. ¿Como cambia esta descripción al introducir el tamaño de estos astros?. En ambos casos, desprecie la interacción del cometa con la atmósfera de la tierra.
6. Obtenga la relación entre  $w_2$  y las variables orbitales descrita por la ecuación (10.154), con un procedimiento similar al realizado para obtener la Ec. (10.153).

## Colisiones y dispersión

Vamos a analizar el problema de dos partículas que interactúan de alguna manera, pero que están aisladas del resto del universo. En algunos casos es suficiente asumir que las dos partículas “chocan” o entran en “contacto directo” entre sí, lo cual implícitamente significa que las dos partículas se pueden aproximar indefinidamente una a la otra. En este caso hablamos de una **colisión** entre las dos partículas. Cuando estudiamos el comportamiento de partículas macroscópicas, ésta constituye una buena aproximación y será el primer caso que trataremos.

Sin embargo, debe tenerse en cuenta que las interacciones son funciones de la distancia, y la fuerza usualmente adquiere un valor singular para distancia cero. Este hecho resulta de particular importancia en la Física atómica y subatómica dado que en estos escenarios la distancia mínima de aproximación entre dos partículas es de un orden de magnitud similar al tamaño mismo de las partículas. Cuando tenemos en cuenta que existe una fuerza a distancia entre las partículas (por ejemplo una interacción eléctrica) entonces el proceso no se interpreta como un contacto directo entre tales partículas sino como la interacción de estas últimas a una distancia muy corta. En este caso hablamos de una **dispersión** (también suele usarse el anglicismo **scattering**). Usualmente, el experimento se prepara de manera que inicialmente las dos partículas estén muy alejadas entre sí de manera que se puede desprestigiar la interacción entre ellas en el instante inicial (y por tanto, la energía potencial). Se “lanzan” entonces las dos partículas de masas  $m_1$  y  $m_2$  con velocidades  $\mathbf{v}_1$  y  $\mathbf{v}_2$  respectivamente. Las dos partículas se acercan entonces e interactúan entre sí para luego dispersarse de modo que el producto saliente son dos partículas de masas  $m'_1$  y  $m'_2$  con velocidades  $\mathbf{v}'_1$  y  $\mathbf{v}'_2$  respectivamente. Obsérvese que hemos supuesto que en los productos finales tanto las velocidades como las masas pueden haber cambiado, incluso el número de partículas salientes podría ser diferente del número de partículas entrantes. Por ejemplo, veremos más adelante que en las reacciones de captura entran dos partículas y sale sólo una, la explosión de una granada se puede ver como un fenómeno de colisión ya que sólo intervienen fuerzas internas; en este caso el estado inicial es de una partícula y el estado final podría ser de muchas. Generalmente, en el experimento se miden los estados finales cuando las partículas han vuelto a alejarse lo suficiente como para desprestigiar de nuevo la interacción entre ellas. Estos estados final e inicial en que las partículas están muy alejadas entre sí, se conocen como estados asintóticos inicial y final respectivamente, son entonces estos valores asintóticos los que usualmente interesan en un experimento de dispersión.

Es de anotar sin embargo, que los términos colisión y dispersión suelen usarse indistintamente. Es así como es común utilizar el término **colisiones atómicas** en lugar de **dispersiones atómicas**.

### 11.1. Colisiones y dispersiones generales

Cuando dos partículas se aproximan y llegan a su región de colisión o dispersión, se apartan de nuevo y lo que analizamos usualmente es su cambio de momento lineal y de energía, independientemente de si consideramos una interacción a distancia o no. Como primera consideración, es obvio que el momento lineal se debe conservar ya que el sistema es aislado, si asumimos que el número de partículas se conserva tendremos

entonces:

$$\begin{aligned} \mathbf{p}_1 + \mathbf{p}_2 &= \mathbf{p}'_1 + \mathbf{p}'_2 \Rightarrow \\ m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 &= m'_1 \mathbf{v}'_1 + m'_2 \mathbf{v}'_2 \end{aligned} \quad (11.1)$$

y si suponemos que la interacción entre las dos partículas es conservativa, se tiene que el principio de conservación de la energía se puede escribir como:

$$T + V(r) = T' + V'(r) \quad (11.2)$$

siendo  $r \equiv \|\mathbf{r}_2 - \mathbf{r}_1\|$  la distancia entre las partículas. En esta sección utilizaremos la notación primada para estados finales, y la no primada para estados iniciales. Como lo que se puede medir con mayor facilidad en un experimento es la energía cinética de los estados iniciales y finales, definiremos el factor  $Q$  de una colisión como:

$$Q \equiv T' - T = V(r) - V'(r) \quad (11.3)$$

en una colisión general las dos partículas intercambian momento lineal y energía pero de tal manera que la suma total de estas dos cantidades no se modifique. Si las energías cinéticas final e inicial no son las mismas, significa que hubo un cambio de signo contrario en la energía potencial de modo que se conserve la energía mecánica dada por (11.2).

Ya hemos mencionado que en el caso macroscópico por lo general consideramos la colisión como un contacto directo entre las partículas como sucede con las bolas de billar. Dado que en tal caso no se considera una interacción a distancia, no necesitamos que en los estados inicial y final las partículas estén muy alejadas la una de la otra para efectuar nuestras mediciones, en este caso el cambio en la energía cinética es debido a la deformación que los cuerpos puedan sufrir en virtud de sus propiedades plásticas o elásticas.

A partir de la definición de momento lineal, la energía cinética se puede escribir como:

$$\begin{aligned} T &= \frac{1}{2} m_1 \mathbf{v}_1^2 + \frac{1}{2} m_2 \mathbf{v}_2^2 = \frac{(m_1 \mathbf{v}_1)^2}{2m_1} + \frac{(m_2 \mathbf{v}_2)^2}{2m_2} \Rightarrow \\ T &= \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} \end{aligned} \quad (11.4)$$

ahora bien, las colisiones se pueden clasificar de acuerdo al valor del factor  $Q$ . Si  $Q = 0$  se dice que la colisión es elástica ya que se conserva la energía cinética en el proceso. Si  $Q < 0$  se dice que la colisión es inelástica de primera clase o endoérgica, este último nombre se debe al hecho de que  $Q < 0$  indica que la energía cinética ha disminuído en el proceso, de manera que la energía potencial aumenta (y por tanto la energía interna). Finalmente, si  $Q > 0$  tenemos una colisión inelástica de segunda clase o exoérgica.

De lo anterior se concluye que el factor  $Q$  nos da información acerca de la naturaleza de las interacciones internas entre las partículas. Teniendo en cuenta (11.3) y (11.4) podemos escribir:

$$\frac{\mathbf{p}'_1{}^2}{2m'_1} + \frac{\mathbf{p}'_2{}^2}{2m'_2} - \frac{\mathbf{p}_1^2}{2m_1} - \frac{\mathbf{p}_2^2}{2m_2} = Q \quad (11.5)$$

si el factor  $Q$  es conocido, la Ec. (11.5) junto con el principio de conservación del momento lineal, resuelven completamente el problema de la colisión entre dos partículas, siempre y cuando sus estados finales no sean de más de dos partículas. Es decir, dados los momentos iniciales podemos predecir el valor de los momentos finales. Sin embargo, el factor  $Q$  no puede obtenerse de primeros principios, ya que depende en general de las intrincadas interacciones microscópicas que suceden durante la colisión. Por tanto, es usual determinar  $Q$  en forma experimental.

Un hecho importante es que el factor  $Q$  así definido es independiente del observador. Para el caso de partículas con interacción central a distancia, esto es inmediato si tomamos en cuenta que la energía potencial

para una fuerza central es independiente del observador, ya que sólo depende de la magnitud de la coordenada relativa entre las partículas, y como  $Q = V(r) - V'(r)$  se concluye que  $Q$  es independiente del observador. Para interacciones de contacto es necesario probar que la diferencia de energía cinética final e inicial es la misma para cualquier observador. El factor  $Q$  viene dado por:

$$Q = T' - T = \frac{1}{2}m'_1\mathbf{v}'_1{}^2 + \frac{1}{2}m'_2\mathbf{v}'_2{}^2 - \frac{1}{2}m_1\mathbf{v}_1^2 - \frac{1}{2}m_2\mathbf{v}_2^2$$

ahora bien, el factor  $Q$  medido por el centro de masa se obtiene teniendo en cuenta que la energía cinética en el laboratorio y en el centro de masa se relacionan como en la Ec. (1.32), Pág. 13<sup>1</sup>. Con lo cual se puede escribir:

$$\begin{aligned} Q_{CM} &= T'_{CM} - T_{CM} = \left[ T' - \frac{1}{2}(m'_1 + m'_2)\mathbf{v}'_{CM}{}^2 \right] - \left[ T - \frac{1}{2}(m_1 + m_2)\mathbf{v}_{CM}^2 \right] \\ &= \frac{1}{2}m'_1\mathbf{v}'_1{}^2 + \frac{1}{2}m'_2\mathbf{v}'_2{}^2 - \frac{1}{2}(m'_1 + m'_2)\mathbf{v}'_{CM}{}^2 - \frac{1}{2}m_1\mathbf{v}_1^2 - \frac{1}{2}m_2\mathbf{v}_2^2 + \frac{1}{2}(m_1 + m_2)\mathbf{v}_{CM}^2 \end{aligned}$$

donde hemos tenido en cuenta que la velocidad del centro de masa es la misma antes y después de la colisión, ya que el sistema es aislado. Adicionalmente, puesto que  $m'_1 + m'_2 = m_1 + m_2 = M$  se concluye que:

$$\begin{aligned} T'_{CM} - T_{CM} &= \frac{1}{2}m'_1\mathbf{v}'_1{}^2 + \frac{1}{2}m'_2\mathbf{v}'_2{}^2 - \frac{1}{2}m_1\mathbf{v}_1^2 - \frac{1}{2}m_2\mathbf{v}_2^2 = T' - T \\ &\Rightarrow Q = Q_{CM} \end{aligned}$$

de modo que  $Q$  es el mismo en el sistema de referencia laboratorio y en el del centro de masa

$$Q = T' - T = T'_{CM} - T_{CM}$$

y dado que el laboratorio es cualquier sistema inercial, esto prueba la invariancia de  $Q$  con respecto al sistema de referencia inercial que lo mide. Veremos a continuación, que la descripción de la colisión vista por el centro de masa es mucho más sencilla que vista desde el laboratorio. Con respecto al centro de masa, el momento lineal total del sistema es cero de modo que:

$$\begin{aligned} \mathbf{p}_{1,CM} &= -\mathbf{p}_{2,CM} \Rightarrow \mathbf{p}'_{1,CM} = \mathbf{p}'_{2,CM} ; \mathbf{p}'_{1,CM} = -\mathbf{p}'_{2,CM} \Rightarrow \mathbf{p}''_{1,CM} = \mathbf{p}''_{2,CM} \\ T_{CM} &= \frac{\mathbf{p}^2_{1,CM}}{2m_1} + \frac{\mathbf{p}^2_{2,CM}}{2m_2} = \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \mathbf{p}^2_{1,CM} = \frac{\mathbf{p}^2_{1,CM}}{2\mu} \\ T'_{CM} &= \frac{1}{2} \left( \frac{1}{m'_1} + \frac{1}{m'_2} \right) \mathbf{p}^2_{1,CM} = \frac{\mathbf{p}^2_{1,CM}}{2\mu'} \end{aligned}$$

siendo  $\mu$  y  $\mu'$  las masas reducidas antes y después de la colisión. Podemos escribir en síntesis las siguientes relaciones válidas en el sistema de referencia centro de masa

$$\begin{aligned} \mathbf{p}_{1,CM} &= -\mathbf{p}_{2,CM} , \mathbf{p}^2_{1,CM} = \mathbf{p}^2_{2,CM} ; \mathbf{p}'_{1,CM} = -\mathbf{p}'_{2,CM} , \mathbf{p}''_{1,CM} = \mathbf{p}''_{2,CM} \\ T_{CM} &= \frac{\mathbf{p}^2_{1,CM}}{2\mu} = \frac{\mathbf{p}^2_{2,CM}}{2\mu} ; T'_{CM} = \frac{\mathbf{p}^2_{1,CM}}{2\mu'} = \frac{\mathbf{p}^2_{2,CM}}{2\mu'} \end{aligned} \quad (11.6)$$

Nótese que aunque la masa total sea la misma antes y después de la colisión, la masa reducida sí puede cambiar cuando las masas individuales cambian su valor. Empleando la definición de  $Q$  en el centro de masa resulta:

$$\frac{\mathbf{p}''^2_{1,CM}}{2\mu'} - \frac{\mathbf{p}^2_{1,CM}}{2\mu} = Q \quad (11.7)$$

<sup>1</sup>La notación primada en la Ec. (1.32), significa “visto por el centro de masa”, en nuestra sección la notación primada significa “después de la colisión”.

si suponemos que la colisión es elástica y que además la masa reducida no cambia en el proceso, se tiene que  $Q = 0$ ,  $\mu' = \mu$  de modo que

$$\frac{\mathbf{p}'_{1,CM}}{2\mu} = \frac{\mathbf{p}_{1,CM}}{2\mu} \quad \text{y} \quad \mathbf{p}'_{1,CM} = \mathbf{p}_{1,CM} \quad (\text{colisión elástica con } \mu = \mu') \quad (11.8)$$

es decir no hay intercambio de energía cinética entre las partículas<sup>2</sup>, y los momentos no cambian de magnitud. No obstante, sí hay intercambio de momento ya que aunque no varía la magnitud de los momentos, su dirección puede variar en el proceso.

### 11.1.1. Caso especial 1: reacción de captura

La reacción de captura es un choque inelástico en el cual, las dos partículas incidentes chocan y salen unidas como una sola partícula (también se conoce como colisión plástica o choque perfectamente inelástico). En tal caso, el estado final es una partícula de masa  $M = m_1 + m_2$ . Como el sistema es aislado, la velocidad del centro de masa es la misma antes y después de la colisión. En consecuencia, la partícula saliente viajará con la velocidad del centro de masa. Por tanto, la energía cinética de la partícula saliente es:

$$\begin{aligned} T' &= \frac{1}{2} M \mathbf{v}_{CM}^2 = \frac{1}{2} M \left( \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2} \right)^2 = \frac{1}{2} (m_1 + m_2) \left( \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2} \right)^2 \\ T' &= \frac{1}{2} \frac{(m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2)^2}{m_1 + m_2} \end{aligned}$$

y el factor  $Q$  es entonces:

$$\begin{aligned} Q &= T' - T = \frac{1}{2} \frac{(m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2)^2}{m_1 + m_2} - \frac{1}{2} m_1 \mathbf{v}_1^2 - \frac{1}{2} m_2 \mathbf{v}_2^2 \\ &= \frac{1}{2} \left[ \frac{(m_1^2 \mathbf{v}_1^2 + m_2^2 \mathbf{v}_2^2 + 2m_1 m_2 \mathbf{v}_1 \cdot \mathbf{v}_2) - m_1 (m_1 + m_2) \mathbf{v}_1^2 - m_2 (m_1 + m_2) \mathbf{v}_2^2}{m_1 + m_2} \right] \\ &= \frac{1}{2} \left[ \frac{2m_1 m_2 \mathbf{v}_1 \cdot \mathbf{v}_2 - m_1 m_2 \mathbf{v}_1^2 - m_1 m_2 \mathbf{v}_2^2}{m_1 + m_2} \right] = -\frac{1}{2} \frac{m_1 m_2}{(m_1 + m_2)} (\mathbf{v}_2 - \mathbf{v}_1)^2 \\ Q &= -\frac{1}{2} \mu \mathbf{v}^2 \quad ; \quad \mathbf{v} \equiv \mathbf{v}_2 - \mathbf{v}_1 \end{aligned}$$

es más inmediato el cálculo desde el centro de masa ya que en tal caso  $T'_{CM} = 0$  puesto que la partícula final sale con la velocidad del centro de masa y estaría en reposo respecto a este último. Luego  $Q$  queda:

$$\begin{aligned} Q &= T'_{CM} - T_{CM} \\ Q &= -T_{CM} = -\frac{1}{2} \mu \mathbf{v}^2 \end{aligned} \quad (11.9)$$

donde hemos usado la Ec. (10.6), Pág. 215 vista desde el centro de masa. Es decir, en el caso de una reacción de captura, el factor  $Q$  es igual a menos la energía cinética del sistema inicial vista por el centro de masa.

### 11.1.2. Caso especial 2, blanco en reposo

Vamos a suponer ahora que la partícula 1 tiene masa  $m_1$  y momento lineal  $\mathbf{p}_1$  y que la partícula dos tiene masa  $m_2$  y está en reposo respecto al laboratorio, es decir  $\mathbf{p}_2 = 0$ . En este caso la partícula 1 puede

<sup>2</sup>Aquí se ha probado que no hay intercambio de energía cinética visto por el centro de masa, pero el lector puede verificar fácilmente que tampoco hay intercambio de energía cinética en el laboratorio.

considerarse como el proyectil, y la partícula 2 como el blanco, el principio de conservación de la energía nos dice que:

$$\mathbf{p}'_1 + \mathbf{p}'_2 = \mathbf{p}_1 \Rightarrow \mathbf{p}'_2 = \mathbf{p}_1 - \mathbf{p}'_1$$

elevando esta ecuación al cuadrado obtenemos:

$$\begin{aligned} \mathbf{p}'_2{}^2 &= (\mathbf{p}_1 - \mathbf{p}'_1)^2 = \mathbf{p}_1^2 + \mathbf{p}'_1{}^2 - 2\mathbf{p}_1 \cdot \mathbf{p}'_1 = \mathbf{p}_1^2 + \mathbf{p}'_1{}^2 - 2\|\mathbf{p}_1\| \|\mathbf{p}'_1\| \cos \theta \\ \mathbf{p}'_2{}^2 &= \mathbf{p}_1^2 + \mathbf{p}'_1{}^2 - 2p_1 p'_1 \cos \theta \end{aligned} \quad (11.10)$$

donde  $\theta$  es el ángulo entre los vectores  $\mathbf{p}_1$  y  $\mathbf{p}'_1$ , es decir el ángulo con que se dispersa la partícula 1. El factor  $Q$  viene dado por:

$$Q = \frac{\mathbf{p}'_1{}^2}{2m'_1} + \frac{\mathbf{p}'_2{}^2}{2m'_2} - \frac{\mathbf{p}_1^2}{2m_1} \quad (11.11)$$

y reemplazando (11.10) en (11.11), se obtiene:

$$\begin{aligned} Q &= \frac{\mathbf{p}'_1{}^2}{2m'_1} + \frac{\mathbf{p}_1^2 + \mathbf{p}'_1{}^2 - 2p_1 p'_1 \cos \theta}{2m'_2} - \frac{\mathbf{p}_1^2}{2m_1} \\ Q &= \frac{\mathbf{p}'_1{}^2}{2} \left( \frac{1}{m'_1} + \frac{1}{m'_2} \right) + \frac{\mathbf{p}_1^2}{2} \left( \frac{1}{m'_2} - \frac{1}{m_1} \right) - \frac{p_1 p'_1}{m'_2} \cos \theta \\ Q &= \frac{\mathbf{p}'_1{}^2}{2m'_1} \left( 1 + \frac{m'_1}{m'_2} \right) + \frac{\mathbf{p}_1^2}{2m_1} \left( \frac{m_1}{m'_2} - 1 \right) - \frac{\sqrt{4m_1 m'_1} \frac{p_1}{2m_1} \frac{p'_1}{2m'_1}}{m'_2} \cos \theta \end{aligned}$$

y recordando que  $T_i = \mathbf{p}_i^2 / (2m_i)$ , el factor  $Q$  viene dado por:

$$Q = T'_1 \left( 1 + \frac{m'_1}{m'_2} \right) - T_1 \left( 1 - \frac{m_1}{m'_2} \right) - \frac{2\sqrt{m_1 m'_1 T_1 T'_1}}{m'_2} \cos \theta \quad (11.12)$$

la cual se conoce como ecuación  $Q$  y expresa dicho factor para el caso de un blanco inicialmente en reposo.

**Example 7** Una granada inicialmente en reposo explota en dos fragmentos  $m'_1$  y  $m'_2$ , hallar la energía cinética final en términos de  $Q$ .

Como la granada está inicialmente en reposo, entonces  $\mathbf{p}_1 = 0$  y por conservación del momentum  $\mathbf{p}'_1 + \mathbf{p}'_2 = 0$ , con lo cual  $\mathbf{p}'_1{}^2 = \mathbf{p}'_2{}^2$  por tanto la energía cinética final es:

$$\begin{aligned} T' &= \frac{\mathbf{p}'_1{}^2}{2} \left( \frac{1}{m'_1} + \frac{1}{m'_2} \right) = \frac{\mathbf{p}'_1{}^2}{2\mu'} \\ Q &= T' - T = T' = \frac{\mathbf{p}'_1{}^2}{2\mu'} \Rightarrow \\ \|\mathbf{p}'_1\| &= \|\mathbf{p}'_2\| = \sqrt{2\mu' Q} \end{aligned}$$



de lo cual queda:

$$\begin{aligned} T'_1 &= \frac{\mathbf{p}_1^2}{2m'_1} = \frac{2\mu'Q}{2m'_1} = \frac{m'_1 m'_2 Q}{m'_1 (m'_1 + m'_2)} \\ T'_1 &= \frac{m'_2 Q}{(m'_1 + m'_2)} \\ T'_2 &= \frac{\mathbf{p}_2^2}{2m'_2} = \frac{\mathbf{p}_1^2}{2m'_2} = \frac{2\mu'Q}{2m'_2} \\ T'_2 &= \frac{m'_1 Q}{(m'_1 + m'_2)} \\ T' &= T'_1 + T'_2 = Q \end{aligned}$$

## 11.2. Dispersión en un campo de fuerzas centrales

Históricamente las fuerzas centrales surgen para estudiar el movimiento planetario, no obstante hay muchos escenarios en donde dichas fuerzas son importantes, un ejemplo notable lo constituye el átomo de Bohr. No obstante, si pretendemos trabajar física microscópica las correcciones cuánticas pueden ser importantes. Sin embargo, el lenguaje así como muchos aspectos fenomenológicos suelen ser similares en mecánica clásica y en mecánica cuántica.

En general, si conocemos la forma de la interacción es posible predecir la forma de la órbita y su dependencia con el parámetro tiempo. Sin embargo, es usual en la búsqueda de nueva física realizar experimentos de dispersión (o scattering) que consisten en enviar proyectiles a un blanco que interactúa con éstos y observar la deflexión en la trayectoria de los proyectiles, en este caso lo que se pretende es determinar experimentalmente la forma de la órbita para deducir la forma (desconocida) de la interacción.

Motivados por la anterior discusión, en esta sección trabajaremos la desviación de un haz de partículas por un centro de fuerzas en la formulación de un cuerpo, lo cual significa las siguientes simplificaciones: (a) el centro dispersor es fijo en el espacio de modo que no hace parte de la dinámica del sistema, equivalente a que su masa sea infinita, y (b) se desprecia la interacción entre las partículas del haz, de modo que cada partícula incidente solo interactúa con el centro dispersor. Consideremos por simplicidad que todas las partículas del haz poseen la misma masa y energía, estos proyectiles pueden ser electrones, núcleos, planetas, etc. En general, consideraremos que este haz uniforme es tal que las partículas vienen de regiones muy lejanas al centro de fuerzas de modo que su interacción con el centro de fuerzas se considera despreciable en el tiempo inicial (que se suele tomar como  $t \rightarrow -\infty$ ). En consecuencia, las trayectorias incidentes son líneas rectas. Una vez que las partículas pasan por las vecindades del centro de fuerza (denominada región de dispersión o de colisión), se alejan de dicho centro y sus trayectorias vuelven a ser líneas rectas cuando las partículas están suficientemente lejos del centro dispersor (usualmente decimos que esto ocurre en  $t \rightarrow \infty$ ), pero la dirección de las partículas salientes no coincide con la dirección de las partículas incidentes. La desviación de las partículas contiene información valiosa sobre la naturaleza de la interacción.

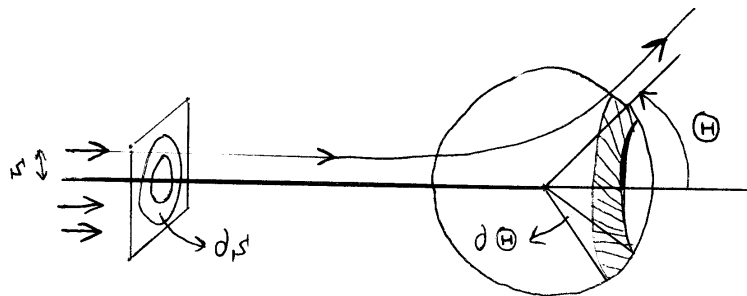


Figura 11.1: Dispersión de un haz de partículas por un centro de fuerzas repulsivo.

El haz incidente se caracteriza a través de su intensidad, es decir, número de partículas por unidad de área por unidad de tiempo que atraviesan una superficie normal a la dirección de propagación del haz (a esta cantidad también se le conoce con el nombre de densidad de flujo). Debido al cambio de dirección del haz cuando ha pasado por el centro de fuerzas, diremos que el haz se ha dispersado o desviado como se puede ver en la Fig. 11.1. Una cantidad físicamente útil para estudiar el fenómeno de la dispersión es la sección eficaz de dispersión en un dirección dada (denotada por  $\sigma(\Omega) d\Omega$ )

$$\sigma(\Omega) d\Omega = \frac{\text{número de partículas dispersadas por unidad de tiempo en un ángulo sólido } d\Omega}{\text{intensidad incidente}} \quad (11.13)$$

donde  $d\Omega$  es un elemento diferencial de ángulo sólido en la dirección  $\Omega$ . La cantidad  $\sigma(\Omega)$  suele también denominarse sección eficaz diferencial de dispersión. Escribiendo el ángulo sólido explícitamente resulta

$$\sigma(\Omega) d\Omega = \sigma(\Theta) \sin \Theta d\Theta d\Phi \quad (11.14)$$

Si la fuerza es central debe haber una simetría total con respecto al eje del haz incidente, lo cual nos conduce a que la sección eficaz diferencial debe ser independiente del ángulo azimutal  $\Phi$  (i.e. el ángulo de rotación alrededor del eje de simetría). Por esta razón se puede hacer una integración parcial del ángulo sólido. Como  $\sigma$  no depende de  $\Phi$  es decir posee simetría azimuthal, tenemos que la integración en  $\Phi$  se puede realizar para obtener

$$\begin{aligned} \sigma(\Omega) d\Omega &= \sigma(\Theta) \sin \Theta d\Theta d\Phi \Rightarrow \sigma(\Omega) d\bar{\Omega} = \sigma(\Theta) \sin \Theta d\Theta \int_0^{2\pi} d\Phi \\ \sigma(\Omega) d\bar{\Omega} &= 2\pi \sigma(\Theta) \sin \Theta d\Theta \end{aligned} \quad (11.15)$$

de modo que podemos definir el diferencial de ángulo sólido en la forma

$$d\bar{\Omega} = 2\pi \sin \Theta d\Theta \quad (11.16)$$

$\Theta$  es el ángulo que hacen la dirección desviada e incidente y se denomina ángulo de dispersión. El término sección eficaz se debe a que  $\sigma(\Omega)$  tiene dimensiones de área. Cuando tenemos una fuerza central y por tanto simetría azimuthal, podemos trabajar con el ángulo sólido  $d\bar{\Omega}$  dado por (11.16) que corresponde al anillo sombreado sobre la superficie esférica en la Fig. (11.1), en lugar del ángulo sólido  $d\Omega = \sin \Theta d\Theta d\Phi$ . De aquí en adelante, volveremos a la notación  $d\Omega$  sobreentendiendo que nos referimos al ángulo sólido dado en (11.16).

Es importante mencionar que lo que se puede medir experimentalmente es la integral de la Ec. (11.14) o de la Ec. (11.15) sobre una cierta porción finita de ángulo sólido, ya que en la realidad solo podemos calcular ángulos de dispersión y contar las partículas que se dispersan dentro de cierto rango finito de esos ángulos de dispersión.

Tomaremos el origen de nuestro sistema coordenado en el centro dispersor, el eje  $X$  irá paralelo a la velocidad del haz incidente y en el mismo sentido del haz. El ángulo  $\theta$  de una determinada posición  $\mathbf{r}$  se mide con respecto al eje  $X$  positivo de modo que para las partículas incidentes (cuando  $t \rightarrow -\infty$ ) se tiene que  $r \rightarrow \infty$  y  $\theta = \pi$ . Para una partícula dada, asumiendo que conocemos su plano de movimiento<sup>3</sup> y teniendo en cuenta que su posición inicial está dada por  $r \rightarrow \infty$  y  $\theta = \pi$ , veremos que las constantes de la órbita y por tanto la dispersión, están determinadas por la energía y el módulo del momento angular  $E$  y  $l$ . Conviene expresar el momento angular en función de la energía y de una cantidad  $s$  denominada parámetro de impacto que es la distancia entre las rectas paralelas a la velocidad del proyectil, que pasan por el proyectil y el centro de fuerza respectivamente (ver Fig. 11.1). Si  $v_0$  es la rapidez incidente de la partícula se tiene que

$$l = r p \sin(\pi - \alpha) = r \sin \alpha m v_0 = m v_0 s = s \sqrt{2m \left( \frac{1}{2} m v_0^2 \right)} = s \sqrt{2mE} \quad (11.17)$$

<sup>3</sup>La determinación del plano de movimiento implica el conocimiento de la variable azimutal  $\Phi$ .

donde  $\alpha$  es el ángulo entre el vector posición incidente y la línea paralela a la dirección de incidencia que pasa por el centro dispersor (i.e. el ángulo entre el vector posición de la partícula incidente y el eje  $X$  negativo)<sup>4</sup>. En esta expresión hemos tenido en cuenta que la partícula incidente está muy lejos del centro de fuerza, de modo que su energía es puramente cinética. Una vez fijados  $E$  y  $s$  queda determinado unívocamente el ángulo de dispersión  $\Theta$ <sup>5</sup>. Las partículas cuyo parámetro de impacto está entre  $s$  y  $s + ds$  deben cruzar la superficie del anillo con radio interno  $s$  y radio externo  $s + ds$  ilustrado en la Fig. 11.1, el área de este anillo es  $2\pi s |ds|$ . Ahora bien, el número de partículas por unidad de tiempo que pasan por dicho anillo es igual a la intensidad incidente  $I$  multiplicada por el área del anillo es decir

$$\left. \frac{dN}{dt} \right|_{\text{anillo } ds} = 2\pi I s |ds| \quad (11.18)$$

por otro lado, el número de partículas por unidad de tiempo que se dispersan en el ángulo sólido  $d\Omega$  (determinado por la región sombreada sobre la superficie esférica en la Fig 11.1)<sup>6</sup> se puede calcular teniendo en cuenta la definición (11.13), de la cual se vé que corresponde simplemente a  $I \sigma(\Theta) d\Omega$ , teniendo en cuenta (11.15) el número de partículas por unidad de tiempo que se dispersa en un ángulo sólido  $d\Omega$  es

$$\left. \frac{dN}{dt} \right|_{\text{angulo sólido } d\Omega} = 2\pi \sigma(\Theta) I \sin \Theta |d\Theta| \quad (11.19)$$

Supondremos de momento que partículas con diferente parámetro de impacto no pueden llegar al mismo ángulo de dispersión. En tal caso, el número de partículas dispersadas por el ángulo sólido  $d\Omega$  comprendido entre  $\Theta$  y  $\Theta + d\Theta$  deberá ser igual al número de partículas incidentes con parámetro de impacto comprendido entre los valores  $s$  y  $s + ds$ . Por tanto en este caso las cantidades definidas en (11.18, 11.19) son iguales y se obtiene

$$2\pi I s |ds| = 2\pi \sigma(\Theta) I \sin \Theta |d\Theta| \quad (11.20)$$

los valores absolutos se introducen para asegurar que el número de partículas sea positivo, ya que a menudo  $s$  y  $\Theta$  varían en sentidos opuestos. Si consideramos a  $s$  como función de la energía y el ángulo de dispersión

$$s = s(\Theta, E) \quad (11.21)$$

y despejando  $\sigma(\Theta)$  en (11.20), la dependencia entre la sección eficaz diferencial y  $\Theta$  vendrá dada por

$$\sigma(\Theta) = \frac{s}{\sin \Theta} \left| \frac{ds}{d\Theta} \right| \quad (11.22)$$

a partir de la ecuación de órbita general (10.51a) se puede obtener directamente una expresión formal para el ángulo de dispersión. Consideraremos por simplicidad una fuerza totalmente repulsiva. Teniendo en cuenta que la órbita debe ser simétrica respecto a la dirección del periápside (recordemos que la órbita es simétrica ante reflexión con respecto a las ápsides, ver discusión de la Ec. 10.50), el ángulo  $\Psi$  del periápside con la dirección incidente es igual al ángulo entre el periápside y la dirección de dispersión (ver Fig. 11.2). En consecuencia, el ángulo de dispersión viene dado por

$$\Theta = \pi - 2\Psi \quad (11.23)$$

El ángulo  $\Psi$  se puede obtener de (10.51a) tomando el límite  $r_0 \rightarrow \infty$  cuando  $\theta_0 \rightarrow \pi$  (dirección incidente).

<sup>4</sup>Dado que la partícula incidente está muy lejos,  $\alpha$  es muy pequeño. Sin embargo,  $r \sin \alpha$  es claramente finito e igual al parámetro de impacto.

<sup>5</sup>En este punto la mecánica cuántica sí difiere de la clásica ya que en la primera no se pueden definir trayectorias sino probabilidades de desviación en distintas direcciones.

<sup>6</sup>Es importante observar que la partícula no necesariamente tiene que cruzar por el anillo sombreado sobre la esfera. Lo que define este anillo es el "espacio angular" por donde debe pasar la partícula. En la Fig. 11.1, la trayectoria dibujada no pasa por el lugar geométrico en el que se dibujó el anillo, pero el ángulo  $\Theta$  de deflexión de esta partícula está dentro del intervalo angular que define el anillo, cuando dicho ángulo se mide desde el origen.

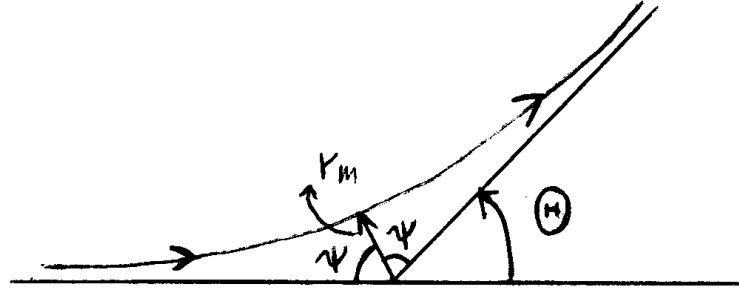


Figura 11.2: Ilustración del ángulo  $\Psi$  entre la dirección incidente y la línea del periápside para una dispersión central repulsiva. La simetría de reflexión conduce a que este ángulo coincide con el formado por la dirección de dispersión y el periápside.

Por otro lado, es claro que  $\theta = \pi - \Psi$  cuando  $r = r_m$  distancia de mayor acercamiento<sup>7</sup>. Integrando la órbita (10.51a) desde la posición incidente hasta el periápside con estas consideraciones se tiene

$$\pi - \Psi = \int_{\infty}^{r_m} \frac{dr}{r^2 \sqrt{\frac{2mE}{l^2} - \frac{2mV}{l^2} - \frac{1}{r^2}}} + \pi$$

quedando finalmente

$$\Psi = \int_{r_m}^{\infty} \frac{dr}{r^2 \sqrt{\frac{2mE}{l^2} - \frac{2mV}{l^2} - \frac{1}{r^2}}} \quad (11.24)$$

y expresando  $l$  en función de  $s$  Ec. (11.17) queda

$$\begin{aligned} \Psi &= \int_{r_m}^{\infty} \frac{dr}{r^2 \sqrt{\frac{2mE}{2mEs^2} - \frac{2mV}{2mEs^2} - \frac{s^2}{s^2r^2}}} = \int_{r_m}^{\infty} \frac{s dr}{r^2 \sqrt{\left(1 - \frac{V}{E} - \frac{s^2}{r^2}\right)}} \\ \Psi &= \int_{r_m}^{\infty} \frac{s dr}{r \sqrt{r^2 \left[1 - \frac{V(r)}{E}\right] - s^2}} \end{aligned} \quad (11.25)$$

de aquí se puede encontrar una expresión formal para calcular  $\Theta$  reemplazando (11.25) en (11.23)

$$\Theta(s) = \pi - 2 \int_{r_m}^{\infty} \frac{s dr}{r \sqrt{r^2 \left[1 - \frac{V(r)}{E}\right] - s^2}} \quad (11.26)$$

o en términos de  $u \equiv 1/r$  con  $du = -dr/r^2 = -u^2 dr$

$$\begin{aligned} \Theta(s) &= \pi - 2 \int_{u_m}^0 \frac{s dr}{\frac{1}{u} \sqrt{\frac{1}{u^2} \left\{ \left[1 - \frac{V(1/u)}{E}\right] - s^2 u^2 \right\}}} = \pi - 2 \int_{u_m}^0 \frac{s u^2 dr}{\sqrt{\left[1 - \frac{V(1/u)}{E}\right] - s^2 u^2}} \\ \Theta(s) &= \pi - 2 \int_0^{u_m} \frac{s (-u^2 dr)}{\sqrt{1 - \frac{V(1/u)}{E} - s^2 u^2}} \end{aligned}$$

<sup>7</sup>Conviene recordar que  $\theta = \theta(t)$  es la posición angular de la partícula en cualquier instante de tiempo, medida desde el eje  $X$  positivo. En cambio  $\Theta$  es el ángulo de dispersión de la partícula el cual es único para cada partícula que se dispersa.

quedando finalmente

$$\Theta(s) = \pi - 2 \int_0^{u_m} \frac{s \, du}{\sqrt{1 - \frac{V(1/u)}{E} - s^2 u^2}} \quad (11.27)$$

Sin embargo, cuando se dispone de una función analítica para la órbita, es frecuente que se pueda encontrar a  $\Theta$  en función de  $s$  por simple inspección. En consecuencia, las Ecs. (11.26, 11.27) se usan muy poco para cálculos analíticos, y usualmente se emplean para el cálculo numérico del ángulo de dispersión.

### 11.2.1. Dispersión de Rutherford

Es de enfatizar que fué un experimento de dispersión el que condujo a Rutherford a modelar la distribución de cargas eléctricas en los átomos. En un experimento clásico realizado por Rutherford, Geiger y Marsden en 1909, se dispersaron partículas alfa (núcleos de Helio) por una lámina de oro de  $10^{-4}$  cm de espesor. El tratamiento de este problema exige consideraciones estadísticas ya que se trata de muchos proyectiles (partículas alfa) incidiendo sobre un enorme conjunto de blancos (átomos de oro) que en primera aproximación se pueden considerar inmóviles. Aquí consideraremos una versión muy simplificada con un solo blanco inmóvil y un haz bien colimado, es decir todas las partículas en la misma dirección. Adicionalmente, despreciamos la interacción de las partículas del haz entre sí y de las partículas del haz con la nube electrónica que rodea al núcleo inmóvil.

Nuestra versión simplificada estudia entonces la repulsión causada por un centro dispersor coulombiano (repulsión entre cargas eléctricas). El campo de fuerzas dispersor es creado por una carga fija  $-Ze$  al ejercerse sobre partículas incidentes de carga  $-Z'e$ , la fuerza viene dada por

$$\mathbf{f}(r) = \frac{ZZ'e^2}{r^2} \mathbf{u}_r$$

es decir una fuerza repulsiva de la forma  $k/r^2$ . Los resultados obtenidos para el problema de Kepler se pueden usar teniendo en cuenta que  $k = -ZZ'e^2$  en la Ec. (10.83)

Ya que la energía potencial es positiva, la energía total  $E$  también lo es, y la órbita será una hipérbola con excentricidad dada por la Ec. (10.90)

$$\varepsilon = \sqrt{1 + \frac{2El^2}{m(ZZ'e^2)^2}} = \sqrt{1 + \left(\frac{2Es}{ZZ'e^2}\right)^2} > 1 \quad (11.28)$$

donde hemos tenido en cuenta la Ec. (11.17). Tomemos  $\theta' = \pi$  en (10.85) y escribamos la ecuación de la órbita

$$\frac{1}{r} = \frac{mk}{l^2} [1 + \varepsilon \cos(\theta - \pi)] = \frac{m(-ZZ'e^2)}{l^2} [1 - \varepsilon \cos \theta]$$

y la ecuación de la órbita queda

$$\frac{1}{r} = \frac{mZZ'e^2}{l^2} (\varepsilon \cos \theta - 1) \quad (11.29)$$

vemos que al tomar  $\theta' = \pi$  en (10.85), el periápside corresponderá a  $\theta = 0$ . Por otro lado, la Ec. (11.29) requiere que  $\varepsilon \cos \theta - 1$  sea no negativo, el hecho de que  $\varepsilon > 1$  garantiza que existen valores de  $\theta$  para los cuales se cumple esta condición. No obstante, tal condición excluye ciertos valores de  $\theta$ . Adicionalmente, dado que el periápside está definiendo el ángulo cero, la dirección  $\Psi$  de la asíntota de incidencia (o dispersión) queda determinada por el valor de  $\theta$  en el límite  $r \rightarrow \infty$  en (11.29)

$$\cos \Psi = \frac{1}{\varepsilon} \quad (11.30)$$

nótese que los valores permitidos de  $\theta$  yacen justamente en el intervalo  $[-\Psi, \Psi]$  como se puede observar de la Fig. 11.2 o teniendo en cuenta que si reemplazamos  $\cos \theta = \cos(\pm\Psi) = 1/\varepsilon$  en (11.29) se anula el término de

<sup>8</sup>Se define  $e$  como la carga negativa del electrón de modo que  $-Ze$  se refiere a una carga positiva.  $Z$  se refiere al número atómico.

la derecha lo cual viene de la condición  $r \rightarrow \infty$ . De acuerdo con la Ec. (11.23)

$$\cos \Psi = \cos \left( \frac{\pi}{2} - \frac{\Theta}{2} \right) = \sin \frac{\Theta}{2} = \frac{1}{\varepsilon} \Rightarrow \quad (11.31)$$

$$\begin{aligned} \varepsilon &= \csc \frac{\Theta}{2} \Rightarrow \varepsilon^2 - 1 = \csc^2 \frac{\Theta}{2} - 1 \\ \cot^2 \frac{\Theta}{2} &= \varepsilon^2 - 1 \end{aligned} \quad (11.32)$$

y usando (11.28) se obtiene

$$\sqrt{\cot^2 \frac{\Theta}{2}} = \frac{2Es}{ZZ'e^2} \quad (11.33)$$

de la Ec. (11.31) y teniendo en cuenta que  $\varepsilon \geq 1$ , se tiene que

$$0 < \frac{1}{\varepsilon} = \sin \frac{\Theta}{2} \leq 1$$

de modo que

$$0 < \frac{\Theta}{2} \leq \frac{\pi}{2} \Rightarrow \cot \frac{\Theta}{2} \geq 0 \quad (11.34)$$

con lo cual la Ec. (11.33) queda de la forma

$$\cot \frac{\Theta}{2} = \frac{2Es}{ZZ'e^2} \quad (11.35)$$

despejando  $s$  en esta ecuación, se obtiene la relación funcional entre el parámetro de impacto y el ángulo de dispersión

$$s = \frac{ZZ'e^2}{2E} \cot \frac{\Theta}{2} \quad (11.36)$$

que es una relación de la forma (11.21) con la cual es posible calcular  $\sigma(\Theta)$  por medio de la Ec. (11.22). Reemplazando (11.36) en (11.22) se encuentra que  $\sigma(\Theta)$  es

$$\begin{aligned} \sigma(\Theta) &= \left[ \frac{ZZ'e^2}{2E} \cot \frac{\Theta}{2} \right] \frac{1}{\sin \Theta} \left| \frac{d}{d\Theta} \left( \frac{ZZ'e^2}{2E} \cot \frac{\Theta}{2} \right) \right| \\ \sigma(\Theta) &= \left[ \left( \frac{ZZ'e^2}{2E} \right)^2 \cot \frac{\Theta}{2} \right] \frac{1}{(2 \sin \frac{\Theta}{2} \cos \frac{\Theta}{2})} \left| -\frac{1}{2} \csc^2 \frac{\Theta}{2} \right| \\ \sigma(\Theta) &= \frac{1}{4} \left( \frac{ZZ'e^2}{2E} \right)^2 \frac{\cos \frac{\Theta}{2}}{\sin \frac{\Theta}{2}} \frac{\csc^2 \frac{\Theta}{2}}{(\sin \frac{\Theta}{2} \cos \frac{\Theta}{2})} = \frac{1}{4} \left( \frac{ZZ'e^2}{2E} \right)^2 \frac{\csc^2 \frac{\Theta}{2}}{\sin^2 \frac{\Theta}{2}} \end{aligned}$$

finalmente

$$\sigma(\Theta) = \frac{1}{4} \left( \frac{ZZ'e^2}{2E} \right)^2 \csc^4 \frac{\Theta}{2} \quad (11.37)$$

La Ec. (11.37) corresponde a la llamada **dispersión de Rutherford** quien la dedujo para la dispersión de partículas  $\alpha$  por núcleos atómicos. Es notable el hecho de que en mecánica cuántica en el límite no relativista, la sección eficaz coincide con este resultado clásico.

### 11.2.2. Características generales de la sección eficaz

En Física atómica tiene también importancia el concepto de sección eficaz total de dispersión  $\sigma_T$  definido por

$$\sigma_T = \int_{4\pi} \sigma(\Omega) d\Omega = 2\pi \int_0^\pi \sigma(\Theta) \sin \Theta d\Theta \quad (11.38)$$

el último paso solo se puede realizar si hay simetría azimutal, lo cual efectivamente ocurre si la fuerza es central. Sin embargo si intentamos calcular la sección eficaz total para la dispersión coulombiana de Rutherford, sustituyendo (11.37) en (11.38) obtenemos un resultado divergente. Es fácil ver la razón por la cual esto es así, la sección eficaz total es el número de partículas que por unidad de intensidad incidente se dispersan en todas direcciones. Ocurre que la interacción coulombiana es de alcance infinito, las desviaciones muy pequeñas solo tienen lugar para partículas con parámetro de impacto muy grande. Por tanto todas las partículas de un haz incidente de sección lateral infinita se desviarían en mayor o menor medida y por tanto deben incluirse en la sección eficaz total de dispersión. En consecuencia, la infinitud de la sección eficaz total en mecánica clásica no es exclusiva del potencial coulombiano y se extiende a todos los potenciales de alcance infinito independiente de lo grandes que sean<sup>9</sup>. Solo si el potencial presenta un **corte** de modo que se anula mas allá de cierta distancia, será finita esta sección eficaz total. Este es el caso del campo coulombiano de un núcleo, el cual presenta un corte debido al apantallamiento de la nube electrónica.

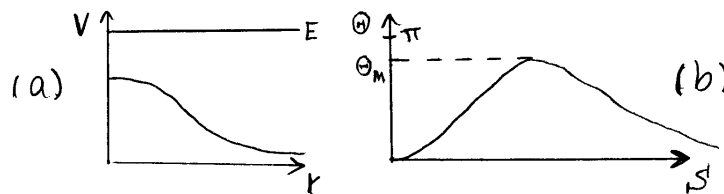


Figura 11.3: (a) perfil de un potencial repulsivo que permanece finito en el centro de fuerza. (b) Gráfica de  $\Theta$  vs  $s$  para el anterior potencial donde se observa que a cada valor permitido de  $\Theta$  le corresponde dos parámetros de impacto diferentes (excepto para  $\Theta = \Theta_m$ ).

Es muy importante tener presente que la validez de la Ec. (11.22) depende de que las partículas con diferente parámetro de impacto no puedan llegar al mismo ángulo de dispersión, ya que de ello depende la validez de (11.20) y (11.22). En la dispersión de Rutherford, esta condición se cumple ya que el ángulo de dispersión es una función monótona suave (y por tanto uno a uno) del parámetro de impacto  $s$ . Esto se puede apreciar en la Ec. (11.36), teniendo en cuenta que  $\Theta$  está entre cero y  $\pi$ . En la Ec. (11.36) se ve que al disminuir  $s$  desde el infinito hacia cero el ángulo  $\Theta$  crece monótonamente a partir de cero (cuando  $s \rightarrow \infty$  es lógico que  $\Theta \rightarrow 0$  ya que la interacción tiende a cero), alcanzando el valor de  $\pi$  cuando  $s = 0$  lo cual corresponde a colisión frontal con el centro de fuerzas, que naturalmente obliga a la partícula a recular en la dirección contraria a la incidencia.

No obstante, existen otros potenciales clásicos cuyo comportamiento requiere reevaluar la expresión (11.22), por ejemplo para un potencial repulsivo como el de la figura 11.3a, tal que el potencial es finito incluso en  $r \rightarrow 0$  y tal que la energía de la partícula es mayor al potencial para todo valor de  $r$ , es fácil ver que la curva de  $\Theta$  versus  $s$  se puede comportar de la forma indicada en la figura 11.3b, la cual no es una función uno a uno. Este comportamiento se puede explicar mediante el siguiente razonamiento físico: Para parámetros de impacto muy grandes, la partícula permanece todo el tiempo muy lejos del centro de fuerzas de tal modo que su ángulo de dispersión es muy pequeño. Si  $s = 0$  la partícula se dirige en línea recta directamente al centro de fuerzas y dado que su energía es superior al máximo valor del potencial, pasará a través del centro de fuerzas sin desviarse de su curso (en vez de recular). Por esta razón, el ángulo de dispersión se anula en los dos extremos de  $s$ , y dado que  $0 \leq \Theta \leq \pi$ , si la gráfica es bien comportada es de esperarse que tenga por lo menos un máximo local  $\Theta_M$ , para algún valor del parámetro de impacto. La Fig. 11.3b muestra que para  $\Theta < \Theta_M$  existen dos valores de  $s$  que corresponden al mismo ángulo de dispersión. Cada uno de ellos contribuye a la sección eficaz de dispersión según un ángulo y por tanto la Ec. (11.22) se debe modificar en la forma

$$\sigma(\Theta) = \sum_i \frac{s_i}{\sin \Theta} \left| \frac{ds}{d\Theta} \right|_i \quad (11.39)$$

<sup>9</sup>En cuántica, todos los potenciales que tienden a cero a grandes distancias más rápidamente que  $1/r^2$  producen un sección eficaz total de dispersión finita.

donde el subíndice  $i$  distingue los distintos valores de  $s$  que dan el mismo valor de  $\Theta$ . A manera de ejemplo, pensemos que la curva de  $\Theta$  vs  $s$  es tal que  $\Theta = 0$  en  $s = 0, \infty$  y que además posee dos máximos locales  $\Theta_1, \Theta_3$  ( $\Theta_1 > \Theta_3$ ) y un mínimo local  $\Theta_2$  tal que  $s_1 < s_2 < s_3$ . En este caso, hay ángulos para los cuales hay cuatro parámetros de impacto asociados ( $\Theta_2 < \Theta < \Theta_3$ ), otros tienen dos parámetros  $s$  asociados ( $\Theta_3 < \Theta < \Theta_1$  y  $\Theta < \Theta_2$ ), cuando  $\Theta = \Theta_2$  ó  $\Theta = \Theta_3$  hay tres parámetros de impacto asociados, y finalmente cuando  $\Theta = \Theta_1$  solo hay uno.

Volviendo al caso descrito por la Fig. 11.3b, resulta de especial interés la sección eficaz correspondiente al ángulo máximo  $\Theta_M$ . Puesto que para este ángulo se anula la derivada de  $\Theta$  con respecto a  $s$ , se vé de la Ec. (11.39) que la sección eficaz debe diverger cuando  $\Theta \rightarrow \Theta_M$ . Pero para todos los ángulos con  $\Theta > \Theta_M$ , la sección eficaz es nula (para el caso de la Fig. 11.3b, solo hay un máximo local de modo que ninguna partícula se dispersa con un ángulo mayor a  $\Theta_M$ ). Este fenómeno de la subida infinita de la sección eficaz seguida de una desaparición brusca, es muy semejante a lo que ocurre en la óptica geométrica de la dispersión de la luz solar por las gotas de lluvia. En virtud de esta semejanza, a este fenómeno se le denomina **dispersión en arco iris**.

Cuando el scattering es debido a fuerzas atractivas, pueden surgir otras complicaciones. Por ejemplo, dado que la partícula es atraída hacia el centro de fuerzas, el ángulo  $\Psi$  entre la dirección incidente y la dirección del periapsis puede ser mayor que  $\pi/2$ , con lo cual a través de la Ec. (11.23) obtenemos un ángulo de dispersión negativo. Esto no supone mucho problema en virtud de que lo que importa para medir  $\sigma(\Theta)$  es el valor absoluto de  $\Theta$ . No obstante, el valor absoluto de dicho ángulo de acuerdo con la Ec. (11.26) ó (11.27), puede ser mayor que  $2\pi$ , lo cual significa que físicamente la partícula podría dar una o más vueltas alrededor del centro de fuerzas, antes de ser “lanzada” hacia su dirección asintótica de dispersión.

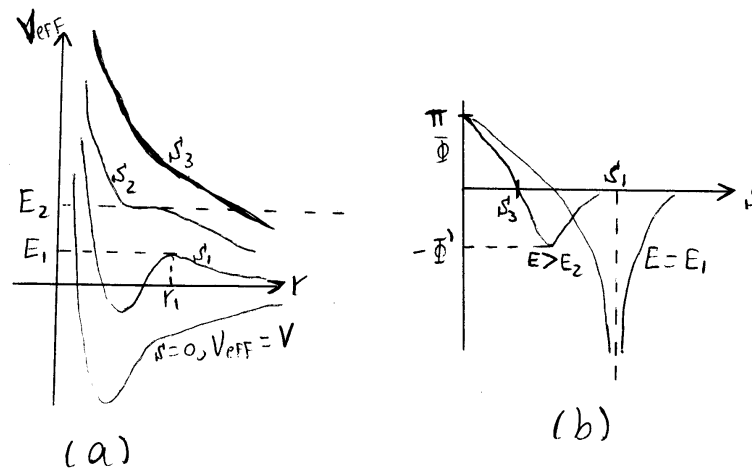


Figura 11.4: (a) Potenciales efectivos asociados a un potencial intermolecular real típico, cada valor del parámetro de impacto conduce a un potencial efectivo diferente. (b) Gráfica del ángulo de desviación  $\Phi$  en función de  $s$ , para el potencial anterior con dos valores diferentes de la energía.

Veamos un ejemplo de como puede ocurrir esto físicamente. En la gráfica 11.4a, se pintan una serie de potenciales efectivos para un potencial intermolecular (real) típico. Cada curva está asociada al mismo potencial real, pero con una barrera centrífuga diferente, lo cual corresponde a un valor distinto del momento angular  $l$  y por tanto también del parámetro de impacto  $s$ . Lo más característico de esta clase de potenciales intermoleculares **reales** es que son atractivos a grandes distancias decayendo a un ritmo mayor a  $1/r^2$ , y repulsivos con magnitud de fuerza rápidamente creciente a cortas distancias<sup>10</sup>. Dado que a grandes distancias el potencial real (atractivo) decae más rápido que  $1/r^2$  la barrera centrífuga que se forma cuando  $s \neq 0$  dominará a grandes distancias y para valores pequeños de  $s$  se formará un máximo local, a medida que el parámetro de impacto

<sup>10</sup>Estas son las características que a grandes rasgos se esperan para que la materia no implusione y al mismo tiempo pueda formar condensados.



aumenta la curva tiende a “aplanarse” y para algún valor del parámetro de impacto  $s_2$ , solo hay un punto de inflexión en  $V_{eff}$  para cierto valor de la energía  $E_2$ . Finalmente, para parámetros de impacto aún mayores no aparecen extremos locales ni puntos de inflexión en el potencial.

Consideremos una partícula con parámetro de impacto  $s_1$  y energía  $E_1$  correspondiente a la energía en el máximo local. Recordemos que la distancia entre  $V_{eff}$  y  $E_1$  es la energía cinética radial. Por tanto, cuando la partícula llega a una distancia  $r_1$  del centro de fuerzas la velocidad radial se anula. Recordemos además que cuando la energía coincide con un máximo local tenemos una órbita circular inestable. En ausencia de perturbaciones la partícula de energía  $E_1$  y parámetro de impacto  $s_1$  quedará realizando una órbita circular de radio  $r_1$  de **forma indefinida**. Si la energía es un poco mayor que  $E_1$  la órbita ya no es circular pero su velocidad radial será muy pequeña en las vecindades de  $r_1$ , y la partícula puede permanecer en esta vecindad un largo tiempo realizando varios giros. En contraste, la velocidad angular  $\dot{\theta}$  no está afectada por la existencia de un máximo y para un valor dado de  $r$  viene dada por (11.17)

$$\dot{\theta} = \frac{l}{mr_1^2} = \frac{s_1}{r_1^2} \sqrt{\frac{2E}{m}}$$

por tanto durante el tiempo en el que la partícula atraviesa la región del máximo, su velocidad angular vendrá dada aproximadamente por este valor y puede hacer que dicha partícula ejecute varios giros. En tales casos, se dice que la dispersión clásica exhibe una orbitalización o espiralización.

Claramente para energías mayores que  $E_2$  (en donde se presenta el punto de inflexión para un  $s_2$  dado), no es posible un movimiento orbital. Allí se puede presentar no obstante otro fenómeno interesante, y es que la combinación de componentes atractivas y repulsivas del potencial efectivo puede incluso conducir a deflexión nula para algún valor finito del parámetro de impacto. Por otro lado, a valores grandes de energía con pequeños parámetros de impacto, los efectos de dispersión dominantes corresponden a repulsión fuerte a cortas distancias, y el scattering se asemeja cualitativamente a la dispersión de Rutherford<sup>11</sup>.

Dado que la partícula se puede deflectar en un ángulo mayor a  $\pi$ , pero el ángulo medido en el laboratorio está entre 0 y  $\pi$ , es útil en este caso diferenciar el ángulo de deflexión  $\Phi$ , que es el que se calcula con las expresiones de la derecha en las Ecs. (11.26) ó (11.27) y el ángulo de dispersión que se observa  $\Theta$ . Para un  $\Phi$  dado el valor  $\Theta$  se calcula con la expresión

$$\Theta = \pm\Phi - 2m\pi \quad , \quad m \text{ entero positivo}$$

el signo y el valor de  $m$  se escogen para que  $\Theta$  quede en el rango entre 0 y  $\pi$ . La suma expresada por (11.39) cubre entonces todos los valores de  $\Phi$  que conducen al mismo  $\Theta$ . En la Fig. 11.4b, se hace una gráfica de  $\Phi$  vs  $s$  para el potencial de la figura 11.4a, para dos energías diferentes. La orbitación que se produce en  $E = E_1$  se muestra como una singularidad en la curva en  $s = s_1$  ya que el sistema da infinitas vueltas<sup>12</sup>. Para  $E > E_2$  no hay orbitación pero hay un efecto arco iris en  $\Phi = -\Phi'$  (aunque la sección eficaz no se anula para ángulos mayores, sino menores a  $-\Phi'$ ). Nótese que  $\Phi$  se anula (y por tanto también  $\Theta$ ) para  $s = s_3$  lo cual significa a partir de (11.22) que la sección eficaz se vuelve infinita en la dirección adelante debido a que el factor  $\sin \Phi = \sin \Theta$  se anula<sup>13</sup>. Similarmente, la sección eficaz puede diverger en la dirección hacia atrás (i.e. para  $\Theta = \pi$ ) si ocurre que

$$s \left| \frac{ds}{d\Theta} \right|$$

permanece finito en  $\Theta = \pi$ . Estos infinitos en la dispersión adelante atrás se conocen como **dispersión gloria**, de nuevo por su análogo al correspondiente fenómeno en óptica meteorológica. El efecto óptico es familiar para

<sup>11</sup>Esto es de esperarse ya que a mayor energía con bajo parámetro de impacto, la partícula puede penetrar la nube electrónica y acercarse mucho al núcleo. En realidad la dispersión de Rutherford se da gracias a este efecto.

<sup>12</sup>Dado que  $\Psi > 0$ , la Ec. (11.23) muestra que  $\Phi \leq \pi$ , de modo que  $\Phi$  solo puede diverger con valores negativos, como se aprecia en la Fig. 11.4b.

<sup>13</sup>El hecho de que exista un parámetro de impacto finito y diferente de cero para el cual no hay deflexión, significa que los efectos de atracción y repulsión a diversas distancias en que estuvo la partícula, tuvieron un efecto de cancelación. Sin embargo, esto no significa que la partícula viajó en línea recta, pues en la región de dispersión la trayectoria pudo ser muy compleja. De hecho las líneas rectas que definen las trayectorias inicial y final son paralelas pero no necesariamente coincidentes.

los viajeros de avión que pueden ver un anillo de luz que envuelve la sombra del avión proyectada sobre una nube subyacente.

Aunque las correcciones cuánticas son en general importantes, a veces dichas correcciones son pequeñas como ocurre con la dispersión a bajas energías en retículos cristalinos de iones. Incluso en ocasiones se recurre a métodos semiclásicos en los cuales es necesario el conocimiento de la órbita clásica.

### 11.3. Transformación del problema de la dispersión entre coordenadas de laboratorio y centro de masa para blanco en reposo

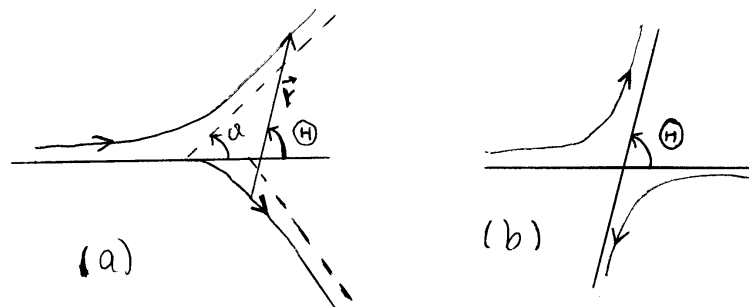


Figura 11.5: (a) *Dispersión de dos partículas vista por el sistema de referencia del laboratorio.* (b) *Dispersión de dos partículas vista por el sistema de referencia del centro de masa.*

Hasta el momento, hemos supuesto que el centro de fuerzas es fijo y por lo tanto trabajamos el problema de un cuerpo. En la realidad, lo que tenemos es un cuerpo mucho más masivo que el otro que en todo caso debe reular. Por otro lado, podemos estar interesados en la colisión de dos partículas con masas similares o casos en que el blanco es móvil. Todo ello conduce al hecho de que el problema de la colisión es un problema de dos cuerpos. Aunque ya hemos demostrado que este problema se puede reducir al problema de un cuerpo con masa reducida  $\mu$  cuando la interacción es central, no podemos simplemente reemplazar los dos cuerpos por este cuerpo equivalente. La dificultad radica en que el ángulo de dispersión se mide realmente en el laboratorio y corresponde al ángulo  $\vartheta$  entre las direcciones inicial y final de la partícula dispersada en las coordenadas de laboratorio. Por otro lado, el ángulo  $\Theta$  calculado a partir del problema equivalente de un cuerpo es el ángulo entre las direcciones inicial y final del vector relativo entre las dos partículas en las coordenadas del centro de masa<sup>14</sup>.

En lo que sigue asumiremos que en el instante inicial ( $t \rightarrow -\infty$ ), uno de los cuerpos está en reposo. En este caso, los valores de  $\Theta$  y  $\vartheta$  solo son iguales si la partícula que estaba en reposo permanece estacionaria en el proceso. Dado que en la realidad la partícula inicialmente en reposo se pone en movimiento gracias a la interacción con el proyectil, los dos ángulos serán diferentes como se vé en la figura 11.5a. En consecuencia, el problema equivalente de un cuerpo no nos da directamente el ángulo de dispersión que mide el laboratorio<sup>15</sup>.

#### 11.3.1. Relación entre el ángulo de dispersión medido por el laboratorio y el medido por el centro de masa

Para encontrar la relación entre  $\Theta$  y  $\vartheta$  debemos primero examinar como ocurre la dispersión para un sistema de referencia que se mueve con el centro de masa (que claramente también es inercial). En este sistema

<sup>14</sup>Estrictamente,  $\Theta$  es el mismo cuando se mide en el sistema de referencia del laboratorio como se vé en la Fig. 11.5a, en virtud de que el vector relativo es idéntico en ambos sistemas de referencia. Esto se puede ver en la Fig. 10.1 de la Pág. 214, la cual muestra claramente que  $\mathbf{r}_2 - \mathbf{r}_1 = \mathbf{r}'_2 - \mathbf{r}'_1 = \mathbf{r}$ , y esta relación es válida para todo tiempo.

<sup>15</sup>Recordemos que las trayectorias que hemos obtenido para un cuerpo bajo fuerzas centrales, tampoco son las órbitas que vería un sistema de referencia inercial, y solo coinciden razonablemente cuando uno de los cuerpos está casi en reposo y es muy masivo con respecto al otro.

de referencia el momento total es cero siempre, así que las dos partículas siempre se mueven con momentos opuestos. Antes de la dispersión las dos partículas se acercan mutuamente y luego de la dispersión se alejan una de otra.

En la Fig. 10.1 de la Pág. 214, podemos ver que el vector relativo entre las partículas pasa por el centro de masa, así mismo los vectores  $\mathbf{r}$ ,  $\mathbf{r}'_1, \mathbf{r}'_2$  son todos colineales para todo tiempo. Esto significa que desde el punto de vista del centro de masa, no solo los momentos son antiparalelos sino que las direcciones incidentes de las dos partículas yacen sobre la misma recta, al igual que las direcciones salientes o reflejadas como se ilustra en la Fig. 11.5b. De aquí se concluye que los ángulos de dispersión de las dos partículas son idénticos ya que son opuestos por el vértice como se vé en la Fig. 11.5b. Por otro lado, vemos que el ángulo de dispersión de  $\mathbf{r}'_i$  sería el mismo que el ángulo de deflexión de  $\mathbf{r}$  ya que son colineales todo el tiempo. Por tanto el ángulo de deflexión de las partículas visto por el CM es el ángulo  $\Theta$  que se muestra en la Fig. 11.5a. La dispersión vista por el centro de masa se ilustra en la Fig. 11.5b.

Ahora bien, la relación entre  $\Theta$  y  $\vartheta$  se obtendrá entonces haciendo la transformación entre el sistema centro de masa y laboratorio. Esta transformación ya fué considerada en la sección 10.1. Derivando en el tiempo la Ec. (10.3) se obtiene

$$\mathbf{v}_1 = \mathbf{V} + \mathbf{v}'_1 \tag{11.40}$$

esta relación es válida para todo tiempo. Por conveniencia, haremos un ligero cambio en la terminología:  $\mathbf{r}_1$ ,  $\mathbf{v}_1$  son la posición y velocidad de la partícula incidente  $m_1$  **después de la dispersión** en el sistema de laboratorio (L).  $\mathbf{r}'_1$ ,  $\mathbf{v}'_1$  son la posición y velocidad de la partícula incidente  $m_1$  **después de la dispersión** en el sistema del centro de masa (CM).  $\mathbf{R}$  y  $\mathbf{V}$  no sufren modificación en su significado antes y después de la colisión.

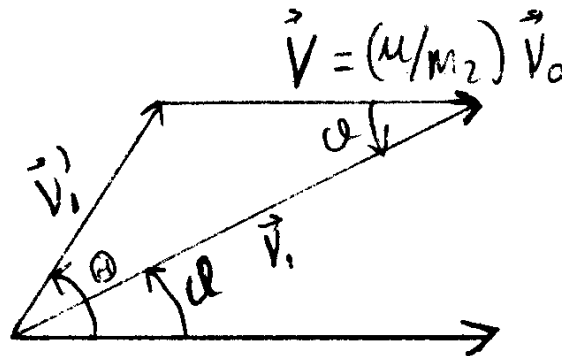


Figura 11.6: Relación entre los vectores  $\mathbf{v}_1, \mathbf{v}'_1$  y  $\mathbf{V}$  después de la colisión.

Dado que el blanco  $m_2$  está inicialmente estacionario en el laboratorio, la velocidad incidente  $\mathbf{v}_0$  del proyectil  $m_1$  en dicho sistema, coincide con la velocidad relativa inicial de las partículas. Por conservación del momento lineal total (medido en el laboratorio) la velocidad constante del centro de masa es

$$\begin{aligned} (m_1 + m_2) \mathbf{V} &= m_1 \mathbf{v}_0 \Rightarrow \mathbf{V} = \frac{m_1 m_2}{m_2 (m_1 + m_2)} \mathbf{v}_0 \\ \mathbf{V} &= \frac{\mu}{m_2} \mathbf{v}_0 \end{aligned} \tag{11.41}$$

La Ec. (11.41) muestra que  $\mathbf{v}_0$  es paralelo a  $\mathbf{V}$ . En consecuencia, si tomamos como eje X el eje de incidencia de la partícula, la velocidad del centro de masa  $\mathbf{V}$  es paralela a este eje. La Fig. 11.6 muestra la relación vectorial (11.40) evaluada **después** de la dispersión, en cuyo tiempo  $\mathbf{v}_1$  y  $\mathbf{v}'_1$  hacen los ángulos  $\vartheta$  y  $\Theta$  respectivamente con el vector  $\mathbf{V}$ , que yace a lo largo de la dirección incidente. Adicionalmente, la figura 11.6 muestra que

$$v_1 \sin \vartheta = v'_1 \sin \Theta \tag{11.42}$$

$$v_1 \cos \vartheta = v'_1 \cos \Theta + V \tag{11.43}$$

el cociente entre estas ecuaciones nos da la relación entre  $\Theta$  y  $\vartheta$  que buscábamos

$$\tan \vartheta = \frac{\sin \Theta}{\cos \Theta + \rho}, \quad \rho \equiv \frac{V}{v'_1} = \frac{\mu}{m_2} \frac{v_0}{v'_1} \quad (11.44)$$

donde hemos usado (11.41). Se puede obtener una expresión alternativa cuando se expresa  $v_1$  en términos de otras velocidades a través del teorema del coseno aplicado al triángulo de la Fig. 11.6.

$$v_1^2 = v'_1{}^2 + V^2 - 2v'_1 V \cos(\pi - \Theta) = v'_1{}^2 + V^2 + 2v'_1 V \cos \Theta \quad (11.45)$$

despejando  $\cos \vartheta$  en (11.43) y usando las Ecs. (11.45, 11.44) se obtiene

$$\begin{aligned} \cos \vartheta &= \frac{v'_1 \cos \Theta + V}{v_1} = \frac{v'_1 \cos \Theta + V}{\sqrt{v'_1{}^2 + V^2 + 2v'_1 V \cos \Theta}} = \frac{\cos \Theta + V/v'_1}{\sqrt{1 + \left(\frac{V}{v'_1}\right)^2 + 2\frac{V}{v'_1} \cos \Theta}} \\ \cos \vartheta &= \frac{\cos \Theta + \rho}{\sqrt{1 + 2\rho \cos \Theta + \rho^2}} \end{aligned} \quad (11.46)$$

vemos que las relaciones entre  $\vartheta$  y  $\Theta$  expresadas en (11.44) y (11.46) involucran el factor  $\rho$  que depende de las rapidez inicial de la partícula 1 vista por (L) y final de la misma partícula pero vista por (CM). Conviene por tanto, caracterizar adecuadamente a este factor.

### 11.3.2. Caracterización del factor $\rho$ de la colisión

Por la definición de centro de masa, la velocidad  $\mathbf{v}'_1$  de la partícula 1 en el CM después de la colisión, está conectada con la velocidad relativa  $\mathbf{v}$  después de la colisión. Esta relación se obtiene utilizando la Ec. (10.4) y derivándola con respecto al tiempo

$$\mathbf{v}'_1 = -\frac{m_2}{m_1 + m_2} \mathbf{v} \Rightarrow \|\mathbf{v}'_1\| = \frac{m_2 m_1}{(m_1 + m_2) m_1} \|\mathbf{v}\| \Rightarrow v'_1 = \frac{\mu}{m_1} v$$

y reemplazando esta expresión en (11.44), el factor  $\rho$  se puede escribir como

$$\rho = \frac{m_1}{m_2} \frac{v_0}{v} \quad (11.47)$$

es necesario recalcar que  $v$  es la rapidez relativa **después** de la colisión.

Si la colisión es inelástica, la energía cinética total de las dos partículas cambia después de la colisión (por ejemplo algo de la energía cinética se puede convertir en energía de excitación interna del blanco). Ya vimos que el cambio de energía cinética resultante de la colisión se expresa por medio del factor  $Q$ . Expresando la energía cinética vista por el laboratorio, su cambio se parametriza en la forma

$$\frac{\mu v^2}{2} + \frac{1}{2} M V^2 = \frac{\mu v_0^2}{2} + \frac{1}{2} M V^2 + Q$$

donde hemos tenido en cuenta que la velocidad del centro de masa permanece constante de modo que

$$\frac{\mu v^2}{2} = \frac{\mu v_0^2}{2} + Q \quad (11.48)$$

El valor  $Q$  del choque inelástico es usualmente negativo<sup>16</sup>, aunque el convenio de signos se toma de manera que sea consistente con el que se usa para las reacciones atómicas y nucleares. Multiplicando por  $2/\mu$  esta ecuación resulta

$$v^2 = v_0^2 + \frac{2}{\mu} Q = v_0^2 + \frac{Q}{\frac{1}{2} \mu v_0^2} v_0^2 = v_0^2 \left( 1 + \frac{Q}{\frac{1}{2} \mu v_0^2} \right) \quad (11.49)$$

<sup>16</sup>En todo caso, es posible observar un incremento en la energía cinética a expensas de una disminución de la energía interna de los cuerpos que colisionan.

teniendo en cuenta que la energía del sistema en el laboratorio es  $(1/2) m_1 v_0^2$  se tiene que

$$\frac{1}{2} \mu v_0^2 = \frac{1}{2} m_1 v_0^2 \left( \frac{m_2}{m_1 + m_2} \right) = \left( \frac{m_2}{m_1 + m_2} \right) E$$

y la Ec. (11.49) queda

$$v^2 = v_0^2 \left[ 1 + \left( \frac{m_1 + m_2}{m_2} \right) \frac{Q}{E} \right]$$

Por tanto, el cociente entre las rapidezces relativas antes y después del choque se puede escribir como

$$\frac{v}{v_0} = \sqrt{1 + \left( \frac{m_1 + m_2}{m_2} \right) \frac{Q}{E}} \quad (11.50)$$

siendo  $E$  el valor de la energía de la partícula incidente (en el sistema de laboratorio). Reemplazando (11.50) en (11.47), se tiene que el factor  $\rho$  para dispersión inelástica, vendrá dado por

$$\rho_{inelast} = \frac{m_1}{m_2 \sqrt{1 + \left( \frac{m_1 + m_2}{m_2} \right) \frac{Q}{E}}} \quad (11.51)$$

Cuando la colisión es elástica ( $Q = 0$ ), la energía cinética total de las dos partículas permanece inalterada y el factor  $\rho$  de la Ec. (11.51) se simplifica a

$$\rho_{elast} = \frac{m_1}{m_2} \quad (11.52)$$

y es independiente de las energías y rapidezces. El mismo resultado se obtiene haciendo  $Q = 0$  en (11.48) de lo cual se obtiene  $v = v_0$  y reemplazando esta igualdad en (11.47), se obtiene (11.52).

Las Ecs. (11.44, 11.46) muestran que los valores de  $\Theta$  y  $\vartheta$  son en general diferentes si  $\rho \neq 0$ , y la Ec. (11.51) muestra que esta condición ocurre para cualquier valor finito de  $Q$ .

### 11.3.3. Sección eficaz en términos de los dos ángulos de dispersión

Los valores de la sección eficaz diferencial dependen de cual de los dos ángulos se tome como variable independiente para  $\sigma$ . Para obtener la relación entre ellas, observemos que en un experimento particular el número de partículas dispersadas por unidad de tiempo a través de un elemento de ángulo sólido dado, debe ser el mismo sin importar si medimos el fenómeno como función de  $\vartheta$  o como función de  $\Theta$ . Esta cantidad expresada para el ángulo  $\Theta$  viene dada por (11.19) y al ser igual para ambos ángulos se obtiene

$$2\pi I \sigma(\Theta) \sin \Theta |d\Theta| = 2\pi I \sigma'(\vartheta) \sin \vartheta |d\vartheta|$$

donde  $\sigma'(\vartheta)$ ,  $\sigma(\Theta)$  son las secciones eficaces diferenciales expresadas en función de cada ángulo de dispersión, ambos medidos en el sistema de laboratorio. Esta relación se puede escribir en la forma

$$\sigma'(\vartheta) = \sigma(\Theta) \frac{\sin \Theta}{\sin \vartheta} \left| \frac{d\Theta}{d\vartheta} \right| = \sigma(\Theta) \left| \frac{d(\cos \Theta)}{d(\cos \vartheta)} \right|$$

La derivada se puede evaluar por medio de la Ec. (11.46) y se obtiene fácilmente

$$\begin{aligned} \cos \vartheta &= \frac{\cos \Theta + \rho}{\sqrt{1 + 2\rho \cos \Theta + \rho^2}} \\ \frac{d(\cos \vartheta)}{d(\cos \Theta)} &= \frac{\sqrt{1 + 2\rho \cos \Theta + \rho^2} - \frac{2\rho}{2} (\cos \Theta + \rho) (1 + 2\rho \cos \Theta + \rho^2)^{-1/2}}{(1 + 2\rho \cos \Theta + \rho^2)} \\ \frac{d(\cos \vartheta)}{d(\cos \Theta)} &= \frac{\sqrt{1 + 2\rho \cos \Theta + \rho^2}}{\sqrt{1 + 2\rho \cos \Theta + \rho^2}} \left[ \frac{\sqrt{1 + 2\rho \cos \Theta + \rho^2} - \rho (\cos \Theta + \rho) (1 + 2\rho \cos \Theta + \rho^2)^{-1/2}}{(1 + 2\rho \cos \Theta + \rho^2)} \right] \\ \frac{d(\cos \vartheta)}{d(\cos \Theta)} &= \frac{(1 + 2\rho \cos \Theta + \rho^2) - \rho (\cos \Theta + \rho)}{(1 + 2\rho \cos \Theta + \rho^2)^{3/2}} \end{aligned}$$

$$\frac{d(\cos \vartheta)}{d(\cos \Theta)} = \frac{1 + \rho \cos \Theta}{(1 + 2\rho \cos \Theta + \rho^2)^{3/2}}$$

de lo cual la sección eficaz  $\sigma'(\vartheta)$  queda finalmente

$$\sigma'(\vartheta) = \sigma(\Theta) \frac{(1 + 2\rho \cos \Theta + \rho^2)^{3/2}}{1 + \rho \cos \Theta} \quad (11.53)$$

es importante enfatizar que  $\sigma(\Theta)$  no es la sección eficaz que mediría un observador en el sistema del centro de masa. Tanto  $\sigma(\Theta)$  como  $\sigma'(\vartheta)$  son secciones eficaces medidas en el laboratorio, pero están expresadas en término de coordenadas angulares diferentes. Un observador fijo en el centro de masa vería una densidad de flujo de partículas incidentes diferente de la que se mide en el sistema de laboratorio y habría que incluir esa transformación de la densidad de flujo si por alguna razón quisiéramos relacionar las secciones eficaces medidas en uno y otro sistema.

La relación entre los dos ángulos de dispersión es particularmente simple en el caso en el cual tenemos una colisión elástica donde las partículas tienen masas iguales, en este caso  $\rho = 1$  y según (11.46) se tiene

$$\cos \vartheta = \sqrt{\frac{1 + \cos \Theta}{2}} = \cos \frac{\Theta}{2}$$

de modo que

$$\vartheta = \frac{\Theta}{2} \quad (\rho = 1) \quad (11.54)$$

en consecuencia, en el caso de dispersión elástica con masas iguales no podrá haber ángulos de dispersión mayores que  $\pi/2$  en el sistema de laboratorio; toda la dispersión tiene lugar en el hemisferio de adelante. Correspondientemente, la sección eficaz de dispersión vendrá dada en función de  $\Theta$  según la Ec. (11.53) en la forma

$$\begin{aligned} \sigma'(\vartheta) &= \sigma(\Theta) \frac{(2 + 2 \cos \Theta)^{3/2}}{1 + \cos \Theta} = \sigma(\Theta) \frac{2^{3/2} (1 + \cos \Theta)^{3/2}}{1 + \cos \Theta} \\ &= \sigma(\Theta) \left[ 2^{3/2} (1 + \cos \Theta)^{1/2} \right] = \sigma(\Theta) \left[ 2^{3/2} \left( 2 \cos^2 \frac{\Theta}{2} \right)^{1/2} \right] \end{aligned}$$

que combinada con (11.54) nos da

$$\sigma'(\vartheta) = 4 \cos \vartheta \sigma(\Theta) \quad ; \quad \vartheta \leq \frac{\pi}{2}, \quad (\text{colisión elástica con } \rho = 1) \quad (11.55)$$

incluso si tenemos dispersión isótropa según  $\Theta$ , en la cual  $\sigma(\Theta)$  es una constante independiente de  $\Theta$ , la sección eficaz en función de  $\vartheta$  varía como el coseno de este ángulo.

De acuerdo con la discusión anterior, incluso en los choques elásticos en los que la energía cinética total permanece constante, una colisión con blanco fijo trae como consecuencia que el proyectil cede parte de su energía cinética al blanco. Es decir, la colisión **frena** a la partícula incidente. Para obtener el grado de frenamiento en un escenario inelástico, podemos emplear la Ec. (11.45) expresando  $v_1'$  y  $V$  en función de  $v_0$  a través de las Ecs. (11.44, 11.41), respectivamente

$$\begin{aligned} v_1^2 &= v_1'^2 + V^2 + 2v_1'V \cos \Theta = \left( \frac{\mu v_0}{m_2 \rho} \right)^2 + \left( \frac{\mu v_0}{m_2} \right)^2 + 2 \left( \frac{\mu v_0}{m_2 \rho} \right) \left( \frac{\mu v_0}{m_2} \right) \cos \Theta \\ v_1^2 &= \left( \frac{\mu v_0}{m_2 \rho} \right)^2 [1 + \rho^2 + 2\rho \cos \Theta] \end{aligned}$$

el cociente entre las rapidezces  $v_1$  y  $v_0$  queda

$$\frac{v_1^2}{v_0^2} = \left( \frac{\mu}{m_2 \rho} \right)^2 [1 + 2\rho \cos \Theta + \rho^2] \quad (11.56)$$

si el choque es elástico, se tiene que  $\rho = m_1/m_2$  con lo cual la Ec. (11.56) se puede simplificar en la forma

$$\begin{aligned}\frac{\frac{1}{2}m_1v_1^2}{\frac{1}{2}m_1v_0^2} &= \left[ \frac{m_1m_2/(m_1+m_2)}{m_2(m_1/m_2)} \right]^2 [1 + 2\rho \cos \Theta + \rho^2] = \left[ \frac{m_2}{(m_1+m_2)} \right]^2 [1 + 2\rho \cos \Theta + \rho^2] \\ \frac{E_1}{E_0} &= \left[ \frac{1}{\left(1 + \frac{m_1}{m_2}\right)} \right]^2 [1 + 2\rho \cos \Theta + \rho^2]\end{aligned}$$

quedando finalmente

$$\frac{E_1}{E_0} = \frac{1 + 2\rho \cos \Theta + \rho^2}{(1 + \rho)^2}, \quad \rho = \frac{m_1}{m_2} \quad (\text{choque elástico}) \quad (11.57)$$

siendo  $E_0$  ( $E_1$ ) la energía cinética inicial (final) del proyectil en el sistema de laboratorio. Es fácil ver que (11.57) corresponde a un frenamiento del proyectil incidente i.e.  $E_1 \leq E_0$

$$\frac{E_1}{E_0} = \frac{1 + 2\rho \cos \Theta + \rho^2}{(1 + \rho)^2} \leq \frac{1 + 2\rho + \rho^2}{(1 + \rho)^2} = 1$$

La relación se simplifica aún más si el choque elástico ocurre entre masas iguales i.e. con  $\rho = 1$

$$\begin{aligned}\frac{E_1}{E_0} &= \frac{1 + \cos \Theta}{2} = \cos^2 \frac{\Theta}{2} \\ \frac{E_1}{E_0} &= \cos^2 \vartheta \quad (\text{choque elástico con } \rho = 1)\end{aligned}$$

donde hemos utilizado (11.54). Por lo tanto, para el ángulo de dispersión máximo ( $\Theta = \pi$ ,  $\vartheta = \pi/2$ ), el proyectil cede toda su energía al blanco y queda detenido, visto por el sistema de referencia del laboratorio.

Esta transferencia de energía cinética por dispersión es el principio básico utilizado en los reactores de neutrones térmicos. Los neutrones rápidos producidos por fisión colisionan elásticamente en forma sucesiva con un moderador, hasta que su energía cinética se convierte en energía térmica, en cuyo caso es más posible que originen fisiones a que sean capturados. A priori, los elementos ligeros deberían ser los mejores moderadores, idealmente el hidrógeno (con el cual  $\rho \cong 1$ ). Sin embargo, por otras razones técnicas, el hidrógeno solo es práctico en los reactores nucleares cuando es parte de una mezcla o compuesto (como el agua). Se utilizan más como moderadores el deuterio y el carbono 12. El hidrógeno que forma parte de la parafina, el agua o los plásticos sí se usa como moderador.

Estos cálculos de transformación del laboratorio al centro de masa y la cesión de energía cinética se han realizado en un contexto clásico. Cálculos análogos se pueden realizar para choques de “contacto” tales como los que ocurren entre bolas de billar (ver Sec. 11.1.2). Lo interesante es que a pesar de su simplicidad estos cálculos son de muy amplio uso ya que usan principios fundamentales como la conservación del momento lineal y la energía. Mientras se usen los principios de conservación y el factor  $Q$ , los detalles de la colisión son irrelevantes, en realidad la región de interacción se considera como una “caja negra” y lo que medimos son los estados asintóticos inicial y final. No interesa mucho que los fenómenos ocurridos en esta caja negra sean cuánticos o clásicos. Como ya mencionamos, las fórmulas obtenidas en este capítulo sirven para analizar fenómenos de naturaleza cuántica tales como la dispersión neutrón protón, siempre que las energías sean lo suficientemente bajas para ignorar los efectos relativistas.

## 11.4. Ejercicios

1. Se ha encontrado experimentalmente que en una colisión frontal (o central) de dos esferas sólidas tales como dos bolas de billar, las velocidades después del choque están relacionadas con las velocidades antes del choque por la expresión

$$v'_1 - v'_2 = -e(v_1 - v_2) \quad (11.58)$$

donde  $e$  es el coeficiente de restitución y tiene un valor entre cero y uno. Esta relación fue propuesta por Newton y tiene validez solamente aproximada. Adicionalmente, se conserva el momento lineal en el choque. Probar lo siguiente:

a) Las velocidades después del choque están dadas por:

$$v'_1 = \frac{v_1(m_1 - m_2e) + v_2m_2(1 + e)}{m_1 + m_2} \quad (11.59)$$

$$v'_2 = \frac{v_2(m_2 - m_1e) + v_1m_1(1 + e)}{m_1 + m_2} \quad (11.60)$$

b) La  $Q$  de la colisión es:

$$-\frac{1}{2}(1 - e^2) \frac{m_1m_2}{m_1 + m_2} (v_1 - v_2)^2 \quad (11.61)$$

c) Cuál debería ser el valor de  $e$  para que la colisión fuera elástica?

- En una colisión frontal perfectamente inelástica o plástica (reacción de captura) los dos cuerpos se mueven juntos después del choque. (a) Calcule el coeficiente de restitución  $e$  definido en (11.58), para esta colisión. (b) Calcule el valor de  $Q$  en términos del coeficiente de restitución  $e$ .
- Demostrar que si la energía y el momento lineal se conservan en un choque (choque elástico), las velocidades  $\mathbf{v}_1$  y  $\mathbf{v}_2$  antes de la colisión, se relacionan con las velocidades  $\mathbf{v}'_1$  y  $\mathbf{v}'_2$  después de la colisión, por medio de la ecuación

$$\mathbf{u} \cdot (\mathbf{v}'_1 - \mathbf{v}'_2) = -\mathbf{u} \cdot (\mathbf{v}_1 - \mathbf{v}_2)$$

siendo  $\mathbf{u}$  un vector unitario en la dirección en la cual el momento lineal de una de las partícula ha cambiado. Esto implica que en la colisión, la componente de la velocidad relativa a lo largo de la dirección de intercambio de momento lineal ha cambiado de sentido. Aplique el resultado a una colisión frontal y compare el resultado con el obtenido para  $e = 1$  en las ecuaciones (11.59, 11.60).

- Una granada de masa  $M$  va con velocidad  $\mathbf{v}_0$  cuando está a una altura  $h$  a la cual explota en dos fragmentos iguales. Inicialmente, los fragmentos se mueven horizontalmente en el sistema de referencia  $C$ . El factor  $Q$  de la explosión es  $Q = Mv_0^2$ . Determinar los puntos de colisión de los fragmentos con el suelo con relación al punto donde ocurre la explosión cuando (a)  $\mathbf{v}_0 = -v_0\mathbf{u}_y$ , y (b) Cuando  $\mathbf{v}_0 = v_0\mathbf{u}_x$ .
- Demuestre que para una fuerza central repulsiva  $f = kr^{-3}$ , la sección eficaz diferencial está dada por

$$\sigma(\Theta) d\Theta = \frac{k}{2E} \frac{(1-x) dx}{x^2(2-x)^2 \sin \pi x} ; \quad x \equiv \frac{\Theta}{\pi}$$

siendo  $E$  la energía total.

- Es común encontrar en los modelos de Física Nuclear, pozos rectangulares de potencial definidos por

$$V = \begin{cases} 0 & \text{si } r > a \\ -V_0 & \text{si } r \leq a \end{cases}$$

demuestre que la dispersión producida por este potencial en Mecánica Clásica es igual a la refracción de rayos luminosos por una esfera de radio  $a$ , e índice relativo de refracción  $n$  dado por

$$n = \sqrt{\frac{E + V_0}{E}}$$



esta equivalencia explica porqué los fenómenos de refracción podían ser explicados tanto por la teoría corpuscular de Newton como por las ondas de Huygens. Demuestre que la sección eficaz diferencial viene dada por

$$\sigma(\Theta) = \frac{n^2 a^2}{4 \cos \frac{\Theta}{2}} \frac{(n \cos \frac{\Theta}{2} - 1)(n - \cos \frac{\Theta}{2})}{(1 + n^2 - 2n \cos \frac{\Theta}{2})^2}$$

y encuentre la sección eficaz total. Esta clase de analogías entre dispersiones mecánicas y fenómenos ondulatorios es muy común en Física. De hecho, también existen numerosas analogías con ondas cuánticas de probabilidad.

# Interludio matemático: Matrices, vectores y tensores cartesianos

En el presente capítulo desarrollaremos desde un punto de vista práctico el álgebra básica de los vectores y matrices en el espacio cartesiano  $\mathbb{R}^n$ . Enfatizaremos en la relación que hay entre matrices y transformaciones lineales. Se asume cierta familiaridad del lector con el álgebra lineal ya que el propósito de este capítulo es el de dar un compendio de propiedades requeridas para los capítulos subsiguientes. Debe mencionarse sin embargo, que además de las propiedades matemáticas de los vectores y matrices se estudiarán aspectos de un interés más físico como son la interpretación pasiva y activa de las matrices, la clasificación de escalares y vectores por sus propiedades bajo paridad, y la caracterización de los tensores cartesianos.

## 12.1. Propiedades básicas de las matrices

Asumiremos que un vector cartesiano (en el espacio  $\mathbb{R}^n$ ) se puede representar como un arreglo de la forma

$$\mathbf{x} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{12.1}$$

una transformación de  $\mathbf{x}$  a otro vector  $\mathbf{x}'$  se denomina una transformación lineal homogénea<sup>1</sup>, si dicha transformación es de la forma

$$x'_i = a_{ij}x_j \quad ; \quad i, j = 1, \dots, n \tag{12.2}$$

donde  $a_{ij}$  son coeficientes constantes y se usa convención de suma sobre índices repetidos. Consideremos el caso de dos transformaciones lineales sucesivas de la forma (12.2). Adicionalmente, definiremos vectores  $\mathbf{x}$  que son generales (no necesariamente vectores posición). Hagamos primero la transformación de  $\mathbf{x}$  a  $\mathbf{x}'$  realizada por el operador  $\mathbf{B}$

$$\mathbf{x}' = \mathbf{B}\mathbf{x} \Rightarrow x'_k = b_{kj}x_j \tag{12.3}$$

seguida por una transformación de  $\mathbf{x}'$  a  $\mathbf{x}''$  realizada por el operador  $\mathbf{A}$

$$\mathbf{x}'' = \mathbf{A}\mathbf{x}' \Rightarrow x''_i = a_{ik}x'_k \tag{12.4}$$

---

<sup>1</sup>Una transformación lineal inhomogénea tiene la forma  $x'_i = a_{ij}x_j + b_i$ , donde las  $b_i$  son constantes. De aquí en adelante hablaremos de una transformación lineal sobreentendiendo que es homogénea, a menos que se indique lo contrario.

para obtener la relación entre  $x_i''$  y las  $x_j$  se pueden combinar las dos ecuaciones para obtener

$$\mathbf{x}'' = \mathbf{A}\mathbf{B}\mathbf{x} \quad \Rightarrow \quad \mathbf{x}'' = \mathbf{C}\mathbf{x} \quad \Rightarrow \quad \mathbf{C} = \mathbf{A}\mathbf{B} \quad (12.5)$$

$$x_i'' = a_{ik}b_{kj}x_j \quad \Rightarrow \quad x_i'' = c_{ij}x_j \quad (12.6)$$

$$c_{ij} \equiv a_{ik}b_{kj} \quad (12.7)$$

vemos que podemos definir una nueva transformación lineal  $\mathbf{C}$  caracterizada por los elementos  $c_{ij}$  que forman un nuevo arreglo matricial que se obtiene de los elementos  $a_{ik}$  y  $b_{kj}$  de  $\mathbf{A}$  y  $\mathbf{B}$  con el algoritmo descrito en (12.7). Esto nos indica que la composición de transformaciones lineales  $\mathbf{A}$  y  $\mathbf{B}$  nos define otra transformación lineal  $\mathbf{C}$ . Es fácil ver que el algoritmo (12.7) nos dice que la composición (“multiplicación”) de operadores no es conmutativa

$$\begin{aligned} (\mathbf{A}\mathbf{B})_{ij} &= c_{ij} = a_{ik}b_{kj} \quad ; \quad (\mathbf{B}\mathbf{A})_{ij} = d_{ij} = b_{ik}a_{kj} \\ &\Rightarrow \quad \mathbf{A}\mathbf{B} \neq \mathbf{B}\mathbf{A} \end{aligned}$$

lo cual se puede comprobar con unos simples casos particulares (ver ejercicio 1, Pág. 335). Esto implica que el resultado de aplicar dos operadores lineales sucesivos depende en general del orden de aplicación de tales operadores. Puede comprobarse sin embargo que esta multiplicación es asociativa (ver ejercicio 2, Pág. 336)

$$(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C})$$

la operación descrita por la Ec. (12.2) se puede escribir en arreglo matricial emulando el algoritmo (12.7) en donde los vectores se escriben como matrices de una columna en la forma de la Ec. (12.1). Por simplicidad, ilustraremos esta operación en tres dimensiones

$$\begin{aligned} \mathbf{x}' &= \mathbf{A}\mathbf{x} \quad ; \quad \mathbf{x}' \equiv \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} \quad ; \quad \mathbf{x} \equiv \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad ; \quad \mathbf{A} \equiv \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ (\mathbf{x}')_i &= (\mathbf{A}\mathbf{x})_i \quad \Rightarrow \quad x'_i = a_{ij}x_j = a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 \quad ; \quad i = 1, 2, 3 \end{aligned}$$

la suma de dos matrices se define como la matriz que se obtiene al sumar los elementos de cada una, claramente esta operación sí es conmutativa

$$(\mathbf{A} + \mathbf{B})_{ij} = a_{ij} + b_{ij} = c_{ij} = (\mathbf{C})_{ij}$$

un proceso importante es la composición inversa de  $\mathbf{A}$  que nos lleva de regreso desde  $\mathbf{x}'$  hacia  $\mathbf{x}$

$$\mathbf{x}' = \mathbf{A}\mathbf{x} \quad ; \quad \mathbf{A}^{-1}\mathbf{x}' = \mathbf{x} \quad \Rightarrow \quad \mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{x} \quad (12.8)$$

como  $\mathbf{x}$  es un vector arbitrario se concluye que

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{1} \quad (12.9)$$

donde  $\mathbf{1}$  es el operador identidad que deja al vector (o al sistema coordenado) inalterado. Por otro lado, multiplicando la segunda de las Ecs. (12.8) por  $\mathbf{A}$  a la izquierda

$$\mathbf{A}\mathbf{A}^{-1}\mathbf{x}' = \mathbf{A}\mathbf{x} \quad \Rightarrow \quad \mathbf{A}\mathbf{A}^{-1}\mathbf{x}' = \mathbf{x}'$$

y teniendo en cuenta que  $\mathbf{x}'$  también es arbitrario se tiene que

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{1} \quad (12.10)$$

de modo que la matriz inversa de  $\mathbf{A}$  debe cumplir

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{1} \quad (12.11)$$

designemos a los elementos de la inversa como  $\bar{a}_{ij}$ . La segunda de las Ecs. (12.8) se escribe como

$$x_i = \bar{a}_{ij}x'_j \quad (12.12)$$

que debe ser consistente con la primera de las Ecs. (12.8)

$$x'_k = a_{ki}x_i \quad (12.13)$$

y sustituyendo  $x_i$  de (12.12) en (12.13)

$$x'_k = a_{ki}\bar{a}_{ij}x'_j$$

y puesto que las componentes de  $\mathbf{x}'$  son independientes, esta ecuación solo es correcta si la suma se reduce idénticamente a  $x'_k$  i.e.

$$a_{ki}\bar{a}_{ij} = \delta_{kj} \quad (12.14)$$

la delta de Kronecker efectivamente es la única representación de los elementos de la matriz identidad como se puede ver del algoritmo (12.7). Podemos llegar también a (12.14) directamente a partir de (12.10) usando el algoritmo (12.7). Si partimos de (12.9) llegamos a la relación

$$\bar{a}_{ij}a_{jk} = \delta_{ik}$$

se puede ver además que la inversa del producto de dos matrices es la inversa del producto aplicado en orden contrario. Por definición de inverso se tiene que

$$\mathbf{ABx} = \mathbf{x}' \Rightarrow (\mathbf{AB})^{-1}\mathbf{x}' = \mathbf{x} \quad (12.15)$$

multiplicando por  $\mathbf{A}^{-1}$  a la izquierda de la primera ecuación (12.15) y posteriormente por  $\mathbf{B}^{-1}$  resulta

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{ABx} &= \mathbf{A}^{-1}\mathbf{x}' \Rightarrow \mathbf{Bx} = \mathbf{A}^{-1}\mathbf{x}' \Rightarrow \mathbf{B}^{-1}\mathbf{Bx} = \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{x}' \\ &\Rightarrow \mathbf{x} = \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{x}' \end{aligned}$$

comparando esta ecuación con la segunda de las Ecs. (12.15) teniendo en cuenta que estas deben ser válidas para  $\mathbf{x}$  y  $\mathbf{x}'$  arbitrarios, resulta

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad (12.16)$$

es importante tener en cuenta también que la matriz identidad deja inalterada a otra matriz cuando se hace el producto entre ellas

$$\mathbf{A}\mathbf{1} = \mathbf{1}\mathbf{A} = \mathbf{A}$$

finalmente, definiremos la traspuesta de una matriz como aquella matriz que se obtiene intercambiando filas por columnas y viceversa en la matriz  $\mathbf{A}$ , esta matriz simbolizada por  $\tilde{\mathbf{A}}$  se define entonces por

$$\left(\tilde{\mathbf{A}}\right)_{ij} \equiv \tilde{a}_{ij} = a_{ji}$$

es obvio que la traspuesta de la traspuesta es la misma matriz

$$\widetilde{(\tilde{\mathbf{A}})} = \mathbf{A} \quad (12.17)$$

es fácil demostrar que la traspuesta del producto de dos matrices es el producto de las matrices traspuestas en orden contrario

$$\left(\widetilde{\mathbf{B}\tilde{\mathbf{A}}}\right)_{ij} = \left(\tilde{\mathbf{B}}\right)_{ik} \left(\tilde{\mathbf{A}}\right)_{kj} = b_{ki}a_{jk} = a_{jk}b_{ki} = (\mathbf{AB})_{ji} = \widetilde{(\mathbf{AB})}_{ij}$$

y como esto vale para un elemento arbitrario  $i, j$  se tiene que

$$\widetilde{(\mathbf{AB})} = \tilde{\mathbf{B}}\tilde{\mathbf{A}} \quad (12.18)$$

es muy fácil verificar que la traspuesta de una suma es la suma de las traspuestas

$$\widetilde{(\mathbf{A} + \mathbf{B})} = \tilde{\mathbf{A}} + \tilde{\mathbf{B}}$$

**Matrices simétricas y antisimétricas**

Si una matriz coincide con su traspuesta

$$a_{ij} = a_{ji} \quad \Leftrightarrow \quad \mathbf{A} = \tilde{\mathbf{A}}$$

se dice que la matriz es **simétrica**. Si la matriz coincide con menos la traspuesta

$$a_{ij} = -a_{ji} \quad \Leftrightarrow \quad \mathbf{A} = -\tilde{\mathbf{A}}$$

entonces tenemos una matriz **antisimétrica**. Claramente, los elementos diagonales de una matriz antisimétrica son nulos.

Nótese que siempre es posible descomponer una matriz en una componente simétrica y otra antisimétrica de la siguiente forma

$$\mathbf{A} = \mathbf{A}_s + \mathbf{A}_a \quad ; \quad \mathbf{A}_s \equiv \frac{1}{2} (\mathbf{A} + \tilde{\mathbf{A}}) \quad ; \quad \mathbf{A}_a \equiv \frac{1}{2} (\mathbf{A} - \tilde{\mathbf{A}})$$

además, la traspuesta de la matriz también es combinación de estas componentes

$$\tilde{\mathbf{A}} = \mathbf{A}_s - \mathbf{A}_a$$

**12.1.1. Determinantes y trazas de matrices**

Otra propiedad importante de las matrices es el determinante. El determinante de una matriz  $\mathbf{A}$  se denota en una de estas dos formas

$$|\mathbf{A}| \quad \text{ó} \quad \det \mathbf{A}$$

usaremos una u otra notación de acuerdo con la conveniencia. El determinante consiste en un número real o complejo asociado a la matriz, su construcción está motivada por el estudio de las ecuaciones lineales simultáneas. Supondremos que el lector está familiarizado con el algoritmo de expansión por cofactores para el cálculo de los determinantes. Hablamos de matrices singulares cuando su determinante es nulo, de lo contrario hablamos de matrices no singulares. Es condición necesaria y suficiente que una matriz sea no singular para que exista su inversa, puesto que dicha inversa depende de  $|\mathbf{A}|^{-1}$ . El determinante de la traspuesta de la matriz coincide con el determinante de la matriz

$$|\tilde{\mathbf{A}}| = |\mathbf{A}| \quad \text{ó} \quad \det \tilde{\mathbf{A}} = \det \mathbf{A} \quad (12.19)$$

y para la matriz conjugada se tiene

$$|\mathbf{A}^*| = |\mathbf{A}|^* \quad \text{ó} \quad \det \mathbf{A}^* = (\det \mathbf{A})^* \quad (12.20)$$

Adicionalmente, se puede demostrar que el determinante del producto, es el producto de los determinantes

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}| \quad \text{ó} \quad \det (\mathbf{AB}) = (\det \mathbf{A}) \cdot (\det \mathbf{B}) \quad (12.21)$$

y dado que el determinante de la identidad es 1 se tiene que

$$1 = |\mathbf{1}| = |\mathbf{AA}^{-1}| = |\mathbf{A}| \cdot |\mathbf{A}^{-1}|$$

de modo que

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1} \quad \text{ó} \quad \det (\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1} \quad (12.22)$$

por otro lado si cualquier fila o columna de una matriz se multiplica por un escalar  $\alpha$ , el determinante queda multiplicado por tal escalar. Por ejemplo en tres dimensiones

$$\left| \begin{pmatrix} \alpha & a_{11} & \alpha & a_{12} & \alpha & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \right| = \left| \begin{pmatrix} a_{11} & \alpha & a_{12} & a_{13} \\ a_{21} & \alpha & a_{22} & a_{23} \\ a_{31} & \alpha & a_{32} & a_{33} \end{pmatrix} \right| = \alpha \left| \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \right| \quad (12.23)$$

de modo que si multiplicamos un escalar por una matriz  $n \times n$ , el nuevo determinante es

$$\det(\alpha \mathbf{A}) = \alpha^n \det \mathbf{A} \quad (12.24)$$

como caso particular

$$\det(-\mathbf{A}) = (-1)^n \det \mathbf{A} \quad (12.25)$$

otra propiedad importante es la traza de la matriz, definida como **la suma de sus elementos diagonales**

$$\text{Tr} \mathbf{A} = a_{ii} \quad (12.26)$$

donde enfatizamos en la suma sobre índices repetidos. Demostraremos que

$$\text{Tr} [\mathbf{AB}] = \text{Tr} [\mathbf{BA}] \quad (12.27)$$

lo cual se puede ver fácilmente en la forma

$$\text{Tr} [\mathbf{AB}] = (\mathbf{AB})_{ii} = a_{ik} b_{ki} = b_{ki} a_{ik} = (\mathbf{BA})_{kk} = \text{Tr} [\mathbf{BA}]$$

Una propiedad muy importante de la traza es que **es un invariante cíclico**, es decir

$$\begin{aligned} \text{Tr} [\mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-2)} \mathbf{A}^{(n-1)} \mathbf{A}^{(n)}] &= \text{Tr} [\mathbf{A}^{(n)} \mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-2)} \mathbf{A}^{(n-1)}] \\ &= \text{Tr} [\mathbf{A}^{(n-1)} \mathbf{A}^{(n)} \mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-2)}] \end{aligned} \quad (12.28)$$

y así sucesivamente. Para demostrarlo, basta con definir

$$\mathbf{B} \equiv \mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-1)}$$

de modo que

$$\text{Tr} [\mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-2)} \mathbf{A}^{(n-1)} \mathbf{A}^{(n)}] = \text{Tr} [\mathbf{BA}^{(n)}] = \text{Tr} [\mathbf{A}^{(n)} \mathbf{B}] = \text{Tr} [\mathbf{A}^{(n)} \mathbf{A}^{(1)} \mathbf{A}^{(2)} \dots \mathbf{A}^{(n-2)} \mathbf{A}^{(n-1)}]$$

y teniendo en cuenta que los índices (1), (2), ... son mudos, vemos que cualquier cambio cíclico es posible. Vale la pena tener en cuenta que la propiedad (12.27) no significa que para calcular la traza se puedan conmutar las matrices, por ejemplo para tres o más matrices la traza no es la misma para cualquier orden de las matrices, solo cambios cíclicos son posibles (ver ejercicio 1). En tal sentido debemos interpretar (12.27) como un cambio cíclico de dos matrices y no como una conmutación.

Veremos más adelante que las trazas y los determinantes son invariantes ante un cambio de base.

### 12.1.2. Matrices rectangulares

Una matriz rectangular es un arreglo de números constituido por  $m$  filas y  $n$  columnas. En ese caso se dice que la matriz es de dimensión  $m \times n$ . Los elementos de esta matriz serían de la forma

$$(\mathbf{A})_{ik} = a_{ik} \quad ; \quad i = 1, \dots, m \quad ; \quad k = 1, \dots, n$$

la traspuesta de esta matriz sería de dimensión  $n \times m$ . Un arreglo vectorial columna (de aquí en adelante lo llamaremos simplemente vector aunque no necesariamente sea un vector en todo el sentido de la palabra) es una matriz rectangular de dimensión  $m \times 1$ , su traspuesta es una matriz de dimensión  $1 \times m$ .

Ahora bien, sería deseable extrapolar el algoritmo (12.7) para realizar productos de matrices rectangulares

$$c_{ij} \equiv a_{ik} b_{kj}$$

Se observa que la extrapolación del producto matricial entre dos matrices rectangulares  $\mathbf{C} = \mathbf{AB}$  solo se puede definir consistentemente si el número de columnas de  $\mathbf{A}$  es el mismo que el número de filas de  $\mathbf{B}$ .

$$\mathbf{AB} = \mathbf{C} \quad \text{si} \quad \mathbf{A} \equiv \mathbf{A}_{m \times n} \quad \text{y} \quad \mathbf{B} \equiv \mathbf{B}_{n \times d} \quad \Rightarrow \quad \mathbf{C}_{m \times d}$$

En particular el producto de un vector columna (matriz  $m \times 1$ ) con una matriz  $m \times m$  en la forma  $\mathbf{x}\mathbf{A}$  no se puede definir. Sin embargo, el producto entre la traspuesta del vector anterior (vector fila) y la matriz  $\mathbf{A}$  en la forma  $\tilde{\mathbf{x}}\mathbf{A}$  se puede definir. Similarmente no se puede definir  $\mathbf{A}\tilde{\mathbf{x}}$  pero si se puede definir  $\mathbf{A}\mathbf{x}$ . De las anteriores consideraciones las cantidades  $\mathbf{A}\mathbf{x}$  y  $\tilde{\mathbf{x}}\mathbf{A}$  corresponden a un nuevo vector columna y un nuevo vector fila respectivamente.

A partir de las dimensiones para las matrices rectangulares

$$\mathbf{A}_{m \times n} \Rightarrow \tilde{\mathbf{A}}_{n \times m} \quad y \quad \mathbf{B}_{n \times d} \Rightarrow \tilde{\mathbf{B}}_{d \times n}$$

el producto  $\mathbf{AB}$  está definido. Sin embargo, se puede notar que sus traspuestas solo se pueden multiplicar en el orden opuesto, i.e. en el orden  $\tilde{\mathbf{B}}\tilde{\mathbf{A}}$ . De por sí es simple demostrar que al igual que con las matrices cuadradas el traspuesto del producto es el producto del traspuesto en orden inverso, Ec. (12.18). Aplicando esta propiedad se puede ver que

$$\widetilde{(\mathbf{A}\mathbf{x})} = \tilde{\mathbf{x}}\tilde{\mathbf{A}} \quad ; \quad \widetilde{(\tilde{\mathbf{x}}\mathbf{A})} = \tilde{\mathbf{A}}\mathbf{x}$$

donde hemos tenido en cuenta que la traspuesta de la traspuesta es la matriz original.

## 12.2. Interpretación activa y pasiva de las transformaciones lineales: cambios de base

Es usual describir un vector dado en dos sistemas coordenados diferentes, digamos  $X_1X_2X_3$  y  $X'_1X'_2X'_3$ . Asumiremos que ambos sistemas tienen un origen común, de modo que uno de ellos se obtiene por rotación del otro. Cada sistema coordenado tiene un conjunto ortonormal de vectores  $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$  y  $\{\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3\}$ .

La matriz de transformación  $\mathbf{A}$  puede ser considerada como un operador que actuando sobre el sistema coordenado no primado lo transforma en el sistema coordenado primado. Simbólicamente esta operación se escribe como:

$$(\mathbf{r})' = \mathbf{A}\mathbf{r} \tag{12.29}$$

el paréntesis indica que el vector  $\mathbf{r}$  como tal NO ha sido transformado, simplemente sus componentes se están escribiendo en un nuevo sistema coordenado. De modo que la matriz  $\mathbf{A}$  opera sobre las componentes del vector en el sistema no primado para dar las componentes **del mismo vector** en el sistema primado. Esta operación se conoce como un **cambio de base**, puesto que  $\mathbf{A}$  actúa solo sobre el sistema coordenado dejando al vector inalterado. En tres dimensiones la transformación de coordenadas es simplemente una rotación de los ejes coordenados. Usualmente estaremos interesados en transformaciones lineales que nos lleven de un sistema coordenado ortonormal a otro sistema coordenado también ortonormal, este tipo de transformaciones se denominan **transformaciones ortogonales**, las cuales caracterizaremos matemáticamente en la sección 12.4. Sin embargo, la discusión que estamos desarrollando no requiere estrictamente que las transformaciones sean ortogonales, la única condición indispensable como veremos más adelante es que la transformación sea invertible.

Por otro lado, es de anotar que sin cambiar el andamiaje matemático construido hasta ahora, podemos reinterpretar al operador  $\mathbf{A}$  como actuando sobre el vector  $\mathbf{r}$  para producir un nuevo vector  $\mathbf{r}'$  donde las componentes de ambos se calculan **en el mismo sistema coordenado**. Esto lo representamos por

$$\mathbf{r}' = \mathbf{A}\mathbf{r} \tag{12.30}$$

esta notación sin paréntesis indica que **el vector como tal ha sido transformado**, con ambos vectores  $\mathbf{r}$  y  $\mathbf{r}'$  descritos por **el mismo sistema coordenado**. Por tanto, en el caso particular de rotación en dos dimensiones en lugar de rotar el sistema coordenado en el sentido antihorario, estaríamos girando el vector posición en sentido horario manteniendo fijos los ejes. Los ángulos de rotación en ambos casos poseen naturalmente la misma magnitud pero signo opuesto. Vale destacar que en la transformación (12.30) las componentes del nuevo vector estarán relacionadas con las del vector original por las mismas transformaciones que en las Ecs. (12.29) aunque la interpretación geométrica sea distinta.

Para entender la diferencia, es esencial comprender que la definición completa de un operador requiere conocer no solo que operación realiza, sino sobre qué objetos actúa, en este sentido matemático estricto los dos operadores (cambio de base y cambio del vector) son en realidad diferentes aunque posean la misma representación matricial. Nótese en particular que el sentido de la rotación cuando se interpreta como cambio en el vector, es inverso con respecto al caso en que se interpreta como rotación del sistema coordinado.

No obstante visto desde un punto de vista práctico, dado que las operaciones matemáticas son las mismas, es posible y a veces conveniente cambiar de una interpretación a otra según las necesidades. Por ejemplo, en el caso de la especificación de la orientación de un cuerpo rígido relativo a un sistema de referencia  $S$ , la interpretación del operador como transformando el sistema coordinado es la más conveniente. Por otro lado la interpretación del operador como agente que cambia al vector tiene muchas aplicaciones. Por tanto, tomaremos una u otra interpretación a lo largo del texto.

Esta dualidad en la interpretación o más bien en el rol de los operadores está presente en otro tipo de transformaciones de coordenadas más generales que las transformaciones ortogonales. En ocasiones se interpreta su rol como el de cambiar el sistema coordinado expresando cierta cantidad o función en términos del nuevo sistema coordinado pero sin alterarla. En otras ocasiones se puede considerar como actuando sobre la cantidad o función en sí misma cambiando a nuevas cantidades en el mismo sistema coordinado. Cuando la transformación es del tipo de un cambio de base i.e. solo sobre el sistema coordinado, hablamos de un rol **pasivo** de la transformación. Por otro lado cuando al operador se le otorga el rol de cambiar a un vector u otra cantidad física, hablamos de un rol **activo** de la transformación. Ya nos habíamos encontrado con esta dualidad en el capítulo sobre transformaciones canónicas y en realidad es aplicable en muchos campos de la Física.

Es importante el hecho de que las propiedades matemáticas de las transformaciones activas y pasivas son idénticas, de modo que en la mayor parte de lo que sigue, no haremos distinción entre  $(\mathbf{r})'$  y  $\mathbf{r}'$  en las operaciones subsecuentes, a menos que se especifique lo contrario (como es el caso de la sección 12.2.1).

### 12.2.1. Transformaciones de similaridad

Veremos que las interpretaciones activa y pasiva de un operador se pueden combinar para encontrar la manera en la cual la representación de un operador activo se transforma cuando se realiza un cambio de base. Pensemos que tenemos una matriz **activa**  $\mathbf{A}$  que actúa sobre un vector  $\mathbf{F}$  (matriz columna) para transformarlo en el vector  $\mathbf{G}$ . Tenemos entonces la transformación

$$\mathbf{G} = \mathbf{A}\mathbf{F} \quad (12.31)$$

ahora pensemos que el sistema coordinado original será transformado a otro a través de una matriz ortogonal  $\mathbf{B}$ . El cambio de base se expresa como

$$(\mathbf{G})' = \mathbf{B}\mathbf{G} \quad (12.32)$$

donde revivimos la notación  $(\mathbf{G})'$  para indicar que representa al mismo vector pero en un sistema coordinado diferente. Combinando (12.31) y (12.32) se obtiene

$$\mathbf{B}\mathbf{G} = \mathbf{B}\mathbf{A}\mathbf{F} \Rightarrow (\mathbf{G})' = (\mathbf{B}\mathbf{A})\mathbf{F}$$

nótese que en la última igualdad los dos vectores  $\mathbf{G}$  y  $\mathbf{F}$  están expresados en bases diferentes. Es deseable que ambos queden escritos en la nueva base (por ejemplo dos vectores solo se pueden sumar componente a componente si ambos están expresados en la misma base). Esto se logra introduciendo un operador identidad en la forma

$$\begin{aligned} (\mathbf{G})' &= (\mathbf{B}\mathbf{A})(\mathbf{B}^{-1}\mathbf{B})\mathbf{F} \Rightarrow (\mathbf{G})' = (\mathbf{B}\mathbf{A}\mathbf{B}^{-1})(\mathbf{B}\mathbf{F}) \\ &\Rightarrow (\mathbf{G})' = (\mathbf{B}\mathbf{A}\mathbf{B}^{-1})(\mathbf{F})' \end{aligned} \quad (12.33)$$

Si comparamos (12.31) con (12.33) teniendo en cuenta que los vectores primados son idénticos a los vectores sin primar salvo que están escritos en bases diferentes, llegamos a la siguiente conclusión: La matriz de



transformación  $\mathbf{A}$  (en la base original) me produce la misma transformación que la matriz de transformación  $\mathbf{BAB}^{-1}$  en la nueva base. En otras palabras,  $\mathbf{A}$  y  $\mathbf{BAB}^{-1}$  son representaciones diferentes del mismo operador, donde se vé que la representación depende de la base que usemos para construirla. En consecuencia podemos definir

$$\mathbf{A}' = \mathbf{BAB}^{-1} \quad (12.34)$$

una transformación de esta forma (entre  $\mathbf{A}$  y  $\mathbf{A}'$ ) se denomina una transformación de similaridad. A las matrices  $\mathbf{A}'$  y  $\mathbf{A}$  se les denomina matrices equivalentes, lo cual es lógico si recordamos la interpretación activa de  $\mathbf{A}$  y pasiva de  $\mathbf{B}$  que nos llevó a observar que  $\mathbf{A}'$  representa en la nueva base al mismo operador activo que representa la matriz  $\mathbf{A}$  en la base antigua. Otra razón para llamar equivalentes a estas matrices es tener en cuenta que la transformación de similaridad forma una relación de equivalencia entre  $\mathbf{A}'$  y  $\mathbf{A}$  (reflexiva, simétrica y transitiva). Se deja como ejercicio al lector demostrar que la similaridad es una relación de equivalencia.

Puede demostrarse que tanto el determinante como la traza de una matriz  $\mathbf{A}$  son invariantes ante una transformación de similaridad. Esto es de gran importancia porque significa que estas cantidades están asociadas a un operador lineal de manera unívoca, sin importar la base en que se exprese tal operador.

Veamos la invarianza del determinante

$$\begin{aligned} |\mathbf{A}'| &= |\mathbf{BAB}^{-1}| = |\mathbf{B}| \cdot |\mathbf{A}| \cdot |\mathbf{B}^{-1}| = |\mathbf{B}| \cdot |\mathbf{A}| \cdot |\mathbf{B}|^{-1} \\ &\Rightarrow |\mathbf{A}'| = |\mathbf{A}| \end{aligned}$$

donde hemos usado (12.21) y (12.22). La invarianza de la traza se demuestra fácilmente en forma explícita

$$Tr[\mathbf{A}'] = Tr[\mathbf{BAB}^{-1}] = \sum_{i=1}^n (\mathbf{BAB}^{-1})_{ii} = \sum_{i,k,l} b_{ik} a_{kl} \bar{b}_{li} = \sum_{i,k,l} \bar{b}_{li} b_{ik} a_{kl} = \sum_{k,l} \delta_{lk} a_{kl} = \sum_k a_{kk} = Tr\mathbf{A}$$

una forma más inmediata de verlo es teniendo en cuenta que la traza es un invariante cíclico (ver Ec. 12.28), de modo que

$$Tr[\mathbf{A}'] = Tr[\mathbf{BAB}^{-1}] = Tr[\mathbf{B}^{-1}\mathbf{BA}] = Tr\mathbf{A}$$

nótese que las propiedades de las transformaciones de similaridad aquí mostradas, no dependen de que la matriz  $\mathbf{B}$  para el cambio de base sea ortogonal. Lo único que se pide es que esta matriz sea no singular para que exista la inversa. En realidad la definición de transformación de similaridad no se restringe al caso en el cual la matriz  $\mathbf{B}$  sea ortogonal, aunque el uso de estas transformaciones nos facilitará realizar una interpretación geométrica, ya que una transformación ortogonal transforma una base ortonormal en otra base ortonormal.

### 12.3. Problema de valores propios

Sea  $\mathbf{A}$  una matriz  $n \times n$  y  $\mathbf{x}$  un vector columna  $n \times 1$ . Para una matriz específica  $\mathbf{A}$  definimos la ecuación

$$\mathbf{Ax} = \lambda \mathbf{x} \quad (12.35)$$

como la ecuación de valores propios asociada a dicha matriz. La idea es solucionar para los posibles valores de  $\lambda$  y vectores asociados  $\mathbf{x}$ . Los valores  $\lambda$  son en general complejos y se conocen como **valores propios** en tanto que los vectores  $\mathbf{x}$  asociados se denominan **vectores propios**. La operación en (12.35) nos muestra que la ecuación de valores propios consiste en la búsqueda de vectores para los cuales el operador  $\mathbf{A}$  nos “alarga” o nos “acorta” el vector en cuestión, eventualmente cambiando su “sentido” pero siempre conservando su “dirección”. El valor propio sería entonces el factor de “contracción” o “dilatación” y en caso de ser negativo invertiría el “sentido” del vector original<sup>2</sup>.

<sup>2</sup>La razón para colocar todos estos términos entre comillas, es que tal interpretación geométrica solo es clara para vectores reales. Si bien la analogía para vectores complejos no está exenta de utilidad.

La ecuación de valores propios o autovalores, se puede escribir de la forma

$$(\mathbf{A} - \lambda \mathbf{1}) \mathbf{x} = 0 \quad (12.36)$$

en forma expandida en tres dimensiones se puede escribir

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + a_{13}x_3 &= 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + a_{23}x_3 &= 0 \\ a_{31}x_1 + a_{32}x_2 + (a_{33} - \lambda)x_3 &= 0 \end{aligned} \quad (12.37)$$

los argumentos que siguen son válidos para matrices de dimensión arbitraria finita, pero por comodidad los escribiremos en tres dimensiones. Este conjunto de ecuaciones homogéneas para  $x_1, x_2, x_3$  solo tiene solución no trivial si el determinante del sistema es nulo de modo que

$$\det(\mathbf{A} - \lambda \mathbf{1}) = |\mathbf{A} - \lambda \mathbf{1}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0 \quad (12.38)$$

condición que se conoce como **ecuación secular o característica** de la matriz. Los valores de  $\lambda$  para los cuales se satisface esta ecuación son los autovalores o valores propios asociados a la matriz. Es de anotar que aún en el caso en el cual existen soluciones no triviales, el conjunto de ecuaciones homogéneas (12.37) no nos da valores definidos para las tres (o para las  $n$ ) componentes sino solo cocientes entre éstas. Este hecho se puede entender por argumentos algebraicos o geométricos. Desde el punto de vista algebraico, esto está relacionado con el hecho de que el producto del autovector  $\mathbf{x}$  con un escalar también sería autovector, lo cual se puede ver inmediatamente de (12.36)<sup>3</sup>. Geométricamente, esto implica que solo la dirección del autovector está determinada pero su magnitud (e incluso sentido) permanecen indeterminados, lo cual se ve particularmente claro en tres dimensiones. Dado que  $\mathbf{A}$  representa una transformación lineal es claro que si  $\mathbf{A}$  conserva la dirección del vector  $\mathbf{x}$  i.e.  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  también conservará la dirección del vector  $\alpha\mathbf{x}$  para  $\alpha$  arbitrario.

Al expandir el determinante (12.38) se observa que en general la solución de la ecuación secular se reduce a encontrar las raíces de un polinomio de grado  $n$ . Sabemos pues que hay un número máximo de  $n$  de éstas raíces y que tales raíces pueden ser complejas. En general, podemos construir a lo más  $n$  vectores linealmente independientes  $\mathbf{x}_k$  cada uno asociado a un valor propio  $\lambda_k$ . Denotaremos la componente  $i$ -ésima del  $k$ -ésimo vector propio en la forma  $x_{ik}$ . Podemos realizar un arreglo matricial con estos vectores, colocándolos como vectores columna en forma adyacente. En tres dimensiones tal arreglo matricial queda en la forma

$$\mathbf{X} \equiv (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3) = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \quad (12.39)$$

las Ecs. (12.36) se escriben para cada valor propio  $\lambda_k$  y su respectivo vector propio  $\mathbf{x}_k$  en la forma

$$(\mathbf{A} - \lambda_k \mathbf{1}) \mathbf{x}_k = 0 \Rightarrow \mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k \quad \text{no suma sobre } k \quad (12.40)$$

escribiendo las Ecs. (12.40) en componentes resulta (para  $n$  dimensiones)

$$\begin{aligned} \sum_{j=1}^n a_{ij}x_{jk} &= \lambda_k x_{ik} \Rightarrow \\ \sum_{j=1}^n a_{ij}x_{jk} &= \sum_{j=1}^n x_{ij}\delta_{jk}\lambda_k \end{aligned} \quad (12.41)$$

<sup>3</sup>Alternativamente, esto se puede ver del hecho de que la ecuación secular solo tiene solución no trivial cuando una o más de las ecuaciones es linealmente dependiente con las demás. En tal caso hay más incógnitas que ecuaciones y tenemos entonces infinitas soluciones.

en las dos ecuaciones anteriores no hay suma sobre el índice repetido  $k$ . Recordemos que  $x_{jk}$  es la componente  $j$ -ésima del vector  $\mathbf{x}_k$ . Ahora bien, la cantidad  $\delta_{jk}\lambda_k \equiv \lambda_{jk}$  se puede asociar a una matriz diagonal que en tres dimensiones se escribe

$$\boldsymbol{\lambda} \equiv \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad (12.42)$$

matricialmente la Ec. (12.41) se escribe como

$$\mathbf{A}\mathbf{X} = \mathbf{X}\boldsymbol{\lambda}$$

si multiplicamos por  $\mathbf{X}^{-1}$  a la izquierda se obtiene

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \boldsymbol{\lambda} \quad (12.43)$$

esta operación es una transformación de similaridad operando sobre  $\mathbf{A}$ . Nótese que además la matriz  $\mathbf{X}$  que se construye con los vectores propios es la matriz de transformación (más estrictamente la matriz de transformación es  $\mathbf{X}^{-1}$ ), dicha matriz  $\mathbf{X}^{-1}$  diagonaliza a  $\mathbf{A}$  a través de la transformación de similaridad y los elementos de la diagonal corresponden a los valores propios ( $\lambda_k$  asociado al vector columna  $\mathbf{x}_k$  en la matriz  $\mathbf{X}$  en la Ec. 12.39).

Es obvio de la Ec. (12.43), que la condición necesaria y suficiente para que la diagonalización de  $\mathbf{A}$  sea posible, es que la matriz  $\mathbf{X}$  admita inversa. A su vez una condición necesaria y suficiente para esto, es que los vectores propios que constituyen a la matriz  $\mathbf{X}$  sean un conjunto linealmente independiente. En otras palabras, la diagonalización de una matriz  $n \times n$  es posible si y solo si, existen  $n$  vectores propios linealmente independientes de dicha matriz. Es decir, los vectores propios linealmente independientes de la matriz deben formar una base de  $\mathbb{R}^n$ .

Surge entonces la pregunta natural de cuales son las condiciones para que dada una matriz  $n \times n$ , exista un conjunto de  $n$  vectores propios linealmente independientes. Ya hemos mencionado que la ecuación secular nos lleva a buscar las raíces de un polinomio de grado  $n$ . En algunas ocasiones hay raíces degeneradas, es decir que se repiten dos o más veces. Cuando existen valores propios degenerados, es decir que un subconjunto de ellos posee el mismo valor, no es siempre posible diagonalizar la matriz  $\mathbf{A}$ . Volveremos sobre este punto en la sección 12.12.

Por otro lado, en el caso en que la diagonalización es posible, el determinante y la traza de  $\mathbf{A}$  se pueden calcular teniendo en cuenta que tales cantidades son invariantes ante una transformación de similaridad, por lo tanto

$$\det \mathbf{A} = \det [\mathbf{X}^{-1}\mathbf{A}\mathbf{X}] = \det \boldsymbol{\lambda} = \lambda_1 \lambda_2 \dots \lambda_n \quad (12.44)$$

$$\text{Tr} \mathbf{A} = \text{Tr} [\mathbf{X}^{-1}\mathbf{A}\mathbf{X}] = \text{Tr} \boldsymbol{\lambda} = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad (12.45)$$

de modo que el determinante y la traza de una matriz diagonalizable son simplemente el producto y la suma de sus valores propios respectivamente.

Finalmente, si la matriz  $\mathbf{A}$  en la Ec. (12.35) es real, podemos ver tomando el conjugado en dicha ecuación, que si  $\mathbf{x}$  es un vector propio de  $\mathbf{A}$  asociado a un valor propio  $\lambda$ , entonces  $\mathbf{x}^*$  también es vector propio de  $\mathbf{A}$  y estará asociado al valor propio  $\lambda^*$ .

### 12.3.1. El problema de la degeneración de valores propios

Dado un valor propio  $\lambda_i$  de una matriz  $\mathbf{A}$ , existen infinitos vectores propios asociados ya que si  $\mathbf{x}$  es vector propio, todos los vectores de la forma  $\alpha\mathbf{x}$  (con  $\alpha$  escalar) también son vectores propios asociados a  $\lambda_i$ . Sin embargo, todos estos vectores son linealmente dependientes entre sí. Ocurre en ocasiones que existen varios vectores linealmente independientes asociados a un mismo valor propio. El grado de **degeneración** de un valor propio  $\lambda_i$  lo definiremos como el máximo número de vectores propios linealmente independientes asociados a  $\lambda_i$ . Si un valor propio  $\lambda_i$  es  $p$ -degenerado, tenemos un conjunto de  $p$ -vectores linealmente independientes

$$\lambda_i \rightarrow \left\{ \mathbf{x}_i^{(1)}, \mathbf{x}_i^{(2)}, \dots, \mathbf{x}_i^{(p)} \right\} \quad (12.46)$$

donde el subíndice indica el valor propio asociado y el superíndice es el índice de degeneración. Si  $\mathbf{A}$  es una matriz  $n \times n$ , sus vectores propios estarán en  $\mathbb{R}^n$  ó  $\mathbb{C}^n$  que genéricamente lo escribiremos como  $\mathcal{E}^n$ . Por tanto  $p \leq n$ , y todas las combinaciones lineales de los vectores del tipo  $\mathbf{x}_i^{(k)}$  generan un subespacio del tipo  $\mathcal{E}^p$ . Es fácil ver que un vector arbitrario no nulo<sup>4</sup> de  $\mathcal{E}^p$  es vector propio de  $\mathbf{A}$  con valor propio  $\lambda_i$ . Para verlo tomemos un vector arbitrario  $\mathbf{x} \in \mathcal{E}^p$ , tal vector es una combinación lineal de los vectores en la Ec. (12.46), y dado que  $\mathbf{A}$  define una transformación lineal tenemos

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{A} \left[ \alpha_m \mathbf{x}_i^{(m)} \right] = \alpha_m \mathbf{A}\mathbf{x}_i^{(m)} = \alpha_m \lambda_i \mathbf{x}_i^{(m)} = \lambda_i \left[ \alpha_m \mathbf{x}_i^{(m)} \right] \Rightarrow \\ \mathbf{A}\mathbf{x} &= \lambda_i \mathbf{x} \end{aligned}$$

que demuestra lo que se pide. Adicionalmente, puesto que (12.46) define el máximo conjunto de vectores propios linealmente independientes asociados a  $\lambda_i$ , es claro que ningún vector propio asociado a  $\lambda_i$  está por fuera de  $\mathcal{E}^p$ , resumimos este resultado en la siguiente forma

**Theorem 8** *Sea  $\lambda_i$  un valor propio asociado a la matriz  $\mathbf{A}$  de dimensión  $n \times n$ . El conjunto  $\mathcal{E}_i$  de todos los vectores propios asociados a este valor propio junto con el cero, forma un subespacio vectorial de  $\mathbb{C}^n$  con dimensión  $p \leq n$ . Siendo  $p$  el grado de degeneración del valor propio  $\lambda_i$ .*

Esto además implica que si el conjunto de vectores dado en (12.46) no es ortonormal, puede ortonormalizarse (por ejemplo con un proceso de ortogonalización de Gram-Schmidt), lo cual equivale a escoger una base ortonormal de vectores de  $\mathcal{E}_i$ , que obviamente son vectores propios de  $\mathbf{A}$  asociados a  $\lambda_i$ .

De lo anterior es fácil ver que si  $\mathbf{A}$  es real y un valor propio dado  $\lambda_i$  es real, siempre es posible escoger una base de vectores **reales** en  $\mathcal{E}_i$ . Tomemos un vector  $\mathbf{x}_i^{(k)}$  de la base (12.46), suprimiremos el supraíndice por comodidad. La ecuación de valores propios y su conjugada nos dan

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad ; \quad \mathbf{A}\mathbf{x}_i^* = \lambda_i \mathbf{x}_i^* \quad \text{no suma sobre } i \quad (12.47)$$

donde hemos tomado en cuenta el carácter real de  $\mathbf{A}$  y  $\lambda$ . Supongamos primero que  $\mathbf{x}_i$  es linealmente dependiente de  $\mathbf{x}_i^*$ . Por lo menos una de las combinaciones lineales dadas por

$$\frac{\mathbf{x}_i + \mathbf{x}_i^*}{2} = \text{Re } \mathbf{x}_i \quad ; \quad \frac{i(\mathbf{x}_i - \mathbf{x}_i^*)}{2} = \text{Im } \mathbf{x}_i \quad (12.48)$$

es un vector no nulo, y será vector propio con componentes reales asociado al mismo valor propio. En todo caso, si ambos vectores son diferentes de cero, serán linealmente dependientes.

Supongamos ahora que  $\mathbf{x}_i$  es linealmente independiente de  $\mathbf{x}_i^*$ , como ambos son vectores propios asociados a  $\lambda_i$ , esto implica que  $\lambda_i$  tiene degeneración por lo menos de orden 2. Es claro que el espacio dos dimensional generado por  $\mathbf{x}_i$  y  $\mathbf{x}_i^*$  consta de vectores propios con valor propio  $\lambda_i$  (si bien no es necesariamente el subespacio más grande que cumple esta condición). La misma combinación lineal definida en (12.48), nos da en este caso dos vectores **reales** y linealmente independientes<sup>5</sup>, que generan el mismo subespacio que  $\mathbf{x}_i$  y  $\mathbf{x}_i^*$ . Procediendo de esta forma con todos los vectores de la Ec. (12.46), podemos encontrar un conjunto linealmente independiente de vectores reales que generen  $\mathcal{E}_i$ . Vale decir sin embargo, que no todos los vectores reales obtenidos de esta manera van a ser necesariamente independientes. Por ejemplo,  $\mathbf{x}_i^{*(1)}$  podría ser linealmente dependiente con  $\mathbf{x}_i^{(2)}$  en esta caso el espacio generado por  $\text{Re}\mathbf{x}_i^{*(1)}$  y  $\text{Im}\mathbf{x}_i^{*(1)}$  contiene a  $\mathbf{x}_i^{(2)}$  aunque no necesariamente a  $\mathbf{x}_i^{*(2)}$ . Por tanto, los vectores

$$\text{Re}\mathbf{x}_i^{*(1)}, \quad \text{Im}\mathbf{x}_i^{*(1)}, \quad \text{Re}\mathbf{x}_i^{*(2)}, \quad \text{Im}\mathbf{x}_i^{*(2)}$$

<sup>4</sup>Un vector propio es por definición diferente de cero.

<sup>5</sup>La independencia lineal de  $\text{Re } \mathbf{x}_i$  e  $\text{Im } \mathbf{x}_i$  se garantiza dado que el conjunto original  $\mathbf{x}_i$  y  $\mathbf{x}_i^*$  es linealmente independiente por hipótesis, y la transformación

$$\begin{pmatrix} \text{Re } \mathbf{x}_i \\ \text{Im } \mathbf{x}_i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \mathbf{x}_i \\ \mathbf{x}_i^* \end{pmatrix}$$

es no singular.

no van a ser todos linealmente independientes. No obstante, es claro que al barrer todos los vectores de la Ec. (12.46) obtendremos una base de  $n$  vectores reales, más posiblemente algunos vectores linealmente dependientes. Por otro lado, esta base se puede ortonormalizar por ejemplo con un proceso de ortonormalización de Gram-Schmidt con combinaciones lineales reales, de modo que la base ortonormal obtenida sea también real. Resumimos estos resultados en la forma

**Theorem 9** Sea  $\mathbf{A}$  una matriz real  $n \times n$  con un valor propio real  $\lambda_i$  que genera el subespacio  $\mathcal{E}_i$  de  $\mathcal{E}^n$ . Entonces existe una base de vectores reales que genera a  $\mathcal{E}_i$ . En particular, existe una base ortonormal real que genera a  $\mathcal{E}_i$ .

Nótese que el teorema no dice que los vectores propios asociados a  $\lambda_i$  no puedan ser complejos, de hecho si  $\mathbf{x}$  es vector propio real, entonces  $e^{i\theta}\mathbf{x}$  es un vector propio complejo. Así mismo, dados varios vectores propios reales de  $\mathcal{E}_i$ , una combinación lineal compleja también pertenece a  $\mathcal{E}_i$  si este espacio es del tipo  $\mathbb{C}^n$ . El teorema tampoco dice que los valores propios de una matriz real sean necesariamente reales, de hecho veremos muy pronto que los valores propios de una matriz real pueden ser reales o complejos.

## 12.4. Propiedades básicas de las matrices ortogonales

Del conjunto de transformaciones lineales

$$x'_i = a_{ij}x_j \quad , \quad i = 1, 2, 3 \quad (12.49)$$

existe un conjunto particularmente interesante para la Física: el conjunto más general de transformaciones lineales homogéneas<sup>6</sup>, que deja invariante la norma de los vectores con componentes reales. Si un vector real arbitrario  $\mathbf{r}$  se escribe en la forma  $\mathbf{r} = x_i\mathbf{u}_i$  siendo  $x_i$  valores reales, entonces las transformaciones lineales (activas o pasivas) de la forma (12.49) que mantengan invariante la norma de vectores reales, deben cumplir la condición

$$x_i x_i = x'_i x'_i \quad (12.50)$$

usando (12.49) esta ecuación se reescribe como

$$\begin{aligned} x_i x_i &= (a_{ij}x_j)(a_{ik}x_k) \\ \delta_{jk}x_j x_k &= a_{ij}a_{ik}x_j x_k \end{aligned}$$

donde hemos tenido en cuenta que los índices repetidos son mudos. Como las coordenadas  $x_k, x_j$  son arbitrarias entonces la relación anterior es cierta si y solo si

$$a_{ij}a_{ik} = \delta_{jk} \quad ; \quad j, k = 1, 2, 3 \quad (12.51)$$

el lector puede demostrar que esta condición tiene la siguiente forma alternativa

$$a_{ji}a_{ki} = \delta_{jk} \quad ; \quad j, k = 1, 2, 3 \quad (12.52)$$

A cualquier transformación lineal de la forma (12.49) que cumpla las condiciones (12.51) (o su forma alternativa las Ecs. 12.52) se le denomina una **transformación ortogonal**. Las Ecs. (12.51) se conocen como **condiciones de ortogonalidad**.

Consideremos las condiciones de ortogonalidad (12.51) y escribámoslas en notación matricial

$$a_{ij}a_{ik} = \delta_{jk} \Rightarrow \tilde{a}_{ji}a_{ik} = \delta_{jk} \Rightarrow \left(\tilde{\mathbf{A}}\mathbf{A}\right)_{jk} = \mathbf{1}_{jk}$$

tenemos que

$$\tilde{\mathbf{A}}\mathbf{A} = \mathbf{1} \Rightarrow \tilde{\mathbf{A}}\mathbf{A}\mathbf{A}^{-1} = \mathbf{1}\mathbf{A}^{-1} \Rightarrow \tilde{\mathbf{A}}\mathbf{1} = \mathbf{A}^{-1}$$

<sup>6</sup>Volvemos a insistir en el carácter homogéneo de las transformaciones lineales, ya que existen transformaciones lineales inhomogéneas (e.g. las traslaciones) que dejan invariante la norma de un vector.

con lo cual se llega a la relación

$$\tilde{\mathbf{A}} = \mathbf{A}^{-1} \Leftrightarrow \tilde{\mathbf{A}}\mathbf{A} = \mathbf{A}\tilde{\mathbf{A}} = \mathbf{1} \quad (12.53)$$

Inversamente se puede demostrar que la condición anterior nos conduce a las condiciones de ortogonalidad. De modo que para las matrices ortogonales la inversa debe coincidir con la traspuesta.

Nótese que la condición  $\tilde{\mathbf{A}}\mathbf{A} = \mathbf{1}$  se escribe explícitamente como la Ec. (12.51), en tanto que la condición  $\mathbf{A}\tilde{\mathbf{A}} = \mathbf{1}$ , nos lleva a la condición de ortogonalidad (equivalente) dada por la Ec. (12.52). Mnemotécnicamente, los elementos matriciales del producto  $\tilde{\mathbf{A}}\mathbf{A}$  ( $\mathbf{A}\tilde{\mathbf{A}}$ ), se obtienen como un producto punto entre los arreglos vectoriales formados por las columnas (filas) de la matriz  $\mathbf{A}$ , (ver Ec. 13.13 en pag 343 y comentarios). Por tanto, en una matriz ortogonal el producto punto entre diferentes filas, o entre diferentes columnas debe ser cero y el producto punto de una fila o columna consigo misma debe ser uno.

De la condición (12.53) se puede ver que el inverso de una matriz ortogonal es también una matriz ortogonal. En el caso de las rotaciones (que son operaciones que claramente conservan la norma de los vectores) esto es de esperarse ya que geoméricamente sabemos que si realizamos una rotación, existe la rotación inversa que me devuelve a la configuración original.

En lo que concierne al determinante de una matriz ortogonal, vemos que al combinar (12.19) (12.21) y (12.53) se obtiene que

$$\begin{aligned} \det(\tilde{\mathbf{A}}\mathbf{A}) &= 1 = (\det \tilde{\mathbf{A}})(\det \mathbf{A}) \Rightarrow \\ (\det \mathbf{A})^2 &= 1 \quad ; \quad \det \mathbf{A} = \pm 1 \end{aligned}$$

vemos entonces que el determinante de una matriz ortogonal es más o menos uno. Por razones que veremos más adelante, las transformaciones ortogonales de determinante +1 se denominan **propias** y las de determinante -1 se denominan **impropias**.

De otra parte, existe otra forma geométrica de caracterizar a las matrices ortogonales reales en un enfoque pasivo. Este es el conjunto más general de matrices reales que me convierte una base ortonormal real en otra base ortonormal real, esto se demostrará como caso especial de un contexto más general en la sección 12.9. Esta característica es muy importante ya que usualmente estamos interesados en pasar de un conjunto de ejes ortogonales a otro conjunto de ejes también ortogonales y adicionalmente conservar la unitariedad de los vectores base.

Cuando el autovalor es complejo, entonces en general su correspondiente vector propio también tendrá componentes complejas. Cuando los vectores son complejos, la magnitud al cuadrado de estos se calcula usualmente de la forma

$$\|\mathbf{x}\|^2 = |x_i| |x_i| = x_i^* x_i \quad (12.54)$$

y la condición de invarianza de esta norma no coincide con la condición de invarianza (12.50). Examinemos en detalle algunos aspectos de las invarianzas expresadas por (12.50) y (12.54).

#### 12.4.1. Matrices ortogonales y norma de vectores complejos (opcional)

Las matrices ortogonales han sido originalmente construídas para mantener invariante la norma de vectores con componentes reales, dicha norma se define usualmente en la forma descrita por la Ec. (12.50)

$$x_i x_i = x'_i x'_i \quad (12.55)$$

debemos tener presente que en la demostración no se usa en ningún momento el carácter real de los  $x_i$  ni tampoco de las componentes de la matriz  $a_{ij}$ . En conclusión, aunque la motivación original involucra solo números reales los resultados nos muestran que podemos hacer un enunciado más general

**Theorem 10** Sea  $\mathbf{x}^T = (x_1, \dots, x_n)$  un arreglo vectorial con componentes complejas. El conjunto de transformaciones lineales homogéneas más general que deja invariante la cantidad  $x_i x_i$ , es el conjunto de matrices complejas  $\mathbf{A}$  que cumplen la condición  $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$ .

El punto es que en la mayoría de los casos, la norma de vectores complejos no se define como  $x_i x_i$  sino en la forma  $x_i x_i^*$  como ya se discutió. No obstante, existe una importante excepción, en relatividad especial cuando usamos el espacio de Minkowski con eje temporal imaginario, definimos los vectores complejos  $(x_1, x_2, x_3, ict)$  y su módulo cuadrado<sup>7</sup> se define como  $x_i x_i = x_1^2 + x_2^2 + x_3^2 - c^2 t^2$ .

Veamos ahora como actúan las matrices ortogonales sobre el módulo cuadrado de un vector complejo cuando tal módulo se define en la forma  $x_i x_i^*$ , siguiendo el algoritmo para multiplicación de matrices rectangulares se tiene que

$$\|\mathbf{x}\|^2 = \widetilde{\mathbf{x}}^* \mathbf{x} \quad ; \quad \mathbf{x}^\dagger \equiv \widetilde{\mathbf{x}}^*$$

donde hemos definido la conjugada transpuesta del vector como el **adjunto** del vector simbolizado  $\mathbf{x}^\dagger$ . Escribamos entonces

$$\|\mathbf{x}\|^2 = \mathbf{x}^\dagger \mathbf{x}$$

usando la propiedad (12.53) se tiene que si  $\mathbf{x}' = \mathbf{A}\mathbf{x}$  siendo  $\mathbf{A}$  una matriz ortogonal, la norma de  $\mathbf{x}'$  se escribe como

$$\|\mathbf{x}'\|^2 = \widetilde{\mathbf{x}'}^* \mathbf{x}' = (\widetilde{\mathbf{A}\mathbf{x}})^* \mathbf{A}\mathbf{x} = \widetilde{\mathbf{x}}^* \widetilde{\mathbf{A}}^* \mathbf{A}\mathbf{x} = \widetilde{\mathbf{x}}^* (\mathbf{A}^{-1})^* \mathbf{A}\mathbf{x} = \widetilde{\mathbf{x}}^* (\mathbf{A}^*)^{-1} \mathbf{A}\mathbf{x}$$

Ahora bien si la matriz es ortogonal **real** se tiene que  $\mathbf{A} = \mathbf{A}^*$  y se obtiene la invarianza de la norma usual para vectores complejos de modo que se llega al siguiente

**Theorem 11** *Bajo una transformación ortogonal real, la magnitud al cuadrado de vectores complejos definida en la forma  $x_i x_i^*$  permanece invariante.*

Sin embargo, las matrices ortogonales reales no son el conjunto más general de matrices que dejan invariante la norma de vectores complejos. Más adelante caracterizaremos al conjunto más general de matrices (complejas) que dejan invariante la norma de números complejos, este conjunto se denomina matrices unitarias y contiene a las matrices ortogonales reales como subconjunto propio.

Por otro lado multiplicando la Ec. (12.35) por su conjugada transpuesta, podemos escribir

$$\begin{aligned} (\widetilde{\mathbf{A}\mathbf{x}})^* (\mathbf{A}\mathbf{x}) &= (\widetilde{\lambda\mathbf{x}})^* (\lambda\mathbf{x}) \Rightarrow \\ \widetilde{\mathbf{x}'}^* \mathbf{x}' &= \lambda\lambda^* \widetilde{\mathbf{x}}^* \mathbf{x} \end{aligned}$$

y si imponemos la invarianza de la norma del vector complejo ante la transformación se tiene que

$$\|\lambda\|^2 = 1 \tag{12.56}$$

en conclusión, todos los valores propios de una matriz que preserve la norma usual de los vectores complejos tienen norma unidad<sup>8</sup>.

#### 12.4.2. Transformaciones ortogonales propias e impropias

Previamente hemos encontrado que las matrices ortogonales solo pueden tener determinante  $\pm 1$ . A continuación mostraremos que una matriz ortogonal que describe a una transformación continua, no puede ser descrita por una matriz de determinante  $-1$ . Consideremos la matriz ortogonal más simple con determinante  $-1$  que es menos la identidad

$$\mathbf{S} = -\mathbf{1}_{3 \times 3}$$

esta transformación tiene la propiedad de cambiar el signo de cada una de las componentes de los ejes coordenados (si la transformación es pasiva). Esta operación nos convierte un sistema coordinado dextrógiro (de mano derecha) en uno levógiro (de mano izquierda) como se aprecia en la Fig. 12.1, esta operación se denomina

<sup>7</sup>Estrictamente, el módulo al cuadrado aquí descrito define una pseudonorma en lugar de una norma, ya que el módulo al cuadrado puede ser negativo en algunos casos.

<sup>8</sup>Nótese que en este caso la suposición de que la matriz sea real no es necesaria, de modo que las matrices complejas más generales que preservan la norma de los vectores complejos (matrices unitarias) también poseen valores propios con norma unidad.

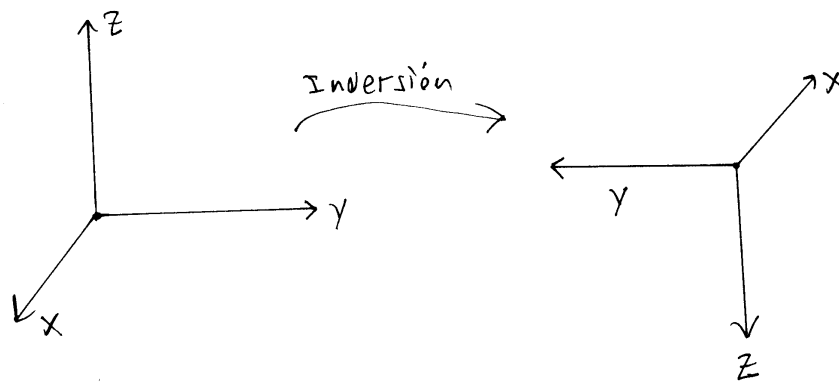


Figura 12.1: Conversión de un sistema coordenado dextrógiro en uno levógiro (cambio de quiralidad). En la figura de la derecha, el eje  $X$  va hacia adentro del papel. Nótese que en el sistema dextrógiro (de mano derecha) se tiene que  $\mathbf{u}_x \times \mathbf{u}_y = \mathbf{u}_z$  en tanto que en el sistema levógiro (de mano izquierda) se tiene que  $\mathbf{u}_x \times \mathbf{u}_y = -\mathbf{u}_z$ .

**inversión de los ejes coordenados**<sup>9</sup>. Una forma de realizar esta operación es rotar alrededor de un cierto eje en  $180^\circ$ , para luego realizar una reflexión con respecto al plano definido por dicho eje y que pasa por el origen de coordenadas. Si tomamos por ejemplo el eje  $Z$  para la operación de girar  $180^\circ$ , vemos que esta operación debe realizar la transformación  $x, y, z \rightarrow -x, -y, z$ ; la cual se puede obtener aplicando sobre el vector columna  $(x, y, z)$  la siguiente matriz

$$\text{rotación } 180^\circ \text{ alrededor de } Z = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

la matriz de reflexión debe convertir  $z \rightarrow -z$  y dejar inalteradas las otras coordenadas

$$\text{reflexión respecto al plano } XY = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

la composición de las dos operaciones (tomada en cualquier orden) nos da la matriz de inversión que cambia de signo todas las coordenadas

$$\text{Inversión} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\mathbf{1}_{3 \times 3}$$

la naturaleza discreta de esta operación nos dice que una inversión de un sistema dextrógiro a uno levógiro (cambio de quiralidad), no puede ser realizado por ningún cambio continuo en la orientación de los ejes coordenados. Por lo tanto, una inversión no puede corresponder a una transformación continua (por ejemplo, una rotación continua de los ejes coordenados o el desplazamiento físico de un cuerpo rígido). Ahora bien, teniendo en cuenta que toda matriz ortogonal de determinante  $-1$  se puede escribir como el producto de  $\mathbf{S} = -\mathbf{1}_{3 \times 3}$  con una matriz ortogonal de determinante  $+1$  (e.g. una rotación más una inversión), la argumentación anterior será válida para todas las matrices ortogonales de determinante  $-1$ . En consecuencia, las transformaciones ortogonales que representan a cualquier transformación continua deben ser de determinante  $+1$ . Otra forma de ver esto es teniendo en cuenta que la matriz que representa a una transformación continua, debe evolucionar

<sup>9</sup>En el sistema dextrógiro  $\mathbf{u}_x \times \mathbf{u}_y = \mathbf{u}_z$  en tanto que en el levógiro  $\mathbf{u}_x \times \mathbf{u}_y = -\mathbf{u}_z$  y lo mismo pasa con todos los productos cruz fundamentales. Esto implica que el algoritmo del producto cruz en el sistema levógiro cambia todos los signos respecto al algoritmo del dextrógiro. El cambio de un sistema coordenado dextrógiro a uno levógiro o viceversa se denomina un **cambio de quiralidad** del sistema coordenado.



en forma continua desde la identidad la cual tiene determinante  $+1$ , el cambio repentino a un determinante  $-1$  no se puede lograr con una transformación continua de los elementos de la matriz.

En virtud de lo anterior, las matrices ortogonales se dividen en dos grandes grupos: Las transformaciones **propias** (determinante  $+1$ ) e **impropias** (determinante  $-1$ ). Las primeras generan transformaciones continuas, en tanto que las otras representan una transformación continua combinada con una transformación discreta de inversión. Un estudio sistemático de las propiedades de transformación de los vectores bajo la operación de inversión (también conocida como transformación de **paridad**) nos lleva a clasificar a los vectores y escalares de acuerdo con tales propiedades de transformación como veremos a continuación.

## 12.5. Vector asociado a una matriz antisimétrica real $3 \times 3$

Para el caso particular de tres dimensiones, las matrices antisimétricas reales pueden parametrizarse vectorialmente. Una matriz antisimétrica real  $3 \times 3$  se puede parametrizar en la forma

$$\mathbf{A}_a = \begin{pmatrix} 0 & a_{12} & -a_{31} \\ -a_{12} & 0 & a_{23} \\ a_{31} & -a_{23} & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{pmatrix} \quad (12.57)$$

esta matriz tiene solo tres grados de libertad independientes, de modo que es razonable construir un arreglo vectorial de la forma

$$\mathbf{v}_A \equiv \begin{pmatrix} a_{23} \\ a_{31} \\ a_{12} \end{pmatrix} \equiv \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad (12.58)$$

Ahora bien, si aplicamos la matriz a un vector arbitrario  $\mathbf{x}$  obtenemos

$$\mathbf{A}_a \mathbf{x} = \begin{pmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} v_3 x_2 - v_2 x_3 \\ v_1 x_3 - v_3 x_1 \\ v_2 x_1 - v_1 x_2 \end{pmatrix} \quad (12.59)$$

las componentes del nuevo vector claramente son idénticas a las obtenidas a través del siguiente producto cruz

$$\mathbf{A}_a \mathbf{x} = \mathbf{x} \times \mathbf{v}_A \quad (12.60)$$

De modo que a una matriz antisimétrica real  $3 \times 3$  aplicada a un vector  $\mathbf{x}$  se le puede asociar un producto cruz entre  $\mathbf{x}$  y el vector asociado a la matriz antisimétrica. Inversamente, cualquier producto vectorial se puede asociar con una matriz antisimétrica actuando sobre un vector, lo cual se puede ver en la siguiente forma

$$(\mathbf{x} \times \mathbf{v}_A)_i = \varepsilon_{ijk} x_j v_k = (\varepsilon_{ijk} v_k) x_j \equiv A_{ij}^a x_j = (\mathbf{A}_a \mathbf{x})_i$$

donde hemos definido

$$A_{ij}^a \equiv \varepsilon_{ijk} v_k \quad (12.61)$$

la antisimetría de esta matriz es clara a partir de la antisimetría del tensor de Levi-Civita  $\varepsilon_{ijk}$ . La relación inversa se obtiene multiplicando la Ec. (12.61) por  $\varepsilon_{ijm}$  con suma sobre índices repetidos y usando la propiedad  $\varepsilon_{ijm} \varepsilon_{ijk} = 2\delta_{mk}$

$$\begin{aligned} \varepsilon_{ijm} A_{ij}^a &\equiv \varepsilon_{ijm} \varepsilon_{ijk} v_k = 2\delta_{mk} v_k = 2v_m \\ v_m &= \frac{1}{2} \varepsilon_{mij} A_{ij}^a \end{aligned} \quad (12.62)$$

por ejemplo

$$(\mathbf{v}_A)_1 \equiv v_1 = \frac{1}{2} (\varepsilon_{123} A_{23}^a + \varepsilon_{132} A_{32}^a) = \frac{1}{2} [\varepsilon_{123} A_{23}^a + (-\varepsilon_{123}) (-A_{23}^a)] = \varepsilon_{123} A_{23}^a = a_{23}$$

donde hemos usado la antisimetría de  $\varepsilon_{ijk}$  y de  $A_{ij}^a$ . Este resultado coincide con la Ec. (12.58).

En resumen, cualquier matriz antisimétrica real  $3 \times 3$  se puede parametrizar como

$$\mathbf{A}_a = \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix} = \begin{pmatrix} 0 & a_{12} & -a_{31} \\ -a_{12} & 0 & a_{23} \\ a_{31} & -a_{23} & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{pmatrix} \quad (12.63)$$

donde los tres grados de libertad se pueden asociar con un vector en la forma

$$\mathbf{v}_A \equiv \begin{pmatrix} a_{23} \\ a_{31} \\ a_{12} \end{pmatrix} \equiv \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} ; \quad (\mathbf{v}_A)_m = v_m = \frac{1}{2} \varepsilon_{mij} a_{ij} \quad (12.64)$$

si aplicamos esta matriz a un vector arbitrario  $\mathbf{x}$  obtenemos

$$\mathbf{A}_a \mathbf{x} = \mathbf{x} \times \mathbf{v}_A \quad (12.65)$$

inversamente, cualquier producto vectorial se puede asociar con una matriz antisimétrica actuando sobre un vector en la forma

$$(\mathbf{x} \times \mathbf{v}_A)_i = A_{ij}^a x_j = (\mathbf{A}_a \mathbf{x})_i \quad ; \quad A_{ij}^a \equiv \varepsilon_{ijk} v_k^A \quad (12.66)$$

Esto se aplica con frecuencia a cantidades como el torque y el momento angular.

## 12.6. Clasificación de vectores y escalares por sus propiedades de paridad

Tomemos una matriz antisimétrica  $\mathbf{A}_a$  y construyamos una matriz equivalente  $\mathbf{A}'_a$  por medio de una transformación de similaridad

$$\mathbf{A}'_a = \mathbf{B} \mathbf{A}_a \mathbf{B}^{-1} \quad (12.67)$$

nos restringiremos a transformaciones de similaridad con matrices ortogonales reales  $\mathbf{B}$ . En una transformación de similaridad ortogonal se conserva la antisimetría i.e.  $\mathbf{A}'_a$  también es antisimétrica, lo cual se puede ver transponiendo la Ec. (12.67)

$$\tilde{\mathbf{A}}'_a = \widetilde{\mathbf{B}^{-1} \mathbf{A}_a \mathbf{B}} = \tilde{\mathbf{B}} \tilde{\mathbf{A}}_a \tilde{\mathbf{B}} = -\mathbf{B} \mathbf{A}_a \mathbf{B}^{-1} = -\mathbf{A}'_a$$

En el caso de tres dimensiones, recordemos que los argumentos que nos llevaron a las Ecs. (12.58-12.60) nos indicaron que a toda matriz  $\mathbf{A}_a$  real antisimétrica y de dimensión  $3 \times 3$  se le puede asociar un arreglo vectorial  $\mathbf{v}_A$ . Esto implica que la matriz  $\mathbf{A}'_a$  también se puede escribir en la parametrización (12.57), y denotamos los nuevos elementos como  $v'_1, v'_2, v'_3$ . Para encontrar la transformación que nos lleva de  $\mathbf{v}_A$  a  $\mathbf{v}'_A$ , comenzaremos escribiendo la Ec. (12.67) en componentes

$$A'_{im} = b_{ij} A_{jk}^a \left( \tilde{b} \right)_{km} = b_{ij} A_{jk}^a b_{mk}$$

y utilizando la Ec. (12.61) podemos escribir esta ecuación en términos de las componentes de los vectores asociados a las matrices antisimétricas

$$\varepsilon_{iml} v'_l = b_{ij} (\varepsilon_{jkr} v_r) b_{mk}$$

multiplicando a ambos lados por  $\varepsilon_{imd}$  y sumando sobre  $i, m$  (índices repetidos), tenemos

$$\begin{aligned} \varepsilon_{imd} \varepsilon_{iml} v'_l &= \varepsilon_{imd} \varepsilon_{jkr} b_{ij} b_{mk} v_r \Rightarrow 2\delta_{ld} v'_l = \varepsilon_{imd} \varepsilon_{jkr} b_{ij} b_{mk} v_r \Rightarrow \\ v'_d &= \frac{1}{2} \varepsilon_{imd} \varepsilon_{jkr} b_{ij} b_{mk} v_r = \frac{1}{2} \varepsilon_{imd} \varepsilon_{jkr} b_{ij} b_{mk} \delta_{rg} v_g = \frac{1}{2} \varepsilon_{imd} \varepsilon_{jkr} b_{ij} b_{mk} (b_{hr} b_{hg}) v_g \end{aligned}$$

en el último paso hemos usado la condición de ortogonalidad de  $\mathbf{B}$ , Ec. (12.51). Tomemos por ejemplo  $d = 3$

$$\begin{aligned} v'_3 &= \frac{1}{2} \varepsilon_{im3} \varepsilon_{jkr} b_{ij} b_{mk} b_{hr} b_{hg} v_g = \frac{1}{2} \varepsilon_{123} \varepsilon_{jkr} b_{1j} b_{2k} b_{hr} b_{hg} v_g + \frac{1}{2} \varepsilon_{213} \varepsilon_{jkr} b_{2j} b_{1k} b_{hr} b_{hg} v_g \\ &= \frac{1}{2} \varepsilon_{jkr} b_{1j} b_{2k} b_{hr} b_{hg} v_g + \frac{1}{2} \varepsilon_{kjr} b_{1k} b_{2j} b_{hr} b_{hg} v_g = \varepsilon_{jkr} b_{1j} b_{2k} b_{hr} b_{hg} v_g \end{aligned}$$

donde hemos usado el hecho de que  $j, k$  son índices mudos. Sumando explícitamente sobre  $h$ , y agrupando las expresiones que se contraen, resulta

$$v'_3 = (\varepsilon_{jkr} b_{1j} b_{2k} b_{1r}) (b_{1g} v_g) + (\varepsilon_{jkr} b_{1j} b_{2k} b_{2r}) (b_{2g} v_g) + (\varepsilon_{jkr} b_{1j} b_{2k} b_{3r}) (b_{3g} v_g) \quad (12.68)$$

los dos primeros sumandos de la derecha se anulan en virtud de la antisimetría de  $\varepsilon_{jkr}$ . Por ejemplo, en  $\varepsilon_{jkr} b_{1j} b_{2k} b_{1r}$  para  $k = 1$ , aparecen  $\varepsilon_{213} b_{12} b_{21} b_{13} + \varepsilon_{312} b_{13} b_{21} b_{11} = 0$ , y similarmente para  $k = 2, 3$ . Tenemos entonces<sup>10</sup>

$$v'_3 = (\varepsilon_{jkr} b_{1j} b_{2k} b_{3r}) (b_{3g} v_g) = (\det \mathbf{B}) (b_{3g} v_g)$$

procediendo de forma similar con las otras componentes obtenemos:

$$v'_i = (\det \mathbf{B}) (b_{ig} v_g) \quad (12.69)$$

esto a su vez se puede interpretar como la propiedad de transformación para el arreglo vectorial  $\mathbf{v}_A$  asociado a la matriz  $\mathbf{A}_a$ . El arreglo vectorial  $\mathbf{v}_A$  transforma bajo una transformación de similaridad ortogonal en la forma

$$\mathbf{v}'_A = (\det \mathbf{B}) \mathbf{B} \mathbf{v}_A \quad (12.70)$$

si comparamos este comportamiento con el del vector posición bajo transformaciones ortogonales

$$\mathbf{r}' = \mathbf{B} \mathbf{r} \quad (12.71)$$

notamos que la transformación de  $\mathbf{v}_A$  es similar a la del vector posición excepto por el factor  $\det \mathbf{B}$ . Para transformaciones propias (i.e.  $\det \mathbf{B} = +1$ ) ambas transformaciones son idénticas y definiremos un vector en  $\mathbb{R}^3$  como una tripla que bajo transformaciones ortogonales propias cambia con la prescripción dada en la Ec. (12.71). En contraste, ambas transformaciones poseen un signo diferente bajo transformaciones ortogonales impropias. Recordando además que una transformación impropia consiste en una transformación propia junto con una inversión, se deduce que bajo la transformación de inversión, el cambio de  $\mathbf{v}_A$  difiere en un signo con respecto al cambio del vector posición  $\mathbf{r}$ . El vector posición  $\mathbf{r}$  invierte su signo bajo inversión en tanto que  $\mathbf{v}_A$  permanece inalterado ante dicha operación. Esto induce a definir dos tipos de vectores

1. Vectores polares: aquellos que bajo el operador de inversión, invierten el signo de sus componentes. El vector posición (con origen fijo) es un prototipo para estos vectores.
2. Vectores axiales o pseudovectores: aquellos que permanecen inalterados ante una operación de inversión, el vector  $\mathbf{v}_A$  asociado a una matriz antisimétrica es el prototipo para esta clase de vectores.

Una manera de establecer sistemáticamente esta clasificación es a través del operador paridad  $\hat{\mathbf{P}}$  definido como aquél que invierte todas las coordenadas  $x_i \rightarrow -x_i$ . Una forma útil de obtener un escalar es a través del producto punto, por ejemplo el producto punto  $\mathbf{r} \cdot \mathbf{r}$  no cambia su signo bajo paridad ya que cada vector cambia su signo. En general, el producto punto entre dos vectores polares no cambia su signo bajo paridad y serán el prototipo de los verdaderos escalares. Por otro lado, el producto punto de un vector polar con uno axial invierte su signo bajo paridad, es decir, difiere del comportamiento de un verdadero escalar y se le denomina

<sup>10</sup>En síntesis, los términos de la forma  $\varepsilon_{jkr} b_{mj} b_{nk} b_{pr}$  serán nulos si por lo menos dos de los índices  $m, n, p$  coinciden. Si  $m, n, p$  son todos distintos, se obtiene el determinante si están en orden cíclico, y menos el determinante si están en orden anti-cíclico.

un **pseudoescalar**. Si  $S$  es un escalar,  $S_A$  un pseudoescalar,  $\mathbf{V}_P$  un vector polar y  $\mathbf{V}_A$  un vector axial, se obtiene

$$\begin{aligned}\widehat{\mathbf{P}}S &= S ; \widehat{\mathbf{P}}S_A = -S_A ; \widehat{\mathbf{P}}\mathbf{V}_P = -\mathbf{V}_P ; \widehat{\mathbf{P}}\mathbf{V}_A = \mathbf{V}_A ; \widehat{\mathbf{P}}(S\mathbf{V}_P) = -S\mathbf{V}_P \\ \widehat{\mathbf{P}}(S_A\mathbf{V}_A) &= -S_A\mathbf{V}_A ; \widehat{\mathbf{P}}(\mathbf{V}_A \cdot \mathbf{V}_P) = -(\mathbf{V}_A \cdot \mathbf{V}_P) ; \widehat{\mathbf{P}}(\mathbf{V}_A \cdot \mathbf{V}_A) = (\mathbf{V}_A \cdot \mathbf{V}_A) \\ \widehat{\mathbf{P}}(\mathbf{V}_P \cdot \mathbf{V}_P) &= (\mathbf{V}_P \cdot \mathbf{V}_P) ; \widehat{\mathbf{P}}(S_AS_P) = -S_AS_P ; \widehat{\mathbf{P}}(S\mathbf{V}_P) = -S\mathbf{V}_P ; \widehat{\mathbf{P}}(S\mathbf{V}_A) = S\mathbf{V}_A\end{aligned}$$

Un ejemplo muy común de vector axial es aquél que proviene del producto cruz de dos vectores polares

$$\mathbf{V}_A = \mathbf{V}_1 \times \mathbf{V}_2$$

lo cual se puede ver teniendo en cuenta que bajo inversión de los ejes coordenados, cada vector  $\mathbf{V}_1$  y  $\mathbf{V}_2$  invierte su signo (por ser polares). Ejemplos comunes son el momento angular y el torque

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} ; \tau = \mathbf{r} \times \mathbf{F}$$

similarmente el producto cruz de un vector axial por uno polar es un vector polar. De lo anterior y teniendo en cuenta la fuerza de Lorentz

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B})$$

se vé que si postulamos que la carga eléctrica es un escalar, el campo magnético debe ser un vector axial ya que  $\mathbf{F}$  y  $\mathbf{v}$  son vectores polares. Este hecho es consistente con la ley de Biot-Savart que nos describe la integral de un vector axial diferencial ya que es el producto cruz de dos vectores polares (de la forma  $d\mathbf{l} \times \hat{\mathbf{r}}/r^2$ ), la integración es entonces un vector axial.

Desde un punto de vista pasivo se puede ver porqué un vector polar invierte su signo bajo inversión. El vector permanece inalterado en tanto que los ejes coordenados se invierten, de modo que las componentes en el nuevo sistema coordenado aparecen invertidas. ¿Que ocurre con un vector axial? ocurre que un vector axial lleva consigo una convención de “quiralidad”. Bajo inversión, un sistema coordenado dextrógiro se convierte en uno levógiro, si tomamos el producto cruz de dos vectores polares

$$V_{Ai} = V_{Pj}^{(1)}V_{Pk}^{(2)} - V_{Pj}^{(2)}V_{Pk}^{(1)} ; i, j, k \text{ en orden cíclico}$$

el orden cíclico requerido en esta ecuación implica un cambio similar desde la regla de la mano derecha hacia la regla de la mano izquierda. Por lo tanto, incluso en una interpretación pasiva, hay un cambio real de dirección de este producto cruz bajo inversión.

## 12.7. Transformaciones ortogonales propias

La anterior discusión nos motiva a restringir nuestro estudio a las transformaciones ortogonales propias dado que estaremos interesados en transformaciones continuas. La propiedad de ortogonalidad usada en la forma (12.53) nos lleva a la propiedad

$$(\mathbf{A} - \mathbf{1})\tilde{\mathbf{A}} = \mathbf{1} - \tilde{\mathbf{A}}$$

tomando el determinante a ambos lados

$$[\det(\mathbf{A} - \mathbf{1})][\det\tilde{\mathbf{A}}] = \det(\mathbf{1} - \tilde{\mathbf{A}}) \Rightarrow [\det(\mathbf{A} - \mathbf{1})] = \det(\widetilde{\mathbf{1} - \tilde{\mathbf{A}}})$$

donde hemos usado  $\det\tilde{\mathbf{A}} = \det\mathbf{A} = 1$ . Puesto que la matriz identidad es simétrica, se tiene que

$$\det(\mathbf{A} - \mathbf{1}) = \det(\mathbf{1} - \mathbf{A}) \Rightarrow \det(\mathbf{A} - \mathbf{1}) = \det[-(\mathbf{A} - \mathbf{1})]$$

y aplicando la propiedad (12.25) para matrices de dimensión impar

$$\begin{aligned}\det(\mathbf{A} - \mathbf{1}) &= -\det(\mathbf{A} - \mathbf{1}) && \text{para } n \text{ impar} \\ \det(\mathbf{A} - \mathbf{1}) &= 0 && \text{para } n \text{ par}\end{aligned}\tag{12.72}$$

y comparando con la ecuación general de valores propios (12.38) se observa que  $\lambda = 1$  es una solución para tal ecuación. En conclusión, para una matriz ortogonal propia de dimensión impar, uno de los valores propios de la matriz es  $\lambda = +1$ . Es importante enfatizar que esta conclusión solo es válida para matrices de dimensión impar en virtud del uso de la Ec. (12.25). Para  $n$  par, básicamente se llega a una identidad trivial y no a (12.72). Adicionalmente, recordemos que si la matriz  $\mathbf{A}$  es real, se tiene que si  $\lambda$  es solución de la ecuación secular, también lo será  $\lambda^*$ .

### 12.7.1. Matrices ortogonales reales propias en tres dimensiones

Para el caso específico de tres dimensiones, es posible obtener más propiedades sobre los valores propios. Dado que la dimensión es impar, tenemos al menos un valor propio igual a la unidad. Por convención designemos  $\lambda_3 = +1$  de modo que queremos obtener información sobre  $\lambda_1$  y  $\lambda_2$ . Ya hemos demostrado que el determinante de una matriz es invariante ante una transformación de similaridad. Por lo tanto, el determinante de  $\mathbf{A}$  coincide con el de la matriz  $\lambda$  en virtud de la relación (12.43), por lo tanto para matrices ortogonales propias se tiene que<sup>11</sup>

$$|\mathbf{A}| = \lambda_1 \lambda_2 \lambda_3 = \lambda_1 \lambda_2 = 1 \quad (12.73)$$

y recordando que los valores propios son de módulo unidad para las matrices unitarias y en particular para las matrices ortogonales reales (ver Ec. 12.56 y su discusión), tenemos

$$\|\lambda_1\| = \|\lambda_2\| = \lambda_3 = +1 \quad (12.74)$$

todo complejo unitario se escribe de la forma  $e^{i\theta}$ . Teniendo en cuenta las Ecs. (12.73) y (12.74) podemos escribir

$$\begin{aligned} \lambda_1 &= e^{i\Phi_1} ; \lambda_2 = e^{i\Phi_2} ; \lambda_3 = 1 \\ \lambda_1 \lambda_2 &= 1 = e^{i\Phi_1} e^{i\Phi_2} \Rightarrow \Phi_1 = -\Phi_2 \equiv \Phi \end{aligned}$$

finalmente, la estructura de valores propios se escribe como

$$\lambda_1 = e^{i\Phi} ; \lambda_2 = e^{-i\Phi} ; \lambda_3 = 1 \quad (12.75)$$

Vemos que  $\lambda_1$  es el complejo conjugado de  $\lambda_2$  lo cual es consistente con el hecho de que si la matriz  $\mathbf{A}$  es real,  $\lambda^*$  es solución de la ecuación secular, siempre que  $\lambda$  lo sea. Esto a su vez implica que  $\lambda_1$  y  $\lambda_2$  son ambos complejos o ambos reales. Son reales cuando  $\Phi = 0, \pm\pi$  y complejos en otros casos. Tenemos entonces tres posibles estructuras

1. Cuando  $\Phi = 0$ , todos los autovalores son  $+1$ . La matriz  $\lambda$  es la identidad y es fácil ver que  $\mathbf{A}$  también lo es (la matriz identidad tiene la misma representación en cualquier base), este caso es la solución trivial
2. Cuando  $\Phi = \pm\pi$ ,  $\lambda_1 = \lambda_2 = -1$ , y  $\lambda_3 = 1$ . La transformación descrita por la matriz  $\lambda$ , se puede considerar como una inversión en dos ejes coordenados manteniendo el tercero inalterado, como se puede ver si usamos estos valores de  $\lambda_i$  en la matriz (12.42) y la aplicamos a un vector arbitrario. Igualmente se puede ver como una rotación de  $\pi$  con respecto al tercer eje. Dado que  $\mathbf{A}$  es equivalente a  $\lambda$ , podemos decir que dada la base  $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$  en que se escribió  $\mathbf{A}$ , podemos llegar a una nueva base  $\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3$  en la cual la representación matricial es  $\lambda$ , y las inversiones y rotaciones ya descritas involucrarían a los ejes  $X'Y'Z'$ .
3.  $\Phi \neq 0, \pm\pi$ , en este caso  $\lambda_1$  y  $\lambda_2$  son complejos y  $\lambda_1 = \lambda_2^*$ .

<sup>11</sup>Esto es válido siempre que la matriz sea diagonalizable. En la sección 12.12, veremos que la matrices ortogonales reales son diagonalizables.

## 12.8. Matriz adjunta

Para definir la norma de un vector complejo en forma matricial podemos escribir

$$\|\mathbf{x}\|^2 = x_i^* x_i \Leftrightarrow \|\mathbf{x}\|^2 = \widetilde{\mathbf{x}}^* \mathbf{x}$$

nótese que hemos tenido que escribir a la izquierda el vector traspuesto (es decir vector fila) a fin de que la operación esté definida según la discusión en la sección (12.1.2). La matriz rectangular  $\widetilde{\mathbf{x}}^*$  se obtiene al conjugar y trasponer la matriz rectangular  $\mathbf{x}$ , vamos a definir a  $\widetilde{\mathbf{x}}^*$  como la adjunta de la matriz  $\mathbf{x}$  y la denotamos con un obelisco  $\dagger$

$$\mathbf{x}^\dagger \equiv \widetilde{\mathbf{x}}^* \quad (12.76)$$

vamos a definir ahora una función compleja de los vectores columna  $\mathbf{x}$  e  $\mathbf{y}$  en la forma

$$(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x}^\dagger \mathbf{y} = x_i^* y_i$$

esta operación cumple con las propiedades de un producto interno i.e.

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})^* ; (\mathbf{x}, \alpha \mathbf{y} + \beta \mathbf{z}) = \alpha (\mathbf{x}, \mathbf{y}) + \beta (\mathbf{x}, \mathbf{z}) ; (\alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z}) = \alpha^* (\mathbf{x}, \mathbf{z}) + \beta^* (\mathbf{y}, \mathbf{z}) ; (\mathbf{x}, \mathbf{x}) = \|\mathbf{x}\|^2 \quad (12.77)$$

vemos que si hacemos  $\mathbf{x} = \mathbf{y}$  se obtiene la norma al cuadrado del vector  $\mathbf{x}$ . Adicionalmente, si los vectores son reales esta operación me genera el producto punto. En general, diremos que un vector  $\mathbf{x}$  es unitario o normalizado si  $(\mathbf{x}, \mathbf{x}) = \|\mathbf{x}\|^2 = 1$ , y dos vectores  $\mathbf{x}$  e  $\mathbf{y}$  serán ortogonales si  $(\mathbf{x}, \mathbf{y}) = 0$ .

Por otra parte, si transformamos el vector  $\mathbf{x}$  a través de cierta matriz

$$\mathbf{x}' \equiv \mathbf{M} \mathbf{x} \quad (12.78)$$

y realizamos el producto escalar o producto interno

$$C = (\mathbf{x}')^\dagger \mathbf{y} = (\mathbf{x}', \mathbf{y})$$

nos preguntamos ahora como debe transformar el vector  $\mathbf{y}$  (sin alterar al vector  $\mathbf{x}$ ) de modo que este producto interno permanezca intacto, es decir de modo que

$$(\mathbf{x}')^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{y}' \Leftrightarrow (\mathbf{x}', \mathbf{y}) = (\mathbf{x}, \mathbf{y}') \quad (12.79)$$

siendo  $\mathbf{y}'$  la correspondiente transformación en  $\mathbf{y}$  denotemos entonces

$$\mathbf{y}' = \mathbf{B} \mathbf{y} \quad (12.80)$$

a la matriz  $\mathbf{B}$  la llamamos la matriz adjunta de  $\mathbf{M}$ . Reemplazando (12.78) y (12.80) en (12.79) resulta

$$(\mathbf{M} \mathbf{x})^\dagger \mathbf{y} = \mathbf{x}^\dagger (\mathbf{B} \mathbf{y}) \Rightarrow \widetilde{(\mathbf{M} \mathbf{x})}^* \mathbf{y} = \mathbf{x}^\dagger (\mathbf{B} \mathbf{y}) \Rightarrow \widetilde{\mathbf{x}}^* \widetilde{\mathbf{M}}^* \mathbf{y} = \mathbf{x}^\dagger (\mathbf{B} \mathbf{y})$$

y definiendo a la conjugada traspuesta de la matriz  $\mathbf{M}$  en forma análoga a la Ec. (12.76)

$$\mathbf{M}^\dagger \equiv \widetilde{\mathbf{M}}^* \quad (12.81)$$

obtenemos

$$\mathbf{x}^\dagger \mathbf{M}^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{B} \mathbf{y}$$

dado que esto es válido para todo  $\mathbf{x}$  e  $\mathbf{y}$ , se concluye que

$$\mathbf{B} = \mathbf{M}^\dagger$$

donde el símbolo  $\mathbf{M}^\dagger$  indica la conjugada traspuesta de la matriz  $\mathbf{M}$ , concepto que se extiende naturalmente de la definición análoga para vectores columna. La matriz adjunta de  $\mathbf{M}$  es entonces su conjugada traspuesta. En notación de producto interno escribimos

$$(\mathbf{M}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{M}^\dagger\mathbf{y}) \quad (12.82)$$

Veremos que varias propiedades interesantes de transformación se traducen en propiedades de la matriz adjunta. Es fácil ver las siguientes propiedades

$$\left(\mathbf{M}^\dagger\right)^\dagger = \mathbf{M} ; \quad (\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger ; \quad \det \mathbf{M}^\dagger = (\det \mathbf{M})^* \quad (12.83)$$

a continuación veremos algunas matrices especiales relacionadas con la adjunta de una matriz.

## 12.9. Matrices unitarias y cambios de base (opcional)

Nos preguntamos ahora por el conjunto más general de matrices que al transformar a ambos vectores deja invariante el producto interno i.e.

$$(\mathbf{x}, \mathbf{y}) = (\mathbf{x}', \mathbf{y}') \Leftrightarrow \mathbf{x}^\dagger\mathbf{y} = (\mathbf{x}')^\dagger\mathbf{y}' \quad \text{con} \quad \mathbf{x}' \equiv \mathbf{U}\mathbf{x} ; \quad \mathbf{y}' \equiv \mathbf{U}\mathbf{y}$$

resulta entonces que

$$\mathbf{x}^\dagger\mathbf{y} = (\mathbf{U}\mathbf{x})^\dagger(\mathbf{U}\mathbf{y}) = \mathbf{x}^\dagger\mathbf{U}^\dagger\mathbf{U}\mathbf{y}$$

al ser válido para todo  $\mathbf{x}$  y para todo  $\mathbf{y}$  resulta

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{1} \quad (12.84)$$

$\mathbf{U}^\dagger$  es entonces la inversa a izquierda de la matriz  $\mathbf{U}$  pero recordando que para matrices cuadradas la inversa a izquierda es igual a la inversa a la derecha resulta

$$\mathbf{U}^\dagger = \mathbf{U}^{-1} \quad (12.85)$$

En particular, si  $\mathbf{x} = \mathbf{y}$  observamos que estas transformaciones dejan invariante la norma del vector  $\mathbf{x}$ . Por lo tanto la propiedad (12.85) define al conjunto más general de transformaciones lineales homogéneas que preservan el producto interno y en particular la norma de los vectores. A las matrices que cumplen esta propiedad se les denomina matrices unitarias. Cuando los elementos matriciales son complejos, la magnitud de los vectores (que en general tendrán también componentes complejas), no será invariante bajo una transformación con matrices ortogonales complejas. Las matrices unitarias son matrices **complejas** que mantienen invariante la norma de vectores **complejos**. Los valores propios de las matrices unitarias poseen módulo unidad como se demostró en (12.56) pues debemos recordar que la única condición que se usó para llegar a (12.56) fué la preservación de la norma de los vectores (reales o complejos) ante la transformación asociada a la matriz (real o compleja). El determinante se obtiene de la propiedad (12.84)

$$\begin{aligned} \det(\mathbf{U}^\dagger\mathbf{U}) &= 1 \Rightarrow \det \tilde{\mathbf{U}}^* \det \mathbf{U} = 1 \Rightarrow (\det \mathbf{U})^* \det \mathbf{U} = 1 \\ |\det \mathbf{U}|^2 &= 1 \Rightarrow \det \mathbf{U} = e^{i\theta} \end{aligned}$$

de modo que el determinante al igual que los valores propios es un número complejo con norma unidad. Esto es consistente con el hecho de que cuando la matriz es diagonalizable<sup>12</sup>, el determinante es el producto de sus valores propios. En este caso todos los valores propios son de norma unidad y por tanto también su producto.

<sup>12</sup>En la sección 12.12, veremos que las matrices unitarias son diagonalizables.

Existe otra propiedad fundamental de las matrices unitarias que con frecuencia la literatura toma como definición. Una matriz es unitaria si y solo si transforma cualquier base ortonormal en otra base ortonormal. Para verlo, sea  $\{\mathbf{u}_i\}$  un conjunto de vectores ortonormales i.e.

$$(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$$

transformemos cada elemento  $\mathbf{u}_i$  de la base con una matriz unitaria i.e.  $\mathbf{u}'_i = \mathbf{U}\mathbf{u}_i$  tenemos que

$$(\mathbf{u}'_i, \mathbf{u}'_j) = (\mathbf{U}\mathbf{u}_i, \mathbf{U}\mathbf{u}_j) = (\mathbf{u}_i, \mathbf{u}_j)$$

donde hemos usado la definición de matriz unitaria como aquella que no altera el valor del producto interno. Por lo tanto

$$(\mathbf{u}'_i, \mathbf{u}'_j) = (\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$$

recíprocamente, pensemos en el conjunto más general de transformaciones lineales homogéneas que nos llevan de un sistema coordinado ortonormal a otro ortonormal, tenemos entonces

$$\begin{aligned} (\mathbf{u}'_i, \mathbf{u}'_j) &= (\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij} \quad ; \quad \mathbf{u}'_i = \mathbf{M}\mathbf{u}_i \Rightarrow \\ (\mathbf{u}'_i, \mathbf{u}'_j) &= (\mathbf{M}\mathbf{u}_i, \mathbf{M}\mathbf{u}_j) = (\mathbf{u}_i, \mathbf{M}^\dagger(\mathbf{M}\mathbf{u}_j)) = (\mathbf{u}_i, (\mathbf{M}^\dagger\mathbf{M})\mathbf{u}_j) \Rightarrow \\ (\mathbf{u}_i, (\mathbf{M}^\dagger\mathbf{M})\mathbf{u}_j) &= (\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij} \end{aligned}$$

y puesto que  $\mathbf{u}_i, \mathbf{u}_j$  pertenecen a una base ortonormal arbitraria, vemos que se llega a la condición  $\mathbf{M}^\dagger\mathbf{M} = \mathbf{1}$ , que caracteriza a las matrices unitarias. Esta propiedad es de suma importancia dado que usualmente estamos interesados en utilizar bases ortonormales incluso cuando cambiamos de base. Si definimos un espacio vectorial donde los escalares son reales los vectores base están asociados al espacio  $\mathbb{R}^n$ , en tanto que si los escalares son complejos hablamos de una base en el espacio unitario  $\mathbb{C}^n$  de dimensión  $n$  donde cada componente representa un número complejo. El espacio  $\mathbb{C}^2$  es por ejemplo de gran importancia en la caracterización del espín en mecánica cuántica y los cambios de base en este espacio se realizan con matrices unitarias de dimensión dos. Para nuestros propósitos estaremos interesados en vectores de  $\mathbb{R}^n$  en cuyo caso será suficiente trabajar el subconjunto propio de matrices unitarias reales, es decir el conjunto de transformaciones ortogonales reales<sup>13</sup>. Los cambios de base se realizarán entonces con matrices ortogonales reales propias<sup>14</sup>. Recordemos entonces que ante un cambio de base, un operador activo  $\mathbf{A}$  cambia su representación matricial en la forma

$$\mathbf{A}' = \mathbf{B}\mathbf{A}\mathbf{B}^{-1} \quad (12.86)$$

y si ambas bases son ortonormales,  $\mathbf{B}$  es una matriz unitaria u ortogonal real según el espacio en el que trabajemos.

La condición de unitariedad (12.85) se puede escribir en componentes en la forma siguiente

$$\begin{aligned} \mathbf{U}\tilde{\mathbf{U}}^* &= \mathbf{1} \quad ; \quad \tilde{\mathbf{U}}^*\mathbf{U} = \mathbf{1} \Rightarrow u_{ij}\tilde{u}_{jk}^* = \delta_{ik} \quad ; \quad \tilde{u}_{ij}^*u_{jk} = \delta_{ik} \\ u_{ij}u_{kj}^* &= \delta_{ik} \quad ; \quad u_{ji}^*u_{jk} = \delta_{ik} \end{aligned} \quad (12.87)$$

el primer conjunto de ecuaciones (12.87) nos indica que los vectores fila que constituyen a la matriz unitaria son ortonormales entre sí. Similarmente, el segundo conjunto de ecuaciones (12.87) muestra la ortonormalidad de los vectores columna que constituyen la matriz unitaria. Es claro que solo un conjunto de estas relaciones es independiente y el otro es redundante. Naturalmente, si las componentes son reales las condiciones de unitariedad (12.87) coinciden con las condiciones de ortogonalidad Ecs. (12.51, 12.52).

Finalmente, cabe enfatizar que un cambio de base en  $\mathbb{C}^n$  ( $\mathbb{R}^n$ ) de la forma (12.86) no requiere que la matriz de transformación sea unitaria u ortogonal. Solo se necesita que la matriz pasiva de transformación sea no

<sup>13</sup>Las matrices unitarias complejas mantienen la norma de los vectores que originalmente son reales, pero me pueden llevar un vector de  $\mathbb{R}^n$  a un vector de  $\mathbb{C}^n$ .

<sup>14</sup>Una matriz ortogonal real impropia nos llevaría a un sistema coordinado izquierdo, y a menos que estemos analizando paridad, este cambio de quiralidad no es deseable.



singular. El problema es que la nueva base no será ortonormal lo cual puede traer varias dificultades que ilustraremos con un sencillo ejemplo, escribamos el producto interno de un cierto conjunto de vectores base en la forma

$$(\mathbf{v}_i, \mathbf{v}_j) = g_{ij} \quad (12.88)$$

y denominamos a  $g_{ij}$  el tensor métrico. Si los vectores unitarios son ortonormales entonces  $g_{ij} = \delta_{ij}$ . Si la base es ortogonal pero los vectores base no están normalizados  $g_{ij}$  es no trivial pero sigue siendo diagonal. Si la base no es ortogonal ni normal este tensor es además no diagonal. Escribimos  $\mathbf{x}$  e  $\mathbf{y}$  en términos de la base y calculamos el producto interno entre ellos

$$\begin{aligned} \mathbf{x} &= x_i \mathbf{v}_i ; \quad \mathbf{y} = y_j \mathbf{v}_j \Rightarrow (\mathbf{x}, \mathbf{y}) = (x_i \mathbf{v}_i, y_j \mathbf{v}_j) = x_i^* y_j (\mathbf{v}_i, \mathbf{v}_j) \\ &\Rightarrow (\mathbf{x}, \mathbf{y}) = x_i^* g_{ij} y_j \end{aligned}$$

donde hemos usado las propiedades (12.77) del producto interno. En notación matricial

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\dagger \mathbf{G} \mathbf{y} \quad (12.89)$$

Sea un operador lineal  $\mathbf{A}$  y veamos que relación tiene su representación matricial con la del operador adjunto  $\mathbf{B}$  en esta base

$$\begin{aligned} (\mathbf{A}\mathbf{x}, \mathbf{y}) &= (\mathbf{x}, \mathbf{B}\mathbf{y}) \Rightarrow (\mathbf{A}\mathbf{x})^\dagger \mathbf{G} \mathbf{y} = \mathbf{x}^\dagger \mathbf{G} (\mathbf{B}\mathbf{y}) \\ &\Rightarrow \mathbf{x}^\dagger \mathbf{A}^\dagger \mathbf{G} \mathbf{y} = \mathbf{x}^\dagger \mathbf{G} \mathbf{B} \mathbf{y} \end{aligned}$$

que nos lleva a la condición

$$\mathbf{A}^\dagger \mathbf{G} = \mathbf{G} \mathbf{B} \Leftrightarrow \mathbf{B} = \mathbf{G}^{-1} \mathbf{A}^\dagger \mathbf{G} \quad (12.90)$$

nótese que la representación del operador adjunto no está dada por la conjugada traspuesta de la representación del operador original.

Es fácil ver la condición para que una matriz represente a un operador activo unitario en esta base

$$\begin{aligned} (\mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y}) &= (\mathbf{x}, \mathbf{y}) \Rightarrow (\mathbf{x}^\dagger \mathbf{U}^\dagger) \mathbf{G} (\mathbf{U}\mathbf{y}) = \mathbf{x}^\dagger \mathbf{G} \mathbf{y} \Rightarrow \mathbf{U}^\dagger \mathbf{G} \mathbf{U} = \mathbf{G} \\ &\Rightarrow \mathbf{G}^{-1} \mathbf{U}^\dagger \mathbf{G} = \mathbf{U}^{-1} \end{aligned} \quad (12.91)$$

En bases no ortonormales, la condición de ortonormalidad entre dos vectores  $\mathbf{a}_i$  y  $\mathbf{a}_j$  se escribe

$$(\mathbf{a}_i, \mathbf{a}_j) = \mathbf{a}_i^\dagger \mathbf{G} \mathbf{a}_j = \delta_{ij} \quad (12.92)$$

Esto nos motiva a introducir una sutil diferencia entre **operadores lineales** adjuntos, unitarios, etc. y **matrices** adjuntas, unitarias, etc. Tomemos de ejemplo los operadores unitarios y matrices unitarias para la discusión. En general la literatura utiliza el término **matriz unitaria** para los arreglos matriciales que cumplen el algoritmo matricial  $\mathbf{U}^{-1} = \mathbf{U}^\dagger$  esta es una definición algebraica. Por otro lado, se define un **operador lineal unitario** como aquél que deja invariante el producto interno entre dos vectores arbitrarios, esta es una definición geométrica y por tanto independiente del sistema coordenado que se use. No obstante, una matriz unitaria en el sentido aquí expuesto solo representa a un operador unitario cuando la base es ortonormal. Cuando la base no es ortonormal la representación matricial de un operador unitario está dada por el algoritmo (12.91) y la llamaremos **representación matricial del operador unitario**. Por supuesto, la misma discusión vale para el concepto de adjunto y otros conceptos que veremos a continuación como hermiticidad, normalidad, etc.

Lo anterior nos ilustra algunas dificultades que aparecen cuando se toman bases no ortonormales, por esta razón es usual que los cambios de base se hagan garantizando que se llega a otra base ortonormal i.e. con matrices unitarias u ortogonales reales. Veremos sin embargo, que las bases no ortonormales nos permiten resolver un problema modificado de valores propios con muchas aplicaciones en Física (ver Sec. 12.14).

## 12.10. Matrices con espectro completo

En lo que sigue asumiremos que la matriz  $\mathbf{A}$  en cuestión es diagonalizable, y estudiaremos las consecuencias de esta hipótesis. Esto nos servirá de motivación para estudiar las condiciones requeridas para que una matriz sea diagonalizable, las cuales se discutirán en la sección 12.12<sup>15</sup>.

La diagonalizabilidad de una matriz  $\mathbf{A}$  es equivalente a decir que si la matriz  $\mathbf{A}$  es  $n \times n$ , posee  $n$  vectores propios linealmente independientes (ver sección 12.3, Pág. 12.3), es decir que los vectores forman una base del espacio  $\mathcal{E}^n$  en el cual actúa la matriz<sup>16</sup>. Denotaremos a los valores propios diferentes como  $\lambda_i$  con  $i = 1, \dots, p$  y a los  $n$  vectores propios linealmente independientes como  $\mathbf{v}_i^{(k)}$  donde  $k = 1, \dots, g_i$ ; siendo  $g_i$  el grado de degeneración de  $\lambda_i$ . Puesto que los vectores propios del conjunto  $\{\mathbf{v}_i^{(k)}\}$  forman una base en  $\mathcal{E}^n$ , podemos expandir cualquier vector  $\mathbf{x} \in \mathcal{E}^n$  en esta base

$$\mathbf{x} = \sum_{i=1}^p \sum_{k=1}^{g_i} c_i^k \mathbf{v}_i^{(k)} = \sum_{k=1}^{g_1} c_1^k \mathbf{v}_1^{(k)} + \sum_{k=1}^{g_2} c_2^k \mathbf{v}_2^{(k)} + \dots + \sum_{k=1}^{g_p} c_p^k \mathbf{v}_p^{(k)} \quad (12.93)$$

de acuerdo con el teorema 8, tenemos que

$$\mathbf{x}_m \equiv \sum_{k=1}^{g_m} c_m^k \mathbf{v}_m^{(k)} \in \mathcal{E}_m, \quad m = 1, 2, \dots, p \quad (12.94)$$

donde  $\mathcal{E}_m$  es el subespacio de  $\mathcal{E}^n$  generado por  $\lambda_m$ , dicho subespacio es de dimensión  $g_m$ . Dado que todos los vectores propios forman una base de  $\mathcal{E}^n$ , se tiene que

$$\sum_{m=1}^p g_m = n$$

ahora bien, una vez que la base  $\{\mathbf{v}_i^{(k)}\}$  se ordena de una manera específica, el conjunto ordenado de coeficientes  $\{c_i^{(k)}\}$  definido en la Ec. (12.93) para un vector dado  $\mathbf{x} \in \mathcal{E}^n$  es único, en virtud de la independencia lineal de la base. Esto a su vez significa que para un vector dado  $\mathbf{x} \in \mathcal{E}^n$ , cada uno de los vectores  $\mathbf{x}_m \in \mathcal{E}_m$  definidos en (12.94) es único. Esto significa que un vector arbitrario de  $\mathcal{E}^n$  tiene una descomposición única de la forma

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_p \quad ; \quad \mathbf{x}_m \in \mathcal{E}_m \quad \forall \mathbf{x} \in \mathcal{E}^n \quad (12.95)$$

en otras palabras, cada vector de  $\mathcal{E}^n$  se descompone de una manera única en una suma de vectores  $\mathbf{x}_m$  que pertenecen a los subespacios  $\mathcal{E}_m$  generados por los valores propios  $\lambda_m$ . Esto se expresa también diciendo que  $\mathcal{E}^n$  se descompone en una suma directa de subespacios  $\mathcal{E}_m$  y se escribe

$$\mathcal{E}^n = \mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \dots \oplus \mathcal{E}_p \equiv \sum_{\oplus m=1}^p \mathcal{E}_m, \quad \lambda_m \leftrightarrow \mathcal{E}_m \quad (12.96)$$

a la componente  $\mathbf{x}_m \in \mathcal{E}_m$  se le llama la proyección del vector  $\mathbf{x}$  sobre el subespacio  $\mathcal{E}_m$ . Las Ecs. (12.95, 12.96) expresan que el espectro (conjunto de valores propios) de  $\mathbf{A}$  “llena” el espacio  $\mathcal{E}^n$ , o que su espectro es **completo**. Esta es otra forma de expresar el hecho de que exista un conjunto de vectores propios de  $\mathbf{A}$  que forman una base de  $\mathcal{E}^n$ .

<sup>15</sup>A pesar de que se hará referencia constante a la sección 12.12, dicha sección puede omitirse en una primera lectura, si aceptamos la hipótesis de que las matrices en cuestión poseen un conjunto completo de vectores propios y comprendemos las consecuencias de esta hipótesis.

<sup>16</sup>Debe anotarse sin embargo, que una combinación lineal de los vectores propios de la base no es necesariamente un vector propio. Esto se debe a que los vectores propios de esta base están en general asociados a diferentes valores propios. Si la combinación lineal involucra solo vectores asociados a un valor propio dado, el vector resultante también será vector propio.

## 12.11. Matrices Hermíticas y simétricas reales

La discusión se orientará asumiendo que las bases utilizadas son ortonormales, a menos que se indique lo contrario. Hemos definido el operador adjunto de  $\mathbf{A}$  denotado por  $\mathbf{A}^\dagger$  como aquel para el cual se cumple la condición

$$(\mathbf{A}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^\dagger\mathbf{y}) \quad , \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{E}^n$$

nos preguntamos por los operadores para los cuales  $\mathbf{A} = \mathbf{A}^\dagger$  i.e. para los cuales se cumple

$$(\mathbf{A}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}\mathbf{y}) \quad , \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{E}^n$$

Dado que nuestra base es ortonormal, esta condición equivale a la igualdad de la matriz representativa con su adjunta (conjugada transpuesta). Se dice que una matriz es hermítica o autoadjunta si dicha matriz coincide con su adjunta i.e.

$$\mathbf{A} = \widetilde{\mathbf{A}}^* \equiv \mathbf{A}^\dagger$$

Es fácil demostrar que para una matriz hermítica de dimensión arbitraria, su naturaleza hermítica se preserva ante una transformación de similaridad **unitaria**. Bajo una transformación de similaridad

$$\mathbf{A}' = \mathbf{B}\mathbf{A}\mathbf{B}^{-1}$$

la matriz conjugada traspuesta i.e. la adjunta, estará dada por

$$\mathbf{A}'^\dagger = (\mathbf{B}^{-1})^\dagger \mathbf{A} \mathbf{B}^\dagger$$

donde hemos usado el carácter hermítico de  $\mathbf{A}$ . Si la matriz  $\mathbf{B}$  es unitaria entonces  $\mathbf{B}^{-1} = \mathbf{B}^\dagger$  y  $(\mathbf{B}^{-1})^\dagger = (\mathbf{B}^\dagger)^\dagger = \mathbf{B}$  de lo cual resulta

$$\mathbf{A}'^\dagger = \mathbf{B}\mathbf{A}\mathbf{B}^{-1} = \mathbf{A}'$$

debe enfatizarse que si la matriz del cambio de base  $\mathbf{B}$  no es unitaria, en general la matriz transformada no será hermítica<sup>17</sup>. Cuando los elementos matriciales son reales las matrices hermíticas se vuelven simétricas y las matrices unitarias se vuelven ortogonales. Por lo tanto las matrices simétricas reales y ortogonales reales son casos particulares de matrices hermíticas y unitarias respectivamente.

Veamos algunas propiedades de las matrices hermíticas: hemos visto que la solución del problema de valores y vectores propios de una cierta matriz  $\mathbf{A}$ , nos permite obtener una matriz que genera un cambio de base que diagonaliza a la matriz  $\mathbf{A}$ . La matriz de transformación (más exactamente su inversa), se obtiene adjuntando los arreglos columna que corresponden a los vectores propios y la matriz diagonalizada consta de los valores propios colocados en la diagonal de la nueva matriz ver Ecs. (12.43, 12.42). Este procedimiento tiene sin embargo algunos problemas en el caso general, por ejemplo si no todos los vectores propios son linealmente independientes la matriz que debe diagonalizar se vuelve singular y por tanto, no existe su inversa. Esto implica que en particular pueden existir problemas con la diagonalización cuando hay valores propios degenerados. Sin embargo, el proceso de diagonalización es en general más sencillo para las matrices Hermíticas.

Una de las primeras propiedades interesantes de las matrices hermíticas es que sus valores propios son reales, aún cuando la matriz sea compleja. Denotaremos a la matriz Hermítica como  $\mathbf{H}$  para enfatizar en su hermiticidad, además  $\mathbf{V}^{(j)}$  denotará el  $j$ -ésimo vector propio, con valor propio  $\lambda_j$ . La ecuación de valores propios se escribe como

$$\mathbf{H}\mathbf{V}^{(j)} = \lambda_j \mathbf{V}^{(j)}$$

<sup>17</sup>Esto a su vez se puede ver teniendo en cuenta que la nueva base no será ortonormal. En la nueva base el operador sigue siendo hermítico puesto que esta condición es intrínseca, pero ya no se cumplirá la condición matricial  $\mathbf{A}' = \widetilde{\mathbf{A}'^*}$ . De la Ec. (12.90) vemos que la condición para que la matriz  $\mathbf{A}'$  represente a un operador hermítico en una base no ortonormal es

$$\mathbf{A}' = \mathbf{G}^{-1} \mathbf{A}'^\dagger \mathbf{G}$$

haciendo el producto por la izquierda con el vector adjunto (conjugado traspuesto) correspondiente al vector propio  $n$ -ésimo tenemos

$$\mathbf{V}^{(n)\dagger} \mathbf{H} \mathbf{V}^{(j)} = \lambda_j \mathbf{V}^{(n)\dagger} \mathbf{V}^{(j)} \quad (12.97)$$

para la ecuación de valores propios correspondiente al  $n$ -ésimo vector propio

$$\mathbf{H} \mathbf{V}^{(n)} = \lambda_n \mathbf{V}^{(n)}$$

podemos tomar la ecuación adjunta

$$\mathbf{V}^{(n)\dagger} \mathbf{H}^\dagger = \lambda_n^* \mathbf{V}^{\dagger(n)}$$

y multiplicamos por el vector (columna)  $\mathbf{V}^{(j)}$  a la derecha de ambos miembros, teniendo en cuenta además que la matriz  $\mathbf{H}$  es hermítica, se obtiene

$$\mathbf{V}^{(n)\dagger} \mathbf{H} \mathbf{V}^{(j)} = \lambda_n^* \mathbf{V}^{(n)\dagger} \mathbf{V}^{(j)} \quad (12.98)$$

restando las Ecs. (12.97, 12.98) resulta

$$(\lambda_j - \lambda_n^*) \mathbf{V}^{(n)\dagger} \mathbf{V}^{(j)} = 0 \quad (12.99)$$

si  $j = n$  se obtiene que

$$(\lambda_n - \lambda_n^*) \left\| \mathbf{V}^{(n)} \right\|^2 = 0$$

de modo que  $\lambda_n = \lambda_n^*$  y los valores propios son reales como se quería demostrar. Ahora bien, si  $j \neq n$  y  $\lambda_j \neq \lambda_n$ , es decir si estos valores propios no son degenerados, se obtiene de (12.99)

$$(\lambda_j - \lambda_n) \mathbf{V}^{(n)\dagger} \mathbf{V}^{(j)} = 0 \Rightarrow \mathbf{V}^{(n)\dagger} \mathbf{V}^{(j)} = \left( \mathbf{V}^{(n)}, \mathbf{V}^{(j)} \right) = 0$$

donde hemos usado el carácter real de los valores propios. Vemos entonces que los vectores propios correspondientes a valores propios diferentes son ortogonales. Este hecho tiene un gran alcance en términos de aplicaciones. Por ejemplo en ecuaciones diferenciales, muchos operadores diferenciales son lineales y hermíticos, tales operadores actúan sobre espacios vectoriales de funciones (usualmente de dimensión infinita) sus valores propios son reales y sus vectores propios son ortogonales al menos en ausencia de degeneración, estos vectores propios permiten generalmente construir una base para el espacio vectorial en cuestión. En consecuencia el problema de Sturm-Liouville se reduce a un problema de valores propios de un operador hermítico que al ser lineal admite una representación matricial (aunque estas matrices son en general de dimensión infinita lo cual introduce problemas que no trataremos aquí). Es necesario enfatizar sin embargo que los vectores propios asociados a valores propios degenerados no son necesariamente ortogonales.

En la sección 12.12 veremos que las matrices hermíticas tienen un espectro completo. Por tanto, si  $\mathbf{H}$  es una matriz hermítica  $n \times n$ , podemos inducir una descomposición del espacio  $\mathcal{E}^n$  en los subespacios generados por los vectores propios  $\lambda_m$  de  $\mathbf{H}$ . Así mismo un vector arbitrario  $\mathbf{x}$  de  $\mathcal{E}^n$  se puede descomponer de manera única en proyecciones a lo largo de cada subespacio  $\mathcal{E}_m$

$$\mathcal{E}^n = \mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \dots \oplus \mathcal{E}_p \quad , \quad \lambda_m \leftrightarrow \mathcal{E}_m \quad (12.100)$$

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_p \quad ; \quad \mathbf{x}_m \in \mathcal{E}_m \quad \forall \mathbf{x} \in \mathcal{E}^n \quad (12.101)$$

Ahora bien, puesto que los vectores propios asociados a valores propios diferentes son ortogonales para las matrices hermíticas, tenemos que todos los vectores de  $\mathcal{E}_i$  son perpendiculares a todos los vectores de  $\mathcal{E}_j$  si  $i \neq j$ . Esto se indica diciendo que los subespacios  $\{\mathcal{E}_k\}$  son perpendiculares entre sí y se denota  $\mathcal{E}_i \perp \mathcal{E}_j$ . Por tanto, tenemos también que  $\mathbf{x}_i \perp \mathbf{x}_j$  para las proyecciones dadas en la Ec. (12.101) y hablamos entonces de **proyecciones ortogonales**.

### 12.11.1. Matrices reales simétricas

Adicionalmente, si la matriz hermítica es real se convierte en una matriz **simétrica real** y dado que los valores propios son todos reales, el teorema 9 nos dice que sobre cada subespacio  $\mathcal{E}_m$  podemos construir una base ortonormal de vectores propios reales, y dado que los subespacios son ortogonales entre sí, la unión de las bases ortonormales **reales** asociadas a cada  $\mathcal{E}_m$ , me generan una base ortonormal real de todo el espacio  $\mathcal{E}^n$ . En general, para matrices hermíticas el espectro será completo en  $\mathbb{C}^n$  ya que los vectores propios de la base serán complejos, pero para matrices simétricas reales el espectro será también completo en  $\mathbb{R}^n$ . El hecho de que se pueda construir una base ortonormal real de vectores propios implica que es posible encontrar una matriz ortogonal real<sup>18</sup> que diagonaliza a la matriz  $\mathbf{H}$ . En consecuencia, si partimos de unos ejes en  $\mathbb{R}^n$  ortogonales y reales, la transformación me lleva a otro conjunto de ejes de  $\mathbb{R}^n$  ortogonales y **reales** (ya que la matriz diagonalizante es ortogonal real) en donde la matriz es diagonal. Llamaremos a éstos últimos los **ejes principales** asociados a la matriz. Los ejes estarán orientados en la dirección de los vectores propios reales independientes asociados a la matriz. Todo ello conduce a que las operaciones con matrices simétricas reales se pueden restringir adecuadamente a espacios vectoriales reales  $\mathbb{R}^n$ .

Tenemos entonces que los valores propios de matrices reales simétricas son reales, y los vectores de una base ortonormal de vectores propios pueden elegirse como reales. Para enfatizar la importancia de estos hechos, recordemos por ejemplo que las matrices ortogonales reales en tres dimensiones, solo poseen un valor propio real excepto en casos triviales y por tanto una sola dirección real asociada a un vector propio (eje de rotación)<sup>19</sup>. Adicionalmente, el hecho de que los valores propios sean reales no garantiza que se puedan elegir reales todos los vectores propios que forman una base. De hecho, las bases ortonormales de matrices hermíticas complejas estarán constituidas en general por vectores complejos que no se pueden llevar a forma real (al menos no todos simultáneamente). Ahora bien, si el espectro o los vectores propios son eminentemente complejos, el problema de valores propios, la diagonalización de la matriz etc, deberán tratarse en el espacio  $\mathbb{C}^n$  y no se pueden restringir al espacio  $\mathbb{R}^n$ . Esto hace que las matrices simétricas reales jueguen un papel esencial en mecánica clásica, en donde en general realizaremos operaciones que se restringen al espacio  $\mathbb{R}^n$ .

### 12.11.2. Problema de valores propios de matrices hermíticas en tres dimensiones

Una propiedad interesante para la matrices hermíticas en tres dimensiones es que dados dos vectores propios ortogonales y sus valores propios correspondientes, el tercer vector propio se puede hallar simplemente construyendo un sistema coordenado dextrógiro con los dos anteriores. Para ello debemos demostrar que si  $\mathbf{u}_1^{(p)}, \mathbf{u}_2^{(p)}$  son vectores propios unitarios ortogonales que generan ejes principales, entonces un vector  $\mathbf{u}_3$  ortogonal a los anteriores también es vector propio que genera el tercer eje principal<sup>20</sup>. Para demostrarlo, tengamos en cuenta que de acuerdo con la ecuación de vectores propios se cumple por hipótesis que

$$\mathbf{H}_p \mathbf{u}_1^{(p)} = \lambda_1 \mathbf{u}_1^{(p)} \quad ; \quad \mathbf{H}_p \mathbf{u}_2^{(p)} = \lambda_2 \mathbf{u}_2^{(p)} \quad (12.102)$$

aplicando  $\mathbf{u}_i^{(p)\dagger}$  a la izquierda

$$\mathbf{u}_i^{(p)\dagger} \mathbf{H}_p \mathbf{u}_1^{(p)} = \lambda_1 \mathbf{u}_i^{(p)\dagger} \mathbf{u}_1^{(p)} \quad ; \quad \mathbf{u}_i^{(p)\dagger} \mathbf{H}_p \mathbf{u}_2^{(p)} = \lambda_2 \mathbf{u}_i^{(p)\dagger} \mathbf{u}_2^{(p)} \quad ; \quad i = 1, 2, 3$$

tendremos entonces

$$(\mathbf{H}_p)_{i1} = \lambda_1 \delta_{i1} \quad ; \quad (\mathbf{H}_p)_{i2} = \lambda_2 \delta_{i2}$$

<sup>18</sup>Este es un buen punto para hacer notar que la matriz de diagonalización no es única, lo cual se vé de inmediato debido a la falta de unicidad del conjunto de vectores propios que forman una base. Esta arbitrariedad se puede utilizar también para hacer que la matriz ortogonal real sea propia.

<sup>19</sup>El que una matriz sea real no garantiza que su espectro sea real. Esto está relacionado con el hecho de que la ecuación secular nos lleva a buscar las raíces de un polinomio con coeficientes reales. No obstante, el teorema fundamental del álgebra nos dice que las raíces pueden ser complejas incluso si los coeficientes son reales.

<sup>20</sup>Cuando los vectores son complejos  $n$  dimensionales, es decir que yacen en el espacio  $\mathbb{C}^n$  donde cada componente es compleja, la relación de ortonormalidad se define por  $(\mathbf{u}_i, \mathbf{u}_j) = \mathbf{u}_i^\dagger \mathbf{u}_j = \delta_{ij}$ . Para el caso de vectores reales, el tercer eje principal se puede generar con  $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$  que garantiza la obtención de ejes a derecha.

de lo cual resulta que

$$(\mathbf{H}_p)_{21} = (\mathbf{H}_p)_{31} = (\mathbf{H}_p)_{12} = (\mathbf{H}_p)_{32} = 0$$

el carácter hermítico de la matriz nos dice que los otros elementos fuera de la diagonal también son cero. De modo que la matriz construída en la base  $\mathbf{u}_1^{(p)}, \mathbf{u}_2^{(p)}, \mathbf{u}_3^{(p)}$  sí es diagonal y en consecuencia el tercer eje es también un eje principal. Nótese que el carácter hermítico de la matriz es necesario para aseverar que el tercer eje también es principal. Si la matriz es simétrica real, los ejes definidos por  $\mathbf{u}_1^{(p)}, \mathbf{u}_2^{(p)}$  y  $\mathbf{u}_3^{(p)}$  también serán ejes principales.

Ya hemos visto que si los valores propios no son todos distintos no podemos garantizar que los vectores propios sean ortogonales entre sí, y debemos en general hacer un proceso de ortogonalización del tipo Gram-Schmidt. Sin embargo, en virtud de lo anterior, este problema es aún más sencillo en tres dimensiones, al menos para ejes reales. Supongamos por ejemplo que  $\lambda_1 = \lambda_2 \neq \lambda_3$ , en este caso siempre será posible hallar al menos un vector propio que satisfaga la ecuación de valores propios asociada al valor propio degenerado  $\lambda_1$ , este vector tiene que ser ortogonal al vector propio asociado a  $\lambda_3$ , estos dos vectores propios ya me determinan dos ejes principales ortogonales de modo que el tercero se construye simplemente definiendo un sistema coordenado dextrógiro. En este caso sin embargo podemos decir algo más, si definimos una nueva base en donde rotamos solo los vectores  $\mathbf{u}_1^{(p)}$  y  $\mathbf{u}_2^{(p)}$  i.e. los vectores asociados a la degeneración, tenemos que

$$\begin{pmatrix} \mathbf{u}'_1 \\ \mathbf{u}'_2 \\ \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{pmatrix}$$

se puede verificar que los nuevos vectores  $\mathbf{u}'_1, \mathbf{u}'_2$  también son vectores propios con valor propio  $\lambda_1 = \lambda_2$ . Esto implica entonces que dados dos valores propios degenerados  $\lambda_1 = \lambda_2$  dos vectores propios linealmente independientes asociados a este valor propio degenerado me definen un plano en donde todos los vectores que yacen en tal plano son vectores propios. Los ejes principales son entonces cualquier par ortogonal de ejes en este plano. Tenemos entonces que esta degeneración está asociado a una simetría axial (alrededor del eje  $X_3$ ).

En nuestro lenguaje anterior, el espacio  $\mathbb{R}^3$  se puede escribir como la suma directa de los subespacios generados por los valores propios diferentes  $\lambda_1$  y  $\lambda_3$ . El subespacio generado por  $\lambda_1$  es de dimensión dos debido a su degeneración y es el plano  $\mathbb{R}^2$  expandido por los vectores propios  $\mathbf{u}_1$  y  $\mathbf{u}_2$  asociados a  $\lambda_1$ . El subespacio generado por  $\lambda_3$  es de dimensión uno y es un eje coordenado  $\mathbb{R}$  expandido por  $\mathbf{u}_3$

$$\mathbb{R}^3 = \mathbb{R}^2 \oplus \mathbb{R} \quad , \quad \lambda_1 \leftrightarrow \mathbb{R}^2 \quad , \quad \lambda_3 \leftrightarrow \mathbb{R}$$

y ya vimos que toda combinación lineal dentro de cada subespacio sigue siendo vector propio con el mismo valor propio. Es por ello que tenemos la libertad de rotar en el plano generado por  $\mathbf{u}_1$  y  $\mathbf{u}_2$  y continuar teniendo vectores propios  $\mathbf{u}'_1$  y  $\mathbf{u}'_2$  asociados a  $\lambda_1$ .

Análogamente, si todos los valores propios son iguales, todas las direcciones del espacio conducen a vectores propios, es decir tenemos una simetría esférica pero en este caso la matriz ya estará diagonalizada y será proporcional a la identidad, de modo que la degeneración total no representa tampoco un problema en tres dimensiones.

### 12.11.3. Matrices simétricas reales en $\mathbb{R}^3$

Finalmente, para matrices simétricas reales (i.e. hermíticas reales) es interesante la interpretación geométrica que se le puede dar a los ejes principales en  $\mathbb{R}^3$ . Tomemos un escalar definido por cierto vector unitario **real** en la forma

$$\tilde{\mathbf{n}}\mathbf{H}\mathbf{n} \equiv H \tag{12.103}$$

si denotamos  $\alpha_1, \alpha_2, \alpha_3$  a las componentes de  $\mathbf{n}$  tenemos

$$\mathbf{n} = \alpha_i \mathbf{u}_i$$

y podemos escribir la Ec. (12.103) usando el carácter simétrico de  $\mathbf{H}$

$$\alpha_i H_{ij} \alpha_j \equiv H \Rightarrow$$

$$H_{11}\alpha_1^2 + H_{22}\alpha_2^2 + H_{33}\alpha_3^2 + 2H_{12}\alpha_1\alpha_2 + 2H_{13}\alpha_1\alpha_3 + 2H_{23}\alpha_2\alpha_3 = H$$

No obstante será más práctico escribir esta ecuación en forma normalizada para lo cual definimos un vector  $\rho$  en la forma

$$\rho \equiv \frac{\mathbf{n}}{\sqrt{H}} \quad (12.104)$$

ahora dividimos por  $H$  a ambos lados de (12.103) con lo cual queda

$$\tilde{\rho} \mathbf{H} \rho = 1 \quad (12.105)$$

usando el carácter simétrico de  $\mathbf{H}$ , esta ecuación se escribe

$$H_{11}\rho_1^2 + H_{22}\rho_2^2 + H_{33}\rho_3^2 + 2H_{12}\rho_1\rho_2 + 2H_{13}\rho_1\rho_3 + 2H_{23}\rho_2\rho_3 = 1 \quad (12.106)$$

si la consideramos como una función en tres variables esta es la ecuación de una superficie en el espacio de configuración con ejes  $\rho_1, \rho_2, \rho_3$ . Podemos realizar un cambio de base a través de una matriz ortogonal real propia  $\mathbf{B}$ , reescribiendo (12.105) de la siguiente forma

$$\begin{aligned} \tilde{\rho} (\mathbf{B}^{-1}\mathbf{B}) \mathbf{H} (\mathbf{B}^{-1}\mathbf{B}) \rho &= 1 \Rightarrow \tilde{\rho} \tilde{\mathbf{B}} (\mathbf{B}\mathbf{H}\mathbf{B}^{-1}) \mathbf{B} \rho = 1 \\ &\Rightarrow \widetilde{(\mathbf{B}\rho)} (\mathbf{B}\mathbf{H}\mathbf{B}^{-1}) (\mathbf{B}\rho) = 1 \end{aligned}$$

y dado que para matrices simétricas reales es posible encontrar una matriz ortogonal real propia que la diagonalice con una transformación de similaridad, escogemos a  $\mathbf{B}$  como la matriz de diagonalización (que se construiría con los vectores propios de  $\mathbf{H}$ ) tenemos entonces que  $\mathbf{B}\mathbf{H}\mathbf{B}^{-1} = \boldsymbol{\lambda}$  siendo  $\boldsymbol{\lambda}$  la matriz diagonal con los valores propios en la diagonal. Resulta entonces

$$\tilde{\rho}' \boldsymbol{\lambda} \rho' = 1 ; \quad \rho' \equiv \mathbf{B}\rho \quad (12.107)$$

Esto implica que en general siempre es posible pasar a un sistema coordenado con ejes  $\rho'_1, \rho'_2, \rho'_3$  en donde la Ec. (12.106) toma su forma normal

$$\lambda_1\rho_1'^2 + \lambda_2\rho_2'^2 + \lambda_3\rho_3'^2 = 1 \quad (12.108)$$

Si todos los coeficientes  $\lambda_i$  son positivos, entonces la Ec. (12.108) define la superficie de un elipsoide, estando los ejes principales del elipsoide a lo largo de los nuevos ejes coordenados. Esto ocurrirá cuando la matriz  $\mathbf{H}$  sea una matriz positiva como se discutirá en la sección 12.14. A manera de ejemplo, se sabe que para el tensor de inercia los momentos principales de inercia (elementos diagonales del tensor) son positivos cualquiera que sea la inclinación de los ejes, lo cual garantiza que la forma diagonal tiene solo coeficientes positivos. La forma cuadrática (12.108) es precisamente la forma que tiene la Ec. (12.106) en un sistema de coordenadas en el cual la matriz simétrica real es diagonal. Es decir los ejes principales que vuelven diagonal a  $\mathbf{H}$  son los mismos que llevan a la ecuación de un elipsoide a su forma normal. Los valores propios en consecuencia determinan las longitudes de los ejes del elipsoide. Si dos valores propios coinciden, el elipsoide tendrá dos ejes iguales de modo que será un elipsoide de revolución manifestando la simetría axial asociada, si la degeneración es total, el elipsoide es una esfera manifestando la simetría esférica del problema.

Por supuesto es posible que los coeficientes en (12.108) y por lo tanto los valores propios de la matriz simétrica asociada no sean positivos, en tal caso estaremos transformando otra clase de superficie (en general cónicas) a su forma normal.

## 12.12. Matrices normales (opcional)

Hemos visto que una matriz  $n \times n$  se puede diagonalizar a través de una transformación de similaridad siempre que se pueda encontrar un conjunto de  $n$  vectores propios linealmente independientes, de tal modo que se pueda construir una matriz no singular de diagonalización. Adicionalmente, hemos visto que es deseable que la diagonalización se pueda realizar a través de una matriz unitaria u ortogonal real a fin de pasar de una sistema ortonormal a otro sistema ortonormal. Sin embargo, la diagonalización por matrices unitarias no es posible para cualquier matriz, queremos indagar entonces cual es el conjunto más general de matrices que pueden ser diagonalizadas por matrices unitarias a través de transformaciones de similaridad.

**Definition 12** Decimos que una matriz es **normal** cuando conmuta con su adjunto, esto lo simbolizamos como

$$\mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A} \equiv [\mathbf{A}, \mathbf{A}^\dagger] = 0$$

En esta sección, la notación de paréntesis cuadrados indicará conmutador y no corchete de Poisson. Puede demostrarse sin embargo, que los conmutadores obedecen un álgebra similar a la de los corchetes de Poisson Ecs. (8.9-8.12). Es inmediato ver que las matrices hermíticas y unitarias son casos particulares de matrices normales, por tanto todos los resultados que se deriven aquí serán válidos para matrices hermíticas y unitarias.

Partamos de la hipótesis de que la matriz  $\mathbf{A}$  es diagonalizable con una matriz unitaria, y veamos las condiciones que se obtienen con esta hipótesis

$$\mathbf{U}\mathbf{A}\mathbf{U}^\dagger = \lambda \quad \Leftrightarrow \quad \mathbf{U}\mathbf{A}^\dagger\mathbf{U}^\dagger = \lambda^*$$

donde la segunda ecuación es simplemente la adjunta de la primera. Multiplicando las dos ecuaciones en diferente orden tenemos

$$\begin{aligned} (\mathbf{U}\mathbf{A}\mathbf{U}^\dagger)(\mathbf{U}\mathbf{A}^\dagger\mathbf{U}^\dagger) &= \lambda\lambda^* \quad ; \quad (\mathbf{U}\mathbf{A}^\dagger\mathbf{U}^\dagger)(\mathbf{U}\mathbf{A}\mathbf{U}^\dagger) = \lambda^*\lambda \\ \mathbf{U}\mathbf{A}\mathbf{A}^\dagger\mathbf{U}^\dagger &= \lambda\lambda^* \quad ; \quad \mathbf{U}\mathbf{A}^\dagger\mathbf{A}\mathbf{U}^\dagger = \lambda\lambda^* \end{aligned}$$

donde hemos usado el carácter unitario de  $\mathbf{U}$  y el hecho de que dos matrices diagonales conmutan entre sí (ver ejercicio 3, Pág. 336). Restando las dos ecuaciones queda

$$\begin{aligned} \mathbf{U}(\mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A})\mathbf{U}^\dagger &= 0 \quad \Rightarrow \quad \mathbf{U}^\dagger[\mathbf{U}(\mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A})\mathbf{U}^\dagger]\mathbf{U} = \mathbf{U}^\dagger 0 \mathbf{U} \\ &\Rightarrow \quad \mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A} = 0 \end{aligned}$$

por tanto para que una matriz sea diagonalizable con una transformación unitaria, es **necesario** que la matriz sea normal. Recíprocamente, es posible probar que si la matriz es normal, entonces es diagonalizable por una transformación unitaria de similaridad. Probaremos el recíproco solo para el caso en el cual el espectro de valores propios no es degenerado.

Veremos antes algunas de las propiedades de las matrices normales. Escribamos la ecuación de valores propios y su correspondiente adjunta

$$\begin{aligned} (\mathbf{A} - \lambda\mathbf{1})\mathbf{x} &= 0 \quad ; \quad \mathbf{x}^\dagger(\mathbf{A}^\dagger - \lambda^*\mathbf{1}) = 0 \quad \Rightarrow \\ \mathbf{B}\mathbf{x} &= 0 \quad ; \quad \mathbf{x}^\dagger\mathbf{B}^\dagger = 0 \quad ; \quad \mathbf{B} \equiv \mathbf{A} - \lambda\mathbf{1} \end{aligned} \tag{12.109}$$

calculemos el siguiente conmutador

$$\begin{aligned} [(\mathbf{A} - \lambda\mathbf{1})^\dagger, (\mathbf{A} - \lambda\mathbf{1})] &= [\mathbf{A}^\dagger - \lambda^*\mathbf{1}, \mathbf{A} - \lambda\mathbf{1}] = [\mathbf{A}^\dagger, \mathbf{A} - \lambda\mathbf{1}] - [\lambda^*\mathbf{1}, \mathbf{A} - \lambda\mathbf{1}] \\ &= [\mathbf{A}^\dagger, \mathbf{A} - \lambda\mathbf{1}] = [\mathbf{A}^\dagger, \mathbf{A}] - [\mathbf{A}^\dagger, \lambda\mathbf{1}] = [\mathbf{A}^\dagger, \mathbf{A}] = 0 \end{aligned}$$



donde hemos usado las propiedades del conmutador Ecs. (8.9-8.12), el carácter normal de la matriz, y el hecho de que la identidad conmuta con cualquier matriz. Tenemos entonces que

$$\left[ (\mathbf{A} - \lambda \mathbf{1})^\dagger, (\mathbf{A} - \lambda \mathbf{1}) \right] \equiv \left[ \mathbf{B}^\dagger, \mathbf{B} \right] = 0 \quad (12.110)$$

de modo que la matriz  $\mathbf{B}$  definida en (12.109) es también normal. Usando (12.109) y (12.110) resulta

$$\begin{aligned} \mathbf{x}^\dagger \mathbf{B}^\dagger \mathbf{B} \mathbf{x} &= 0 = \mathbf{x}^\dagger \mathbf{B} \mathbf{B}^\dagger \mathbf{x} \Rightarrow (\mathbf{B}^\dagger \mathbf{x})^\dagger (\mathbf{B}^\dagger \mathbf{x}) = 0 \Rightarrow \\ \|\mathbf{B}^\dagger \mathbf{x}\|^2 &= 0 \Rightarrow \mathbf{B}^\dagger \mathbf{x} = 0 \Rightarrow (\mathbf{A}^\dagger - \lambda^* \mathbf{1}) \mathbf{x} = 0 \end{aligned}$$

por tanto, un autovector de  $\mathbf{A}$  es también autovector de  $\mathbf{A}^\dagger$  pero con autovalor  $\lambda^*$ . Partiendo de la ecuación de valores propios de  $\mathbf{A}^\dagger$  se puede demostrar análogamente que los autovectores de  $\mathbf{A}^\dagger$  también son autovectores de  $\mathbf{A}$ . Entonces, para matrices normales,  $\mathbf{A}^\dagger$  tiene los mismos autovectores de  $\mathbf{A}$  pero sus autovalores son los conjugados de los autovalores de  $\mathbf{A}$ .

Usando lo anterior, tomemos ahora la ecuación de valores propios para  $\mathbf{A}^\dagger$  y  $\mathbf{A}$  en la forma

$$\mathbf{A}^\dagger \mathbf{x}_i = \lambda_i^* \mathbf{x}_i \quad ; \quad \mathbf{A} \mathbf{x}_j = \lambda_j \mathbf{x}_j$$

tomando la adjunta de la primera ecuación y multiplicando la segunda ecuación por  $\mathbf{x}_i^\dagger$  se tiene

$$\mathbf{x}_i^\dagger \mathbf{A} = \lambda_i \mathbf{x}_i^\dagger \quad ; \quad \mathbf{x}_i^\dagger \mathbf{A} \mathbf{x}_j = \lambda_j \mathbf{x}_i^\dagger \mathbf{x}_j$$

multiplicando la primera ecuación por  $\mathbf{x}_j$  a la derecha y restando las dos ecuaciones resulta

$$(\lambda_i - \lambda_j) \mathbf{x}_i^\dagger \mathbf{x}_j = 0$$

para  $\lambda_i \neq \lambda_j$  se tiene que

$$\mathbf{x}_i^\dagger \mathbf{x}_j = (\mathbf{x}_i, \mathbf{x}_j) = 0$$

de modo que los autovectores de una matriz normal correspondientes a autovalores diferentes son ortogonales.

Ahora asumamos que el espectro de valores propios es no degenerado. Si  $\mathbf{A}$  es una matriz normal  $n \times n$ , la no degeneración indica que las  $n$  raíces de la ecuación secular son todas diferentes, y dado que los vectores propios asociados a valores propios diferentes son ortogonales, hemos obtenido  $n$  vectores propios ortogonales y por tanto linealmente independientes. El espectro será entonces completo. Lo anterior nos dice además que el subespacio generado por cada valor propio es unidimensional.

Adicionalmente, esto significa que si normalizamos los  $n$  vectores propios ortogonales, al construir la matriz de diagonalización sus columnas serán ortonormales (en el sentido del producto interno) y por tanto formarán una matriz unitaria.

La dificultad cuando hay degeneración, radica en que el número de raíces diferentes  $p$ , es menor a la dimensionalidad del espacio  $n$ . Podemos garantizar la construcción de  $p$  vectores ortogonales (linealmente independientes), pero no podemos garantizar a priori que se puedan construir  $n$  vectores propios linealmente independientes.

Vamos a tomar la hipótesis de **espectro completo** como válida para el caso degenerado y exploraremos sus consecuencias. En este caso, los valores propios degenerados deben generar subespacios de más de una dimensión, a fin de que el espacio completo  $\mathcal{E}^n$  se pueda escribir como suma directa de estos subespacios, como se aprecia en la Ec. (12.96). De acuerdo con la discusión de la sección 12.10, la base de vectores propios que genera a cada subespacio  $\mathcal{E}_i$  se puede ortonormalizar, y al igual que en la sección 12.11, los espacios  $\mathcal{E}_i$  y  $\mathcal{E}_j$  con  $i \neq j$ , son ortogonales ya que están asociados a valores propios diferentes. Por tanto, la ortonormalización de cada  $\mathcal{E}_i$  conduce a una base ortonormal para todo el espacio  $\mathcal{E}^n$  y con ello podemos formar una matriz unitaria de diagonalización.

Cuando existe degeneración, hay una mayor libertad para escoger los vectores ortonormales asociados a los valores propios degenerados. Pues dada una base ortonormal en un  $\mathcal{E}_i$  dado (de más de una dimensión),

podemos aplicarle una transformación unitaria que nos lleva a otra base ortonormal de  $\mathcal{E}_i$ , que igual nos sirve como parte de la base ortonormal de todo el espacio. Nótese que en ausencia de degeneración, cada subespacio  $\mathcal{E}_i$  es unidimensional y dada una base ortonormal, lo más que se puede hacer es multiplicar a cada vector por una fase compleja (transformación unitaria unidimensional), si la base es real y deseamos que continúe siendo real, los más que se puede hacer es multiplicar cada vector por  $\pm 1$ .

Tenemos entonces que una matriz es normal si y solo si puede diagonalizarse con una matriz unitaria. Este hecho se conoce usualmente como **teorema espectral**. Nótese que dicho teorema no prohíbe que otras matrices diferentes a las normales puedan ser diagonalizables. Sin embargo, si una matriz no es normal puede no ser diagonalizable, y si lo es, esta diagonalización debe obtenerse con una transformación de similaridad no unitaria.

Ya hemos dicho que las matrices unitarias y hermíticas son normales. Por tanto, una matriz unitaria puede ser diagonalizada por otra matriz unitaria ¿que significado puede tener la diagonalización de una matriz unitaria por otra matriz unitaria?. Esto tendrá sentido geométrico si la matriz que se diagonaliza es activa (por ejemplo una matriz que representa la rotación de un vector en  $\mathbb{C}^n$  o en  $\mathbb{R}^n$ ) en tanto que la matriz que diagonaliza debe ser pasiva (reorientación de los ejes en  $\mathbb{C}^n$  o en  $\mathbb{R}^n$ ).

### 12.13. Matrices positivas y definidas positivas

Un conjunto importante de matrices hermíticas son las matrices positivas y definidas positivas

**Definition 13** Una matriz hermítica  $\mathbf{H}$  de dimensión  $n \times n$  se denomina positiva si se cumple la relación

$$\mathbf{a}^\dagger \mathbf{H} \mathbf{a} \geq 0 \quad ; \quad \forall \mathbf{a} \in \mathbb{C}^n \quad (12.111)$$

adicionalmente, si existe por lo menos un vector  $\mathbf{a} \neq \mathbf{0}$  para el cual esta forma bilineal se anule, se dice que la matriz es positiva singular. Si  $\mathbf{a} = \mathbf{0}$  es el único vector que anula a la forma bilineal, se dice que la matriz es definida positiva. Si solo sabemos que se cumple la condición (12.111), diremos simplemente que la matriz es positiva.

La denominación de matriz positiva está relacionada con la estructura de su espectro

**Theorem 14** Una matriz hermítica  $\mathbf{H}$  de dimensión  $n \times n$  es positiva si y solo si sus valores propios son no-negativos.

**Demostración:** Supongamos primero que  $\mathbf{H}$  es positiva. Tomemos  $\mathbf{a} = \mathbf{x}_j$  siendo  $\mathbf{x}_j$  un vector propio asociado al valor propio  $\lambda_j$ . De (12.111) se tiene que

$$\mathbf{x}_j^\dagger \mathbf{H} \mathbf{x}_j \geq 0 \quad \Rightarrow \quad \lambda_j \mathbf{x}_j^\dagger \mathbf{x}_j = \lambda_j \|\mathbf{x}_j\|^2 \geq 0 \quad , \quad \text{no suma sobre } j$$

por tanto  $\lambda_j \geq 0$ .

Ahora suponemos que los valores propios de  $\mathbf{H}$  son no-negativos, y llegamos a que la matriz es positiva. Para ello recurrimos al hecho de que para una matriz hermítica es posible encontrar una base ortonormal completa  $\{\mathbf{x}_j\}$  de vectores propios, donde  $\mathbf{x}_j$  está asociado a un  $\lambda_j$  dado<sup>21</sup>. Por tanto, todo vector  $\mathbf{a} \in \mathbb{C}^n$  se puede escribir como combinación lineal de  $n$  vectores propios  $\mathbf{x}_j$  ortonormales de  $\mathbf{H}$ , en consecuencia

$$\begin{aligned} \mathbf{a}^\dagger \mathbf{H} \mathbf{a} &= (c_i \mathbf{x}_i)^\dagger \mathbf{H} (c_j \mathbf{x}_j) = c_i^* c_j \mathbf{x}_i^\dagger \mathbf{H} \mathbf{x}_j = c_i^* c_j \lambda_j \mathbf{x}_i^\dagger \mathbf{x}_j = c_i^* c_j \lambda_j \delta_{ij} \\ \mathbf{a}^\dagger \mathbf{H} \mathbf{a} &= \sum_i |c_i|^2 \lambda_i \geq 0 \end{aligned} \quad (12.112)$$

donde en el último paso usamos el carácter no-negativo del espectro. **QED.**

El siguiente teorema justifica la introducción de la terminología definida positiva y singular positiva

<sup>21</sup> Aquí no usamos índice de degeneración, simplemente  $j = 1, 2, \dots, n$ , y puede que algunos  $\lambda$ 's tengan el mismo valor.

**Theorem 15** *Una matriz hermítica es definida positiva si y solo si todos sus valores propios son estrictamente positivos. Una matriz hermítica es singular positiva si y solo si todos sus valores propios son no-negativos, y al menos uno de ellos es nulo. En particular, esto implica que las matrices singulares positivas no son invertibles, en tanto que las definidas positivas son invertibles.*

**Demostración:** Supongamos que la matriz hermítica tiene valores propios estrictamente positivos. Si  $\mathbf{a} \neq \mathbf{0}$  en la Ec. (12.112), entonces uno o más de los coeficientes  $c_i$  es no nulo y dado que cada  $\lambda_i > 0$ , es claro que  $\mathbf{a}^\dagger \mathbf{H} \mathbf{a}$  en (12.112), es mayor que cero, de modo que la matriz es definida positiva. Recíprocamente, si  $\mathbf{H}$  es definida positiva y hacemos  $\mathbf{a} = \mathbf{x}_i$  vector propio con valor propio  $\lambda_i$ , obtenemos

$$\mathbf{x}_i^\dagger \mathbf{H} \mathbf{x}_i = \|c_i\|^2 \lambda_i > 0 \quad \text{no suma sobre } i$$

donde hemos usado el hecho de que  $\mathbf{x}_i \neq \mathbf{0}$  y el carácter definido positivo de  $\mathbf{H}$ .

La negación de estas implicaciones conduce al resto del teorema<sup>22</sup>. Sin embargo, lo ilustraremos explícitamente para estudiar el comportamiento de las matrices singulares positivas.

Partimos primero de que la matriz es singular positiva, entonces existe por lo menos un vector  $\mathbf{a} \neq \mathbf{0}$  tal que

$$\mathbf{a}^\dagger \mathbf{H} \mathbf{a} = 0 \quad ; \quad \mathbf{a} \neq \mathbf{0}$$

escribiendo  $\mathbf{a} = c_i \mathbf{x}_i$  como combinación lineal de vectores propios  $\mathbf{x}_i$  de  $\mathbf{H}$ , y haciendo un procedimiento análogo al que nos llevó a (12.112), resulta

$$\mathbf{a}^\dagger \mathbf{H} \mathbf{a} = \|c_i\|^2 \lambda_i = 0 \quad ; \quad \mathbf{a} \neq \mathbf{0}$$

puesto que  $\mathbf{a} \neq \mathbf{0}$ , hay uno o varios  $c_i$  no-nulos. Dado que  $\lambda_i \geq 0$ , es necesario que todos los valores propios asociados a coeficientes  $c_i$  no nulos sean cero. Por tanto, si  $\mathbf{a} = c_i \mathbf{x}_i \neq \mathbf{0}$ , es un vector que genera una forma bilineal nula con  $\mathbf{H}$ , los valores propios asociados a todos los vectores propios que intervienen en esta combinación lineal deben ser nulos. Pueden existir por supuesto, mas valores propios nulos (i.e. un mayor grado de degeneración de  $\lambda = 0$ ), si existen varios vectores propios  $\mathbf{a}_k \neq \mathbf{0}$ , linealmente independientes que anulen a la forma bilineal. El grado de degeneración del valor propio nulo será entonces la dimensión del subespacio generado por  $\lambda = 0$ .

Recíprocamente, supongamos que cada  $\lambda_i \geq 0$ , y que existe por lo menos un  $\lambda_k = 0$ . Tomando  $\mathbf{a} = \mathbf{x}_k$  i.e. como uno de los vectores propios asociado a  $\lambda_k$ , la forma bilineal queda

$$\mathbf{x}_k^\dagger \mathbf{H} \mathbf{x}_k = \lambda_k \mathbf{x}_k^\dagger \mathbf{x}_k = \lambda_k \|\mathbf{x}_k\|^2 = 0 \quad ; \quad \mathbf{x}_k \neq \mathbf{0}$$

de modo que hay por lo menos un vector diferente de cero que anula la forma bilineal.

Finalmente, puesto que las matrices positivas son diagonalizables con una transformación de similaridad, el determinante es el producto de los valores propios. Por tanto, el determinante de una matriz definida positiva es positivo y la matriz es invertible, en tanto que el determinante de una matriz singular positiva es cero y por tanto es no invertible. **QED.**

### 12.13.1. Matrices simétricas reales que son positivas

Cuando la matriz hermítica se vuelve real, es decir cuando tenemos una matriz simétrica real, la condición de positividad se puede expresar de una manera más sencilla.

**Theorem 16** *Una matriz simétrica real  $\mathbf{M}$  es positiva si y solo si*

$$\tilde{\mathbf{v}} \mathbf{M} \mathbf{v} \geq 0 \quad ; \quad \forall \mathbf{v} \in \mathbb{R}^n \quad (12.113)$$

*Adicionalmente,  $\mathbf{M}$  es definida positiva si y solo si el único vector de  $\mathbb{R}^n$  que anula a la forma bilineal es  $\mathbf{v} = \mathbf{0}$ . Así mismo,  $\mathbf{M}$  es positiva singular si y solo si existe al menos un vector no nulo de  $\mathbb{R}^n$  que anula a la forma bilineal.*

<sup>22</sup>En forma un tanto mas precisa, hay que negar las dos implicaciones siguientes: Una matriz positiva es definida positiva si y solo si todos sus valores propios son estrictamente positivos.

**Demostración:** En la forma bilineal  $\mathbf{a}^\dagger \mathbf{M} \mathbf{a}$ , el vector  $\mathbf{a}$  es en general complejo. Separaremos  $\mathbf{a}$  en sus partes real e imaginaria

$$\mathbf{a} = \mathbf{v} + i\mathbf{w} \quad (12.114)$$

de modo que la forma bilineal se escribe como

$$\mathbf{a}^\dagger \mathbf{M} \mathbf{a} = (\mathbf{v}^\dagger - i\mathbf{w}^\dagger) \mathbf{M} (\mathbf{v} + i\mathbf{w}) = \tilde{\mathbf{v}} \mathbf{M} \mathbf{v} + \tilde{\mathbf{w}} \mathbf{M} \mathbf{w} + i(\tilde{\mathbf{v}} \mathbf{M} \mathbf{w} - \tilde{\mathbf{w}} \mathbf{M} \mathbf{v}) \quad (12.115)$$

donde hemos usado el hecho de que por construcción los vectores  $\mathbf{v}$ ,  $\mathbf{w}$  son reales y por tanto su adjunto coincide con su traspuesto. La parte imaginaria en (12.115) se anula debido a la simetría de  $\mathbf{M}$ , de modo que el factor  $\mathbf{a}^\dagger \mathbf{M} \mathbf{a}$  es real y queda en la forma

$$\mathbf{a}^\dagger \mathbf{M} \mathbf{a} = \tilde{\mathbf{v}} \mathbf{M} \mathbf{v} + \tilde{\mathbf{w}} \mathbf{M} \mathbf{w} \quad ; \quad \mathbf{a} = \mathbf{v} + i\mathbf{w} \quad ; \quad \mathbf{v}, \mathbf{w} \in \mathbb{R}^n \quad (12.116)$$

De esta ecuación se deducen fácilmente los resultados del teorema. **QED.**

Este teorema muestra una vez más que para matrices simétricas reales podemos restringirnos a trabajar en  $\mathbb{R}^n$  de forma consistente, incluso si trabajamos originalmente en  $\mathbb{C}^n$  (este proceso de restringirnos a  $\mathbb{R}^n$  a partir de  $\mathbb{C}^n$ , se conoce como un endorfismo).

## 12.14. Problema de valores propios modificado para matrices positivas

En el tratamiento de los osciladores acoplados, se llega a una ecuación matricial de la forma

$$\mathbf{V} \mathbf{a} = \lambda \mathbf{T} \mathbf{a} \quad (12.117)$$

Donde  $\mathbf{T}$  y  $\mathbf{V}$  son matrices,  $\lambda$  un número complejo y  $\mathbf{a}$  un vector columna. Nótese que la Ec. (12.117), es una ecuación de valores propios con estructura diferente a la mostrada en (12.35), ya que al operar  $\mathbf{V}$  sobre  $\mathbf{a}$  no obtenemos  $\mathbf{a}$  por una constante sino una constante multiplicada por  $\mathbf{T} \mathbf{a}$ . Adicionalmente, asumiremos que las matrices  $\mathbf{T}$  y  $\mathbf{V}$  en la Ec. (12.117) son simétricas y reales y además  $\mathbf{T}$  es definida positiva en tanto que  $\mathbf{V}$  es positiva. Con estas hipótesis de trabajo demostraremos las siguientes propiedades:

1. Los valores propios  $\lambda$  son reales no negativos<sup>23</sup>. Si además  $\mathbf{V}$  es definida positiva, los valores propios  $\lambda$  serán estrictamente positivos.
2. Las componentes de los autovectores  $\mathbf{a}$  son reales, excepto por una posible fase global.
3. Si  $\mathbf{V}$  es positiva singular, tendremos un valor propio nulo  $m$ -degenerado, donde  $m$  es el número de vectores propios reales  $\mathbf{a}_k$  linealmente independientes tales que  $\tilde{\mathbf{a}}_k \mathbf{V} \mathbf{a}_k = 0$ .
4. Los autovectores  $\mathbf{a}$  correspondientes a valores propios diferentes son ortogonales, aunque bajo una métrica diferente como veremos más adelante.
5. La matriz de autovectores  $\mathbf{A}$  diagonaliza tanto a  $\mathbf{T}$  como a  $\mathbf{V}$ . Esta diagonalización se realiza con una transformación de congruencia que definiremos más adelante. La matriz  $\mathbf{T}$  se diagonaliza a la identidad y la matriz  $\mathbf{V}$  se diagonaliza a una matriz cuya diagonal contiene los valores propios.

Para demostrar estas propiedades comencemos por rotular los valores propios  $\lambda_k$  y los correspondientes vectores propios  $\mathbf{a}_{(k)}$  de la matriz  $\mathbf{V}$  de dimensión  $n \times n$ . El paréntesis indica que estamos rotulando un vector propio y no una componente del vector. Es claro que  $k = 1, \dots, n$  y algunos de los  $n$  valores propios pueden tener el mismo valor.

Reescribimos la Ec. (12.117) en la forma

$$\mathbf{V} \mathbf{a}_{(k)} = \lambda_k \mathbf{T} \mathbf{a}_{(k)} \quad (12.118)$$

<sup>23</sup>Los valores  $\lambda$  en la Ec. (12.117) se denominarán valores propios y los vectores  $\mathbf{a}$  en dicha ecuación se denominarán vectores propios, a pesar de que no lo son en todo el sentido de la palabra.

la ecuación adjunta resultante (para cualquier otro valor y vector propio) es

$$\mathbf{a}_{(m)}^\dagger \mathbf{V} = \lambda_m^* \mathbf{a}_{(m)}^\dagger \mathbf{T} \quad (12.119)$$

donde hemos usado el carácter real y simétrico de  $\mathbf{T}$  y  $\mathbf{V}$ <sup>24</sup>. Multiplicando (12.119) por  $\mathbf{a}_{(k)}$  a la derecha, multiplicando (12.118) por  $\mathbf{a}_{(m)}^\dagger$  a la izquierda y sustrayendo se obtiene

$$\begin{aligned} 0 &= \lambda_k \mathbf{a}_{(m)}^\dagger \mathbf{T} \mathbf{a}_{(k)} - \lambda_m^* \mathbf{a}_{(m)}^\dagger \mathbf{T} \mathbf{a}_{(k)} \\ 0 &= (\lambda_k - \lambda_m^*) \mathbf{a}_{(m)}^\dagger \mathbf{T} \mathbf{a}_{(k)} \end{aligned} \quad (12.120)$$

cuando  $k = m$ , la Ec. (12.120) queda de la forma

$$(\lambda_k - \lambda_k^*) \mathbf{a}_{(k)}^\dagger \mathbf{T} \mathbf{a}_{(k)} = 0 \quad (12.121)$$

y dado que  $\mathbf{T}$  es definido positivo, si  $\mathbf{a} \neq 0$  tenemos

$$\mathbf{a}_{(k)}^\dagger \mathbf{T} \mathbf{a}_{(k)} > 0 \quad (12.122)$$

combinando (12.122) con (12.121) se obtiene que los  $\lambda_k$  son reales.

Ahora bien, con un procedimiento análogo al que se hizo para demostrar el teorema 9, Pág. 293, es posible demostrar que si  $\mathbf{T}$ ,  $\mathbf{V}$  y un valor propio dado  $\lambda_i$  son todos reales, siendo  $\mathcal{E}_i$  el subespacio generado por  $\lambda_i$ , entonces existe una base de vectores reales que genera a  $\mathcal{E}_i$ . En particular, existe una base ortonormal real que genera a  $\mathcal{E}_i$ . Adicionalmente, dado que todos los valores propios son reales, obtenemos que es siempre posible obtener una base ortonormal real de  $\mathcal{E}^n$ . Esto implica que si el espacio vectorial original es  $\mathbb{C}^n$ , es posible realizar un endorfismo que me restrinja consistentemente el problema de valores propios modificado (12.117) al espacio vectorial  $\mathbb{R}^n$ . Hay por supuesto aún un grado de arbitrariedad ya que el valor de una componente particular de los  $a_{j(k)}$  se puede escoger a voluntad sin violar la Ec. (12.117). Podemos entonces requerir que esta componente sea real y el carácter real de  $\lambda_k$  asegura entonces el carácter real de todas las otras componentes<sup>25</sup>.

Multipliquemos ahora la Ec. (12.118) por  $\tilde{\mathbf{a}}_{(k)}$  a la izquierda y resolvemos para  $\lambda_k$

$$\begin{aligned} \tilde{\mathbf{a}}_{(k)} \mathbf{V} \mathbf{a}_{(k)} &= \lambda_k \tilde{\mathbf{a}}_{(k)} \mathbf{T} \mathbf{a}_{(k)} \\ \lambda_k &= \frac{\tilde{\mathbf{a}}_{(k)} \mathbf{V} \mathbf{a}_{(k)}}{\tilde{\mathbf{a}}_{(k)} \mathbf{T} \mathbf{a}_{(k)}} \end{aligned} \quad (12.123)$$

y como los vectores  $\mathbf{a}_{(k)}$  son reales no nulos<sup>26</sup>, el teorema 16 nos dice que el numerador es no negativo y el denominador es estrictamente positivo en virtud de la carácter positivo de  $\mathbf{V}$  y definido positivo de  $\mathbf{T}$ . Por tanto los valores propios  $\lambda_k$  **son no negativos**. Si adicionalmente  $\mathbf{V}$  es definida positiva, el numerador es también positivo y por lo tanto los valores propios son estrictamente **positivos**. Por otro lado, si  $\mathbf{V}$  es singular positiva, existe un conjunto  $m$  de vectores no nulos linealmente independientes tales que  $\tilde{\mathbf{a}}_{(k)} \mathbf{V} \mathbf{a}_{(k)} = 0$ , por tanto tendremos  $m$  valores propios nulos.

Volvamos a la Ec. (12.120) para la cual hemos examinado el caso en que  $k = m$  Ec. (12.121). Estudiemos entonces el caso en el cual  $m \neq k$ , teniendo en cuenta que los vectores y valores propios son reales

$$(\lambda_k - \lambda_m) \tilde{\mathbf{a}}_{(m)} \mathbf{T} \mathbf{a}_{(k)} = 0 \quad (12.124)$$

<sup>24</sup>Estrictamente, solo estaríamos usando su carácter hermítico.

<sup>25</sup>Es obvio de la Ec. (12.117) que si un vector real  $\mathbf{a}$  es vector propio entonces también lo será  $e^{i\alpha} \mathbf{a}$ , es decir que cada componente del vector tiene la misma fase.

<sup>26</sup>Nótese que si los vectores tienen una fase global, el cociente dado en la Ec. (12.123) se debe escribir como

$$\lambda_k = \frac{\mathbf{a}_{(k)}^\dagger \mathbf{V} \mathbf{a}_{(k)}}{\mathbf{a}_{(k)}^\dagger \mathbf{T} \mathbf{a}_{(k)}}$$

y claramente se cancela dicha fase.

si las raíces  $\lambda_k$  y  $\lambda_m$  de la ecuación secular son distintas, la Ec. (12.124) solo se cumple si

$$\tilde{\mathbf{a}}_{(m)} \mathbf{T} \mathbf{a}_{(k)} = 0 \quad ; \quad m \neq k \quad (12.125)$$

ya se ha mencionado en varias ocasiones que los valores de los  $a_{j(k)}$  no están completamente fijados por la ecuación de valores propios (12.118). Podemos remover esta indeterminación exigiendo la condición

$$\tilde{\mathbf{a}}_{(k)} \mathbf{T} \mathbf{a}_{(k)} = \mathbf{1} \quad (12.126)$$

hay  $n$  de estas ecuaciones y por tanto queda fija la componente que antes era arbitraria en cada  $\mathbf{a}_{(k)}$ , excepto por un signo global ya que si cambiamos  $\mathbf{a}_{(k)} \rightarrow -\mathbf{a}_{(k)}$  aún se siguen cumpliendo (12.125) y (12.126). Si organizamos los vectores propios en columnas adyacentes para formar una matriz  $\mathbf{A}$  con elementos  $a_{jk}$  las Ecs. (12.125) y (12.126) se pueden sintetizar en una sola

$$\tilde{\mathbf{A}} \mathbf{T} \mathbf{A} = \mathbf{1} \quad (12.127)$$

si existe degeneración ya no será válido el argumento que nos llevó a la Ec. (12.125). Sin embargo, de manera similar al problema de valores propios ordinario, es posible establecer que una combinación lineal arbitraria de vectores propios asociados a un mismo valor propio es también un vector propio con el mismo valor propio. De esta forma, el conjunto de todos los vectores propios reales de  $\mathbf{V}$  asociados a un valor propio dado  $\lambda_k$  junto con el cero, forman un subespacio vectorial de  $\mathbb{R}^n$ , i.e. el subespacio generado por el valor propio, cuya dimensión  $p$  es el grado de degeneración de  $\lambda_k$ . Dentro de este subespacio podemos escoger cualquier conjunto de  $p$  vectores linealmente independientes como base del subespacio. En particular, podemos escoger una base ortonormal en el sentido dado por las Ecs. (12.125), para lo cual se puede utilizar un proceso generalizado de ortogonalización de Gram-Schmidt (ver ejercicio 8 Pág. 336).

En conclusión, incluso en presencia de degeneración es siempre posible encontrar un conjunto de coeficientes  $a_{jk}$  que satisfagan la ecuación de valores propios (12.118) y las Ecs. (12.125, 12.126) con lo cual se cumple la Ec. (12.127), dicha ecuación nos inspira para definir una relación semejante a la transformación de similaridad. Hablamos de la transformación de **congruencia** de  $\mathbf{C}$  por medio de  $\mathbf{A}$  cuando tenemos la relación

$$\mathbf{C}' = \tilde{\mathbf{A}} \mathbf{C} \mathbf{A} \quad (12.128)$$

si  $\mathbf{A}$  es ortogonal se tiene que  $\tilde{\mathbf{A}} = \mathbf{A}^{-1} \equiv \mathbf{B}$  y la transformación de congruencia se convierte en transformación de similaridad. La Ec. (12.127) nos dice entonces que  $\mathbf{A}$  transforma a  $\mathbf{T}$  por medio de una transformación de congruencia en una matriz diagonal, en particular en la matriz identidad  $n \times n$ .

En componentes, la ecuación de valores propios (12.118) se escribe

$$\begin{aligned} V_{ij} a_{j(k)} &= \lambda_k T_{ij} a_{j(k)} \\ V_{ij} a_{j(k)} &= T_{ij} a_{j(m)} \delta_{mk} \lambda_k \end{aligned}$$

no hay suma sobre  $k$ . Podemos definir una matriz diagonal  $\boldsymbol{\lambda}$  que contiene a los valores propios en la diagonal, de la forma  $\lambda_{mk} = \delta_{mk} \lambda_k$ . Si además escribimos  $a_{j(k)}$  como un arreglo matricial  $\mathbf{A}$  de los vectores propios se tiene que

$$V_{ij} a_{jk} = T_{ij} a_{jm} \lambda_{mk}$$

que matricialmente se escribe

$$\mathbf{V} \mathbf{A} = \mathbf{T} \mathbf{A} \boldsymbol{\lambda} \quad (12.129)$$

multiplicando (12.129) por  $\tilde{\mathbf{A}}$  a la izquierda

$$\tilde{\mathbf{A}} \mathbf{V} \mathbf{A} = \tilde{\mathbf{A}} \mathbf{T} \mathbf{A} \boldsymbol{\lambda}$$

que teniendo en cuenta (12.127) se convierte en

$$\tilde{\mathbf{A}} \mathbf{V} \mathbf{A} = \boldsymbol{\lambda} \quad (12.130)$$

las Ecs. (12.127) y (12.130) nos dicen entonces que la misma transformación de congruencia (ambas con la matriz  $\mathbf{A}$  de vectores propios) diagonaliza tanto a  $\mathbf{T}$  como a  $\mathbf{V}$ . La primera se diagonaliza a la identidad y la segunda a una matriz diagonal donde los elementos diagonales son los valores propios.

Reescribiendo (12.117) en la forma

$$(\mathbf{V} - \lambda \mathbf{T}) \mathbf{a} = 0$$

vemos que lo que tenemos es un conjunto homogéneo de ecuaciones para los coeficientes  $a_k$  que solo tiene solución no trivial cuando el determinante del sistema es nulo

$$|\mathbf{V} - \lambda \mathbf{T}| = 0 \tag{12.131}$$

por tanto, los valores propios son las raíces de esta ecuación secular (modificada). En coordenadas cartesianas y usando  $\mathbf{T} = \mathbf{1}$ , (12.131) se convierte en la ecuación secular usual

$$|\mathbf{V} - \lambda \mathbf{1}| = 0 \tag{12.132}$$

### 12.14.1. Diagonalización simultánea de dos formas cuadráticas

Una aplicación interesante de los resultados anteriores es la diagonalización simultánea de dos formas cuadráticas. Tomemos dos formas cuadráticas en las variables  $\eta_i$  y  $\chi_i$  con coeficientes reales de la forma

$$T = \sum_{i=1}^n \sum_{j \geq i}^n K_{ij} \chi_i \chi_j \quad ; \quad V = \sum_{i=1}^n \sum_{j \geq i}^n Z_{ij} \eta_i \eta_j \tag{12.133}$$

es conveniente escribir estas formas cuadráticas como una estructura bilineal mediada por matrices simétricas reales en la forma<sup>27</sup>

$$\begin{aligned} T &= \frac{1}{2} \chi_i T_{ij} \chi_j \quad , \quad V = \frac{1}{2} \eta_i V_{ij} \eta_j \quad ; \quad T_{ij} = T_{ji} = (1 + \delta_{ij}) K_{ij} \quad , \quad V_{ij} = V_{ji} = (1 + \delta_{ij}) Z_{ij} \\ T &= \frac{1}{2} \tilde{\chi} \mathbf{T} \chi \quad ; \quad \mathbf{V} = \frac{1}{2} \tilde{\eta} \mathbf{V} \eta \end{aligned} \tag{12.134}$$

Impongamos la condición de que la matriz simétrica real  $\mathbf{T}$  definida por los elementos  $T_{ij}$ , sea definida positiva y que la matriz simétrica real  $\mathbf{V}$  definida por los  $V_{ij}$  sea una matriz positiva. Las formas cuadráticas son entonces positivas. Hagamos una transformación de las coordenadas  $\chi$  y  $\eta$ , a otras coordenadas  $\psi$  y  $\zeta$  por medio de la matriz  $\mathbf{A}$  que se construye con los vectores propios de la Ec. (12.118)

$$\psi = \mathbf{A}^{-1} \chi \quad ; \quad \zeta = \mathbf{A}^{-1} \eta \tag{12.135}$$

$$\chi = \mathbf{A} \psi \quad ; \quad \eta = \mathbf{A} \zeta \tag{12.136}$$

por razones que veremos más adelante, llamaremos a  $\psi_i$  y a  $\zeta_i$  las **coordenadas normales** asociadas a las formas cuadráticas  $T$  y  $V$  respectivamente. Usando (12.136), la forma cuadrática positiva  $V$  Ec. (12.134) en términos de las coordenadas normales queda

$$\begin{aligned} V &= \frac{1}{2} \tilde{\eta} \mathbf{V} \eta = \frac{1}{2} (\tilde{\mathbf{A}} \zeta) \mathbf{V} (\mathbf{A} \zeta) = \frac{1}{2} (\tilde{\zeta} \tilde{\mathbf{A}}) \mathbf{V} (\mathbf{A} \zeta) \\ V &= \frac{1}{2} \tilde{\zeta} (\tilde{\mathbf{A}} \mathbf{V} \mathbf{A}) \zeta \end{aligned}$$

<sup>27</sup>Por supuesto, son posibles otras matrices para definir la forma bilineal, la condición es que se cumpla la igualdad

$$T = \sum_{i=1}^n \sum_{j \geq i}^n K_{ij} \chi_i \chi_j = \frac{1}{2} \chi_i T_{ij} \chi_j$$

que para valores fijos no diagonales de  $i$  y  $j$  con  $j > i$ , se cumple para todo  $\chi_i \chi_j$  si y solo si

$$K_{ij} = \frac{1}{2} \chi_i T_{ij} \chi_j + \frac{1}{2} \chi_j T_{ji} \chi_i \quad \text{no suma}$$

en virtud de que hay un doble conteo para cada par  $(i, j)$  no diagonal de la forma bilineal. Por ejemplo, se puede escoger  $T_{ij} = 2K_{ij}$  para  $i > j$  y  $K_{ij} = 0$  para  $i < j$  o viceversa. Sin embargo, es conveniente elegir la matriz simétrica y real a fin de que sea hermítica.

pero recordemos que  $\mathbf{A}$  diagonaliza a  $\mathbf{V}$  a través de una transformación de congruencia Ec. (12.130) de modo que la estructura de la forma cuadrática positiva se simplifica en estas coordenadas.

$$V = \frac{1}{2} \tilde{\boldsymbol{\zeta}} \boldsymbol{\lambda} \boldsymbol{\zeta} = \frac{1}{2} \lambda_k \zeta_k \zeta_k \quad (12.137)$$

la forma cuadrática  $T$  definida positiva, adquiere una forma aún más simple que la de  $V$  en las nuevas coordenadas  $\boldsymbol{\psi}$ , puesto que  $\mathbf{A}$  diagonaliza a  $\mathbf{T}$  y la convierte en la identidad, como se vé en (12.127)

$$\begin{aligned} T &= \frac{1}{2} \tilde{\boldsymbol{\chi}} \mathbf{T} \boldsymbol{\chi} = \frac{1}{2} (\widetilde{\mathbf{A}\boldsymbol{\psi}}) \mathbf{T} (\mathbf{A}\boldsymbol{\psi}) = \frac{1}{2} (\tilde{\boldsymbol{\psi}} \tilde{\mathbf{A}}) \mathbf{T} (\mathbf{A}\boldsymbol{\psi}) \\ T &= \frac{1}{2} \tilde{\boldsymbol{\psi}} (\tilde{\mathbf{A}} \mathbf{T} \mathbf{A}) \boldsymbol{\psi} \end{aligned}$$

y usando (12.127) esta forma cuadrática se reduce finalmente a

$$T = \frac{1}{2} \tilde{\boldsymbol{\psi}} \boldsymbol{\psi} = \frac{1}{2} \psi_i \psi_i \quad (12.138)$$

las Ecs. (12.137, 12.138) muestran que en coordenadas normales  $\boldsymbol{\psi}$  y  $\boldsymbol{\zeta}$ , tanto  $V$  como  $T$  se simplifican ya que se convierten en formas cuadráticas desacopladas, i.e. formas cuadráticas sin términos cruzados. Esto es consecuencia de que  $\mathbf{A}$  produce una transformación a ejes principales similar a la transformación a ejes principales que deja a una matriz hermítica en forma diagonal. Los ingredientes nuevos en este caso son: (a) la transformación es de congruencia (y no de similitud), (b) La misma transformación de congruencia diagonaliza simultáneamente a dos matrices  $\mathbf{T}$  y  $\mathbf{V}$ , como se aprecia en las Ecs. (12.136).

Esta técnica resulta muy poderosa para desacoplar sistemas de ecuaciones algebraicas y diferenciales. En la sección 15.2 aplicaremos esta técnica para encontrar los modos normales de una red de osciladores acoplados.

## 12.15. Interpretación geométrica de la diagonalización simultánea de una matriz positiva y otra definida positiva (opcional)<sup>28</sup>

### 12.15.1. Argumentación por geometría analítica

A priori la diagonalización simultánea de las dos matrices parece un incidente afortunado. Sin embargo, a través de un análisis geométrico alterno, veremos que en realidad esto es una consecuencia inmediata de la naturaleza positiva de  $\mathbf{V}$  y definida positiva de  $\mathbf{T}$ . Este análisis nos llevará además a la conclusión de que no podemos en general diagonalizar tres formas cuadráticas en forma simultánea.

Considerando que los vectores  $\boldsymbol{\eta}$  y  $\boldsymbol{\chi}$  yacen en  $\mathbb{R}^n$ , y teniendo en cuenta que  $T$  es simétrica y real, se sigue que siempre existe una transformación ortogonal real  $\mathbf{B}$  que lleve a un nuevo sistema de coordenadas cartesianas  $\mathbf{y}$  en el cual  $\mathbf{T}$  sea diagonal

$$\mathbf{y} = \mathbf{B}\boldsymbol{\chi} \quad (12.139)$$

la matriz  $\mathbf{B}$  debe transformar a  $\mathbf{T}$  mediante una transformación de similitud en una matriz diagonal  $\mathbf{C}$

$$\mathbf{B}\mathbf{T}\mathbf{B}^{-1} = \mathbf{B}\tilde{\mathbf{T}}\mathbf{B} = \mathbf{C} \quad (12.140)$$

la transformación inversa se escribe como

$$\boldsymbol{\chi} = \tilde{\mathbf{B}}\mathbf{y} \quad ; \quad \tilde{\boldsymbol{\chi}} = \tilde{\mathbf{y}}\mathbf{B} \quad (12.141)$$

usando (12.140) y (12.141), la forma cuadrática  $T$  dada por (12.134) se transforma de la siguiente manera

$$2T = \tilde{\boldsymbol{\chi}} \mathbf{T} \boldsymbol{\chi} = \tilde{\mathbf{y}} \mathbf{B} \tilde{\mathbf{T}} \mathbf{B} \mathbf{y} = \tilde{\mathbf{y}} \mathbf{C} \mathbf{y} = C_i y_i y_i \quad (12.142)$$

<sup>28</sup>Esta sección no es muy útil para cálculos prácticos, pero aporta una gran claridad conceptual sobre la diagonalización simultánea de dos formas cuadráticas. Puede omitirse en una primera lectura.



la forma cuadrática  $C_i y_i y_i$  solo puede ser definida positiva si cada término  $C_i$  es estrictamente positivo, lo cual se puede ver haciendo todas las  $y_i$  cero excepto una. Esto también se puede ver teniendo en cuenta que la diagonalización por una transformación de similaridad nos lleva a la matriz diagonal de los valores propios  $C_i$  de  $\mathbf{T}$ , los cuales son estrictamente positivos en virtud del teorema 15. De esta forma la ecuación para las  $y_i$

$$\frac{C_i}{2T} y_i y_i = 1 \quad (12.143)$$

define un elipsoide de dimensión  $n$  cuyos ejes principales van a lo largo de los ejes definidos por las nuevas coordenadas  $y_i$ , la Ec. (12.143) se puede convertir de forma inmediata en la ecuación de una esfera para un nuevo conjunto de variables  $z_i$

$$z_i = y_i \sqrt{C_i} \quad \text{no suma} \quad (12.144)$$

$$\frac{1}{2T} z_i z_i = 1 \quad (12.145)$$

podemos reescribir (12.144) en la forma

$$\begin{aligned} z_i &= \sqrt{C_j} \delta_{ij} y_j = D_{ij} y_j \\ D_{ij} &\equiv \delta_{ij} \sqrt{C_j} \quad \text{no suma} \end{aligned} \quad (12.146)$$

en forma matricial podemos escribir

$$\mathbf{z} = \mathbf{D} \mathbf{y} \quad ; \quad \mathbf{D} \equiv \mathbf{1}_{n \times n} \mathbf{C}_R \quad ; \quad \mathbf{C}_R \equiv \begin{pmatrix} \sqrt{C_1} \\ \sqrt{C_2} \\ \vdots \\ \sqrt{C_n} \end{pmatrix} \quad (12.147)$$

y la Ec. (12.145) para la forma cuadrática  $T$  queda

$$2T = z_i z_i = \tilde{\mathbf{z}} \mathbf{1}_{n \times n} \mathbf{z}$$

es decir  $\mathbf{T}$  ha sido transformado a la identidad a través de dos transformaciones sucesivas: primero una transformación de similaridad con una matriz ortogonal real Ec. (12.140) y luego un reescalamiento (12.144). La primera transformación comporta una rotación de ejes a los ejes principales en donde el elipsoide asociado a  $T$  queda en su forma normal, la segunda transformación no involucra una rotación de ejes, sino un cambio de escala de los ejes  $y_i$ . Lo que hace tal reescalamiento es estirar o comprimir los ejes principales de modo que el elipsoide se convierta en una esfera.

Ahora nos preguntamos que le hace estas transformaciones a la forma cuadrática positiva  $V$ . Hagamos las mismas transformaciones sobre las coordenadas  $\eta$  asociadas a  $V$

$$\mathbf{w} \equiv \mathbf{B} \eta \quad ; \quad \mathbf{x} \equiv \mathbf{D} \mathbf{w} \quad \Rightarrow \quad \mathbf{x} = \mathbf{D} \mathbf{B} \eta \quad (12.148)$$

la inversa será

$$\begin{aligned} \eta &= (\mathbf{D} \mathbf{B})^{-1} \mathbf{x} \quad \Rightarrow \quad \eta = \tilde{\mathbf{B}} \mathbf{D}^{-1} \mathbf{x} \\ \tilde{\eta} &= \tilde{\mathbf{x}} \mathbf{D}^{-1} \mathbf{B} \\ (D^{-1})_{ij} &= \frac{1}{\sqrt{C_j}} \delta_{ij} \end{aligned}$$

donde hemos usado que  $\mathbf{D}$  es diagonal y por tanto es simétrica y su inversa se obtiene con el recíproco de los elementos diagonales ver Ec. (12.146). La composición de estas dos transformaciones no diagonaliza en general a  $\mathbf{V}$ , pero la forma cuadrática  $V$  queda de la forma

$$2V = \tilde{\eta} \mathbf{V} \eta = (\tilde{\mathbf{x}} \mathbf{D}^{-1} \mathbf{B}) \mathbf{V} (\tilde{\mathbf{B}} \mathbf{D}^{-1} \mathbf{x}) = \tilde{\mathbf{x}} \left[ \mathbf{D}^{-1} \mathbf{B} \mathbf{V} \tilde{\mathbf{B}} \mathbf{D}^{-1} \right] \mathbf{x}$$

y dado que  $\mathbf{D}^{-1}$  y  $\mathbf{V}$  son simétricas es inmediato que la matriz

$$\mathbf{M} \equiv \mathbf{D}^{-1} \mathbf{B} \mathbf{V} \tilde{\mathbf{B}} \mathbf{D}^{-1} \quad (12.149)$$

es simétrica, adicionalmente todas las matrices involucradas son reales luego  $\mathbf{M}$  es simétrica y real. Nótese que en general el producto  $\mathbf{B} \mathbf{V} \tilde{\mathbf{B}}$  no necesariamente es diagonal, ya que la diagonalización de  $\mathbf{V}$  con la misma transformación de congruencia, solo está garantizada si tal transformación diagonaliza a  $\mathbf{T}$  a la identidad. Sin embargo, la matriz  $\mathbf{B}$  diagonaliza a  $\mathbf{T}$  pero no la convierte en la identidad (esto solo ocurre si  $\mathbf{T}$  ya es la identidad). No obstante, el hecho de que  $V$  se escriba

$$2V = \tilde{\mathbf{x}} \mathbf{M} \mathbf{x}$$

siendo  $\mathbf{M}$  simétrica y real nos garantiza que existe una transformación ortogonal real  $\mathbf{F}$  que diagonaliza a  $\mathbf{M}$  por medio de una transformación de similaridad

$$\boldsymbol{\zeta} = \mathbf{F} \mathbf{x} \Rightarrow \mathbf{x} = \tilde{\mathbf{F}} \boldsymbol{\zeta} \Rightarrow \tilde{\mathbf{x}} = \tilde{\boldsymbol{\zeta}} \mathbf{F} \quad (12.150)$$

$$2V = \tilde{\mathbf{x}} \mathbf{M} \mathbf{x} = \tilde{\boldsymbol{\zeta}} \mathbf{F} \mathbf{M} \tilde{\mathbf{F}} \boldsymbol{\zeta} = \tilde{\boldsymbol{\zeta}} \boldsymbol{\lambda} \boldsymbol{\zeta} = \lambda_i \zeta_i \zeta_i \quad (12.151)$$

como la idea es diagonalizar a las dos matrices  $\mathbf{T}$  y  $\mathbf{V}$  con la misma transformación, debemos realizar esta tercera transformación en las variables asociadas a  $\mathbf{T}$  (que son las  $z_i$  luego de las dos transformaciones anteriores), para lo cual definimos unas nuevas variables  $\psi_i$  asociadas a  $T$

$$\boldsymbol{\psi} = \mathbf{F} \mathbf{z} \quad (12.152)$$

$$2T = \tilde{\mathbf{z}} \mathbf{1} \mathbf{z} = \tilde{\boldsymbol{\psi}} \mathbf{F} \mathbf{1} \tilde{\mathbf{F}} \boldsymbol{\psi} = \tilde{\boldsymbol{\psi}} \mathbf{1} \boldsymbol{\psi}$$

vemos que esta última rotación no afecta la estructura de la forma cuadrática definida positiva  $T$ , de modo que el tensor sigue siendo representado por la identidad. Geométricamente, esto se vé con el hecho de que la esfera asociada no establece ninguna diferencia cuando se rotan los ejes. Aquí tenemos entonces la explicación de porqué una matriz definida positiva y otra positiva se pueden diagonalizar simultáneamente: la superficie asociada a la forma cuadrática definida positiva es un elipsoide que se puede transformar en una esfera mediante dos transformaciones sucesivas, una rotación ortogonal de ejes seguido de un reescalamiento de ejes. Aunque esta transformación no diagonaliza a  $\mathbf{V}$ , nos permite hacer una transformación ortogonal adicional ya que cualquier rotación adicional sigue definiendo ejes principales para la esfera. Esta libertad para rotar los ejes sin romper la estructura cuadrática de  $T$  (i.e. sin cambiar la representación para  $\mathbf{T}$ ) es la que permite diagonalizar a  $\mathbf{V}$ . En resumen, la diagonalización de las dos formas cuadráticas se realiza comenzando con una transformación ortogonal real  $\mathbf{B}$  que diagonaliza a  $\mathbf{T}$  seguida de un reescalamiento de ejes  $\mathbf{D}$  que lo trivializa, y otra transformación ortogonal  $\mathbf{F}$  que no cambia la representación de  $\mathbf{T}$ . Por el lado de  $\mathbf{V}$  las dos primeras transformaciones no diagonalizan a  $\mathbf{V}$  pero la matriz resultante resulta ser simétrica y real lo cual posibilita la diagonalización de esa matriz a través de la última transformación  $\mathbf{F}$ .

En aras de organizar la información resumiremos los resultados aquí obtenidos: hemos partido de dos formas cuadráticas en términos de las variables  $\boldsymbol{\chi}$ ,  $\boldsymbol{\eta}$  descritas por la Ec. (12.134)

$$T = \frac{1}{2} \tilde{\boldsymbol{\chi}} \mathbf{T} \boldsymbol{\chi} \quad ; \quad V = \frac{1}{2} \tilde{\boldsymbol{\eta}} \mathbf{V} \boldsymbol{\eta}$$

donde la matriz  $\mathbf{T}$  es definida positiva y la matriz  $\mathbf{V}$  es positiva, ambas son simétricas y reales. Hemos reescrito estas formas cuadráticas en términos de nuevas variables  $\boldsymbol{\psi}$ ,  $\boldsymbol{\zeta}$  que denominamos coordenadas normales en virtud de que en ellas la forma cuadrática se desacopla según se ve en las Ecs. (12.137, 12.138)

$$V = \frac{1}{2} \tilde{\boldsymbol{\zeta}} \boldsymbol{\lambda} \boldsymbol{\zeta} = \frac{1}{2} \lambda_k \zeta_k \zeta_k \quad ; \quad T = \frac{1}{2} \tilde{\boldsymbol{\psi}} \boldsymbol{\psi} = \frac{1}{2} \psi_i \psi_i$$

la transformación que nos lleva de las coordenadas originales a las normales se puede rastrear en las Ecs. (12.139, 12.147, 12.152) y las Ecs. (12.148, 12.150)

$$\boldsymbol{\psi} = \mathbf{F} \mathbf{D} \mathbf{B} \boldsymbol{\chi} \quad ; \quad \boldsymbol{\zeta} = \mathbf{F} \mathbf{D} \mathbf{B} \boldsymbol{\eta} \quad (12.153)$$

Y comparando (12.153) con (12.135), vemos que la matriz de transformación que nos lleva desde las coordenadas originales a las normales viene dada por

$$\mathbf{A}^{-1} = \mathbf{FDB} \Rightarrow \mathbf{A} = \widetilde{\mathbf{B}}\mathbf{D}^{-1}\widetilde{\mathbf{F}} \quad (12.154)$$

donde hemos usado el hecho de que  $\mathbf{B}$  y  $\mathbf{F}$  son ortogonales. Por otro lado la matriz que transforma a  $\mathbf{T}$  y  $\mathbf{V}$  se obtiene fácilmente de (12.149, 12.151)

$$\lambda = \mathbf{FM}\widetilde{\mathbf{F}} = \mathbf{FD}^{-1}\mathbf{BV}\widetilde{\mathbf{B}}\mathbf{D}^{-1}\widetilde{\mathbf{F}} = \mathbf{FD}^{-1}\mathbf{BV}\widetilde{\mathbf{B}}\widetilde{\mathbf{D}}^{-1}\widetilde{\mathbf{F}} = (\mathbf{FD}^{-1}\mathbf{B})\mathbf{V}(\widetilde{\mathbf{FD}^{-1}\mathbf{B}}) \quad (12.155)$$

y a partir de (12.130) se tiene que

$$\lambda = \widetilde{\mathbf{A}}\mathbf{VA} \quad (12.156)$$

comparando (12.155) con (12.156) vemos que

$$\mathbf{A} = (\widetilde{\mathbf{FD}^{-1}\mathbf{B}}) = \widetilde{\mathbf{B}}\mathbf{D}^{-1}\widetilde{\mathbf{F}}$$

en concordancia con la Ec. (12.154). La matriz de transformación de  $\mathbf{T}$  y  $\mathbf{V}$  es entonces

$$\widetilde{\mathbf{A}} \equiv \mathbf{FD}^{-1}\mathbf{B} \quad ; \quad \widetilde{\mathbf{A}}\mathbf{VA} = \lambda \quad ; \quad \widetilde{\mathbf{A}}\mathbf{TA} = \mathbf{1} \quad (12.157)$$

es fácil demostrar que  $\mathbf{A}$  es ortogonal si y solo si  $\mathbf{D}^{-1}$  es ortogonal. Puesto que en general  $\mathbf{D}^{-1}$  no es ortogonal, tampoco lo será  $\mathbf{A}$ , con lo cual las transformaciones de congruencia (12.127, 12.130) no son en general transformaciones de similaridad. Es fácil también ver las condiciones necesarias y suficientes para que  $\mathbf{D}^{-1}$  sea ortogonal. Imponiendo la condición de ortogonalidad (12.51) a  $\mathbf{D}^{-1}$  resulta

$$(\mathbf{D}^{-1})_{ij}(\mathbf{D}^{-1})_{ik} = \delta_{jk} \Rightarrow \frac{\delta_{ij}}{\sqrt{C_i}} \frac{\delta_{ik}}{\sqrt{C_i}} = \delta_{jk} \Rightarrow \delta_{jk} = C_i \delta_{jk} \Rightarrow C_i = 1$$

es decir  $\mathbf{D}^{-1}$  es ortogonal si y solo si  $C_i = 1$  para todo  $i$ , esto implica que todos los valores propios de  $\mathbf{T}$  son iguales a la unidad, de modo que  $\mathbf{T}$  debe ser la identidad. Esto también puede verse teniendo en cuenta que si  $\mathbf{A}$  es ortogonal, entonces  $\mathbf{T}$  es equivalente a la identidad (ya que  $\mathbf{1}$  y  $\mathbf{T}$  estarían conectados por una transformación de similaridad), y la identidad solo es equivalente consigo misma.

Si  $\mathbf{T} = C\mathbf{1}$  es fácil ver que  $\mathbf{D}^{-1} = \mathbf{1}/\sqrt{C}$ , de modo que todos los ejes se reescalan de la misma forma. Geométricamente, puesto que  $\mathbf{T}$  es proporcional a la identidad, la forma cuadrática  $T$  ya es una esfera en la base de coordenadas original<sup>29</sup>, de modo que solo tenemos que “normalizar” el radio de la esfera para que  $\mathbf{T}$  se diagonalice a la identidad, por lo cual el factor de escala es el mismo para todos los ejes. Nótese que esta transformación no es de similaridad, ya que los valores propios de  $\mathbf{T} = C\mathbf{1}$  son diferentes a los de la identidad a menos que  $C = 1$ .

Nótese que si las dos formas cuadráticas son positivas pero ninguna es definida positiva, entonces la forma normal de las cuadráticas es también de la forma (12.142), pero en virtud del teorema 15 por lo menos uno de los coeficientes es nulo, esto imposibilita que en general esta superficie defina un elipsoide que se pueda convertir en una esfera con un reescalamiento. Es por esto que por lo menos una de las formas cuadráticas debe ser definida positiva. Por ejemplo, si uno o más de los coeficientes en (12.142) es cero, la ecuación define un elipsoide de menor dimensión que solo se puede convertir en una esfera también de menor dimensión. Tomemos el caso tridimensional, si uno de los coeficientes es cero entonces la ecuación define una elipse que se convierte en círculo, pero un círculo no permanece invariante ante una rotación arbitraria de ejes **en tres dimensiones**.

### Diagonalización simultánea de tres formas cuadráticas

Los razonamientos anteriores nos explican también porqué en general no se puede diagonalizar tres formas cuadráticas digamos  $T$ ,  $V_1$  y  $V_2$  donde la primera es definida positiva y las otras dos son positivas singulares.

<sup>29</sup>La simetría esférica a su vez se manifiesta en la total degeneración de los valores propios de  $\mathbf{T}$  i.e.  $\lambda_1 = \lambda_2 = \lambda_3 = C$ .

La tercera rotación se puede usar para diagonalizar digamos a  $V_1$ , pero  $V_2$  queda en general no diagonal, si intentamos una cuarta rotación para diagonalizar a  $V_2$  ocurre en general que se estropea la diagonalización de  $V_1$ . Esto debido a que no se puede convertir a  $V_1$  ó  $V_2$  en una esfera de la dimensión del espacio, que quede invariante bajo rotaciones.

Si tenemos dos formas definidas positivas  $T_1$ ,  $T_2$  y una positiva  $V$ , es evidente que se podrán diagonalizar las tres formas simultáneamente cuando las superficies elipsoidales están orientadas de la misma forma y sean semejantes (i.e. los cocientes entre sus ejes tomados digamos en orden ascendente son todos iguales). Lo primero asegura que la misma reorientación de ejes lleva ambos elipsoides a su forma normal y lo segundo asegura que el mismo reescalamiento convierte en esferas a las dos formas cuadráticas definidas positivas. Al tener dos esferas para  $T_1$  y  $T_2$  tenemos libertad para rotar y diagonalizar la tercera forma cuadrática sin romper la diagonalización de las otras dos. En términos algebraicos, podemos diagonalizar simultáneamente a dos matrices definidas positivas  $T_1, T_2$  y una matriz positiva  $V$ , cuando (a)  $T_1$  es diagonal si y solo si  $T_2$  es diagonal. (b) Si  $\Lambda_i^{(1)}$ ,  $\Lambda_i^{(2)}$  son los valores propios ordinarios de  $T_1$  y  $T_2$  respectivamente, entonces<sup>30</sup>

$$\frac{\Lambda_i^{(2)}}{\Lambda_i^{(1)}} = C \quad \text{para } i = 1, \dots, n$$

siendo  $C$  una constante independiente de  $i$ . Estas dos condiciones son equivalentes a la ecuación

$$\mathbf{T}_2 = C\mathbf{T}_1 \quad (12.158)$$

Para trivializar a  $T_1$ , usamos la matriz de vectores propios de la ecuación

$$\mathbf{V}\mathbf{a}_{(k)} = \lambda_k \mathbf{T}_1 \mathbf{a}_{(k)}$$

y la transformación de congruencia producirá las diagonalizaciones

$$\tilde{\mathbf{A}}\mathbf{V}\mathbf{A} = \boldsymbol{\lambda} \quad ; \quad \tilde{\mathbf{A}}\mathbf{T}_1\mathbf{A} = \mathbf{1} \quad ; \quad \tilde{\mathbf{A}}\mathbf{T}_2\mathbf{A} = C\mathbf{1} \quad (12.159)$$

a pesar de que en el análogo geométrico  $C$  sería positiva, es obvio que esto es válido para una constante arbitraria, incluso compleja.

### 12.15.2. Argumentación por geometría de Riemann

Podemos darle una interpretación geométrica a la ecuación de valores propios modificada (12.118), teniendo en cuenta que la Ec. (12.127) se asemeja a una condición de ortonormalidad (en coordenadas cartesianas  $\mathbf{T} = \mathbf{1}$  y obtenemos la condición de ortonormalidad). Supondremos que los espacios vectoriales en cuestión son reales, así como todas las componentes de vectores y matrices. En realidad, la ecuación (12.127) se puede interpretar como una condición de ortonormalidad que podemos tomar con dos enfoques.

En un enfoque pasivo,  $\mathbf{T}$  puede considerarse el tensor métrico asociado a una base de vectores de  $\mathbb{R}^n$  que no es ortonormal (ver sección 12.9). De modo que el producto interno está mediado por dicho tensor métrico según la Ec. (12.89)

$$(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}}\mathbf{T}\mathbf{y} \quad (12.160)$$

con lo cual claramente la Ec. (12.127) o equivalentemente las Ecs. (12.125, 12.126), expresarían la ortonormalidad con el producto interno (12.160). Dado que la norma es un invariante ante un cambio de base y es positiva para todos los vectores no nulos se tiene que

$$\|\mathbf{x}\|^2 = (\mathbf{x}, \mathbf{x}) = \tilde{\mathbf{x}}\mathbf{T}\mathbf{x} \geq 0 \quad , \quad \forall \mathbf{x} \in \mathbb{R}^n$$

y solo es cero para  $\mathbf{x} = \mathbf{0}$ . Por tanto, el tensor métrico tiene que ser definido positivo como efectivamente supusimos para  $\mathbf{T}$ .

<sup>30</sup>Esta notación se utiliza para no confundir estos valores propios con los valores propios modificados de  $\mathbf{V}$  dados por  $\mathbf{V}\mathbf{a} = \lambda\mathbf{T}\mathbf{a}$ .

En un enfoque activo, los vectores propios  $\mathbf{a}_{(k)}$  son ortonormales, pero en un espacio de Riemann que no es necesariamente cartesiano (un espacio que no es plano sino curvo)<sup>31</sup>. En un espacio de Riemann se define el elemento de longitud de camino en la forma

$$ds^2 = g_{ik} dx_i dx_k$$

donde los  $g_{ik}$  son elementos del **tensor métrico**  $\mathbf{G}$  del espacio. En dicho espacio se define un producto escalar inducido por el tensor métrico (ver Ec. 12.88)

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = x_i g_{ij} y_j = \tilde{\mathbf{x}} \mathbf{G} \mathbf{y} \quad ; \quad g_{ij} \equiv (\mathbf{v}_i, \mathbf{v}_j)$$

siendo  $\mathbf{v}_i$  los vectores de la base elegida. Este tensor induce a su vez una norma o módulo para los vectores

$$\|\mathbf{x}\|^2 \equiv (\mathbf{x}, \mathbf{x}) = \tilde{\mathbf{x}} \mathbf{G} \mathbf{x} \geq 0 \quad (12.161)$$

y por las mismas razones anteriores, esto conduce a que el tensor métrico debe estar representado por una matriz definida positiva. Adicionalmente, se vé de su definición que es una matriz simétrica real. Si el espacio se vuelve plano (cartesiano), el tensor métrico se reduce a la identidad.

Aunque usualmente nos referiremos a la interpretación activa, cada parte del desarrollo subsecuente tiene también su interpretación en el enfoque pasivo. Podemos decir que la matriz  $\mathbf{T}$  cumple con las propiedades de un tensor métrico ya que el módulo de un vector sería no negativo y es nulo solo si corresponde al vector nulo.  $\mathbf{T}$  es por definición una matriz simétrica real y definida positiva como lo es el tensor métrico. En otras palabras, en un espacio de configuración en el cual  $\mathbf{T}$  sea el tensor métrico, el término  $\tilde{\mathbf{a}} \mathbf{T} \mathbf{a}$  define una norma al cuadrado para un vector en el espacio de configuraciones. En tal espacio, la Ec. (12.127) o las correspondientes ecuaciones por vectores componentes (12.125, 12.126), representan una condición de ortonormalidad.

Veamos primero como hacer el cambio de una base  $\{\mathbf{v}_i^{\mathbf{T}}\}$  que no es ortonormal a otra base  $\{\mathbf{v}_i^{\mathbf{I}}\}$  que sí es ortonormal. Escribamos el producto interno para el espacio de Riemann con tensor métrico  $\mathbf{T}$

$$(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}} \mathbf{T} \mathbf{y} = \tilde{\mathbf{x}} (\mathbf{B}^{-1} \mathbf{B}) \mathbf{T} (\mathbf{B}^{-1} \mathbf{B}) \mathbf{y}$$

elijamos a  $\mathbf{B}$  como una matriz ortogonal real. Dado que  $\mathbf{T}$  es una matriz simétrica real, siempre es posible encontrar una matriz ortogonal real que la diagonalice, por tanto para una elección apropiada de  $\mathbf{B}$  se tiene

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}} \tilde{\mathbf{B}} (\mathbf{B} \mathbf{T} \mathbf{B}^{-1}) \mathbf{B} \mathbf{y} = \widetilde{(\mathbf{B} \mathbf{x})} \mathbf{C} (\mathbf{B} \mathbf{y}) = \tilde{\mathbf{x}}^a \mathbf{C} \mathbf{y}^a \Rightarrow \\ (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}} \mathbf{T} \mathbf{y} = \tilde{\mathbf{x}}^a \mathbf{C} \mathbf{y}^a \quad ; \quad \mathbf{x}^a \equiv \mathbf{B} \mathbf{x} \end{aligned} \quad (12.162)$$

donde  $\mathbf{C}$  es una matriz diagonal. La Ec. (12.162) nos muestra que la transformación de coordenadas inducida por  $\mathbf{B}$  nos lleva a un tensor métrico  $\mathbf{C}$  diagonal

$$C_{ij} = (\mathbf{v}_i^a, \mathbf{v}_j^a) = \delta_{ij} C_i$$

donde los  $C_i$  son los valores propios de  $\mathbf{T}$ . Esto quiere decir que los vectores que forman la nueva base son ortogonales pero no están normalizados, es decir su norma no es la unidad. Esto es lógico, ya que si la base original no estaba normalizada, una transformación ortogonal real la puede reorientar para volverla ortogonal pero no puede cambiar su norma. Esto nos indica que el proceso de normalización no se puede llevar a cabo con una transformación ortogonal real.

Si cada  $C_i$  es diferente, tenemos que cada vector base tiene norma diferente (que corresponde a  $\sqrt{C_i}$ ). Pasemos ahora a una base normalizada

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}}^a \mathbf{C} \mathbf{y}^a = x_i^a C_{ij} y_j^a = x_i^a (\mathbf{v}_i^a, \mathbf{v}_j^a) y_j^a = \left( \sqrt{C_i} x_i^a \right) \left( \frac{\mathbf{v}_i^a}{\sqrt{C_i}}, \frac{\mathbf{v}_j^a}{\sqrt{C_i}} \right) \left( \sqrt{C_i} y_j^a \right) \\ (\mathbf{x}, \mathbf{y}) &= x_i^b \left( \hat{\mathbf{v}}_i^b, \hat{\mathbf{v}}_j^b \right) y_j^b \end{aligned}$$

<sup>31</sup>En este caso, un enfoque pasivo se refiere a un cambio de base dentro del mismo espacio  $\mathbb{R}^n$ , y un enfoque activo se refiere a una transformación del espacio como tal.

donde hemos definido

$$x_i^b \equiv \sqrt{C_i} x_i^a \quad ; \quad \hat{\mathbf{v}}_i^b \equiv \frac{\mathbf{v}_i^a}{\sqrt{C_i}} \quad (12.163)$$

claramente los vectores  $\hat{\mathbf{v}}_i^b$  son ahora unitarios y ortogonales, el producto interno se escribe entonces en la forma

$$(\mathbf{x}, \mathbf{y}) = x_i^b \delta_{ij} y_j^b = x_i^b y_i^b$$

Esta relación se puede construir también por componentes

$$(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}}^a \mathbf{C} \mathbf{y}^a = x_i^a C_{ij} y_j^a = x_i^a \delta_{ij} C_{ij} y_j^a = x_i^a C_{ii} y_i^a = \left( \sqrt{C_i} x_i^a \right) \left( \sqrt{C_i} y_i^a \right) = x_i^b y_i^b$$

las Ecs. (12.163), se puede escribir matricialmente en la forma

$$\mathbf{x}^b \equiv \mathbf{D} \mathbf{x}^a \quad ; \quad \hat{\mathbf{v}}_i^b \equiv \mathbf{D}^{-1} \mathbf{v}_i^a$$

donde  $\mathbf{D}$  es la matriz de reescalamiento descrita en las Ecs. (12.146, 12.147). Tenemos que el tensor métrico es ahora la identidad

$$(\mathbf{x}, \mathbf{y}) = x_i^b y_i^b = \tilde{\mathbf{x}}^b \mathbf{y}^b = \tilde{\mathbf{x}}^b \mathbf{1} \mathbf{y}^b$$

hemos partido de una base que no es ortogonal ni normal para llegar a una base ortonormal. Esto se hizo con dos transformaciones: la primera es una transformación asociada a una matriz ortogonal que convierte la base original en una base ortogonal pero que no está normalizada. La segunda es una transformación de escala que se usa para normalizar los vectores de la base ortogonal, los factores de escala son efectivamente las normas de los vectores de la base ortogonal. El tensor métrico asociado es ahora la identidad. Nótese que aquí podríamos realizar una tercera transformación con cualquier matriz ortogonal real y el tensor seguiría siendo la identidad ya que esta matriz me convertiría una base ortonormal en otra base ortonormal. Hagamos entonces una transformación ortogonal real  $\mathbf{F}$  que más adelante se elegirá de forma adecuada.

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}}^b \mathbf{1} \mathbf{y}^b = \tilde{\mathbf{x}}^b \tilde{\mathbf{F}} \mathbf{1} \mathbf{F} \mathbf{y}^b = \widetilde{(\mathbf{F} \mathbf{x}^b)} \mathbf{1} (\mathbf{F} \mathbf{y}^b) \\ (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}}^c \mathbf{1} \mathbf{y}^c \quad ; \quad \mathbf{x}^c \equiv \mathbf{F} \mathbf{x}^b \end{aligned}$$

rastreado las definiciones tenemos

$$\mathbf{B} \mathbf{x} = \mathbf{x}^a \quad , \quad \mathbf{D} \mathbf{B} \mathbf{x} = \mathbf{x}^b \quad \Rightarrow \quad \mathbf{N} \mathbf{x} = \mathbf{x}^b \quad ; \quad \mathbf{N} \equiv \mathbf{D} \mathbf{B} \quad (12.164)$$

$$\mathbf{F} \mathbf{D} \mathbf{B} \mathbf{x} = \mathbf{x}^c \quad \Rightarrow \quad \mathbf{R} \mathbf{x} = \mathbf{x}^c \quad ; \quad \mathbf{R} \equiv \mathbf{F} \mathbf{D} \mathbf{B} \quad (12.165)$$

el producto interno se puede escribir en los pasos sucesivos

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &= \tilde{\mathbf{x}} \mathbf{T} \mathbf{y} = \tilde{\mathbf{x}} \tilde{\mathbf{B}} \mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}} \mathbf{y} = \tilde{\mathbf{x}}^a \mathbf{C} \mathbf{y}^a = \tilde{\mathbf{x}}^a \mathbf{D} \mathbf{D}^{-1} \mathbf{C} \mathbf{D}^{-1} \mathbf{D} \mathbf{y}^a = \tilde{\mathbf{x}}^a \tilde{\mathbf{D}} (\mathbf{D}^{-1} \mathbf{C} \mathbf{D}^{-1}) \mathbf{D} \mathbf{y}^a \\ &= \tilde{\mathbf{x}}^b \mathbf{1} \mathbf{y}^b = \tilde{\mathbf{x}}^b \tilde{\mathbf{F}} \mathbf{F} \mathbf{1} \tilde{\mathbf{F}} \mathbf{y}^b = \tilde{\mathbf{x}}^c \mathbf{F} \mathbf{1} \tilde{\mathbf{F}} \mathbf{y}^c = \tilde{\mathbf{x}}^c \mathbf{1} \mathbf{y}^c \end{aligned}$$

donde hemos usado el carácter simétrico de  $\mathbf{D}$ . Nótese que dado que  $\mathbf{D}$  y  $\mathbf{C}$  son ambos diagonales, entonces conmutan entre sí de modo que  $\mathbf{D}^{-1} \mathbf{C} \mathbf{D}^{-1} = \mathbf{D}^{-2} \mathbf{C} = \mathbf{1}$ . Se puede ver entonces que el algoritmo de trivialización del tensor métrico es

$$\begin{aligned} \mathbf{T} &\rightarrow \mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}} \rightarrow \mathbf{D}^{-1} (\mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}}) \mathbf{D}^{-1} \rightarrow \mathbf{F} \left[ \mathbf{D}^{-1} (\mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}}) \mathbf{D}^{-1} \right] \tilde{\mathbf{F}} = \mathbf{F} \left[ \mathbf{D}^{-1} (\mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}}) \tilde{\mathbf{D}}^{-1} \right] \tilde{\mathbf{F}} \\ &= \mathbf{F} \left[ (\mathbf{D}^{-1} \mathbf{B}) \mathbf{T} (\tilde{\mathbf{D}}^{-1} \tilde{\mathbf{B}}) \right] \tilde{\mathbf{F}} \equiv \mathbf{F} \left[ \mathbf{K} \tilde{\mathbf{T}} \tilde{\mathbf{K}} \right] \tilde{\mathbf{F}} = (\mathbf{F} \mathbf{K}) \mathbf{T} (\tilde{\mathbf{F}} \tilde{\mathbf{K}}) \equiv \mathbf{W} \tilde{\mathbf{T}} \tilde{\mathbf{W}} \end{aligned}$$

que podemos escribir como

$$\mathbf{T} \rightarrow \mathbf{T}_a \rightarrow \mathbf{T}_b \rightarrow \mathbf{T}_c \quad ; \quad \mathbf{T}_a \equiv \mathbf{B} \tilde{\mathbf{T}} \tilde{\mathbf{B}} \quad , \quad \mathbf{T}_b \equiv \mathbf{K} \tilde{\mathbf{T}} \tilde{\mathbf{K}} \quad , \quad \mathbf{T}_c \equiv \mathbf{W} \tilde{\mathbf{T}} \tilde{\mathbf{W}} \quad (12.166)$$

$$\text{con} \quad \mathbf{K} \equiv \mathbf{D}^{-1} \mathbf{B} \quad ; \quad \mathbf{W} \equiv \mathbf{F} \mathbf{D}^{-1} \mathbf{B} \quad (12.167)$$

como esto es válido para  $\mathbf{x}$ ,  $\mathbf{y}$  arbitrarios se tiene que después de las dos primeras transformaciones donde el tensor métrico ya se ha trivializado, el producto interno se escribe

$$(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}}^T \mathbf{T} \mathbf{y} = \tilde{\mathbf{x}}^b \mathbf{1} \mathbf{y}^b = \tilde{\mathbf{x}}^T \tilde{\mathbf{N}} \left( \mathbf{K} \mathbf{T} \tilde{\mathbf{K}} \right) \mathbf{N} \mathbf{y}$$

y luego de la tercera transformación donde el tensor continúa trivializado

$$(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}}^T \mathbf{T} \mathbf{y} = \tilde{\mathbf{x}}^c \mathbf{1} \mathbf{y}^c = \tilde{\mathbf{x}}^T \tilde{\mathbf{R}} \left( \mathbf{W} \mathbf{T} \tilde{\mathbf{W}} \right) \mathbf{R} \mathbf{y}$$

$\mathbf{K}$  es entonces la primera transformación que trivializa al tensor métrico. Veamos como se comporta  $\mathbf{V}$  ante esta misma transformación

$$\mathbf{V} \rightarrow \mathbf{K} \mathbf{V} \tilde{\mathbf{K}}$$

en virtud de la simetría de  $\mathbf{V}$  y del hecho de que todas las matrices que intervienen son reales, es claro que esta nueva matriz es simétrica y real. Por tanto puede ser diagonalizada por alguna matriz ortogonal real. Esigamos entonces la matriz ortogonal real  $\mathbf{F}$  de modo que diagonalize a  $\mathbf{V}$ . Recordemos que esta transformación continúa dejando trivial al tensor métrico de modo que las dos matrices  $\mathbf{T}$  y  $\mathbf{V}$  quedan diagonalizadas una vez se ejecuta la tercera transformación

$$\begin{aligned} \mathbf{V} &\rightarrow \mathbf{F} \left( \mathbf{K} \mathbf{V} \tilde{\mathbf{K}} \right) \tilde{\mathbf{F}} = \mathbf{W} \mathbf{V} \tilde{\mathbf{W}} = \lambda \\ &\Rightarrow \tilde{\mathbf{A}} \mathbf{V} \mathbf{A} = \lambda ; \quad \tilde{\mathbf{A}} \equiv \mathbf{W} = \mathbf{F} \mathbf{D}^{-1} \mathbf{B} \end{aligned} \quad (12.168)$$

la matriz de diagonalización está entonces dada por la Ec. (12.168) que coincide con la Ec. (12.157). Nótese que la diagonalización de  $\mathbf{V}$  a través de  $\mathbf{A}$  no nos da los valores propios de la ecuación **normal** de valores propios ya que esta matriz contiene una transformación de reescalamiento. En consecuencia, la transformación de congruencia no es de similitud y no se conservan los valores propios **usuales** en el proceso. Solo en el caso en el cual  $\mathbf{T} = \mathbf{1}$ , obtendremos los vectores y valores propios usuales.

Con base en lo anterior, podemos dar una interpretación geométrica a la diagonalización simultánea de  $\mathbf{T}$  y  $\mathbf{V}$  a través de  $\mathbf{A}$ . La transformación (12.127) corresponde a una transformación lineal de un sistema de ejes que no son ortogonales ni normales para pasar a un sistema de ejes ortogonales y normales, a través de tres pasos. En el primer paso se hace una transformación ortogonal que me produce una base ortogonal pero no normal (tensor métrico diagonal), en el segundo paso se realiza un reescalamiento que me normaliza los vectores. Con estos pasos se consigue una base ortonormal lo cual equivale a trivializar al tensor métrico  $\mathbf{T}$ . En cuanto a la matriz  $\mathbf{V}$  estas dos transformaciones no la diagonalizan pero la mantienen simétrica y real. Esto implica que una transformación ortogonal adicional (tercer paso) puede diagonalizar a  $\mathbf{V}$ , y al mismo tiempo mantener inalterada la forma del tensor métrico, ya que una matriz ortogonal me lleva de una base ortonormal a otra base ortonormal. Obtenemos entonces un conjunto de ejes principales comunes para  $\mathbf{T}$  y  $\mathbf{V}$ .

Cuando se presenta degeneración el número de ecuaciones independientes entre los valores propios resulta insuficiente para determinar los cocientes entre componentes de un vector propio. Por ejemplo, si  $\lambda$  es una raíz doble, se podrán elegir de forma arbitraria dos componentes cualesquiera  $a_j$ , y las demás se determinan a través de las ecuaciones de valores propios. Recordemos además que degeneración corresponde a una simetría del sistema, cuanto más degenerados son los valores propios, más libertad existe de escoger los vectores propios, es decir aumentan las transformaciones que dejan invariante al sistema. De nuevo esto también se puede ver teniendo en cuenta que a mayor degeneración de un  $\lambda$  dado, mayor es la dimensionalidad del subespacio generado por  $\lambda$ , y por tanto mayor el número de vectores propios linealmente independientes. Esto también implica mayor libertad para escoger una base de vectores propios que genere al mencionado subespacio.

Finalmente, incluso si nos restringimos a trabajar en el espacio plano  $\mathbb{R}^n$ , podemos tener una interpretación activa de la ecuación de valores propios modificada. Por ejemplo, en el problema de los osciladores acoplados se llega a la ecuación de valores propios modificada en la cual  $\mathbf{T}$  representa la matriz de energía cinética y  $\mathbf{V}$  representa al potencial, en este problema ambas matrices son activas.

## 12.16. Tensores cartesianos

Hemos caracterizado las matrices ortogonales y la manera en que transforman a ciertos vectores en  $\mathbb{R}^3$ . Utilizando al vector posición y al vector asociado a una matriz real antisimétrica como prototipos, hemos definido los conceptos de vector polar y vector axial. En la presente sección redefiniremos estas cantidades utilizando el concepto de grupo. Definiremos además el concepto de tensor cartesiano de rango  $n$ . El uso del lenguaje de grupos nos permitirá comprender la diferencia entre los tensores cartesianos y otros tipos de tensores utilizados en Física. Es de anotar que la formulación más precisa para definir tensores requiere el desarrollo de la teoría de representaciones de grupos, lo cual está fuera de los propósitos del presente documento. Presentaremos aquí una visión simplificada que nos brinda una primera aproximación a la relación entre el concepto de tensor y el concepto de grupo. Desarrollaremos además algunas propiedades de los tensores cartesianos necesarias para nuestros desarrollos subsecuentes.

## 12.17. Los grupos $O(3)$ y $SO(3)$ y la definición de tensores cartesianos

Basados en nuestra experiencia con las transformaciones ortogonales, definiremos el conjunto de todas las matrices reales  $3 \times 3$  que cumplen las condiciones

$$\widetilde{\mathbf{A}} = \mathbf{A}^{-1} \quad ; \quad \det \mathbf{A} = \pm 1 \quad (12.169)$$

A este conjunto lo denominaremos  $O(3)$  para denotar que consta de matrices **O**rtogonales en **3** dimensiones. Veremos algunas propiedades de este conjunto. Sean  $\mathbf{A}, \mathbf{B} \in O(3)$ , la transpuesta del producto de estas matrices es

$$\widetilde{\mathbf{C}} \equiv \widetilde{\mathbf{AB}} = \widetilde{\mathbf{B}}\widetilde{\mathbf{A}} = \mathbf{B}^{-1}\mathbf{A}^{-1} = (\mathbf{AB})^{-1} = \mathbf{C}^{-1}$$

el determinante de este producto es

$$\det \mathbf{C} = \det (\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B} = \pm 1$$

de modo que el producto de dos matrices ortogonales reales es también ortogonal real. Por otro lado, es bien conocido que el producto matricial es asociativo, por lo tanto lo es en particular para las matrices ortogonales reales. Adicionalmente, es obvio que la matriz identidad es una matriz ortogonal, de modo que la identidad está contenida en el conjunto. Finalmente, dado que el determinante de estas matrices no es cero, existe la matriz inversa de cada una, y es fácil demostrar que  $\mathbf{A}^{-1}$  también es ortogonal real<sup>32</sup>

$$\widetilde{(\mathbf{A}^{-1})} = \widetilde{(\widetilde{\mathbf{A}})} = \mathbf{A} = (\mathbf{A}^{-1})^{-1} \quad ; \quad \det \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} = \pm 1$$

vale anotar que el producto de dos matrices ortogonales

$$\mathbf{C} = \mathbf{AB} \quad ; \quad C_{ij} = \sum_{k=1}^3 A_{ik} B_{kj}$$

es en general no-conmutativo. Geométricamente, esto tendrá como consecuencia el hecho de que las rotaciones en tres dimensiones no conmutan entre sí, como veremos más adelante. Lo anterior se puede sintetizar en la siguiente forma

**Theorem 17** *Sea  $O(3) \equiv \{\mathbf{A}_i\}$  el conjunto de todas las matrices reales  $3 \times 3$  que cumplen con las condiciones (12.169). Si definimos al producto matricial como ley de combinación entre matrices de  $O(3)$ , tenemos que*

1. Si  $\mathbf{A}_i \in O(3)$  y  $\mathbf{A}_j \in O(3)$  entonces  $\mathbf{A}_i \mathbf{A}_j \in O(3)$
2.  $(\mathbf{A}_i \mathbf{A}_j) \mathbf{A}_k = \mathbf{A}_i (\mathbf{A}_j \mathbf{A}_k)$  para todo  $\mathbf{A}_i, \mathbf{A}_j, \mathbf{A}_k \in O(3)$

<sup>32</sup>El inverso de una matriz real (cuando existe) también es real como se puede ver con el algoritmo de construcción de la inversa.



3. Existe  $\mathbf{1} \in O(3)$  tal que  $\mathbf{1}\mathbf{A}_i = \mathbf{A}_i\mathbf{1} = \mathbf{A}_i$  para todo  $\mathbf{A}_i \in O(3)$

4. Para cada  $\mathbf{A}_i \in O(3)$  existe  $\mathbf{A}_i^{-1} \in O(3)$  tal que  $\mathbf{A}_i\mathbf{A}_i^{-1} = \mathbf{A}_i^{-1}\mathbf{A}_i = \mathbf{1}$

Lo anterior significa que el conjunto  $O(3)$  es un **grupo** cuando usamos la multiplicación de matrices como ley de combinación, ya que los ítems arriba mencionados corresponden a los axiomas que definen a un grupo matemático. Es fácil demostrar que el subconjunto de todas las matrices ortogonales de determinante  $+1$ , también cumple con las propiedades de grupo (i.e. es un subgrupo de  $O(3)$ ) y se denomina el **grupo ortogonal especial en tres dimensiones  $SO(3)$** . En contraste, el subconjunto de las transformaciones ortogonales de determinante  $-1$  no forma un grupo. Nótese que  $SO(3)$  da cuenta de todas las transformaciones ortogonales continuas (rotaciones), en tanto que  $O(3)$  da cuenta además de las transformaciones de inversión y de la combinación entre inversión y rotación.

## 12.18. Tensores de $SO(3)$

Hemos visto que el vector posición  $\mathbf{r}$  en el espacio cartesiano  $\mathbb{R}^3$  puede asociarse a una tripla  $(x_1, x_2, x_3)$  que bajo una transformación ortogonal especial pasiva  $\mathbf{A}$  (rotación de ejes), cambia sus componentes en la forma

$$x'_i = a_{ij}x_j \quad , \quad \mathbf{x}' = \mathbf{A}\mathbf{x} \quad ; \quad \mathbf{A} \in SO(3)$$

aunque todas las transformaciones que consideraremos en esta sección son pasivas, omitiremos la notación  $(\mathbf{r})'$  y la escribiremos simplemente como  $\mathbf{r}'$ . Vamos a tomar al vector posición como un prototipo para la definición de vectores cartesianos

**Definition 18** Una cantidad física  $\mathbf{v}$  se denomina un vector cartesiano (o euclidiano) si se puede representar por una tripla  $(v_1, v_2, v_3) \in \mathbb{R}^3$  de números reales, que bajo una transformación pasiva  $\mathbf{A}$  del grupo  $SO(3)$ , cambia sus componentes en la forma

$$v'_i = a_{ij}v_j \quad , \quad \mathbf{v}' = \mathbf{A}\mathbf{v} \quad ; \quad \forall \mathbf{A} \in SO(3) \quad (12.170)$$

Nótese que la definición de vector cartesiano está referida a un grupo de simetría del espacio tiempo. Si adjuntamos dos vectores  $\mathbf{v}$ ,  $\mathbf{w}$  la transformación de dicha yuxtaposición se escribe de la forma

$$v'_i w'_j = (a_{ik}v_k)(a_{jm}w_m) = a_{ik}a_{jm}v_k w_m$$

similarmente podemos yuxtaponer un conjunto de  $N$  vectores, los cuales transformarían de la forma

$$v_{i_1}^{(1)'} v_{i_2}^{(2)'} \dots v_{i_N}^{(N)'} = a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_N j_N} v_{j_1}^{(1)} v_{j_2}^{(2)} \dots v_{j_N}^{(N)} \quad (12.171)$$

donde  $\mathbf{v}^{(i)}$  denota al  $i$ -ésimo vector. Esta yuxtaposición tendría  $3^N$  componentes (en 3 dimensiones).

**Definition 19** Por extensión, diremos que un tensor cartesiano de  $SO(3)$  (o euclidiano)  $\mathbf{T}$  de rango  $N$  en el espacio cartesiano tridimensional, es una cantidad constituida por  $3^N$  componentes reales que denotamos por  $T_{i_1 i_2 \dots i_N}$  (i.e. con  $N$  índices), donde  $i_k = 1, 2, 3$ ; y que bajo una transformación ortogonal  $\mathbf{A} \in SO(3)$ , transforma bajo el esquema

$$T'_{i_1 i_2 \dots i_N} = a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_N j_N} T_{j_1 j_2 \dots j_N} \quad (12.172)$$

lo cual es una emulación directa de la propiedad de transformación (12.171) que posee una yuxtaposición de  $N$  vectores cartesianos. Sin embargo, el tensor  $T_{i_1 i_2 \dots i_N}$  es una entidad independiente definida por la propiedad de transformación (12.172) y no necesariamente proviene de una yuxtaposición de vectores.

Por la forma en que ha sido construída esta definición, es inmediato que un vector cartesiano o euclidiano es un tensor de primer rango.

Es bien sabido que el producto punto entre dos vectores de  $\mathbb{R}^3$  es un número real. Nos preguntamos ahora por sus propiedades de transformación bajo  $SO(3)$ . Para verlas, usamos las propiedades de transformación de los vectores cartesianos bajo dicho grupo

$$\begin{aligned} c &\equiv \mathbf{v} \cdot \mathbf{w} \equiv v_i w_i \quad ; \quad c' = \mathbf{v}' \cdot \mathbf{w}' = v'_i w'_i = a_{ik} v_k a_{im} w_m = (a_{ik} a_{im}) v_k w_m = \delta_{km} v_k w_m \\ c' &= v_k w_k = c \end{aligned}$$

el producto punto es invariante ante una transformación de  $SO(3)$ . Esto nos induce a construir la siguiente definición

**Definition 20** *Denominaremos tensor cartesiano de rango cero o escalar, a una cantidad física real que es invariante bajo transformaciones del grupo  $SO(3)$  (rotaciones). El ejemplo más claro es el producto punto entre dos vectores cartesianos y en particular la norma de un vector.*

Las nueve componentes de un tensor de segundo rango se transforman como

$$T'_{ij} = a_{ik} a_{jl} T_{kl}$$

en rigor, es necesario distinguir entre un tensor de segundo rango  $\mathbf{T}$  y la matriz cuadrada formada por sus componentes. Un tensor está definido solo por sus propiedades de transformación ante una transformación ortogonal de coordenadas. En contraste, una matriz no está restringida de ninguna forma por los tipos de transformación que pueda experimentar y se puede considerar totalmente independiente de sus propiedades bajo un tipo particular de transformación. Una matriz es simplemente un arreglo de elementos con un algoritmo de suma y multiplicación. Tales arreglos junto con sus algoritmos sirven para expresar de manera sintética cualquier transformación lineal aplicada sobre cierto espacio (no necesariamente vectorial). De la misma forma, un arreglo en columna no es necesariamente un vector Euclidiano, pero está bien definido el producto del arreglo matricial con el arreglo vectorial. Incluso las matrices se pueden definir por sí solas con sus propiedades de suma, producto, inversa etc. formando estructuras matemáticas tales como grupos, anillos etc. En contraste, los tensores siempre están ligados a un espacio vectorial o a una variedad.

Vale decir sin embargo, que si una matriz representa una transformación activa sobre un espacio vectorial, los componentes tensoriales y matriciales se manipulan en la misma forma al menos dentro del dominio restringido a las transformaciones ortogonales. Para toda ecuación tensorial habrá una ecuación matricial correspondiente y viceversa. La Ec. (12.34) nos indica la manera en que transforma una matriz que representa a un operador lineal en un espacio vectorial, cuando hacemos un cambio de base (cambio de sistema coordenado)

$$\mathbf{T}' = \mathbf{A} \mathbf{T} \mathbf{A}^{-1}$$

siendo  $\mathbf{A}$  la matriz que genera el cambio de coordenadas (transformación pasiva). La nueva matriz  $\mathbf{T}'$  representa al mismo operador en una base diferente. Si la transformación es ortogonal se tiene

$$\mathbf{T}' = \mathbf{A} \mathbf{T} \tilde{\mathbf{A}}$$

que en componentes se escribe

$$T'_{ij} = a_{ik} T_{kl} \tilde{a}_{lj} = a_{ik} T_{kl} a_{jl} = a_{ik} a_{jl} T_{kl}$$

que resulta idéntica a la transformación de un tensor de segundo rango al menos bajo las transformaciones ortogonales específicamente.

Nótese sin embargo que si la matriz  $\mathbf{T}$  tiene el rol pasivo de representar un cambio de coordenadas de un sistema específico a otro también específico, entonces no tiene sentido preguntarse por sus propiedades de transformación. Algorítmicamente hablando, la operación  $\mathbf{A} \mathbf{T} \mathbf{A}^{-1}$  es posible y sigue siendo una transformación de similitud pero ya no tiene sentido geométrico. En el caso de matrices activas que representan un operador lineal en un espacio vectorial, la transformación  $\mathbf{A} \mathbf{T} \mathbf{A}^{-1}$  es automáticamente inducida por el cambio de base,

en cambio no hay una manera natural geométrica de inducir la misma transformación en una matriz que juega el rol de cambio de base. En este punto queda quizás mas claro el hecho de que las matrices y sus algoritmos se pueden trabajar sin saber qué están representando, pero para conocer sus eventuales propiedades de transformación es necesario saber en qué espacio actúan y sobre qué objetos (e.g. los vectores o el sistema coordinado).

En virtud de la analogía entre las propiedades de los tensores de segundo rango y las matrices activas, podemos deducir que la misma terminología derivada del álgebra matricial (traspuesta, antisimétrica etc.) será de gran utilidad también en el álgebra de tensores de segundo rango. De igual modo, en el álgebra de tensores de primer rango, podemos usar formas de vector fila y columna para las manipulaciones algebraicas.

### 12.19. Tensores de $O(3)$ y el concepto de pseudotensor

El grupo  $O(3)$  incluye además de las transformaciones ortogonales continuas o propias (de  $SO(3)$ ) las transformaciones discretas de paridad y el producto de operaciones de paridad con transformaciones propias.

Utilizaremos de nuevo al vector posición  $\mathbf{r}$  como prototipo. Ya conocemos sus propiedades de transformación bajo operadores de  $SO(3)$  de modo que nos centraremos en sus propiedades de transformación bajo paridad. En la sección 12.6 vimos que ante una inversión  $\mathbf{S}$  de los ejes coordinados (pasiva) se invierten las componentes del vector posición, es decir

$$\mathbf{r}' = \mathbf{S}\mathbf{r} = -\mathbf{1}\mathbf{r} = -\mathbf{r}$$

sea  $\mathbf{B}_p$  una matriz ortogonal propia. Toda matriz ortogonal impropia se puede escribir como  $-\mathbf{B}_p$  de manera que bajo una combinación de rotación y paridad (transformación impropia  $\mathbf{B}_i$ ) el vector  $\mathbf{r}$  transforma como

$$\mathbf{r}' = -\mathbf{B}_p\mathbf{r} = \mathbf{B}_i\mathbf{r}$$

y como la transformación de  $\mathbf{r}$  bajo una matriz ortogonal propia es simplemente  $\mathbf{r}' = \mathbf{B}_p\mathbf{r}$ , podemos escribir genéricamente

$$\mathbf{r}' = \mathbf{B}\mathbf{r} \quad ; \quad \forall \mathbf{B} \in \mathbb{O}(3) \tag{12.173}$$

a toda tripla de  $\mathbb{R}^3$  cuyas propiedades de transformación bajo el grupo  $O(3)$  sean las dadas por la Ec. (12.173), la denominaremos **vector cartesiano polar** o simplemente **vector polar**.

Ahora si tomamos como prototipo al vector asociado a una matriz real antisimétrica  $3 \times 3$ , vemos que sus propiedades de transformación bajo un elemento del grupo  $O(3)$  vienen dadas por la Ec. (12.70)

$$\mathbf{v}'_A = (\det \mathbf{B}) \mathbf{B}\mathbf{v}_A \quad ; \quad \forall \mathbf{B} \in \mathbb{O}(3) \tag{12.174}$$

a los vectores que transforman como  $\mathbf{v}_A$  bajo  $O(3)$ , los denominamos **vectores cartesianos axiales**. Bajo  $SO(3)$ , los vectores polares y axiales transforman de manera idéntica, y su diferencia está solo en el comportamiento bajo paridad de modo que solo se manifiesta bajo  $O(3)$ . Los vectores polares cambian su signo con la transformación de paridad en tanto que los vectores axiales son invariantes ante una transformación de paridad.

Por otro lado, inspirados en la definición de tensores de  $SO(3)$  de rango  $N$ , podemos definir tensores de  $O(3)$  como una yuxtaposición de vectores de  $O(3)$ . En este caso es sin embargo necesario estudiar el comportamiento de estos tensores bajo paridad. Comenzaremos estudiando los tensores de rango 2. Si yuxtaponemos dos vectores polares es claro que el tensor resultante es invariante bajo paridad, ya que cada vector de la yuxtaposición invierte el signo bajo paridad. Similarmente, la yuxtaposición de dos vectores axiales es invariante bajo paridad. Si en cambio yuxtaponemos un vector polar con uno axial, el tensor resultante invierte su signo bajo paridad.

Definiremos un **tensor cartesiano polar de rango  $N$** , como aquel que bajo  $O(3)$  se comporta como la yuxtaposición de vectores polares. Por tanto un tensor polar de rango  $N$  se comporta bajo  $O(3)$  en la forma

$$T_{i_1 i_2 \dots i_N}^{P1} = a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_N j_N} T_{j_1 j_2 \dots j_N}^P \quad ; \quad \forall \mathbf{A} \in \mathbb{O}(3)$$

Un **tensor cartesiano axial o pseudotensor** será aquel que transforme bajo paridad con el signo opuesto con respecto a un tensor polar. Esto equivale a decir que bajo una transformación impropia el signo debe

cambiar con respecto al de un tensor polar, pero bajo una transformación propia debe transformar de manera idéntica a un tensor polar. Por tanto, la regla de transformación de un pseudotensor es

$$T'_{i_1 i_2 \dots i_N} = (\det \mathbf{A}) a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_N j_N} T_{j_1 j_2 \dots j_N}^S \quad ; \quad \forall \mathbf{A} \in \mathbb{O}(3)$$

podemos escribir las propiedades de transformación de un tensor de  $O(3)$  de rango  $N$  en la siguiente forma sintética

**Definición 21** *Un tensor cartesiano de  $O(3)$  (o euclidiano) y de rango  $N$  en el espacio cartesiano tridimensional, es una cantidad constituida por  $3^N$  componentes reales que denotamos por  $T_{i_1 i_2 \dots i_N}$  (i.e. con  $N$  índices), donde  $i_k = 1, 2, 3$ ; y que bajo una transformación ortogonal  $\mathbf{A} \in O(3)$ , transforma bajo el esquema*

$$T'_{i_1 i_2 \dots i_N} = (\det \mathbf{A})^p a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_N j_N} T_{j_1 j_2 \dots j_N} \quad , \quad p \equiv \begin{cases} 0 & \text{si } T \text{ es polar} \\ 1 & \text{si } T \text{ es pseudotensor} \end{cases} \quad (12.175)$$

cuando  $p = 0$  decimos que el tensor es polar y cuando  $p = 1$  decimos que es un tensor axial o pseudotensor.

Sea un tensor cartesiano formado con la yuxtaposición de  $N$  vectores, donde  $k$  vectores de la yuxtaposición son axiales ( $k \leq N$ ). Es fácil probar que el tensor será polar si  $k = 2n$ , y será axial si  $k = 2n + 1$ , siendo  $n$  un entero no-negativo. Debemos recordar sin embargo que un tensor no necesariamente proviene de una yuxtaposición de vectores.

Veamos ahora los escalares. Tomaremos como prototipo de escalar al producto punto entre dos vectores polares. Bajo paridad dicho producto punto es invariante. Es natural entonces definir un pseudoescalar como un número que bajo paridad invierta su signo. El prototipo inmediato de pseudoescalar es entonces el producto punto entre un vector polar y un vector axial.

De acuerdo con la definición 21, los vectores axiales son pseudotensores de primer rango y los pseudoescalares son pseudotensores de rango cero.

Vemos entonces que tensores cartesianos se refiere a tensores asociados al grupo  $SO(3)$  o al grupo  $O(3)$  si queremos distinguir quiralidad. Existen tensores definidos para otros grupos de simetría del espacio-tiempo, tales como los tensores de Minkowski, que se relacionan con el grupo de Lorentz. En el espacio  $\mathbb{R}^n$  con coordenadas cartesianas, no existe diferenciación entre los índices covariantes y contravariantes, de modo que no se emplea esta terminología propia de los tensores de Minkowski. Los tensores desarrollados en esta sección están contruidos en una base cartesiana de  $\mathbb{R}^n$  y deben llamarse tensores Cartesianos. Sin embargo, emplearemos la simple denominación de tensor mientras no se presente ambigüedad.

### 12.19.1. Los tensores de Kronecker $\delta_{ij}$ y de Levi-Civita $\varepsilon_{ijk}$

Un ejemplo de pseudotensor muy importante en Física es el tensor de Levi Civita. A priori, esto parece sorprendente ya que el tensor de Levi-Civita se ha definido con valores numéricos muy específicos que no dependen de la base usada, en tanto que los tensores transforman de manera no trivial ante un cambio de base. Supongamos que tenemos un pseudotensor de tercer rango  $T_{ijk}$  que en una base particular coincide numéricamente con el tensor de Levi Civita, cuando cambiamos a otro sistema coordenado ortogonal (comenzando con el sistema en donde  $T_{ijk} = \varepsilon_{ijk}$ ) se tiene que

$$T'_{ijk} = (\det \mathbf{A}) a_{ip} a_{jq} a_{kr} T_{pqr} = (\det \mathbf{A}) a_{ip} a_{jq} a_{kr} \varepsilon_{pqr} \quad (12.176)$$

es fácil verificar explícitamente que

$$a_{ip} a_{jq} a_{kr} \varepsilon_{pqr} = \begin{cases} 0 & \text{si al menos dos índices } i, j, k \text{ son iguales} \\ \det \mathbf{A} & \text{si } i, j, k \text{ están en orden cíclico} \\ -\det \mathbf{A} & \text{si } i, j, k \text{ están en orden anticíclico} \end{cases}$$

por tanto si  $i, j, k$  están en orden cíclico tenemos

$$T'_{ijk} = (\det \mathbf{A}) a_{ip} a_{jq} a_{kr} \varepsilon_{pqr} = (\det \mathbf{A})^2 = 1 = \varepsilon_{ijk} \quad ; \quad (i, j, k) \rightarrow (1, 2, 3), (2, 3, 1), (3, 1, 2)$$

en orden anticíclico

$$T'_{ijk} = (\det \mathbf{A}) a_{ip} a_{jq} a_{kr} \varepsilon_{pqr} = -(\det \mathbf{A})^2 = -1 = \varepsilon_{ijk} \quad ; \quad (i, j, k) \rightarrow (1, 3, 2), (2, 1, 3), (3, 2, 1)$$

además  $T'_{ijk} = \varepsilon_{ijk} = 0$ , cuando hay índices repetidos tales como  $(1, 1, 2)$ ,  $(2, 2, 2)$ ,  $(3, 1, 3)$  etc. Concluimos que

$$T'_{ijk} = T_{ijk} = \varepsilon_{ijk}$$

lo cual demuestra que si un pseudotensor cartesiano de tercer rango  $T_{ijk}$  coincide numéricamente con  $\varepsilon_{ijk}$  en una base dada, coincidirá con él en cualquier otra base<sup>33</sup>. Podemos ver que el factor  $\det \mathbf{A}$  en (12.176) es crucial para obtener la misma forma numérica del Levi Civita cuando se cambia de base.

Es mucho más fácil demostrar que  $\delta_{ij}$  es un tensor polar de segundo rango. El pseudotensor de Levi-Civita y el tensor polar de Kronecker son ejemplos un tanto paradójicos de tensores, ya que sus propiedades de transformación (no triviales) ante  $O(3)$ , garantizan su invarianza ante una transformación de dicho grupo.

## 12.20. Diadas y afinores

En virtud de la forma en que se construyó el concepto de tensor podemos definir estructuras algebraicas que permiten realizar fácilmente las operaciones. Definiremos una diada, como la yuxtaposición de dos vectores en la forma  $\mathbf{AB}$ ,  $\mathbf{CD}$  etc. Podemos definir por ejemplo el producto punto entre una diada y un vector en la forma

$$\mathbf{A} \cdot (\mathbf{BC}) \equiv (\mathbf{A} \cdot \mathbf{B}) \mathbf{C} \quad ; \quad (\mathbf{BC}) \cdot \mathbf{A} \equiv \mathbf{B} (\mathbf{C} \cdot \mathbf{A}) \tag{12.177}$$

nótese que el orden es importante en este producto punto. No es igual operar por la izquierda que por la derecha. Además el producto punto entre una diada y un vector es otro vector (ya no podemos llamarlo producto escalar). Cada vector se puede escribir en una base, de modo que una diada se escribe como

$$\mathbf{AB} = (a_i \mathbf{u}_i) (b_j \mathbf{u}_j) = a_i b_j \mathbf{u}_i \mathbf{u}_j \tag{12.178}$$

esto nos lleva a construir el concepto de afinor, como una combinación lineal de diadas

$$a\mathbf{AB} + b\mathbf{CD} + \dots \tag{12.179}$$

y la diada se puede ver como un afinor construido con la yuxtaposición de los vectores base. Hemos supuesto implícitamente propiedades de distributividad y de producto por escalar (linealidad). Se puede definir el producto punto doble de la siguiente forma

$$(\mathbf{AB}) : (\mathbf{CD}) \equiv (\mathbf{A} \cdot \mathbf{C}) (\mathbf{B} \cdot \mathbf{D}) \tag{12.180}$$

quizás una notación mas apropiada es

$$\mathbf{A} \cdot (\mathbf{CD}) \cdot \mathbf{B} \equiv (\mathbf{A} \cdot \mathbf{C}) (\mathbf{B} \cdot \mathbf{D}) \tag{12.181}$$

este doble producto punto sí es un producto escalar. La diada identidad se construye en la forma

$$\mathbf{1} \equiv \mathbf{u}_i \mathbf{u}_i \tag{12.182}$$

donde  $\mathbf{u}_i$  son los vectores unitarios de una base ortonormal. Es fácil ver su carácter de identidad con los algoritmos que acabamos de establecer

$$\mathbf{A} \cdot \mathbf{1} = (a_k \mathbf{u}_k) \cdot (\mathbf{u}_i \mathbf{u}_i) = a_k (\mathbf{u}_k \cdot \mathbf{u}_i) \mathbf{u}_i = a_k \delta_{ki} \mathbf{u}_i = a_k \mathbf{u}_k = \mathbf{A}$$

similarmente se puede demostrar que

$$\mathbf{A} \cdot \mathbf{1} = \mathbf{1} \cdot \mathbf{A} = \mathbf{A} \quad ; \quad \mathbf{A} \cdot \mathbf{1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}$$

---

<sup>33</sup>Bajo  $O(3)$  también se incluyen cambios de base en los cuales cambia la quiralidad de los ejes.

al igual que las matrices, las diadas y afinores son algoritmos de cálculo que permiten clarificar las operaciones con tensores de segundo rango. Sin embargo, es el tensor el que tiene significado geométrico y el algoritmo puede cambiar si usamos por ejemplo bases diferentes a las cartesianas (en donde los conceptos de covarianza y contravarianza adquieren sentido) o definimos tensores en espacios más generales. Por el momento, notemos que si definimos la diada a través de dos índices (cada uno correspondiente a un vector) i.e.

$$\mathbf{AB} \rightarrow A_i B_j \equiv T_{ij}$$

podemos escribir el álgebra de diadas y afinores en términos de índices

$$\begin{aligned} \mathbf{K} \cdot (\mathbf{AB}) &= (\mathbf{K} \cdot \mathbf{A}) \mathbf{B} = K_i A_i B_j \Rightarrow \\ \mathbf{K} \cdot \mathbf{T} &= K_i T_{ij} \end{aligned}$$

similarmemente

$$\mathbf{T} \cdot \mathbf{K} = T_{ij} K_j \quad ; \quad \mathbf{D} \cdot \mathbf{T} \cdot \mathbf{K} = D_i T_{ij} D_j$$

en todos los algoritmos hay suma sobre índices repetidos.

## 12.21. Contracción de tensores, reglas del cociente

Se puede observar que las operaciones con diadas son totalmente análogas a los algoritmos contruidos para matrices y vectores fila y columna reales. Así mismo, son los algoritmos para trabajar tensores cartesianos. En el álgebra de tensores recurriremos con frecuencia a sumas sobre índices repetidos tales como

$$R_j \equiv K_i T_{ij} \quad , \quad R_{km} \equiv K_i T_{ikm} \quad , \quad P_k \equiv T_{ijk} M_{ij}$$

la suma sobre índices repetidos entre tensores se conoce como **contracción**. Los índices que se suman se “eliminan”, quedando algunos índices libres que nos dirán el rango del tensor que resulta después de la contracción. Nótese que cuando un número  $n$  de índices de un tensor se contrae con índices de otro, el rango del tensor está dado por  $N - 2n$ , siendo  $N$  el número total de índices y  $n$  el número total de contracciones. Por ejemplo  $T_{ij} K_j$  tiene en total tres índices y una contracción, su resultado es un tensor de primer rango ( $3 - 2 \cdot 1$ ). El producto cruz entre dos vectores  $\varepsilon_{ijk} A_j B_k$  tiene  $N = 5$  y  $n = 2$  de modo que tiene rango 1. Nótese que  $\varepsilon_{ijk}$  es un pseudotensor y si  $\mathbf{A}$  y  $\mathbf{B}$  son vectores polares, el producto será un pseudovector como ya vimos que ocurre con el producto cruz entre dos vectores polares. El producto  $\mathbf{D} \cdot \mathbf{T} \cdot \mathbf{K}$  donde  $\mathbf{T}$  es tensor de segundo rango y  $\mathbf{D}$ ,  $\mathbf{K}$  son de primer rango, da como resultado un escalar (forma bilineal). Si uno (o tres) de ellos es un pseudotensor obtendremos un pseudoescalar, si dos de ellos son pseudotensores (o si todos son tensores polares) obtenemos un escalar.

Estas reglas de contracción nos dictaminan el número de índices libres. Sin embargo, el tener un número  $k$  de índices libres no nos garantiza que la cantidad que se obtiene después de la contracción también sea un tensor. Debemos probar que la cantidad resultante tenga propiedades de transformación de tensor. Esto se puede demostrar apelando directamente a las propiedades de transformación de los tensores originales. Por ejemplo, si  $T_{ij}$  es un tensor cartesiano de segundo rango y  $K_i$  un tensor cartesiano de primer rango (vector euclidiano) y hacemos la contracción

$$R_j \equiv K_i T_{ij} \tag{12.183}$$

demostraremos que la cantidad  $R_j$  es efectivamente un tensor cartesiano de primer rango, para ello **NO** basta con saber que es una cantidad que depende de un solo índice, es necesario que se examinen sus propiedades de transformación ante una transformación ortogonal  $\mathbf{A}$

$$R'_j = K'_i T'_{ij} = (a_{ik} K_k) (a_{im} a_{jn} T_{mn}) = (a_{ik} a_{im} a_{jn}) (K_k T_{mn})$$

donde hemos usado el hecho de que  $K_i$  y  $T_{ij}$  son tensores cartesianos. Usando las condiciones de ortogonalidad (12.51) resulta

$$\begin{aligned} R'_j &= (\delta_{km} a_{jn}) K_k T_{mn} = a_{jn} K_m T_{mn} = a_{jn} R_n \\ \Rightarrow R'_j &= a_{jn} R_n \end{aligned} \tag{12.184}$$

donde hemos usado la definición de  $R_j$  Ec. (12.183). La Ec. (12.184) nos muestra que efectivamente, la contracción definida en (12.183) corresponde a un tensor de primer rango. En todo lo anterior hemos asumido implícitamente tensores de  $SO(3)$ . Si por ejemplo  $T_{ij}$  es un pseudotensor de  $O(3)$  y  $K_i$  es un vector polar de  $O(3)$ , se puede demostrar con una transformación ortogonal impropia, que el resultado de la contracción es un pseudovector o vector axial. Se deja al lector la demostración general para la contracción de tensores de rango arbitrario, y la demostración de que la contracción es un tensor polar (axial) si el número de pseudotensores en la contracción es par (impar).

Por otro lado, consideremos el conjunto de ecuaciones

$$K_i C_i = B \quad (12.185)$$

$$K_{ij} C_j = B_i \quad (12.186)$$

$$K_{ij} C_{jk} = B_{ik} \quad (12.187)$$

$$K_{ijkl} C_{ij} = B_{kl} \quad (12.188)$$

$$K_{ij} C_k = B_{ijk} \quad (12.189)$$

donde en cada expresión,  $\mathbf{C}$  y  $\mathbf{B}$  son tensores cartesianos de rangos indicados por los índices, y  $\mathbf{C}$  es arbitrario. En cada caso  $\mathbf{K}$  es una cantidad desconocida. La llamada **regla del cociente** establece que si la ecuación en cuestión es válida en todo sistema (rotado) cartesiano de coordenadas,  $\mathbf{K}$  es un tensor cartesiano del rango indicado por sus índices.

Probaremos la regla del cociente (12.186) a manera de ilustración. Sea  $\mathbf{A}$  la matriz ortogonal que nos lleva de un sistema “no primado” de coordenadas a un sistema “primado” de coordenadas. En el sistema cartesiano primado tenemos

$$K'_{ij} C'_j = B'_i = a_{ik} B_k$$

donde hemos usado las propiedades tensoriales de  $\mathbf{B}$ . Puesto que la ecuación es válida en todo sistema cartesiano rotado, podemos asumir su validez en el sistema coordenado “no primado”

$$a_{ik} B_k = a_{ik} (K_{kl} C_l)$$

y transformando  $\mathbf{C}$  de nuevo al sistema cartesiano primado (por medio de  $\mathbf{A}^{-1} = \tilde{\mathbf{A}}$ ) tenemos

$$\begin{aligned} K'_{ij} C'_j &= a_{ik} K_{kl} C_l = a_{ik} K_{kl} (a_{jl} C'_j) \\ \Rightarrow (K'_{ij} - a_{ik} a_{jl} K_{kl}) C'_j &= 0 \end{aligned}$$

esto debe ser válido para cada índice  $i$  y para todo sistema coordenado primado, y dado que  $C'_j$  es arbitrario<sup>34</sup> resulta

$$K'_{ij} = a_{ik} a_{jl} K_{kl} \quad (12.190)$$

que es la definición de un tensor de segundo rango. Vale mencionar sin embargo que es importante que  $\mathbf{B}$  sea diferente de cero. Las otras reglas de cociente se demuestran en forma similar.

## 12.22. Ejercicios

1. Para las matrices dadas por

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 1 \\ 0 & 2 & 7 \\ -1 & 6 & 4 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} 1 & 4 & -2 \\ 3 & 0 & 9 \\ -1 & 1 & -2 \end{pmatrix}; \quad \mathbf{C} = \begin{pmatrix} 6 & 2 & -1 \\ 9 & -5 & 1 \\ 3 & 2 & -4 \end{pmatrix}$$

<sup>34</sup>Por ejemplo si tomamos  $C'_j = \delta_{1j}$  se sigue de inmediato que  $K'_{i1} = a_{ik} a_{1l} K_{kl}$ . Otras escogencias adecuadas nos dan el resto de Ecs. (12.190).

encuentre los productos  $\mathbf{AB}$  y  $\mathbf{BA}$  y muestre que  $\mathbf{AB} \neq \mathbf{BA}$ . Calcule  $Tr(\mathbf{AB})$  y  $Tr(\mathbf{BA})$ . Demuestre explícitamente que

$$\begin{aligned} Tr(\mathbf{ABC}) &= Tr(\mathbf{CAB}) = Tr(\mathbf{BCA}) \\ Tr(\mathbf{BAC}) &= Tr(\mathbf{CBA}) = Tr(\mathbf{ACB}) \end{aligned}$$

pero

$$Tr(\mathbf{ABC}) \neq Tr(\mathbf{BAC})$$

explique este hecho.

- Demuestre que la multiplicación matricial es una operación asociativa.
- Demuestre que el producto de dos matrices diagonales es conmutativo.
- Demuestre que la delta de Kronecker es la única representación posible de la matriz identidad en cualquier base.
- Si un operador activo se representa por la matriz  $\mathbf{A}$  en un sistema coordenado  $XYZ$ , entonces en otro sistema coordenado  $X'Y'Z'$  el mismo operador activo está representado por la matriz  $\mathbf{A}'$  dada por

$$\mathbf{A}' = \mathbf{BAB}^{-1}$$

donde  $\mathbf{B}$  es la matriz que representa al operador pasivo que nos lleva de  $XYZ$  hacia  $X'Y'Z'$ . (a) Demuestre que la relación entre  $\mathbf{A}$  y  $\mathbf{A}'$ , es una relación de equivalencia. (b) Discuta la necesidad de que esta relación deba ser de equivalencia.

- Demuestre que el conmutador de dos matrices definido por

$$[\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA}$$

cumple las propiedades dadas por las Ecs. (8.9-8.12), en donde los paréntesis cuadrados significan conmutador en lugar de corchete de Poisson. Este hecho nos muestra la similaridad entre el álgebra de los conmutadores y la de los corchetes de Poisson.

- Para la matriz compleja  $2 \times 2$

$$\mathbf{A} = \begin{pmatrix} 2+i & 3-2i \\ 1 & 3 \end{pmatrix}$$

demuestre que los valores propios  $\lambda_1, \lambda_2$  y sus vectores propios asociados  $\mathbf{x}_1, \mathbf{x}_2$  vienen dados por

$$\begin{aligned} \mathbf{x}_1 &= \begin{pmatrix} \frac{1}{2}i - \frac{1}{2} - \frac{1}{2}\sqrt{(12-10i)} \\ 1 \end{pmatrix} ; & \lambda_1 &= \frac{5}{2} - \frac{1}{2}\sqrt{(12-10i)} + \frac{1}{2}i \\ \mathbf{x}_2 &= \begin{pmatrix} \frac{1}{2}\sqrt{(12-10i)} - \frac{1}{2} + \frac{1}{2}i \\ 1 \end{pmatrix} ; & \lambda_2 &= \frac{5}{2} + \frac{1}{2}\sqrt{(12-10i)} + \frac{1}{2}i \end{aligned}$$

encuentre una matriz que diagonaliza a  $\mathbf{A}$ , y demuestre que aunque la matriz es diagonalizable, no es posible diagonalizarla con una matriz unitaria.

- Sea  $V \subseteq \mathcal{E}^n$  un subespacio vectorial de  $\mathcal{E}^n$  de dimensión  $k$ . Sea  $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$  una base del subespacio  $V$ , (no necesariamente ortonormal). En el proceso de ortogonalización de Gram-Schmidt construimos un conjunto ortonormal  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  que genera al mismo espacio  $V$  con el siguiente algoritmo

- Construimos una base ortogonal (no necesariamente normalizada)  $\{\mathbf{d}_1, \dots, \mathbf{d}_k\}$ , para lo cual elegimos  $\mathbf{d}_1 = \mathbf{e}_1$  y  $\mathbf{d}_2$  se construye en la forma

$$\mathbf{d}_2 = \mathbf{e}_2 - \frac{(\mathbf{e}_2, \mathbf{d}_1)}{(\mathbf{d}_1, \mathbf{d}_1)} \mathbf{d}_1$$



b) El vector  $\mathbf{d}_3$  se construye como

$$\mathbf{d}_3 = \mathbf{e}_3 - \frac{(\mathbf{e}_3, \mathbf{d}_2)}{(\mathbf{d}_2, \mathbf{d}_2)} \mathbf{d}_2 - \frac{(\mathbf{e}_3, \mathbf{d}_1)}{(\mathbf{d}_1, \mathbf{d}_1)} \mathbf{d}_1$$

y el vector  $m$ -ésimo se construye como

$$\mathbf{d}_m = \mathbf{e}_m - \sum_{k=1}^{m-1} \frac{(\mathbf{e}_m, \mathbf{d}_k)}{(\mathbf{d}_k, \mathbf{d}_k)} \mathbf{d}_k \quad (12.191)$$

c) Estos vectores se normalizan

$$\mathbf{u}_m \equiv \frac{\mathbf{d}_m}{\sqrt{(\mathbf{d}_m, \mathbf{d}_m)}} = \frac{\mathbf{d}_m}{\|\mathbf{d}_m\|}$$

(a) Demuestre que todos los  $k$  vectores del conjunto  $\{\mathbf{d}_i\}$  son no nulos, y que el conjunto es efectivamente ortogonal. (b) Demuestre que un vector dado  $\mathbf{d}_m$  ( $m > 1$ ) es ortogonal al conjunto  $\mathbf{e}_1, \dots, \mathbf{e}_{m-1}$  y que los tres conjuntos  $\{\mathbf{e}_i\}$ ,  $\{\mathbf{d}_i\}$  y  $\{\mathbf{u}_i\}$  generan el mismo espacio vectorial  $V$ . (c) Realice una interpretación geométrica de la sustracción descrita en la Ec. (12.191). (d) ¿Se modifica este algoritmo si el producto interno está mediado por un tensor métrico no trivial? (e.g. Ec. 12.160 Pág. 324). Nota: La ventaja de utilizar notación de producto interno en lugar de producto punto, consiste en que el algoritmo vale para espacios vectoriales reales o complejos.

9. Demuestre que utilizando la multiplicación matricial como ley de combinación, el subconjunto de  $O(3)$  de matrices con determinante  $+1$  forma un subgrupo de  $O(3)$ , pero el subconjunto de matrices de  $O(3)$  con determinante  $-1$ , no forma un subgrupo.
10. Demuestre que el delta de Kronecker  $\delta_{ij}$  es un tensor polar de segundo rango.
11. Demuestre que la contracción de tensores cartesianos arbitrarios, da como resultado un tensor de rango igual al número de índices libres que quedan después de la contracción. Demuestre que el tensor resultante es polar (axial) si el número de pseudotensores en la contracción es par (impar).
12. Demuestre las reglas del cociente Ecs. (12.185-12.189).
13. Un vector de  $O(3)$  se comporta bajo paridad en la forma

$$\widehat{\mathbf{P}}\mathbf{V} = \mp \mathbf{V}$$

según si el vector es polar o axial. Puede decirse que un vector de  $O(3)$  es un vector de  $SO(3)$  que además es autovector del operador paridad, con valor propio  $-1$  si el vector es polar, y  $+1$  si el vector es axial. (a) Demuestre que el operador paridad es un operador lineal hermítico y autoinverso ya que cumple la condición  $\widehat{\mathbf{P}}^2 = \mathbf{1}$  o equivalentemente  $\widehat{\mathbf{P}} = \widehat{\mathbf{P}}^{-1}$ . (b) Demuestre que los valores propios permitidos para un operador lineal hermítico autoinverso son  $\lambda = \pm 1$ .

14. Demuestre que si  $\mathbf{V}_1$  es polar y  $\mathbf{V}_2$  es axial entonces  $\mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2$  no es un autovector del operador paridad. En otras palabras,  $\mathbf{V}$  no es un vector de  $O(3)$  (aunque si es un vector de  $SO(3)$ ). Esto implica en particular, que el conjunto de todos los vectores de  $O(3)$  más el vector nulo, NO forman un espacio vectorial. En una teoría que conserva paridad, los vectores observables son vectores de  $O(3)$ . Por esta razón, la suma  $\mathbf{E} + \mathbf{B}$  de un campo eléctrico más un campo magnético (vector polar más vector axial) no aparece como observable en la teoría electromagnética. En general, la suma de un tensor polar de rango  $n$  con un tensor axial de rango  $n$ , no representará un observable en una teoría que conserve paridad, de manera que solo la suma entre tensores polares o entre tensores axiales formará un observable en éstas teorías<sup>35</sup>.
15. Demuestre que el triple producto escalar  $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$  de tres vectores polares es un pseudoescalar.

<sup>35</sup>Por ejemplo,  $\mathbf{E}^2 + \mathbf{B}^2$  sí representa un observable importante, y claramente forma un escalar polar. Hay teorías que no conservan paridad como es la teoría de la interacción débil en partículas elementales, para éstas teorías es posible sumar tensores con pseudotensores y obtener observables.

## Cinemática del cuerpo rígido

Ya hemos definido un cuerpo rígido como un conjunto de partículas cuyas distancias relativas son constantes. La condición de cuerpo rígido se traduce en un conjunto de ligaduras holónomas de la forma

$$r_{ij} = c_{ij} \quad ; \quad i < j, \quad i = 1, \dots, N \quad (13.1)$$

siendo  $r_{ij}$  la distancia entre las partículas  $i$ -ésima y  $j$ -ésima, y  $c_{ij}$  son valores constantes. Aunque esto constituye una idealización, la mecánica del cuerpo rígido es de fundamental importancia en muchos campos de la Física. En este capítulo discutiremos la cinemática del cuerpo rígido, es decir las características de su movimiento, para lo cual será necesario desarrollar ciertas herramientas matemáticas, que además serán aplicables en muchos otros campos de la Física.

De particular importancia será el movimiento de rotación de un cuerpo rígido, el cual estudiaremos con el formalismo de la matrices ortogonales. Dicho tipo de movimiento introducirá consideraciones geométricas relativas a la rata de cambio de un vector con el tiempo con respecto a un sistema de referencia rotante. Posteriormente se estudiará la dinámica del cuerpo rígido, es decir la respuesta que estos sistemas producen cuando se someten a torques y fuerzas externas.

### 13.1. Coordenadas independientes de un cuerpo rígido

Como punto de partida, es necesario clarificar cuales son los grados de libertad independientes de un cuerpo rígido, es decir las coordenadas necesarias para especificar completamente su configuración en un instante dado de tiempo. La intuición nos indica que para ubicar la configuración completa de un sólido rígido es necesario especificar la posición de un punto de referencia del cuerpo (por ejemplo su centro de masa, un punto fijo de éste, etc.), y la orientación relativa de este cuerpo, todo ello con respecto a un sistema inercial fijo de coordenadas que lo llamaremos el sistema de laboratorio ( $S$ ) ver Fig. 13.1. Si trasladamos paralelamente los ejes de  $S$  hasta el punto de referencia que se tomará como nuevo origen (sistema  $S_0$  en la Fig. 13.1), la orientación relativa se obtiene usando un sistema  $S'$  cuyos ejes están fijos al cuerpo rígido y que por tanto, siguen su dinámica. El sistema  $S'$  está en general rotado con respecto a  $S_0$  pero con origen común con éste. Por consiguiente,  $S'$  se puede obtener por rotación pura del sistema  $S_0$  (sin traslación).

El número de grados de libertad necesarios para pasar de  $S_0$  a  $S'$  se puede obtener con el siguiente argumento geométrico: A partir de los ejes  $X_1X_2X_3$  del sistema  $S_0$ , se puede construir el eje  $X'_1$  asociado con  $S'$  usando dos grados de libertad (e.g. dos ángulos que definen al vector unitario  $\mathbf{u}'_1$  con respecto a  $S_0$ ). Puesto que el origen es común para  $S_0$  y  $S'$ , el eje  $X'_2$  yace en el plano perpendicular a  $X'_1$  y que pasa por el origen. En consecuencia, solo un grado de libertad adicional es necesario (e.g. un ángulo) para orientar a  $X'_2$  en este plano. Finalmente,  $X'_3$  ya queda determinado si queremos construir ejes a derecha. Hemos requerido entonces tres grados de libertad para pasar de  $S_0$  a  $S'$ .

Lo anterior nos conduce entonces a **seis grados de libertad** para definir la configuración del cuerpo rígido a partir de  $S$ : 3 para definir el punto de referencia (origen de  $S_0$  y de  $S'$ ) y 3 para definir la orientación relativa

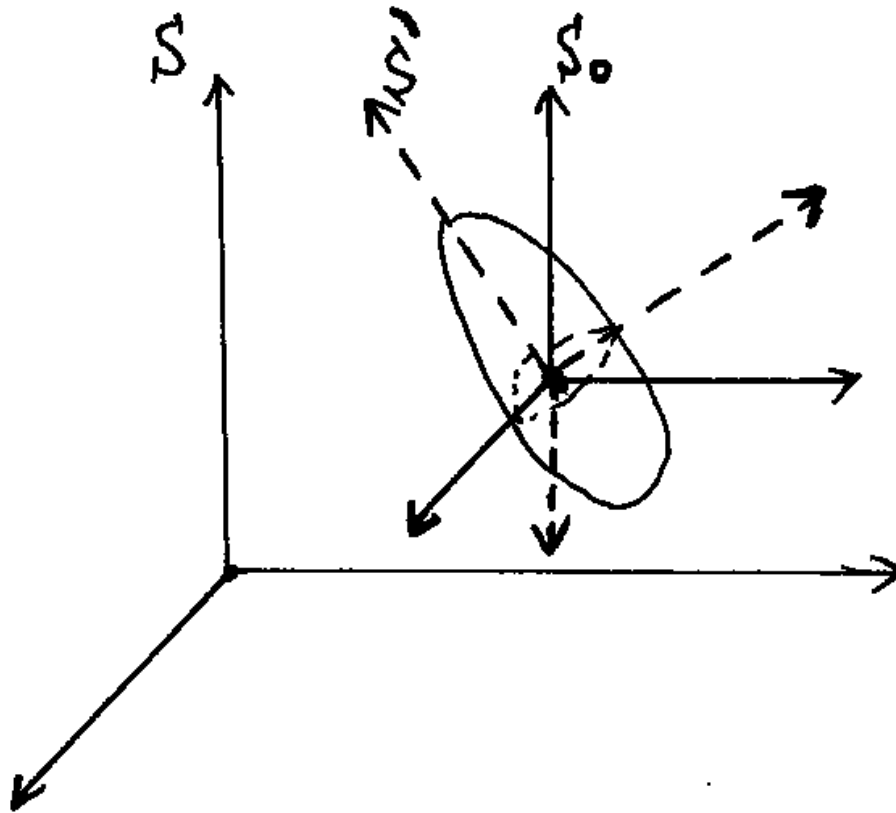


Figura 13.1: Orientación relativa de un cuerpo rígido. El sistema  $S$  es inercial,  $S_0$  es un sistema cuyo origen está fijo en algún punto de referencia del cuerpo y cuyos ejes son paralelos a los de  $S$ . El sistema  $S'$  está fijo al cuerpo y tiene origen común con  $S_0$ .

de  $S'$  con respecto a  $S$  (o lo que es lo mismo, con respecto a  $S_0$ ). Nótese que dicho número de grados de libertad es independiente del número de partículas del sólido e incluso de su naturaleza discreta o continua.

Veamos como conciliar este número con las ligaduras enunciadas en (13.1). Hay  $3N$  grados de libertad superficiales debidos a las  $N$  partículas del sólido. Para un número  $N$  de partículas el número de ecuaciones de ligadura de la forma (13.1) viene dado por  $N(N-1)/2$ . Si intentamos simplemente sustraer el número de ligaduras de los  $3N$  grados de libertad, obtenemos valores negativos (para  $N \geq 8$ ) que crecen rápidamente en magnitud cuando aumenta  $N$ . El punto es que estas ligaduras no son todas independientes. Para fijar un punto en el cuerpo rígido, no es necesario especificar las distancias a todos los demás puntos del cuerpo, es suficiente con especificar su distancia a otros 3 puntos fijos al cuerpo y no colineales como se vé en la figura 13.2. Por lo tanto, una vez fijadas las posiciones de tres partículas no colineales del cuerpo, las ligaduras fijan las posiciones de las otras. Por el momento esto nos deja con 9 grados de libertad, sin embargo, una vez fijada la partícula 1 de referencia (3 grados de libertad) la ubicación de la partícula 2 solo requiere dos grados de libertad pues al estar fija su distancia a la primera, debe yacer en la superficie de una esfera centrada en la partícula 1. Fijadas las partículas 1 y 2, la tercera debe yacer en una circunferencia cuyo centro pasa por el eje definido por el vector relativo  $\mathbf{r}_{12}$  siendo dicho vector perpendicular al círculo y la distancia entre un punto de la circunferencia y las partículas 1 y 2 fijadas por las ligaduras (13.1), esto nos deja con un grado de libertad necesario para fijar esta posición (por ejemplo un ángulo). Aparecen entonces seis grados de libertad en consistencia con lo que se obtuvo con el argumento anterior. Esto se puede ver a su vez de forma más simple teniendo en cuenta que a los 9 grados de libertad para fijar las tres partículas se les deben restar 3 ligaduras

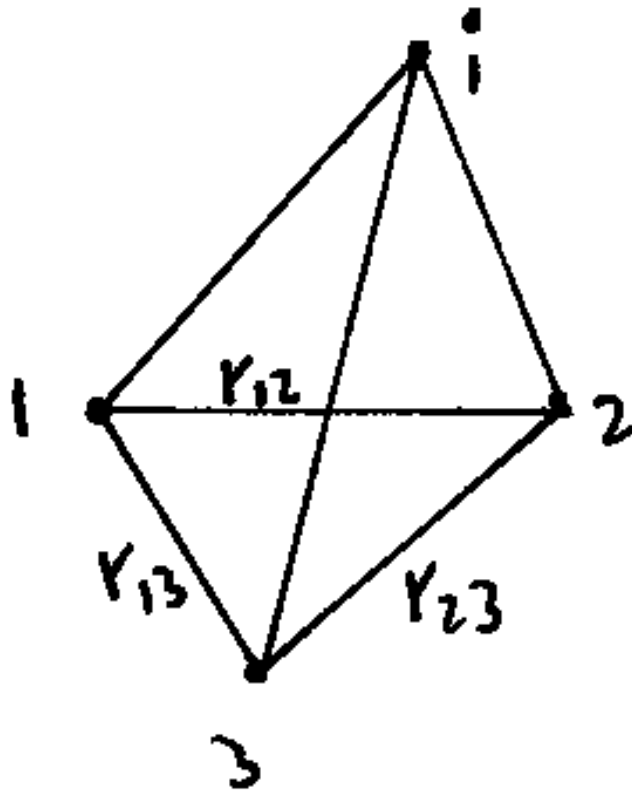


Figura 13.2: Tres puntos de referencia no colineales del cuerpo rígido permiten localizar cualquier partícula  $i$ -ésima en dicho cuerpo.

independientes

$$r_{12} = c_{12} \quad ; \quad r_{23} = c_{23} \quad ; \quad r_{13} = c_{13}$$

por supuesto pueden haber ligaduras adicionales tales como condiciones de rodadura, existencia de puntos fijos etc. que pueden disminuir aún más el número de grados de libertad independientes.

## 13.2. Asignación de los grados de libertad de un cuerpo rígido

Ya mencionamos que una forma de determinar la configuración del cuerpo rígido en un instante dado con respecto a un sistema de referencia inercial  $S$ , consiste en construir un sistema de referencia  $S_0$  con ejes paralelos a  $S$  y con su origen en un punto de referencia del sólido. Adicionalmente construimos otro sistema de ejes  $S'$  fijos al cuerpo y con origen en el punto de referencia del sólido. Tres grados de libertad fijan el origen común de  $S_0$  y  $S'$  con respecto a  $S$  (i.e. la posición del punto de referencia con respecto a  $S$ ), en tanto que otros 3 fijan la orientación relativa de  $S'$  con respecto a  $S_0$  o lo que es lo mismo con respecto a  $S$ .

Como la determinación de la posición del origen de  $S_0$  con respecto a  $S$  no representa particular dificultad, nos restringiremos a estudiar la forma de parametrizar la orientación relativa de  $S'$  con respecto a  $S_0$ . Hay varias maneras de realizar esta parametrización. Una de las más fructíferas es la determinación de los **cosenos directores** de los ejes primados (de  $S'$ ) con respecto a los ejes no primados (de  $S_0$ ). Por simplicidad usaremos una notación de índices en lugar de la nomenclatura  $x, y, z$  i.e.  $x_1, x_2, x_3$

La dirección del eje  $X'_1$  se puede especificar a través de los ángulos que este eje hace con cada uno de los ejes  $X_1, X_2, X_3$ . Los cosenos directores se definen de la forma

$$\cos \theta_{ij} = \cos (\mathbf{u}'_i, \mathbf{u}_j) = \mathbf{u}'_i \cdot \mathbf{u}_j = \mathbf{u}_j \cdot \mathbf{u}'_i \quad ; \quad i, j = 1, 2, 3 \quad (13.2)$$

siendo  $\mathbf{u}_j$  los vectores cartesianos unitarios a lo largo de los ejes  $X_j$ , y  $\mathbf{u}'_i$  los análogos para el sistema  $S'$ . En  $\theta_{ij}$  el primer índice se refiere al índice primado y el segundo al índice no primado. La proyección de un vector  $\mathbf{u}'_i$  a lo largo de  $\mathbf{u}_1$  está dada por  $\mathbf{u}'_i \cdot \mathbf{u}_1$  de modo que la componente de  $\mathbf{u}'_i$  a lo largo de  $\mathbf{u}_1$  viene dada por  $(\mathbf{u}'_i \cdot \mathbf{u}_1) \mathbf{u}_1$  de manera similar se obtienen las componentes de  $\mathbf{u}'_i$  a lo largo de  $\mathbf{u}_2$  y  $\mathbf{u}_3$  con lo cual se obtiene

$$\mathbf{u}'_i = (\mathbf{u}'_i \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{u}'_i \cdot \mathbf{u}_2) \mathbf{u}_2 + (\mathbf{u}'_i \cdot \mathbf{u}_3) \mathbf{u}_3$$

pero recordando la definición de los cosenos directores y usando de aquí en adelante convención de suma sobre índices repetidos, se obtiene

$$\mathbf{u}'_i = (\cos \theta_{ij}) \mathbf{u}_j \quad (13.3)$$

con lo cual se muestra que los 9 cosenos directores determinan completamente la orientación del nuevo sistema coordenado  $X'_1, X'_2, X'_3$  con respecto a  $X_1, X_2, X_3$  de  $S_0$ . El proceso es fácilmente invertible para obtener a los  $\mathbf{u}_i$  en función de los  $\mathbf{u}'_j$ .

$$\mathbf{u}_i = (\mathbf{u}_i \cdot \mathbf{u}'_j) \mathbf{u}'_j = (\cos \theta_{ji}) \mathbf{u}'_j \quad (13.4)$$

Podemos escribir entonces

$$\mathbf{r} = x_i \mathbf{u}_i = x'_j \mathbf{u}'_j$$

puesto que el vector posición en esencia no cambia (ya que el origen es común), solo cambian las componentes y bases con que se mide. Claramente

$$\begin{aligned} x'_i &= \mathbf{r} \cdot \mathbf{u}'_i = (x_j \mathbf{u}_j) \cdot \mathbf{u}'_i = x_j (\mathbf{u}_j \cdot \mathbf{u}'_i) \\ x'_i &= \cos \theta_{ij} x_j \end{aligned} \quad (13.5)$$

los cosenos directores en consecuencia permiten obtener la relación entre las coordenadas de un punto en un sistema con respecto a las del otro (componentes de  $\mathbf{r}$ ). Claramente el análisis anterior es extensible a otros vectores, diferentes al vector posición<sup>1</sup>. La componente de un cierto vector  $\mathbf{G}$  a lo largo de  $X'_i$  viene dada por

$$G_{x'_i} = \mathbf{G} \cdot \mathbf{u}'_i = (G_j \mathbf{u}_j) \cdot \mathbf{u}'_i = G_j (\mathbf{u}_j \cdot \mathbf{u}'_i) = \cos \theta_{ij} G_j$$

de modo que el conjunto de los 9 cosenos directores nos da la transformación completa entre los dos sistemas coordenados.

Al ser los ejes primados fijos al cuerpo, los 9 cosenos directores son en general funciones del tiempo ya que el cuerpo puede cambiar su orientación en el curso del movimiento. En tal sentido, los cosenos directores se pueden considerar como coordenadas que especifican la orientación instantánea del cuerpo rígido con respecto a  $S_0$ . Ya hemos visto sin embargo, que solo tres coordenadas son necesarias para orientar el cuerpo, de modo que los 9 cosenos directores no pueden ser independientes. De hecho, es necesario que existan seis ecuaciones de ligadura para obtener los tres grados de libertad requeridos. Veamos como traducir las ligaduras sobre los cosenos directores en forma de ecuaciones. Tanto los sistemas primado como sin primar son ortogonales

$$\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij} \quad ; \quad \mathbf{u}'_i \cdot \mathbf{u}'_j = \delta_{ij} \quad (13.6)$$

teniendo en cuenta las ecuaciones (13.3)

$$\begin{aligned} \mathbf{u}'_i \cdot \mathbf{u}'_k &= [(\cos \theta_{ij}) \mathbf{u}_j] \cdot [(\cos \theta_{km}) \mathbf{u}_m] = \cos \theta_{ij} \cos \theta_{km} (\mathbf{u}_j \cdot \mathbf{u}_m) \\ \mathbf{u}'_i \cdot \mathbf{u}'_k &= \cos \theta_{ij} \cos \theta_{km} \delta_{jm} = \cos \theta_{im} \cos \theta_{km} \end{aligned}$$

<sup>1</sup>Nótese que estrictamente hablando, la posición no es un vector en el sentido Euclidiano, pues a un punto en el espacio no se le puede asignar magnitud ni dirección. Un vector debe ser un objeto independiente del sistema coordenado con que se mire, por ejemplo esto lo cumple el vector desplazamiento. El "vector" posición se puede pensar como un vector relativo entre el punto en cuestión y un punto de referencia (origen).

que combinado con las relaciones de ortonormalidad en el sistema primado nos da

$$\cos \theta_{im} \cos \theta_{km} = \delta_{ik} \quad (13.7)$$

se puede obtener una expresión análoga partiendo de expresar  $\mathbf{u}_i \cdot \mathbf{u}_k$  en términos de cosenos directores

$$\cos \theta_{mi} \cos \theta_{mk} = \delta_{ik} \quad (13.8)$$

en cualquiera de sus formas, estas relaciones nos brindan seis ecuaciones independientes (por ejemplo se puede imponer que  $i \geq k$  en estas ecuaciones). Esto nos deja con tres grados de libertad independientes como esperábamos.

En vista de lo anterior, los 9 cosenos directores no se pueden usar como coordenadas generalizadas en un formalismo Lagrangiano o Hamiltoniano para encontrar las ecuaciones de movimiento, es necesario encontrar tres funciones independientes de estos cosenos directores. Discutiremos más adelante algunos de ellos siendo los más importantes los ángulos de Euler. No obstante, los cosenos directores son de gran utilidad para estudiar la conexión entre los dos sistemas cartesianos y expresar de una forma más simple y general varios teoremas relativos a los cuerpos rígidos. De igual forma permite el uso de algoritmos extendibles a la relatividad especial y la mecánica cuántica. Por lo tanto, seguiremos trabajando en el lenguaje de los cosenos directores.

### 13.3. Transformaciones ortogonales

Con el fin de acortar la notación definiremos

$$a_{ij} \equiv \cos \theta_{ij}$$

las ecuaciones (13.5) constituyen un conjunto de **transformaciones lineales** siendo  $a_{ij}$  un conjunto de coeficientes constantes (independientes de  $x, x'$ ). En la nueva notación podemos sintetizar estas transformaciones como

$$x'_i = a_{ij} x_j \quad , \quad i = 1, 2, 3 \quad (13.9)$$

dado que los cosenos directores no son todos independientes, estas ecuaciones no definen la transformación lineal homogénea más general. Las ligaduras obtenidas anteriormente se pueden ver fácilmente a la luz de la nueva notación, teniendo en cuenta que el vector posición tiene la misma magnitud en ambos sistemas de referencia  $S'$  y  $S_0$  en virtud de su origen común. Esto nos conduce a la condición

$$x_i x_i = x'_i x'_i \quad (13.10)$$

que en el caso de vectores reales corresponde a la invarianza de su norma, y como se puede ver en la sección 12.4, la transformación lineal más general que deja invariante la cantidad  $x_i x_i$ , conduce a las condiciones

$$a_{ij} a_{ik} = \delta_{jk} \quad ; \quad j, k = 1, 2, 3 \quad (13.11)$$

que resultan idénticas a las ligaduras expresadas en (13.8). Naturalmente, la forma alternativa de la condición de ortogonalidad Ec. (13.7), se escribe en esta notación en la forma

$$a_{ji} a_{ki} = \delta_{jk} \quad ; \quad j, k = 1, 2, 3 \quad (13.12)$$

La transformación lineal (13.9) se denomina **transformación ortogonal** si los coeficientes  $a_{ij}$  de la transformación satisfacen las Ecs. (13.11) conocidas como **condiciones de ortogonalidad**. De modo que la transición entre las coordenadas de  $S_0$  y las de  $S'$  se realizan a través de una transformación ortogonal. Los coeficientes de la transformación se pueden escribir en un arreglo matricial de la forma

$$\mathbf{A} \equiv \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

que denominaremos matriz de la transformación siendo los  $a_{ij}$  elementos matriciales de la transformación. Nótese que la condición de ortogonalidad (13.11) se puede ver mnemotécnicamente como si la matriz estuviera constituida por 3 vectores columna ortonormales entre sí i.e.

$$\mathbf{A} \equiv (\mathbf{V}_1 \ \mathbf{V}_2 \ \mathbf{V}_3) \quad ; \quad \mathbf{V}_k \equiv \begin{pmatrix} a_{1k} \\ a_{2k} \\ a_{3k} \end{pmatrix} \quad ; \quad k = 1, 2, 3$$

$$\mathbf{V}_j \cdot \mathbf{V}_k = a_{ij}a_{ik} = \delta_{jk} \quad (13.13)$$

la cual se puede ver como una relación de ortonormalidad entre los “vectores” columna arriba descritos<sup>2</sup>. Sin embargo, es importante enfatizar que esta es una regla mnemotécnica ya que los  $\mathbf{V}_k$  no definen ningún vector en el espacio. La representación de un vector en términos de componentes debe estar referida a una **base específica**. Estos coeficientes son realmente enlaces entre dos bases diferentes y constituyen una transformación que genera un **cambio de base**. El producto punto definido aquí es solo un algoritmo establecido para el arreglo vectorial pero no se está realizando el producto entre vectores unitarios bien definidos.

A manera de ejemplo, veamos el caso del movimiento en un plano. Si por simplicidad, nos restringimos a rotación en el plano  $X_1X_2$  entonces el eje de rotación es el eje  $X_3$  de modo que  $\mathbf{u}'_3 = \mathbf{u}_3$ . Pensemos que en un instante dado de tiempo el ángulo entre  $X_1$  y  $X'_1$  es  $\phi$ , se tiene entonces que

$$\begin{aligned} a_{11} &= \cos \theta_{11} = \mathbf{u}'_1 \cdot \mathbf{u}_1 = \cos \phi \quad ; \quad a_{12} = \mathbf{u}'_1 \cdot \mathbf{u}_2 = \cos \left( \frac{\pi}{2} - \phi \right) = \sin \phi \quad ; \quad a_{13} = \mathbf{u}'_1 \cdot \mathbf{u}_3 = \mathbf{u}'_1 \cdot \mathbf{u}'_3 = 0 \\ a_{21} &= \mathbf{u}'_2 \cdot \mathbf{u}_1 = \cos \left( \frac{\pi}{2} + \phi \right) = -\sin \phi \quad ; \quad a_{22} = \mathbf{u}'_2 \cdot \mathbf{u}_2 = \cos \phi \quad ; \quad a_{23} = \mathbf{u}'_2 \cdot \mathbf{u}_3 = \mathbf{u}'_2 \cdot \mathbf{u}'_3 = 0 \\ a_{31} &= \mathbf{u}'_3 \cdot \mathbf{u}_1 = \mathbf{u}_3 \cdot \mathbf{u}_1 = 0 \quad ; \quad a_{32} = \mathbf{u}'_3 \cdot \mathbf{u}_2 = \mathbf{u}_3 \cdot \mathbf{u}_2 = 0 \quad ; \quad a_{33} = \mathbf{u}'_3 \cdot \mathbf{u}_3 = 1 \end{aligned}$$

la matriz de transformación queda de la forma

$$\mathbf{A}_{x_1, x_2} = \begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (13.14)$$

y la transformación de coordenadas queda

$$\begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

$$x'_1 = x_1 \cos \phi + x_2 \sin \phi \quad ; \quad x'_2 = -x_1 \sin \phi + x_2 \cos \phi \quad ; \quad x'_3 = x_3$$

puede observarse que las condiciones de ortogonalidad (13.11) se cumplen en esta matriz. Es claro que en este caso solo hay una coordenada independiente cual es el ángulo  $\phi$  de rotación del eje  $X_3$ , los otros dos grados de libertad requeridos para determinar el vector unitario a lo largo del eje de rotación ya han sido considerados.

Para continuar trabajando sobre la cinemática del cuerpo rígido, es necesario utilizar las propiedades matriciales que desarrollamos en el capítulo 12. Insistimos nuevamente que estas propiedades son idénticas en cualquiera de las dos interpretaciones activa o pasiva, de modo que en general no haremos distinción entre  $(\mathbf{r})'$  y  $\mathbf{r}'$  en las operaciones subsecuentes, a menos que se especifique lo contrario.

## 13.4. Ángulos de Euler

Hemos visto que los 9 cosenos directores no forman un conjunto independiente de modo que no pueden ser usados como coordenadas generalizadas para un formalismo Lagrangiano o Hamiltoniano. Es necesario en

<sup>2</sup>Similarmente, se puede ver a la matriz como compuesta de tres vectores fila que también son ortonormales entre sí. Esto se puede ver de la forma alternativa de la condición de ortogonalidad Ec. (13.12).

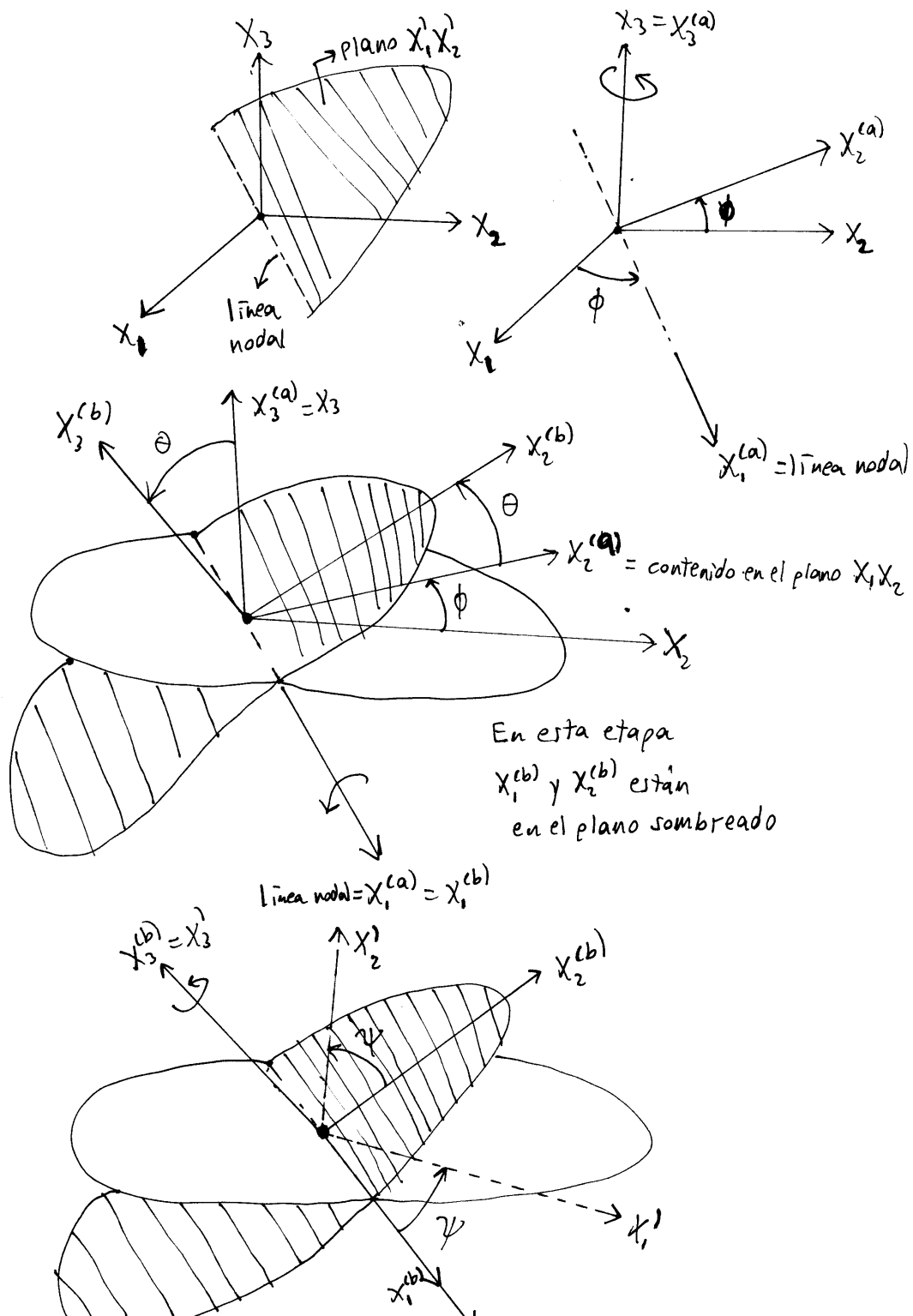


Figura 13.3: Conjunto de rotaciones para llegar desde los ejes  $X_1X_2X_3$  hasta los ejes  $X'_1X'_2X'_3$ .

consecuencia buscar tres parámetros independientes con los cuales se pueda escribir la matriz ortogonal de rotación. Estos parámetros se podrán usar entonces como coordenadas generalizadas independientes para un formalismo Lagrangiano o Hamiltoniano. Por otro lado, dado que las rotaciones son transformaciones continuas, se sigue que además de las condiciones de ortogonalidad, las matrices ortogonales asociadas deben ser **propias** i.e. su determinante debe ser +1.



Hay muchas escogencias posibles para los tres parámetros independientes, siendo las más populares los ángulos de Euler que describimos a continuación: Queremos pasar de un conjunto de ejes coordenados ortogonales  $X_1X_2X_3$  a otro conjunto de ejes ortogonales  $X'_1X'_2X'_3$  donde ambos sistemas de ejes tienen origen común. Es claro que si determinamos las direcciones de los ejes  $X'_1X'_2$  el tercer eje está unívocamente determinado si la transformación es continua (ya que en este caso se conserva la quiralidad del sistema de ejes). Para ello necesitamos determinar el plano en donde yace  $X'_1X'_2$  y algún ángulo que oriente a estos ejes en dicho plano. La figura 13.3 muestra al plano generado por  $X_1X_2$  y al plano (sombreado) generado por  $X'_1X'_2$  los cuales forman un ángulo diedro, y se intersectan en una **línea nodal** como indica la figura 13.3. Para pasar del sistema  $X_1X_2X_3$  al sistema  $X'_1X'_2X'_3$  debemos llevar a los ejes  $X'_1X'_2$  a sus posiciones finales lo cual se puede realizar esquemáticamente en tres pasos (a) Rotar el sistema para que el nuevo eje  $X_1$  quede en el plano sombreado (plano generado por  $X'_1X'_2$ ), (b) rotar el sistema para que el nuevo eje  $X_2$  entre también en el plano sombreado, una vez que los nuevos ejes  $X_1 X_2$  ya estén en el plano de  $X'_1X'_2$  el paso final es (c) hacer una rotación de los ejes en el plano sombreado para que tengan su orientación final.

Veamos en detalle el proceso: (a) En primer lugar pasamos del sistema  $X_1X_2X_3$  al sistema  $X_1^{(a)}X_2^{(a)}X_3^{(a)}$  haciendo una rotación alrededor del eje  $X_3$  en un ángulo  $\phi$  de modo que el nuevo eje  $X_1^{(a)}$  quede a lo largo de la línea nodal, es decir  $\mathbf{X}_1^{(a)}$  queda dentro del plano sombreado, en este caso es claro que  $X_3^{(a)} = X_3$ . (b) En el siguiente paso, pasamos del sistema  $X_1^{(a)}X_2^{(a)}X_3^{(a)}$  al sistema  $X_1^{(b)}X_2^{(b)}X_3^{(b)}$  de modo que el nuevo eje  $X_2^{(b)}$  quede dentro del plano sombreado, pero sin sacar al eje que ya se introdujo en ese plano. Por tanto, se deduce que la rotación debe realizarse alrededor de  $X_1^{(a)}$  para que este eje no se salga del plano sombreado. Haremos entonces una rotación en un ángulo  $\theta$  alrededor de  $X_1^{(a)}$ , siendo  $\theta$  el ángulo adecuado para que el eje  $X_2^{(b)}$  quede dentro del plano sombreado. En este caso es claro que  $X_1^{(b)} = X_1^{(a)}$  y con este paso hemos conseguido que los nuevos ejes  $X_1^{(b)}$  y  $X_2^{(b)}$  estén en el plano generado por  $X'_1X'_2$ , solo falta en consecuencia realizar una rotación dentro de este plano que nos lleve desde  $X_1^{(b)} X_2^{(b)}$  hacia los ejes  $X'_1X'_2$  lo cual se logra con (c) una rotación alrededor del eje  $X_3^{(b)}$  en un ángulo  $\psi$ . En este caso  $X_3^{(b)} = X_3^{(a)}$ .

En consecuencia, una rotación arbitraria del sistema coordenado  $X_1X_2X_3$  al sistema coordenado  $X'_1X'_2X'_3$  se puede realizar de la forma que se vé en la Fig. 13.3: comenzamos haciendo una rotación antihoraria en un ángulo  $\phi$  alrededor de  $X_3$ , al sistema coordenado resultante de esta rotación lo denotaremos por  $X_1^{(a)}, X_2^{(a)}, X_3^{(a)}$  (naturalmente,  $X_3^{(a)}$  coincide con  $X_3$ ). En el segundo paso, rotamos al nuevo sistema coordenado con respecto a  $X_1^{(a)}$  en sentido antihorario en una cantidad  $\theta$  como muestra la figura 13.3, denotamos al sistema resultante como  $X_1^{(b)}, X_2^{(b)}, X_3^{(b)}$ . El eje  $X_1^{(b)}$  (que coincide con  $X_1^{(a)}$ ), se forma con la intersección entre los planos  $X_1X_2$  y  $X_1^{(b)}X_2^{(b)}$  y se conoce como **línea nodal**. Finalmente, se hace una rotación antihoraria  $\psi$  alrededor de  $X_3^{(b)}$  para llegar al sistema coordenado final  $X'_1X'_2X'_3$ . Los elementos de una transformación completa se puede obtener por composición de las tres operaciones aquí descritas. La rotación inicial (alrededor de  $X_3$ ) transforma a un vector (o a sus componentes) en la forma

$$\mathbf{x}^{(a)} = \mathbf{D}\mathbf{x} \quad ; \quad \mathbf{D} = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.15)$$

donde  $\mathbf{D}$  es una matriz que describe un cambio de base de la forma  $X_1X_2X_3 \rightarrow X_1^{(a)}X_2^{(a)}X_3^{(a)}$ . La segunda rotación alrededor de  $X_1^{(a)}$  describe un cambio desde la base  $X_1^{(a)}X_2^{(a)}X_3^{(a)}$  hacia la base  $X_1^{(b)}X_2^{(b)}X_3^{(b)}$  y nos da

$$\mathbf{x}^{(b)} = \mathbf{C}\mathbf{x}^{(a)} \quad ; \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \quad (13.16)$$

y finalmente describimos el cambio de base  $X_1^{(b)}X_2^{(b)}X_3^{(b)} \rightarrow X'_1X'_2X'_3$  con la rotación alrededor de  $X_3^{(b)}$ :

$$\mathbf{x}' = \mathbf{B}\mathbf{x}^{(b)} \quad ; \quad \mathbf{B} = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.17)$$

utilizando de forma sucesiva estas transformaciones, llegamos a la transformación que nos lleva desde  $\mathbf{x}$  hasta  $\mathbf{x}'$

$$\mathbf{x}' = \mathbf{B}\mathbf{x}^{(b)} = \mathbf{B}(\mathbf{C}\mathbf{x}^{(a)}) = \mathbf{B}\mathbf{C}\mathbf{x}^{(a)} = \mathbf{B}\mathbf{C}(\mathbf{D}\mathbf{x})$$

de modo que llegamos desde  $\mathbf{x}$  hasta  $\mathbf{x}'$  con la transformación:

$$\mathbf{x}' = \mathbf{A}\mathbf{x} \quad ; \quad \mathbf{A} \equiv \mathbf{B}\mathbf{C}\mathbf{D} \quad (13.18)$$

haciendo el producto de las tres matrices (ver ejercicio 1 Pág. 371), se obtiene la matriz de rotación más general en términos de los ángulos de Euler

$$\mathbf{A} = \begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \psi \sin \theta \\ -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \cos \psi \sin \theta \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix} \quad (13.19)$$

Se puede demostrar que  $\mathbf{A}^{-1} = \tilde{\mathbf{A}}$  y que  $\det \mathbf{A} = +1$ . Es decir que  $\mathbf{A}$  es efectivamente una matriz ortogonal real propia (ejercicio 2 Pág. 371).

$$\mathbf{A}^{-1} = \tilde{\mathbf{A}} = \begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & \sin \theta \sin \phi \\ \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & -\sin \theta \cos \phi \\ \sin \psi \sin \theta & \cos \psi \sin \theta & \cos \theta \end{pmatrix} \quad (13.20)$$

el rango de los ángulos en la matriz de rotación viene dado por

$$0 \leq \phi < 2\pi \quad , \quad 0 \leq \theta \leq \pi \quad , \quad 0 \leq \psi < 2\pi$$

donde el rango de  $\theta$  está relacionado con el hecho de que el ángulo diedro entre los planos  $X_1X_2$  y  $X'_1X'_2$  está en el intervalo  $[0, \pi]$ . Existe por supuesto una arbitrariedad en la secuencia de rotaciones que se puede escoger, la primera rotación se puede realizar con respecto a cualquiera de los ejes y en las dos rotaciones posteriores, la única limitación es que no se puede hacer dos rotaciones sucesivas con respecto al mismo eje. Por lo tanto hay un total de 12 convenciones posibles para un sistema dextrógiro. Las dos convenciones más usadas en Física son: nuestra convención (denominada  $x$ -convención), y la llamada  $y$ -convención.

En la  $y$ -convención, la primera rotación (alrededor de  $X_3$ ) se hace de modo que el eje  $X_2^{(a)}$  (y no el eje  $X_1^{(a)}$ ) coincida con la línea nodal. La segunda rotación se hace entonces con respecto a  $X_2^{(a)}$  a fin de introducir al nuevo eje  $X_1^{(b)}$  en el plano  $X'_1X'_2$ . La tercera operación rota a los ejes  $X_1^{(b)}X_2^{(b)}$  hasta su orientación final en el plano  $X'_1X'_2$ . Hay una tercera convención muy usada en aplicaciones en ingeniería conocida como  $xyz$ -convención. Es muy importante conocer la convención usada en la construcción de la matriz a la hora de comparar resultados.

### 13.5. Parámetros de Cayley-Klein (Opcional)

Hemos visto que se requieren 3 cantidades independientes para especificar la inclinación de un sistema de ejes coordenados con respecto a otro, o la rotación de un cuerpo rígido. Sin embargo, en algunas ocasiones es conveniente utilizar un sistema de más de tres coordenadas para caracterizar las rotaciones, a pesar de que tal sistema no formaría un conjunto de coordenadas generalizadas independientes. Felix Klein introdujo un sistema de cuatro parámetros que fuesen adecuados para resolver problemas giroscópicos complejos. Los ángulos de Euler son con frecuencia difíciles de manipular en la práctica, debido a la gran cantidad de funciones trigonométricas implicadas. Los llamados parámetros de Cayley-Klein son más adecuados para su utilización en los computadores. Adicionalmente, los parámetros de Cayley-Klein son muy útiles en desarrollos formales en varias ramas de la Física que involucran rotaciones o simetrías rotacionales. En particular, son muy útiles para el estudio de los grupos de simetría  $SU(2)$  y  $SO(3)$ , así como la relación entre ellos. La teoría de representaciones de estos grupos permite a su vez, formalizar el concepto de espín.

En la presente sección solo mostraremos las características más sobresalientes de esta parametrización, ya que no la utilizaremos en la descripción de la cinemática o la dinámica del cuerpo rígido.

Los cuatro parámetros de Cayley-Klein son cuatro números complejos  $\alpha, \beta, \gamma, \delta$ . Es también conveniente definir  $\alpha$  y  $\beta$  en términos de parámetros reales de la siguiente forma

$$\alpha = e_0 + ie_3 \quad ; \quad \beta = e_2 + ie_1 \quad (13.21)$$

los parámetros de Cayley-Klein tienen las restricciones

$$\beta = -\gamma^* \quad , \quad \delta = \alpha^* \quad , \quad |\alpha|^2 + |\beta|^2 = e_0^2 + e_1^2 + e_2^2 + e_3^2 = 1 \quad (13.22)$$

En términos de esta parametrización la matriz de rotación se escribe como

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2}(\alpha^2 - \gamma^2 + \delta^2 - \beta^2) & \frac{i}{2}(\gamma^2 - \alpha^2 + \delta^2 - \beta^2) & \gamma\delta - \alpha\beta \\ \frac{i}{2}(\alpha^2 + \gamma^2 - \beta^2 - \delta^2) & \frac{1}{2}(\alpha^2 + \beta^2 + \gamma^2 + \delta^2) & -i(\alpha\beta + \gamma\delta) \\ \beta\delta - \alpha\gamma & i(\alpha\gamma + \beta\delta) & \alpha\delta + \beta\gamma \end{pmatrix} \quad (13.23)$$

es fácil ver que esta matriz es real, si la escribimos en términos de los parámetros reales  $e_0, e_1, e_2, e_3$

$$\mathbf{A} = \begin{pmatrix} e_0^2 + e_1^2 - e_2^2 - e_3^2 & 2(e_1e_2 + e_0e_3) & 2(e_1e_3 - e_0e_2) \\ 2(e_1e_2 - e_0e_3) & e_0^2 - e_1^2 + e_2^2 - e_3^2 & 2(e_2e_3 + e_0e_1) \\ 2(e_1e_3 + e_0e_2) & 2(e_2e_3 - e_0e_1) & e_0^2 - e_1^2 - e_2^2 + e_3^2 \end{pmatrix}$$

es fácil demostrar que esta es una matriz ortogonal propia. Adicionalmente, se puede demostrar que su relación con los ángulos de Euler viene dada por

$$\begin{aligned} \alpha &= \exp \left[ i \left( \frac{\psi + \phi}{2} \right) \right] \cos \frac{\theta}{2} & ; & \quad \beta = i \exp \left[ i \left( \frac{\psi - \phi}{2} \right) \right] \sin \frac{\theta}{2} \\ \gamma &= -\beta^* = i \exp \left[ -i \left( \frac{\psi - \phi}{2} \right) \right] \sin \frac{\theta}{2} & ; & \quad \delta = \alpha^* = \exp \left[ -i \left( \frac{\psi + \phi}{2} \right) \right] \cos \frac{\theta}{2} \end{aligned}$$

de estas ecuaciones se pueden extraer también las relaciones entre los ángulos de Euler y los cuatro parámetros de Euler reales  $e_\mu$

$$\begin{aligned} e_0 &= \cos \frac{\phi + \psi}{2} \cos \frac{\theta}{2} & ; & \quad e_2 = \sin \frac{\phi - \psi}{2} \sin \frac{\theta}{2} \\ e_1 &= \cos \frac{\phi - \psi}{2} \sin \frac{\theta}{2} & ; & \quad e_3 = \sin \frac{\phi + \psi}{2} \cos \frac{\theta}{2} \end{aligned}$$

un tratamiento detallado de los parámetros de Cayley-Klein se puede encontrar en la sección 4.5 de la segunda edición de la Ref. [1].

## 13.6. Teorema de Euler para el movimiento del cuerpo rígido

Hemos visto que la orientación relativa de un cuerpo rígido en cualquier instante se puede especificar con una transformación ortogonal real propia. Tres parámetros debidamente escogidos determinan la matriz de la transformación. Ahora bien, la orientación del cuerpo puede cambiar en el tiempo, y por lo tanto la matriz de transformación se escribe como  $\mathbf{A}(t)$ . En particular, se pueden escoger los ejes iniciales de  $S'$  coincidentes con los de  $S_0$  con lo cual

$$\mathbf{A}(0) = \mathbf{1}$$

como el movimiento del cuerpo es continuo, la matriz debe evolucionar en forma continua desde la identidad hasta su valor en un tiempo dado  $\mathbf{A}(t)$ .

A continuación, demostraremos que una orientación arbitraria puede ser generada especificando un eje de rotación (vector unitario) y el ángulo de rotación alrededor de tal eje. El siguiente teorema es un enunciado formal de tal hecho

**Theorem 22** *Teorema de Euler: El movimiento general de un cuerpo rígido con un punto fijo es una rotación alrededor de algún eje que pasa por dicho punto.*

El teorema establece que para toda rotación con un punto fijo es siempre posible encontrar un eje que pasa por el punto fijo orientado de una forma particular (con respecto a  $S_0$ ), tal que la rotación con un cierto ángulo  $\psi$  alrededor de este eje, emula la rotación general. En este caso podemos pensar en los parámetros angulares  $\theta, \phi$  (que orientan al eje con respecto a  $S_0$ ) y el parámetro de rotación  $\psi$  como los tres parámetros independientes.

Es claro que una rotación con un punto fijo en un intervalo dado de tiempo nos lleva de una configuración inicial a una configuración final. Si bien el teorema nos afirma que a partir de la configuración inicial es posible llegar a la configuración final por medio de una rotación con un eje fijo para cierto ángulo, esto no quiere decir que el movimiento real haya sido con eje fijo, de hecho durante el movimiento es posible que la orientación del eje de rotación haya cambiado en el tiempo. Esto es análogo al caso traslacional: el movimiento de una partícula puntual nos lleva de una posición inicial a una posición final, para ir de esta posición inicial a la posición final es siempre posible realizarlo a través de una línea recta (dirección constante) recorriendo cierta distancia, pero esto no significa que el movimiento real haya sido en línea recta, aunque en ambos casos la posición inicial y la final coincidan.

Si el punto es fijo con respecto a un sistema inercial (el punto fijo no es necesariamente el centro de masa del cuerpo), entonces podemos hacer que  $S$  coincida con  $S_0$  y tal punto se toma como origen de  $S, S_0$  y  $S'$  para todo tiempo, el desplazamiento del cuerpo rígido no involucra traslación de los ejes del cuerpo respecto a  $S$ , solo cambia su orientación. El teorema entonces establece que los ejes coordenados  $X'_1 X'_2 X'_3$  se pueden obtener para cualquier tiempo  $t$ , por medio de una única rotación alrededor de un eje dado del sistema  $X_1 X_2 X_3$ . Esto implica formalmente que el operador representado por la matriz  $\mathbf{A}$  que describe el movimiento físico del cuerpo rígido es una rotación. Es característico de una rotación, que la dirección del eje de rotación permanezca inalterada en el proceso. Por tanto, un vector que vaya a lo largo de tal eje debe tener las mismas componentes en  $S_0$  y  $S'$ . La otra condición necesaria para una rotación es que la magnitud de los vectores sea un invariante, está garantizada por el carácter ortogonal real de la matriz como vimos en la sección 12.4. En consecuencia, el teorema de Euler será probado si podemos mostrar que para una rotación no trivial (diferente a la identidad) existe al menos un vector no nulo  $\mathbf{x}$  que tiene las mismas componentes en ambos sistemas coordenados  $S$  y  $S'$ , pero que todo vector linealmente independiente de  $\mathbf{x}$  tiene componentes distintas en  $S$  y  $S'$ . Lo último garantiza que la dirección del eje de rotación está unívocamente determinada, ya que geoméricamente una rotación no trivial debe transformar de manera no trivial a cualquier vector que no sea colineal con el eje de rotación.

En notación matricial, debe existir  $\mathbf{x}$  no nulo tal que

$$\mathbf{x}' = \mathbf{A}\mathbf{x} = \mathbf{x} \quad (13.24)$$

comparando esta ecuación con la ecuación de valores propios (12.35), se observa que el problema se reduce a demostrar que  $\lambda = 1$  es un valor propio no degenerado de la matriz  $\mathbf{A}$ , cuando  $\mathbf{A}$  es no trivial. La no degeneración de  $\lambda = 1$  para una rotación no trivial, garantiza que ningún vector linealmente independiente de  $\mathbf{x}$ , puede ser vector propio de  $\mathbf{A}$  con valor propio  $\lambda = 1$ . Por otro lado, es claro que si la rotación es la identidad, cualquier eje que pase por el punto fijo sirve como eje de rotación. Esto se refleja en el hecho de que  $\lambda = 1$  es totalmente degenerado para la rotación identidad.

Podemos entonces reescribir el teorema de Euler en la siguiente forma:

*La matriz ortogonal real propia que especifica el movimiento de un cuerpo rígido con un punto fijo, o bien es la identidad, o bien es tal que  $\lambda = 1$  es un valor propio no degenerado de la matriz.*

En la sección 12.7 encontramos que para toda matriz **ortogonal real propia de dimensión impar** siempre hay por lo menos un valor propio igual a la unidad. Recordando además que las rotaciones solo pueden ser descritas por matrices ortogonales propias y que nuestra dimensión es tres, se obtiene la existencia de por lo menos un vector no nulo que permanece invariante bajo la rotación. Nótese que la validez del teorema de Euler dependió del carácter impar de la dimensión del espacio. Por ejemplo, en dos dimensiones

no existe un vector **en el espacio bidimensional** que quede invariante ante una rotación, el eje de rotación es perpendicular al plano y está por tanto fuera de dicho espacio.

Finalmente, debemos determinar el grado de degeneración de  $\lambda = 1$ . Examinando los valores propios de una matriz ortogonal real propia de tres dimensiones Sec. 12.7.1 Ecs. (12.75), y recordando las transformaciones asociadas a las posibles estructuras de valores propios (ver Pág. 301), se puede ver que los posibles valores de  $\lambda$  son 1,  $e^{i\Phi}$  y  $e^{-i\Phi}$ . Si  $\Phi \neq 0$ , el valor propio unidad es no degenerado, en tanto que  $\Phi = 0$  corresponde a la identidad (degeneración total de  $\lambda = 1$ ). Esto completa la demostración del teorema de Euler.

La Ec. (13.24) nos muestra que el autovector asociado a  $\lambda = 1$  determina la dirección del eje de rotación, y dado que si la rotación es no trivial solo hay un valor propio  $\lambda = 1$  podemos determinar unívocamente dicho eje. Los cosenos directores del eje de rotación se pueden obtener entonces haciendo  $\lambda = 1$  en las ecuaciones de valores propios (12.37) y resolviendo para  $x_1, x_2, x_3$  cuyas soluciones forman un autovector asociado a  $\lambda = 1$  (recordemos que estrictamente solo se obtienen cocientes entre tales componentes, si imponemos unitariedad podemos definir casi completamente un vector en la dirección del eje excepto por su sentido, cualquier múltiplo de este autovector también es autovector asociado al mismo autovalor<sup>3</sup>).

Una vez determinado el eje de rotación procedemos a encontrar el ángulo de rotación alrededor de dicho eje. A través de una transformación de similaridad (cambio de base), es posible transformar la matriz  $\mathbf{A}$  a una matriz equivalente i.e. que describe al mismo operador pero en otra base

$$\mathbf{A}' = \mathbf{B}\mathbf{A}\mathbf{B}^{-1}$$

en particular podemos escoger el cambio de base  $\mathbf{B}$  de tal forma que el nuevo eje  $X'_3$  coincida con el eje de rotación. En tal sistema coordenado  $\mathbf{A}'$  representa una rotación alrededor de  $X'_3$  i.e. en el plano  $X'_1X'_2$ , en un ángulo  $\alpha$ . En consecuencia, la matriz  $\mathbf{A}'$  adquiere la forma

$$\mathbf{A}' = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.25)$$

la traza de  $\mathbf{A}'$  es de la forma

$$\text{Tr}\mathbf{A}' = 1 + 2 \cos \alpha \quad (13.26)$$

y recordando que la traza es invariante ante una transformación de similaridad tenemos que

$$\begin{aligned} \text{Tr}\mathbf{A}' &= \text{Tr}\mathbf{A} \Rightarrow \\ \sum_{i=1}^3 a_{ii} &= 1 + 2 \cos \alpha \end{aligned} \quad (13.27)$$

siendo  $a_{ii}$  los elementos diagonales (conocidos) de la matriz  $\mathbf{A}$ . Por tanto,  $\alpha$  se puede despejar en términos de estos elementos. Por otra parte, asumamos otra transformación de similaridad que nos lleve a una forma diagonal equivalente  $\boldsymbol{\lambda}$  (ver Ecs. 12.42, 12.43). De nuevo usando la invarianza de la traza y la estructura de valores propios (12.75)

$$\text{Tr}\mathbf{A} = \text{Tr}\boldsymbol{\lambda} = 1 + e^{i\Phi} + e^{-i\Phi} = 1 + 2 \cos \Phi \quad (13.28)$$

e igualando (13.27) con (13.28) resulta

$$\begin{aligned} 1 + 2 \cos \Phi &= 1 + 2 \cos \alpha \\ \alpha &= \pm \Phi \end{aligned} \quad (13.29)$$

de modo que el ángulo de rotación es igual a una de las fases complejas asociadas a los valores propios.

<sup>3</sup>En la práctica es más sencillo dar un valor fijo a una de las componentes del autovector. Por ejemplo, podemos hacer  $x_3 = 1$  y resolver para  $x_1$  y  $x_2$  con dos de las ecuaciones linealmente independientes, para finalmente normalizar. Es posible que no se pueda encontrar una solución consistente, en cuyo caso debemos hacer  $x_3 = 0$ , para obtener la solución correcta. Finalmente, es fácil ver de las Ecs. (12.37), que si  $x_1, x_2, x_3$  es una solución, también lo será  $\alpha x_1, \alpha x_2, \alpha x_3$ ; en particular cuando  $\alpha = -1$ .

Ya hemos mencionado que si  $\mathbf{x}_k$  es un autovector asociado a  $\lambda_k$  entonces  $\alpha\mathbf{x}_k$  también lo es. De modo que no solo es ambigua la magnitud del vector sino también su sentido ( $\alpha$  puede ser positivo o negativo). Similarmente, la solución del problema de valores propios también nos da una ambigüedad en el sentido de giro del eje de rotación según se vé en (13.29). La ambigüedad también se refleja en la estructura de valores propios de una matriz ortogonal propia  $3 \times 3$  la cual es de la forma  $\lambda_k = 1$ ,  $e^{\pm i\Phi}$  es decir invariante ante un cambio de signo de la fase  $\Phi$ .

La anterior ambigüedad es a su vez consistente con el hecho de que la solución de valores propios no fija unívocamente a la matriz ortogonal  $\mathbf{A}$ . Para verlo podemos observar que la ecuación secular para el determinante (12.38) nos muestra que la matriz traspuesta (y por tanto la inversa en una matriz ortogonal) tiene los mismos autovalores que  $\mathbf{A}$ . Adicionalmente, usando el carácter ortogonal de  $\mathbf{A}$  se puede ver que

$$\begin{aligned} \mathbf{A}\mathbf{x}_k &= \lambda_k\mathbf{x}_k \Leftrightarrow \tilde{\mathbf{A}}\mathbf{A}\mathbf{x}_k = \lambda_k\tilde{\mathbf{A}}\mathbf{x}_k \Leftrightarrow \mathbf{x}_k = \lambda_k\tilde{\mathbf{A}}\mathbf{x}_k \Rightarrow \\ \mathbf{A}\mathbf{x}_k &= \lambda_k\mathbf{x}_k \Leftrightarrow \tilde{\mathbf{A}}\mathbf{x}_k = \lambda_k^{-1}\mathbf{x}_k \end{aligned} \quad (13.30)$$

donde hemos usado la no degeneración de  $\lambda_k$  para rotaciones no triviales. La Ec. (13.30) muestra que para una matriz ortogonal  $\mathbf{A}$ ,  $\mathbf{x}_k$  es un vector propio de  $\mathbf{A}$  con valor propio  $\lambda_k$  si y solo si  $\mathbf{x}_k$  es vector propio de  $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$  con valor propio  $\lambda_k^{-1}$ . Esto es lógico geoméricamente<sup>4</sup>, ya que si  $\lambda_k = 1$ , entonces  $\mathbf{x}_k$  va en la dirección del eje de rotación y por tanto queda invariante tanto bajo  $\mathbf{A}$  como bajo  $\mathbf{A}^{-1}$ . Si  $\lambda_k = e^{i\Phi}$  la rotación inversa corresponde al mismo eje de rotación pero con el ángulo de rotación cambiado de signo es decir asociado a  $e^{-i\Phi} = \lambda_k^{-1}$ .

Nótese sin embargo, que el vector propio  $\mathbf{x}_k$  asociado a  $\lambda_k = e^{\pm i\Phi}$  con  $\pm\Phi \neq 0$ , no es necesariamente un vector real. En consecuencia, la transformación que lleva a la matriz  $\mathbf{A}$  a su forma canónica no es necesariamente ortogonal real, sino que será en general una matriz unitaria  $\mathbf{U}$

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \boldsymbol{\lambda} \equiv \begin{pmatrix} e^{i\Phi} & 0 & 0 \\ 0 & e^{-i\Phi} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.31)$$

donde  $\mathbf{U}$  se forma con la yuxtaposición de los vectores propios debidamente ortonormalizados<sup>5</sup>. Hemos usado la convención de que el eje  $X'_3$  de los ejes principales sea el eje de rotación (el único vector propio que se puede garantizar como real es el asociado al eje de rotación y a  $\lambda_k = 1$ ). Si sacamos el inverso a ambos lados de (13.31), obtenemos

$$\mathbf{U}^{-1}\tilde{\mathbf{A}}\mathbf{U} = \boldsymbol{\lambda}^{-1} \equiv \begin{pmatrix} e^{-i\Phi} & 0 & 0 \\ 0 & e^{i\Phi} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.32)$$

de modo que la misma transformación de similaridad que diagonaliza a  $\mathbf{A}$  diagonaliza a  $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$ , y la matriz diagonal obtenida es la que se obtiene de  $\mathbf{A}$  intercambiando el signo de los ángulos.

Aún así persiste la ambigüedad del orden en el que aparecen los valores propios en  $X'_1$  y en  $X'_2$ . Las ambigüedades pueden ser al menos aminoradas asumiendo que  $\Phi$  se asocia a  $\mathbf{A}$  y  $-\Phi$  se asocia a  $\mathbf{A}^{-1}$ , y fijando el sentido de los ejes de rotación por la regla de la mano derecha. En otras palabras, asociando la forma canónica de  $\mathbf{A}$  a la matriz diagonal descrita por la Ec. (13.31), de modo que la forma canónica de  $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$  es la matriz diagonal asociada a la Ec. (13.32).

Como corolario del teorema de Euler aparece el **Teorema de Chasles**: El desplazamiento más general de un cuerpo rígido consiste en una traslación mas una rotación. Esto se vé simplemente teniendo en cuenta que si el punto de referencia ya no es fijo entonces podemos descomponer el movimiento visto por  $S$  como la traslación pura de  $S_0$  mas la rotación pura de  $S'$  con respecto a  $S_0$ . En realidad, Chasles demuestra una versión más fuerte que dice que es posible escoger el origen de  $S'$  de modo que la traslación esté en la misma dirección que el eje de rotación.

<sup>4</sup>Si  $\{\lambda_k\}$  es el conjunto de valores propios de  $\mathbf{A}$  entonces  $\{\lambda_k^{-1}\}$  es el conjunto de valores propios de  $\tilde{\mathbf{A}} = \mathbf{A}^{-1}$ . Pero dado que  $\{\lambda_k\} = \{1, e^{i\Phi}, e^{-i\Phi}\}$ , es claro que ambos conjuntos de valores propios son iguales, en concordancia con la observación de que  $\tilde{\mathbf{A}}$  posee los mismos valores propios de  $\mathbf{A}$ .

<sup>5</sup>Las matrices ortogonales reales son normales, de modo que el teorema espectral garantiza que se pueden diagonalizar con una transformación de similaridad unitaria (ver sección 12.12).

### 13.7. Rotaciones finitas

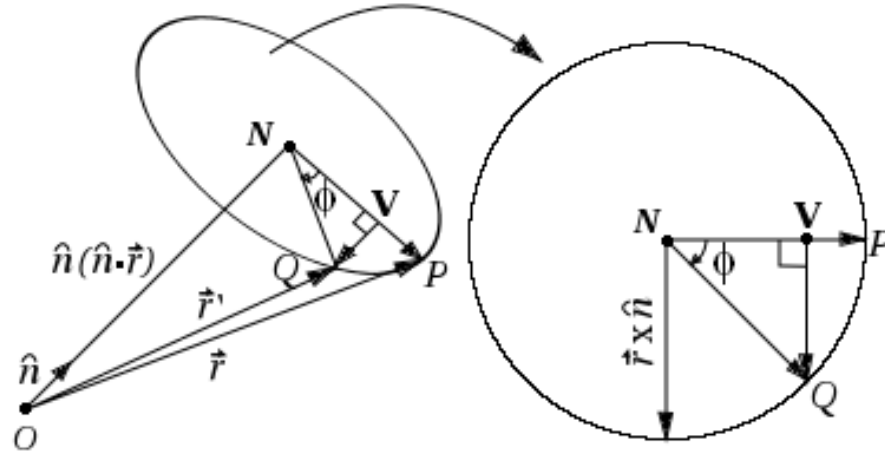


Figura 13.4: Rotación finita en un ángulo  $\Phi$ . El vector unitario  $\mathbf{n}$  determina el eje de rotación en el sentido horario (regla de la mano izquierda).

Hemos trabajado hasta aquí dos parametrizaciones para determinar la orientación relativa de un sistema coordinado. Por un lado los ángulos de Euler me permiten construir cualquier orientación relativa de  $S'$  a partir de la orientación de  $S_0$  mediante tres rotaciones sucesivas. Por otro lado, el teorema de Euler nos dice que la transformación de coordenadas necesaria para cualquier cambio de orientación se puede llevar a cabo con una sola rotación en torno a una dirección adecuada. Es natural entonces tratar de construir una representación en términos de los grados de libertad de dicha rotación como son el ángulo de rotación y los cosenos directores del eje (probablemente con condición de unitariedad del vector director).

Es más conveniente en este caso utilizar un punto de vista activo de la rotación de modo que el sistema coordinado permanece inalterado y se hace una rotación de los vectores. Hay que recordar que una rotación antihoraria para el sistema coordinado corresponde a una rotación horaria para los vectores.

En la figura 13.4, el vector  $\overrightarrow{OP}$  es la posición inicial  $\mathbf{r}$  del vector y la posición final  $\mathbf{r}'$  será  $\overrightarrow{OQ}$ . El vector unitario  $\mathbf{n}$  define la dirección del eje de rotación y su sentido sigue la regla de la mano izquierda, los vectores  $\overrightarrow{NP}$  y  $\overrightarrow{NQ}$  son perpendiculares al eje de rotación y pasan por él. Es claro que la distancia  $ON$  tiene valor  $\mathbf{n} \cdot \mathbf{r}$ , por lo tanto podemos escribir

$$\overrightarrow{ON} = \mathbf{n} (\mathbf{n} \cdot \mathbf{r}) \quad (13.33)$$

Del diagrama también se observa que

$$\begin{aligned} \overrightarrow{ON} + \overrightarrow{NP} &= \overrightarrow{OP} \Rightarrow \overrightarrow{NP} = \overrightarrow{OP} - \overrightarrow{ON} \\ \overrightarrow{NP} &= \mathbf{r} - \mathbf{n} (\mathbf{n} \cdot \mathbf{r}) \end{aligned} \quad (13.34)$$

y así mismo se observa que

$$\|\overrightarrow{NP}\| = \|\overrightarrow{NQ}\| = \|\mathbf{r} \times \mathbf{n}\| \quad (13.35)$$

estas magnitudes son iguales al radio de la circunferencia en la figura. La figura 13.4 muestra a la derecha el diagrama de vectores en el plano  $NQP$  perpendicular al eje de rotación. Con estas relaciones ya podemos escribir la relación entre  $\mathbf{r}$  y  $\mathbf{r}'$  en función de  $\mathbf{n}$  y  $\Phi$ . De este diagrama es fácil ver que

$$\mathbf{r}' = \overrightarrow{OQ} = \overrightarrow{ON} + \overrightarrow{NV} + \overrightarrow{VQ} \quad (13.36)$$

siendo el vector  $\overrightarrow{VQ}$  perpendicular a  $\overrightarrow{NP}$ . El vector  $\mathbf{r} \times \mathbf{n}$  es perpendicular al plano  $ONP$ . Por otro lado la figura también muestra que

$$\overrightarrow{NV} = \|\overrightarrow{NV}\| \frac{\overrightarrow{NP}}{\|\overrightarrow{NP}\|} = \|\overrightarrow{NQ}\| \cos \Phi \frac{\overrightarrow{NP}}{\|\overrightarrow{NP}\|} = \|\overrightarrow{NP}\| \cos \Phi \frac{\overrightarrow{NP}}{\|\overrightarrow{NP}\|} = \overrightarrow{NP} \cos \Phi$$

donde hemos usado (13.35). Ahora usando (13.34) resulta

$$\overrightarrow{NV} = [\mathbf{r} - \mathbf{n}(\mathbf{n} \cdot \mathbf{r})] \cos \Phi \quad (13.37)$$

similarmente escribamos el vector  $\overrightarrow{VQ}$  usando (13.35)

$$\begin{aligned} \overrightarrow{VQ} &= \|\overrightarrow{VQ}\| \frac{(\mathbf{r} \times \mathbf{n})}{\|\mathbf{r} \times \mathbf{n}\|} = \|\overrightarrow{NQ}\| \sin \Phi \frac{(\mathbf{r} \times \mathbf{n})}{\|\overrightarrow{NQ}\|} \\ \overrightarrow{VQ} &= (\mathbf{r} \times \mathbf{n}) \sin \Phi \end{aligned} \quad (13.38)$$

reemplazando (13.33), (13.37) y (13.38) en (13.36) resulta

$$\mathbf{r}' = \mathbf{n}(\mathbf{n} \cdot \mathbf{r}) + [\mathbf{r} - \mathbf{n}(\mathbf{n} \cdot \mathbf{r})] \cos \Phi + (\mathbf{r} \times \mathbf{n}) \sin \Phi$$

ahora simplemente reagrupamos términos para escribir el resultado final

$$\mathbf{r}' = \mathbf{r} \cos \Phi + \mathbf{n}(\mathbf{n} \cdot \mathbf{r}) [1 - \cos \Phi] + (\mathbf{r} \times \mathbf{n}) \sin \Phi \quad ; \quad \text{regla de la mano izquierda} \quad (13.39)$$

a la ecuación (13.39) se le conoce como **fórmula de rotación**. Esta ecuación es válida para cualquier rotación de cualquier magnitud, y por tanto constituye una versión para la rotación finita (en sentido horario, es decir con el vector unitario definido por la regla de la mano izquierda) de la expresión que se encontró en la página 74, Ec. (5.9) para rotaciones infinitesimales en un vector.

### 13.7.1. Forma matricial de la fórmula de rotación

La fórmula (13.39) para la rotación se puede colocar en forma más adecuada para una formulación matricial introduciendo un escalar  $e_0$  y un vector de componentes  $e_1, e_2, e_3$  en la forma

$$e_0 = \cos \frac{\Phi}{2} \quad ; \quad \mathbf{e} = \mathbf{n} \sin \frac{\Phi}{2} \quad (13.40)$$

es evidente que estos parámetros están relacionados por

$$e_0^2 + \mathbf{e}^2 = e_0^2 + e_1^2 + e_2^2 + e_3^2 = 1 \quad (13.41)$$

usando la identidad

$$\begin{aligned} \cos^2 \frac{\Phi}{2} - \sin^2 \frac{\Phi}{2} &= \cos \Phi = e_0^2 - \mathbf{e}^2 \\ \cos \Phi &= e_0^2 - e_1^2 - e_2^2 - e_3^2 \end{aligned} \quad (13.42)$$

y de la identidad

$$\begin{aligned} 2 \sin \frac{\Phi}{2} \cos \frac{\Phi}{2} &= \sin \Phi = 2e_0 \mathbf{e} \\ \mathbf{n} \sin \Phi &= 2e_0 \mathbf{e} \end{aligned} \quad (13.43)$$



usando las identidades (13.40), (13.41), (13.42), (13.43) en la Ec. (13.39) resulta

$$\begin{aligned}
 \mathbf{r}' &= \mathbf{r} (e_0^2 - e_1^2 - e_2^2 - e_3^2) + \mathbf{n} (\mathbf{n} \cdot \mathbf{r}) \left[ 2 \sin^2 \frac{\Phi}{2} \right] + 2 (\mathbf{r} \times \mathbf{e}) e_0 \\
 \mathbf{r}' &= \mathbf{r} (e_0^2 - e_1^2 - e_2^2 - e_3^2) + 2 \mathbf{n} \sin \frac{\Phi}{2} \left( \mathbf{n} \sin \frac{\Phi}{2} \cdot \mathbf{r} \right) + 2 (\mathbf{r} \times \mathbf{e}) e_0 \\
 \mathbf{r}' &= \mathbf{r} (e_0^2 - e_k e_k) + 2 \mathbf{e} (\mathbf{e} \cdot \mathbf{r}) + 2 (\mathbf{r} \times \mathbf{e}) e_0
 \end{aligned} \tag{13.44}$$

esta ecuación establece una relación lineal entre  $\mathbf{r}'$  y  $\mathbf{r}$ , en una forma vectorial. Para construir la matriz ortogonal  $\mathbf{A}$  en estos parámetros, se procede a escribir la ecuación para cada componente  $x'_i$  en términos de una combinación lineal de los  $x_j$ , los coeficientes de la combinación nos dan la matriz  $\mathbf{A}$ . Por ejemplo, para una componente primada se tiene

$$\begin{aligned}
 x'_i &= x_i (e_0^2 - e_k e_k) + 2 e_i (e_j x_j) + 2 \varepsilon_{ijk} e_0 x_j e_k \\
 x'_i &= x_j \delta_{ij} (e_0^2 - e_k e_k) + 2 e_i e_j x_j + 2 \varepsilon_{ijk} e_0 e_k x_j \\
 x'_i &= [\delta_{ij} (e_0^2 - e_k e_k) + 2 e_i e_j + 2 \varepsilon_{ijk} e_0 e_k] x_j \equiv a_{ij} x_j
 \end{aligned}$$

la matriz asociada es entonces

$$a_{ij} = \delta_{ij} (e_0^2 - e_k e_k) + 2 e_i e_j + 2 \varepsilon_{ijk} e_0 e_k$$

que explícitamente se escribe

$$\mathbf{A} = \begin{pmatrix} (e_0^2 + e_1^2 - e_2^2 - e_3^2) & 2(e_1 e_2 + e_0 e_3) & 2(e_1 e_3 - e_0 e_2) \\ 2(e_1 e_2 - e_0 e_3) & (e_0^2 - e_1^2 + e_2^2 - e_3^2) & 2(e_2 e_3 + e_0 e_1) \\ 2(e_1 e_3 + e_0 e_2) & 2(e_2 e_3 - e_0 e_1) & (e_0^2 - e_1^2 - e_2^2 + e_3^2) \end{pmatrix} \tag{13.45}$$

a manera de consistencia veremos que la traza de la matriz (13.45) es idéntica a la traza expresada en la Ec. (13.28), puesto que la traza es invariante ante un cambio de base.

$$\begin{aligned}
 Tr \mathbf{A} &= 3e_0^2 - e_1^2 - e_2^2 - e_3^2 = 3 \cos^2 \frac{\Phi}{2} - \sin^2 \frac{\Phi}{2} = 2 \cos^2 \frac{\Phi}{2} + \left( \cos^2 \frac{\Phi}{2} - \sin^2 \frac{\Phi}{2} \right) \\
 &= (1 + \cos \Phi) + \cos \Phi = 1 + 2 \cos \Phi
 \end{aligned}$$

que coincide con (13.28). Nótese la similitud entre los parámetros reales de Euler definidos en la sección 13.5, por medio de las Ecuaciones (13.21, 13.22) y los definidos por las Ecs. (13.40, 13.41). Compárese también las matrices de rotación definidas por las Ecs. (13.23) y (13.45).

### 13.7.2. Relación entre la parametrización eje-ángulo $(\mathbf{n}, \Phi)$ y los ángulos de Euler

La relación entre los ángulos de Euler y el ángulo de rotación  $\Phi$  alrededor del eje de rotación, se puede obtener igualando las trazas de la matrices en (13.19) y (13.45). La traza de (13.19) se escribe como

$$\begin{aligned}
 Tr_1 \mathbf{A} &= \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi - \sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi + \cos \theta \\
 &= (1 + \cos \theta) \cos \psi \cos \phi - (1 + \cos \theta) \sin \phi \sin \psi + \cos \theta \\
 &= (1 + \cos \theta) (\cos \psi \cos \phi - \sin \phi \sin \psi) + \cos \theta \\
 &= (1 + \cos \theta) \cos (\phi + \psi) + \cos \theta \\
 &= \left( 2 \cos^2 \frac{\theta}{2} \right) \left[ \cos^2 \left( \frac{\phi + \psi}{2} \right) - \sin^2 \left( \frac{\phi + \psi}{2} \right) \right] + \left( \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right)
 \end{aligned}$$

$$\begin{aligned}
Tr_1 \mathbf{A} &= \left(2 \cos^2 \frac{\theta}{2}\right) \left[2 \cos^2 \left(\frac{\phi + \psi}{2}\right) - \cos^2 \left(\frac{\phi + \psi}{2}\right) - \sin^2 \left(\frac{\phi + \psi}{2}\right)\right] + \left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}\right) \\
&= \left(2 \cos^2 \frac{\theta}{2}\right) \left[2 \cos^2 \left(\frac{\phi + \psi}{2}\right) - 1\right] + \left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}\right) \\
&= \left(4 \cos^2 \frac{\theta}{2}\right) \cos^2 \left(\frac{\phi + \psi}{2}\right) - \left(2 \cos^2 \frac{\theta}{2}\right) + \left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}\right) \\
&= 4 \cos^2 \left(\frac{\phi + \psi}{2}\right) \cos^2 \frac{\theta}{2} - \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \\
&= 4 \cos^2 \left(\frac{\phi + \psi}{2}\right) \cos^2 \frac{\theta}{2} - 1
\end{aligned}$$

en tanto que la traza de (13.45) se escribe como

$$\begin{aligned}
Tr_2 \mathbf{A} &= 3e_0^2 - e_1^2 - e_2^2 - e_3^2 = 3 \cos^2 \frac{\Phi}{2} - \sin^2 \frac{\Phi}{2} \\
Tr_2 \mathbf{A} &= 3 \cos^2 \frac{\Phi}{2} - \left(1 - \cos^2 \frac{\Phi}{2}\right) = 4 \cos^2 \frac{\Phi}{2} - 1
\end{aligned}$$

igualando ambas trazas se obtiene

$$\begin{aligned}
4 \cos^2 \frac{\Phi}{2} - 1 &= 4 \cos^2 \left(\frac{\phi + \psi}{2}\right) \cos^2 \frac{\theta}{2} - 1 \\
\cos \left(\frac{\Phi}{2}\right) &= \cos \left(\frac{\phi + \psi}{2}\right) \cos \left(\frac{\theta}{2}\right)
\end{aligned} \tag{13.46}$$

### 13.8. Rotaciones infinitesimales

A lo largo de este capítulo hemos aprendido que para generar rotaciones tridimensionales solo se requieren tres elementos independientes, con lo cual una representación matricial 3x3 necesariamente contiene muchos elementos redundantes y se han tenido que agregar ligaduras que garanticen el número correcto de grados de libertad. El hecho de que las componentes independientes sean tres y que hayamos podido encontrar un eje de rotación que permita realizar el cambio de orientación con una sola rotación, nos conduce a la posibilidad de construir un vector que denota a tal transformación. Una posibilidad mas bien evidente sería elegir un vector a lo largo del eje de rotación cuya magnitud sea una función del ángulo de rotación. Veremos sin embargo, que esta construcción no es consistente: si realizamos dos transformaciones sucesivas asociadas a vectores  $\mathbf{A}$  y  $\mathbf{B}$  la composición de ellos debe hacerse por la suma de éstos la cual es conmutativa

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

hemos visto que la composición de dos matrices en general no es conmutativa y esto es válido en particular para las matrices ortogonales. Es fácil mostrar un ejemplo físico en el cual las rotaciones sucesivas no conmutan, la Fig. 13.5 muestra dos rotaciones sucesivas de un bloque realizadas en diferente orden: en el primer caso se comienza con una rotación antihoraria alrededor de  $x$  con  $\theta_x = \pi/2$  seguida por otra rotación antihoraria de  $\pi/2$  alrededor de  $y$ . En el segundo caso se invierte el orden de las rotaciones y se observa una marcada diferencia entre las dos configuraciones.

En consecuencia, si un vector representara satisfactoriamente a una rotación, y la adición de vectores me representara la composición de rotaciones, dicha operación de adición vectorial no debería ser conmutativa, pero esta es una propiedad fundamental de un espacio vectorial. Luego, las rotaciones finitas no pueden ser representadas por vectores.

Veremos a continuación que las transformaciones infinitesimales conmutan a primer orden y que por tanto pueden ser representadas por un vector. Visto de manera pasiva, consideraremos un cambio de orientación de los ejes coordenados en la cual los ejes  $X'_1, X'_2, X'_3$  solo difieren infinitesimalmente de  $X_1, X_2, X_3$  respectivamente.

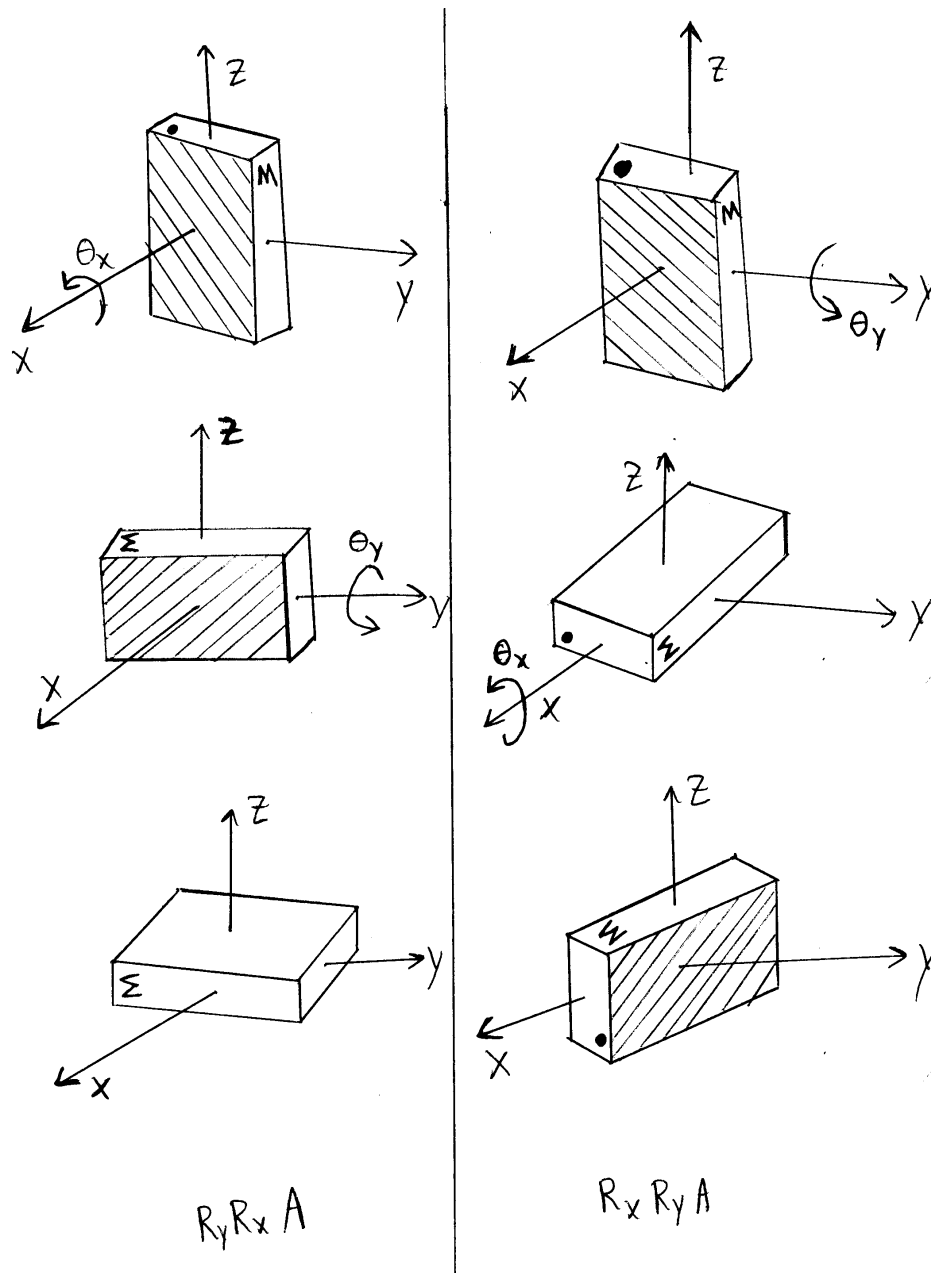


Figura 13.5: Dos rotaciones finitas sucesivas realizadas en diferente orden: Este ejemplo ilustra la no conmutatividad de las operaciones finitas de rotación.

La transformación solo difiere de la identidad por cantidades infinitesimales. Las nuevas componentes  $x'_i$  de un vector vendrán dadas por

$$x'_i = x_i + \varepsilon_{ij} x_j$$

donde los parámetros  $\varepsilon_{ij}$  son infinitesimales de modo que para efectos de cálculo solo se conservarán términos de primer orden en tales parámetros. La relación anterior se escribe

$$x'_i = (\delta_{ij} + \varepsilon_{ij}) x_j$$

que en notación matricial adquiere la forma:

$$\mathbf{r}' = (\mathbf{1} + \boldsymbol{\varepsilon}) \mathbf{r} \tag{13.47}$$

que manifiesta el hecho ya mencionado de que la matriz de transformación solo difiere infinitesimalmente de la identidad. Veamos ahora como actúa la composición de dos transformaciones infinitesimales

$$\mathbf{A}_1 \mathbf{A}_2 = (\mathbf{1} + \boldsymbol{\varepsilon}_1)(\mathbf{1} + \boldsymbol{\varepsilon}_2) = \mathbf{1} + \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_2$$

sin embargo, considerando solo términos de primer orden

$$\mathbf{A}_1 \mathbf{A}_2 = \mathbf{1} + \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 \quad (13.48)$$

esta operación es claramente conmutativa (solo el término de segundo orden que se despreció es no conmutativo). En consecuencia, es posible representar estas operaciones infinitesimales mediante vectores.

Es muy sencillo obtener la matriz inversa de una transformación infinitesimal si tenemos en cuenta la regla de composición a primer orden dada por (13.48), en esta composición se obtiene la identidad si y solo si  $\boldsymbol{\varepsilon}_1 = -\boldsymbol{\varepsilon}_2$ , por lo tanto la inversa se escribiría con la siguiente asociación

$$\mathbf{A} = (\mathbf{1} + \boldsymbol{\varepsilon}) \Leftrightarrow \mathbf{A}^{-1} = (\mathbf{1} - \boldsymbol{\varepsilon}) \quad (13.49)$$

nótese que esto se asemeja a la relación escalar de primer orden  $(1+x)^{-1} \cong 1-x$  para  $x \ll 1$ . Recordando que la condición de ortogonalidad equivale a que la inversa sea idéntica a la traspuesta, se obtiene la siguiente relación<sup>6</sup>

$$\begin{aligned} \mathbf{1} + \tilde{\boldsymbol{\varepsilon}} &= \mathbf{1} - \boldsymbol{\varepsilon} \\ \tilde{\boldsymbol{\varepsilon}} &= -\boldsymbol{\varepsilon} \end{aligned}$$

de modo que la matriz infinitesimal  $\boldsymbol{\varepsilon}$  debe ser antisimétrica a fin de que  $\mathbf{A}$  permanezca ortogonal real. Dada la antisimetría de  $\boldsymbol{\varepsilon}$ , los elementos diagonales de dicha matriz son nulos y tenemos tres grados de libertad como corresponde a toda matriz antisimétrica real  $3 \times 3$ . Esto a su vez implica que la diagonal de  $\mathbf{A}$  sigue teniendo elementos unidad y al expandir el determinante se observa que este continúa siendo +1 a primer orden. Naturalmente, no era de esperarse un cambio brusco del determinante a  $-1$  ya que estas transformaciones son continuas e infinitesimales. Visto de otro modo, no era de esperarse que ocurriera una transformación discreta de inversión cuando lo que tenemos son parámetros continuos infinitesimales. En conclusión estas transformaciones ortogonales son propias y pueden asociarse a una rotación.

Ya hemos visto que una matriz antisimétrica se puede parametrizar como en la ecuación (12.57). Aplicando esta parametrización para  $\boldsymbol{\varepsilon}$  escribimos

$$\boldsymbol{\varepsilon} = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \quad (13.50)$$

donde la notación enfatiza el carácter infinitesimal de los parámetros. Teniendo en cuenta (13.47), el cambio infinitesimal en las componentes de un vector se escribe como

$$\mathbf{r}' - \mathbf{r} \equiv d\mathbf{r} = \boldsymbol{\varepsilon} \mathbf{r} \quad (13.51)$$

Aplicando además las Ecs. (12.57, 12.58, 12.59, 12.60) vemos que un diferencial del vector  $\mathbf{r}$  se escribe

$$d\mathbf{r} = \mathbf{r} \times d\boldsymbol{\Omega} \quad ; \quad d\boldsymbol{\Omega} \equiv (d\Omega_1, d\Omega_2, d\Omega_3)^T \quad (13.52)$$

Esta ecuación nos dice que el cambio infinitesimal del vector  $\mathbf{r}$  es perpendicular a  $\mathbf{r}$  mismo y al vector  $d\boldsymbol{\Omega}$  el cual como veremos más adelante va a lo largo del eje de rotación. Lo primero se puede ver por el hecho de que si hubiera una componente de  $d\mathbf{r}$  a lo largo de  $\mathbf{r}$  implicaría un cambio de magnitud de  $\mathbf{r}$  que no puede ser producido por una transformación ortogonal. Ahora bien, ya hemos visto que el vector asociado a una matriz

<sup>6</sup>Lo discutido hasta aquí es válido para cualquier transformación lineal infinitesimal, ya que no se ha usado la relación de ortogonalidad en ninguna forma.

antisimétrica es un vector axial (ver sección 12.6). Nótese que la ecuación (13.52) es consistente con este hecho, ya que  $\mathbf{r}$  y  $d\mathbf{r}$  son vectores polares, por lo tanto es necesario que  $d\boldsymbol{\Omega}$  sea axial para que su producto cruz con un vector polar sea polar. Como todo vector axial,  $d\boldsymbol{\Omega}$  posee las propiedades de transformación dadas por (12.69)

$$d\Omega'_i = (\det \mathbf{B}) b_{ij} d\Omega_j$$

la forma explícita de  $d\boldsymbol{\Omega}$  en términos de los parámetros de rotación se puede encontrar tomando la fórmula de rotación Ec. (13.39) y aplicándola a un ángulo infinitesimal para el cual  $\cos d\Phi \rightarrow 1$ ,  $\sin d\Phi \rightarrow d\Phi$ , resulta

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + (\mathbf{r} \times \mathbf{n}) d\Phi \Rightarrow \\ \mathbf{r}' - \mathbf{r} &= d\mathbf{r} = (\mathbf{r} \times \mathbf{n}) d\Phi \end{aligned} \quad (13.53)$$

comparando esta expresión con (13.52) obtenemos que

$$d\boldsymbol{\Omega} = \mathbf{n} d\Phi \quad (13.54)$$

Nótese que hemos obtenido lo que esperábamos: el vector  $d\boldsymbol{\Omega}$  que caracteriza a la rotación infinitesimal va en la dirección del eje de rotación y su magnitud es el ángulo de rotación infinitesimal  $d\Phi$ . Vale la pena anotar que este vector diferencial no proviene de ninguna diferencia entre vectores finitos, ya que las rotaciones finitas no pueden ser asociadas a un vector como ya se vió. Geométricamente, una inversión pasiva nos cambia de un sistema dextrógiro a uno levógiro para definir el sentido de  $\mathbf{n}$  con lo cual dicho vector será axial, y dado que  $d\boldsymbol{\Omega}$  también es axial, se deduce que  $d\Phi$  debe ser un escalar (y no un pseudoescalar). La relación (13.54) también se puede encontrar muy fácilmente considerando un ángulo infinitesimal de rotación en la figura 13.6. El signo de  $d\Phi$  se considera positivo en la dirección horaria es decir con convención quiral izquierda.

### 13.9. Rotaciones finitas e infinitesimales en convención quiral derecha

Debemos tener en cuenta que las interpretaciones pasiva o activa de las transformaciones se han tomado a conveniencia. Originalmente se tomó la interpretación pasiva que nos permite orientar los ejes coordenados que se mueven con el cuerpo rígido. Vimos además que una interpretación activa de la rotación de un vector en un sistema coordenado fijo implica una rotación en el sentido contrario i.e. en el sentido horario. Sin embargo, es más usual considerar tales transformaciones activas en la convención antihoraria (en tal caso  $\mathbf{n}$  y  $d\boldsymbol{\Omega}$  sí estarían definidos por la regla de la mano derecha). Por esta razón, haremos un resumen de las fórmulas útiles encontradas pero en la convención de sentido antihorario de rotación. **De aquí de adelante se usarán estas fórmulas en esta convención antihoraria es decir con regla de la mano derecha**

La fórmula de rotación (13.39) en convención antihoraria es

$$\mathbf{r}' = \mathbf{r} \cos \Phi + \mathbf{n} (\mathbf{n} \cdot \mathbf{r}) [1 - \cos \Phi] + (\mathbf{n} \times \mathbf{r}) \sin \Phi \quad ; \quad \text{regla de la mano derecha} \quad (13.55)$$

la rotación infinitesimal correspondiente (13.53, 13.54) es

$$\mathbf{r}' - \mathbf{r} = d\mathbf{r} = (\mathbf{n} \times \mathbf{r}) d\Phi = d\boldsymbol{\Omega} \times \mathbf{r} \quad ; \quad \text{regla de la mano derecha} \quad (13.56)$$

la matriz antisimétrica de rotación (13.50) se escribirá como<sup>7</sup>

$$\begin{aligned} \boldsymbol{\varepsilon} &= \begin{pmatrix} 0 & -d\Omega_3 & d\Omega_2 \\ d\Omega_3 & 0 & -d\Omega_1 \\ -d\Omega_2 & d\Omega_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} d\Phi \quad ; \quad \text{regla de la mano derecha} \\ \boldsymbol{\varepsilon} &\equiv \mathbf{N} d\Phi \end{aligned} \quad (13.57)$$

<sup>7</sup>El cambio de signo en la Ec. (13.57) con respecto a la convención (12.63) se realiza con el fin de que la relación (12.65) se convierta en nuestro contexto en  $\boldsymbol{\varepsilon} \mathbf{r} = d\boldsymbol{\Omega} \times \mathbf{r}$ .

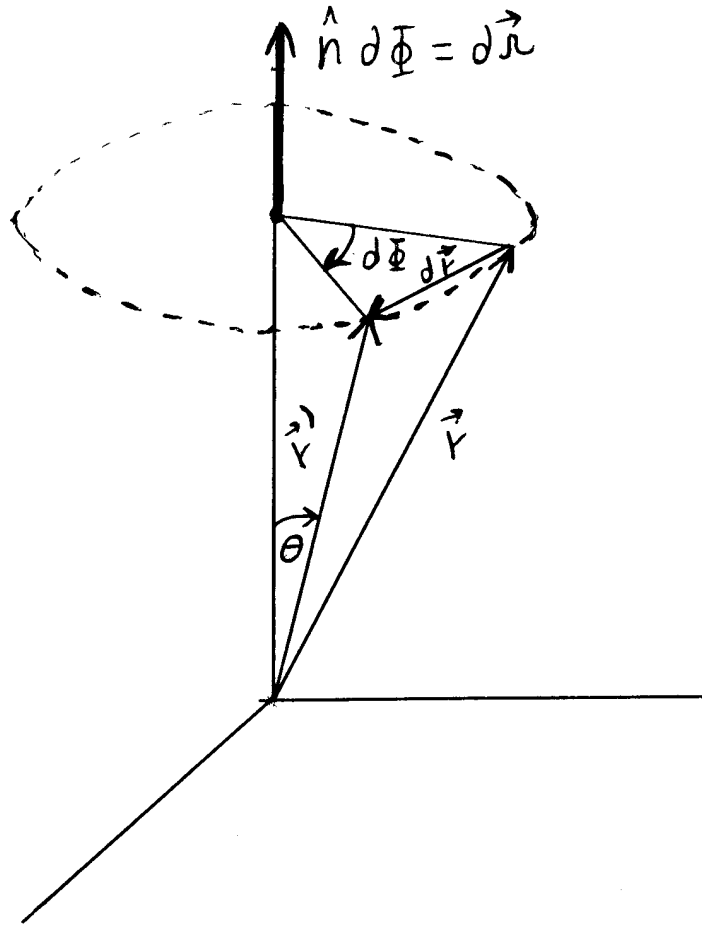


Figura 13.6: Ilustración de una rotación infinitesimal. Se usa regla de la mano izquierda para la determinación del vector unitario  $\mathbf{n}$  que determina al eje de rotación.

$n_i$  son las componentes del vector unitario  $\mathbf{n}$  que define al eje de rotación.  $d\mathbf{r}$  es el cambio infinitesimal desde  $\mathbf{r}$  hasta  $\mathbf{r}'$ . Teniendo en cuenta (13.51), así como (13.57) se tiene que

$$d\mathbf{r} = \boldsymbol{\varepsilon} \mathbf{r} = (\mathbf{N} d\Phi) \mathbf{r}$$

resultando entonces una ecuación diferencial matricial con respecto al ángulo de rotación

$$\frac{d\mathbf{r}}{d\Phi} = \mathbf{N} \mathbf{r} \tag{13.58}$$

los elementos de la matriz  $\mathbf{N}$  definida en (13.57), están dados por

$$N_{ij} = -\varepsilon_{ijk} n_k = \varepsilon_{jik} n_k \tag{13.59}$$

adicionalmente, la matriz  $\boldsymbol{\varepsilon}$  en (13.57) se puede representar de la siguiente forma

$$\boldsymbol{\varepsilon} = n_i \mathbf{M}_i d\Phi ; \mathbf{M}_1 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} ; \mathbf{M}_2 \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} ; \mathbf{M}_3 \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{13.60}$$

las matrices  $\mathbf{M}_i$  se conocen como los generadores de rotaciones infinitesimales y su combinación posee la siguiente propiedad

$$\mathbf{M}_i \mathbf{M}_j - \mathbf{M}_j \mathbf{M}_i \equiv [\mathbf{M}_i, \mathbf{M}_j] = \varepsilon_{ijk} \mathbf{M}_k \tag{13.61}$$

la diferencia entre los dos productos matriciales se denomina conmutador o corchete de Lie. La Ec. (13.61) define un álgebra de Lie del grupo de rotaciones parametrizado en términos del ángulo de rotación y el vector unitario a lo largo del eje de rotación.

### 13.9.1. Construcción de rotaciones finitas por integración de rotaciones infinitesimales

Veremos ahora como se puede construir un operador de rotación finita, a partir de infinitas rotaciones infinitesimales sucesivas que parten desde la identidad. El proceso es similar al realizado para las transformaciones canónicas sección 8.13 aunque aquí partiremos en el orden inverso, comenzando con transformaciones infinitesimales para construir la transformación finita<sup>8</sup>. Partiendo de un cierto eje de rotación el ángulo de rotación comienza en  $\Phi = 0$  hasta un valor final  $\Phi$ , podemos hacer la transformación en  $k$  etapas sucesivas y en cada una el ángulo aumenta en una cantidad  $\Delta\Phi = \Phi/k$ . Si  $k$  es suficientemente grande se puede considerar que cada una de las  $k$  rotaciones es infinitesimal de modo que la primera transformación es

$$\mathbf{r}(0) + \Delta\mathbf{r} = (\mathbf{1} + \varepsilon) \mathbf{r}(0) = (1 + n_i \mathbf{M}_i \Delta\Phi) \mathbf{r}(0) = \left[ 1 + (\mathbf{n} \cdot \mathbf{M}) \frac{\Phi}{k} \right] \mathbf{r}(0)$$

donde hemos usado (13.60). El “producto punto” definido por  $\mathbf{n} \cdot \mathbf{M}$  es solo una forma sintética para denotar la suma  $n_i \mathbf{M}_i$  la cual a su vez es una matriz. La segunda transformación es

$$\mathbf{r}(0) + \Delta\mathbf{r} + \Delta\mathbf{r} = (\mathbf{1} + \varepsilon)(\mathbf{1} + \varepsilon) \mathbf{r}(0) = \left[ 1 + (\mathbf{n} \cdot \mathbf{M}) \frac{\Phi}{k} \right]^2 \mathbf{r}(0)$$

y así sucesivamente se requieren  $k$  transformaciones para llegar al vector final

$$\mathbf{r}(0) + k\Delta\mathbf{r} = \mathbf{r}(\Phi) = \left[ 1 + (\mathbf{n} \cdot \mathbf{M}) \frac{\Phi}{k} \right]^k \mathbf{r}(0)$$

para que cada transformación sea realmente infinitesimal debe tomarse el límite cuando  $k \rightarrow \infty$

$$\mathbf{r}(\Phi) = \lim_{k \rightarrow \infty} \left[ 1 + (\mathbf{n} \cdot \mathbf{M}) \frac{\Phi}{k} \right]^k \mathbf{r}(0) \quad (13.62)$$

y recordando que en funciones ordinarias esta es precisamente una representación de la exponencial

$$e^x = \lim_{k \rightarrow \infty} \left( 1 + \frac{x}{k} \right)^k$$

se tiene entonces que la Ec. (13.62) se puede reescribir como

$$\mathbf{r}(\Phi) = e^{[(\mathbf{n} \cdot \mathbf{M}) \Phi]} \mathbf{r}(0) \quad (13.63)$$

es decir el operador de rotación  $\hat{\mathbf{O}}$  se puede escribir como una función de operadores

$$\mathbf{r}(\Phi) = \hat{\mathbf{O}} \mathbf{r}(0) \quad ; \quad \hat{\mathbf{O}} \equiv \exp [(\mathbf{n} \cdot \mathbf{M}) \Phi]$$

donde  $\mathbf{n}$  y  $\Phi$  se denominan parámetros de la rotación y las tres matrices  $\mathbf{M}$  definidas en (13.60) son los generadores. Nótese que las matrices  $\mathbf{M}$  son las mismas para cada rotación en tanto que los parámetros varían de una rotación a otra.

Hay por supuesto otras representaciones de los generadores  $\mathbf{M}$ , pero todas ellas deben cumplir el álgebra definida por los corchetes de Lie en (13.61). Nótese la similitud entre (13.63) y (8.69), la cual no es coincidental ya que ambas representan transformaciones continuas que se construyen a partir de transformaciones infinitesimales partiendo de la identidad. Nótese finalmente que (13.63) es solución de la ecuación diferencial (13.58) con  $\mathbf{N} = \mathbf{n} \cdot \mathbf{M}$ .

<sup>8</sup>Por supuesto, ambos procedimientos son posibles tanto para las transformaciones canónicas como para las rotaciones.

### 13.10. Velocidad angular en términos de los ángulos de Euler

Para propósitos futuros, es conveniente expresar el vector velocidad angular  $\boldsymbol{\omega}$  en términos de los ángulos de Euler y sus derivadas temporales. La rotación infinitesimal asociada a  $\boldsymbol{\omega}$  se puede construir como tres rotaciones infinitesimales sucesivas cada una asociada a velocidades angulares  $\omega_\phi = \dot{\phi}$ ,  $\omega_\theta = \dot{\theta}$ ,  $\omega_\psi = \dot{\psi}$ . Dadas las propiedades vectoriales de las rotaciones infinitesimales, el vector  $\boldsymbol{\omega}$  se puede obtener como una suma de cada uno de los tres vectores asociados a las rotaciones sucesivas. Sin embargo, las direcciones  $\boldsymbol{\omega}_\phi$ ,  $\boldsymbol{\omega}_\theta$  y  $\boldsymbol{\omega}_\psi$  no están dispuestas en forma simétrica.  $\boldsymbol{\omega}_\phi$  está dispuesta a lo largo del eje  $X_3$  del sistema sin rotar  $S_0$ ,  $\boldsymbol{\omega}_\theta$  está a lo largo del eje  $X_1^{(a)}$  (línea de nodos), y  $\boldsymbol{\omega}_\psi$  está a lo largo del eje  $X_3^{(b)}$ . Por esta razón, es más viable recurrir a las matrices  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  definidas en las Ecs. (13.15, 13.16, 13.17) para obtener las componentes de estos vectores a lo largo de cualquier conjunto de ejes coordenados.

El sistema más conveniente para discutir las ecuaciones de movimiento en un cuerpo rígido es el sistema  $S'$  fijo al cuerpo. Es importante enfatizar que  $\boldsymbol{\omega}$  es la velocidad angular medida por  $S_0$  pero se escribirá en la base de vectores unitarios de  $S'$  es decir

$$\boldsymbol{\omega} = \omega_i \mathbf{u}_i = \omega'_i \mathbf{u}'_i$$

de hecho la velocidad angular medida por  $S'$  es cero. No podemos entonces confundir un vector medido por  $S'$ , con un vector medido por  $S_0$  pero escrito en la base de  $S'$ .

Obtendremos entonces las componentes de  $\boldsymbol{\omega}$  con respecto a los ejes de  $S'$ , aunque omitiremos la notación primada para las componentes. Comencemos entonces con el vector  $\boldsymbol{\omega}_\phi$ . Dado que  $\boldsymbol{\omega}_\phi$  es paralelo al eje  $X_3$  del sistema  $S_0$ , podemos escribirlo fácilmente en la base de los ejes  $X_1 X_2 X_3$  en la forma  $(0, 0, \omega_\phi)^T$ . Para obtener sus componentes a lo largo de los ejes del cuerpo  $X'_1 X'_2 X'_3$  debemos en consecuencia transformar el vector de la base  $X_1 X_2 X_3$  a la base  $X'_1 X'_2 X'_3$  para lo cual hay que aplicar la transformación ortogonal completa  $\mathbf{A} = \mathbf{BCD}$ , Ecs. (13.18, 13.19)

$$(\boldsymbol{\omega}_\phi)_{S'} = \begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \psi \sin \theta \\ -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \cos \psi \sin \theta \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \omega_\phi \end{pmatrix}$$

resultando

$$(\boldsymbol{\omega}_\phi)_{X'_1} = \dot{\phi} \sin \theta \sin \psi \quad ; \quad (\boldsymbol{\omega}_\phi)_{X'_2} = \dot{\phi} \sin \theta \cos \psi \quad ; \quad (\boldsymbol{\omega}_\phi)_{X'_3} = \dot{\phi} \cos \theta$$

nótese que  $\dot{\phi}$  tiene la proyección  $\dot{\phi} \sin \theta$  en el plano  $X'_1 X'_2$  y es perpendicular a la línea de nodos.

Ahora debemos escribir las componentes de  $\boldsymbol{\omega}_\theta$  con respecto a los ejes  $X'_1 X'_2 X'_3$ . Dado que  $\boldsymbol{\omega}_\theta$  va a lo largo de la línea de nodos, la cual está definida por el eje  $X_1^{(a)} = X_1^{(b)}$ , vemos que el vector  $\boldsymbol{\omega}_\theta$  se puede escribir en el sistema coordenado  $X_1^{(b)} X_2^{(b)} X_3^{(b)}$  en la forma  $(\omega_\theta, 0, 0)^T$ . Para escribir sus componentes con respecto a los ejes de  $S'$  requerimos en consecuencia pasar del sistema  $X_1^{(b)} X_2^{(b)} X_3^{(b)}$  al sistema  $X'_1 X'_2 X'_3$  lo cual se obtiene aplicando únicamente la transformación ortogonal final  $\mathbf{B}$  definida en (13.17)

$$(\boldsymbol{\omega}_\theta)_{S'} = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_\theta \\ 0 \\ 0 \end{pmatrix}$$

con lo que se obtiene

$$(\boldsymbol{\omega}_\theta)_{X'_1} = \dot{\theta} \cos \psi \quad ; \quad (\boldsymbol{\omega}_\theta)_{X'_2} = -\dot{\theta} \sin \psi \quad ; \quad (\boldsymbol{\omega}_\theta)_{X'_3} = 0$$

no es necesaria ninguna transformación para obtener  $\boldsymbol{\omega}_\psi$ , el cual yace a lo largo de  $X_3^{(b)} = X'_3$  ( $\omega_\psi = \dot{\psi}$ ) de modo que en la base  $X'_1 X'_2 X'_3$  ya se escribe en la forma  $(0, 0, \dot{\psi})$ . Sumando las componentes de cada velocidad angular, se obtienen las componentes de  $\boldsymbol{\omega}$  con respecto a los ejes fijos al cuerpo i.e. de  $S'$ :

$$\omega_{X'_1} = \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \quad ; \quad \omega_{X'_2} = \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \quad ; \quad \omega_{X'_3} = \dot{\phi} \cos \theta + \dot{\psi} \quad (13.64)$$

con un procedimiento similar se puede encontrar las componentes de  $\boldsymbol{\omega}$  a lo largo de los ejes de  $S_0$  en términos de los ángulos de Euler.



## 13.11. Razón de cambio de un vector visto por sistemas rotantes

Consideremos la dinámica de un vector tal como el vector posición, momento angular etc. El vector puede estar cambiando en el tiempo de cierta manera. No obstante, el cambio observado dependerá en general del sistema coordenado elegido para medir dicho cambio. Definiremos sistemas de referencia  $S$ ,  $S_0$ ,  $S_{rot}$  en donde  $S$  es un sistema inercial,  $S_0$  un sistema con ejes paralelos a  $S$  y  $S_{rot}$  es un sistema rotante con respecto a  $S_0$  pero que tiene origen común con él.  $S_{rot}$  tiene por tanto un eje de rotación instantáneo  $\mathbf{n}(t)$ . Nos centraremos en la forma en que  $S_0$  y  $S_{rot}$  ven la razón de cambio de un cierto vector  $\mathbf{V}$ , ya que la transformación desde  $S_0$  hasta  $S$  es muy sencilla ( $S_0$  posee traslación pura con respecto a  $S$ ).

### 13.11.1. Razón de cambio por argumentos vectoriales

Es importante notar que para ciertos vectores que llamaremos “primitivos” ambos sistemas  $S_0$  y  $S_{rot}$  ven el mismo vector, solo que lo escriben en componentes diferentes como ya se ha mencionado de las transformaciones pasivas. El caso más importante de este tipo de vector es el vector posición teniendo en cuenta que ambos sistemas de referencia tienen origen común. Antes de hacer el formalismo general tomemos un ejemplo: Supongamos que el sistema  $S_{rot}$  está fijo a un cuerpo rígido de tal forma que el vector posición de un punto del cuerpo es claramente constante visto por  $S_{rot}$ . Es claro también que si el cuerpo rígido tiene cierta dinámica, dicho vector no será constante con respecto a  $S_0$ . Nótese entonces que  $S_0$  y  $S_{rot}$  miden el mismo vector en cada instante, digamos en  $\mathbf{r}(t)$  y en  $\mathbf{r}(t + \Delta t)$ . No obstante, ambos sistemas de referencia miden diferentes cambios en tal vector medidos en el mismo intervalo entre  $t$  y  $t + \Delta t$ , ya que es evidente que para  $S_{rot}$  el cambio del vector en cualquier intervalo es nulo, en tanto que el cambio visto por  $S_0$  es diferente de cero. Lo anterior nos lleva a concluir que aunque en ambos sistemas de referencia  $S_0$  y  $S_{rot}$  los vectores son los mismos en cada instante de tiempo, tales sistemas de referencia van a medir razones de cambio diferentes de dichos vectores, veamos ahora el formalismo general.

Tomemos un vector “primitivo” pero por lo demás arbitrario  $\mathbf{V}$ . En un instante  $t$  dicho vector se escribe en ambas bases de la forma

$$\mathbf{V} = V_i \mathbf{u}_i = V_i^{rot} \mathbf{u}_i^{rot} \quad (13.65)$$

un diferencial de este vector se escribirá en la forma

$$d\mathbf{V} = (dV_i) \mathbf{u}_i = (dV_i^{rot}) \mathbf{u}_i^{rot} + V_i^{rot} d\mathbf{u}_i^{rot} \quad (13.66)$$

donde hemos empleado la convención de que los vectores unitarios cartesianos asociados a  $S_0$  son constantes (es decir, son medidos desde  $S_0$  mismo), lo cual es razonable ya que esta misma base se usaría en el sistema inercial  $S$ . Los vectores unitarios  $\mathbf{u}_i^{rot}$  pueden en cambio cambiar de dirección con respecto a  $S_0$ . Dado que los  $\mathbf{u}_i^{rot}$  son unitarios, no pueden cambiar su magnitud y por tanto su cambio es una transformación infinitesimal ortogonal y propia (rotación infinitesimal) alrededor del eje de rotación instantáneo  $\mathbf{n}(t)$ . En consecuencia podemos tomar el resultado expresado en (13.56).

$$d\mathbf{u}_i^{rot} = (\mathbf{n} \times \mathbf{u}_i^{rot}) d\Phi = d\boldsymbol{\Omega} \times \mathbf{u}_i^{rot}$$

a partir de esta expresión la Ec. (13.66) queda

$$(dV_i) \mathbf{u}_i = (dV_i^{rot}) \mathbf{u}_i^{rot} + (d\boldsymbol{\Omega} \times V_i^{rot} \mathbf{u}_i^{rot}) \quad (13.67)$$

ambos miembros de esta ecuación se refieren al cambio del vector visto por  $S_0$ <sup>9</sup>. En contraste, el cambio del vector visto por  $S_{rot}$  corresponde solo al primer término de la derecha en la Ec. (13.67), ya que este sistema vé que los vectores  $\mathbf{u}_i^{rot}$  son estacionarios, es decir que visto por  $S_{rot}$  se tiene que  $d\mathbf{u}_i^{rot} = 0$  y por tanto  $d\boldsymbol{\Omega} \times \mathbf{u}_i^{rot} = 0$ . Finalmente, de acuerdo con (13.65) tenemos que  $V_i^{rot} \mathbf{u}_i^{rot}$  es simplemente  $\mathbf{V}$

$$(d\mathbf{V})_{S_0} = (d\mathbf{V})_{rot} + d\boldsymbol{\Omega} \times \mathbf{V} \quad (13.68)$$

<sup>9</sup>Aunque  $S_0$  no es necesariamente inercial, solo tiene traslación pura con respecto a un sistema inercial. Por tanto, todas las propiedades puramente rotacionales que aparezcan en  $S_0$  se verán idénticas en un sistema inercial  $S$ . Incluso podría permitirse que los ejes de  $S_0$  estuvieran inclinados con respecto a los de  $S$ , siempre que esta inclinación no cambie en el tiempo.

Nótese que dado que  $\mathbf{V}$  como tal es igual en ambos sistemas, no escribimos un rótulo para designar al sistema que lo mide. Finalmente la razón de cambio en un intervalo de tiempo  $dt$  sería

$$\left(\frac{d\mathbf{V}}{dt}\right)_{S_0} = \left(\frac{d\mathbf{V}}{dt}\right)_{rot} + \frac{d\boldsymbol{\Omega}}{dt} \times \mathbf{V}$$

definimos la velocidad angular instantánea asociada al eje de rotación como

$$\boldsymbol{\omega} \equiv \frac{d\boldsymbol{\Omega}}{dt} = \mathbf{n} \frac{d\Phi}{dt} \quad (13.69)$$

este vector está orientado a lo largo del eje de rotación con su sentido definido por la regla de la mano derecha. Su magnitud corresponde al ángulo barrido por unidad de tiempo en el intervalo  $dt$ . La razón de cambio se escribe

$$\left(\frac{d\mathbf{V}}{dt}\right)_{S_0} = \left(\frac{d\mathbf{V}}{dt}\right)_{rot} + \boldsymbol{\omega} \times \mathbf{V} \quad (13.70)$$

Es importante diferenciar las rotaciones “estáticas” con las cuales se realiza un cambio de base, de las rotaciones “dinámicas” que estamos trabajando aquí. Las rotaciones “estáticas” nos llevan de un conjunto de ejes ortogonales  $X_1X_2X_3$  a otro conjunto de ejes  $X'_1X'_2X'_3$  inclinado con respecto al primero, pero esta transformación no ocurre en el tiempo, y la inclinación es la misma en todo instante (i.e. los ángulos de Euler que caracterizan la rotación son constantes). En las rotaciones dinámicas, el sistema  $S'$  no solo está rotado con respecto a  $S$  sino que también es rotante, es decir la rotación evoluciona en el tiempo y tanto la velocidad de rotación como la dirección del eje son funciones del tiempo (o equivalentemente los ángulos de Euler son funciones del tiempo).

### 13.11.2. Razón de cambio por argumentos algebraicos

También podemos realizar una derivación con base en el formalismo de las matrices ortogonales. Denotamos  $\mathbf{A}$  a la matriz que produce la transformación desde la base en  $S_0$  a la base en  $S^{rot}$ . Las componentes de  $\mathbf{V}$  en ambas bases se escriben como

$$\begin{aligned} (\mathbf{V})_{rot} &= \mathbf{A} (\mathbf{V})_{S_0} \Rightarrow \mathbf{A}^{-1} (\mathbf{V})_{rot} = (\mathbf{V})_{S_0} \\ (\mathbf{V})_{S_0} &= \tilde{\mathbf{A}} (\mathbf{V})_{rot} \end{aligned} \quad (13.71)$$

de aquí en adelante acortaremos la notación para escribir  $(\mathbf{V})_S$  y  $(\mathbf{V})_r$  en lugar de  $(\mathbf{V})_{S_0}$  y  $(\mathbf{V})_{rot}$ . En componentes la Ec. (13.71) se escribe

$$V_i^S = a_{ji} V_j^r \quad (13.72)$$

la matriz  $\mathbf{A}$  (i.e. sus componentes) evoluciona en forma continua con el tiempo. Por tanto, una transformación infinitesimal (en un intervalo  $dt$ ) produce cambios infinitesimales en sus elementos. El diferencial del cambio del vector será

$$dV_i^S = a_{ji} dV_j^r + da_{ji} V_j^r \quad (13.73)$$

por otro lado, siempre es posible elegir los ejes de modo que  $S_0$  coincida con  $S^{rot}$  en un cierto instante  $t$ . De esta forma se tiene que

$$\mathbf{A}(t) = \mathbf{1} \quad ; \quad \mathbf{A}(t+dt) = \mathbf{1} + \boldsymbol{\varepsilon} \quad (13.74)$$

como corresponde a las transformaciones infinitesimales, Ec. (13.47). A partir de (13.74) se observa que

$$a_{ji}(t+dt) = \delta_{ji} + \varepsilon_{ji} = a_{ji}(t) + \varepsilon_{ji} \quad ; \quad da_{ji} = \varepsilon_{ji} = -\varepsilon_{ij} \quad (13.75)$$

donde hemos usado la antisimetría de  $\boldsymbol{\varepsilon}$ . Como los ejes de ambos sistemas coordenados son coincidentes en  $t$ , las componentes de los vectores coinciden instantáneamente, pero los diferenciales no serán idénticos ya que los dos sistemas se mueven el uno relativo al otro. Se tiene entonces que

$$V_j^S = V_j^r \quad ; \quad a_{ji} dV_j^r = (\delta_{ji} + \varepsilon_{ji}) dV_j^r = dV_i^r \quad (13.76)$$

donde solo se conservan diferenciales de primer orden. Reemplazando las Ecs. (13.76, 13.75) en (13.73) se obtiene

$$dV_i^S = dV_i^r - \varepsilon_{ij} V_j^r \quad (13.77)$$

Por otro lado, usando las Ecs. (13.54), (13.57) y (13.59), podemos escribir los elementos de la matriz antisimétrica  $\varepsilon_{ij}$  en términos del tensor cíclico de Levi civitá  $\varepsilon_{ijk}$ <sup>10</sup>. Hay sin embargo un aspecto delicado cual es el hecho de que estamos trabajando una transformación pasiva, recordemos que una transformación activa en convención antihoraria equivale a una pasiva en convención horaria, por tanto debemos trabajar con la convención original de ejes a izquierda, para ello basta con cambiar de signo uno de los miembros de la Ec. (13.59) manteniendo intactas las definiciones en (13.54), (13.57)

$$\begin{aligned} \varepsilon_{ij} &= N_{ij} d\Phi = \varepsilon_{ijk} n_k d\Phi = -\varepsilon_{ikj} d\Omega_k \Rightarrow \\ -\varepsilon_{ij} &= \varepsilon_{ikj} d\Omega_k \end{aligned} \quad (13.78)$$

y reemplazando (13.78) en (13.77) se obtiene

$$\begin{aligned} dV_i^S &= dV_i^r + \varepsilon_{ikj} d\Omega_k V_j^r \\ (d\mathbf{V})_i^S &= (d\mathbf{V})_i^r + (d\boldsymbol{\Omega} \times \mathbf{V}^r)_i \end{aligned}$$

recordando además que las componentes del vector como tal coinciden entonces  $V_j^r = V_j^S \equiv V_j$  y se obtiene finalmente

$$(d\mathbf{V})_S = (d\mathbf{V})_r + d\boldsymbol{\Omega} \times \mathbf{V} \quad (13.79)$$

que coincide con la expresión (13.68). Nótese que la Ec. (13.79) no es tanto una ecuación acerca de un vector particular sino más bien una aseveración acerca de la transformación infinitesimal entre dos sistemas coordenados. Ahora bien, dado que (13.70) es válida para un vector o pseudovector arbitrario, podemos escribir una ecuación de operadores de la forma

$$\left[ \frac{d(\dots)}{dt} \right]_S = \left[ \frac{d(\dots)}{dt} \right]_r + \boldsymbol{\omega} \times (\dots) \quad (13.80)$$

y la ecuación vectorial resultante puede ser resuelta a lo largo de cualquier conjunto ortogonal de ejes, fijo o móvil. Nótese de nuevo que la razón de cambio en el tiempo es solo relativa a un conjunto coordenado específico. Cuando la derivada temporal de un vector es con respecto a un sistema coordenado, las componentes pueden ser tomadas a lo largo de otro sistema coordenado solo después de realizar la diferenciación.

### 13.11.3. Segunda derivada en el sistema rotante

Para consideraciones futuras será útil también calcular la segunda derivada del vector arbitrario  $\mathbf{V}$  en  $S_0$  y  $S^{rot}$ . Para simplificar la notación definimos

$$\mathbf{Z}_{S,r} \equiv \left( \frac{d\mathbf{V}}{dt} \right)_{S,r} \quad (13.81)$$

la segunda derivada de  $\mathbf{V}$  en el sistema  $S_0$  (denotada por  $\mathbf{W}_S$ ) se escribe como

$$\mathbf{W}_S \equiv \left[ \frac{d\mathbf{Z}_S}{dt} \right]_S \quad (13.82)$$

y aplicando el operador (13.80) sobre el vector  $\mathbf{Z}_S$  resulta

$$\mathbf{W}_S = \left[ \frac{d\mathbf{Z}_S}{dt} \right]_S = \left[ \frac{d\mathbf{Z}_S}{dt} \right]_r + \boldsymbol{\omega} \times \mathbf{Z}_S \quad (13.83)$$

<sup>10</sup>Dado que el tensor de Levi Civitá tiene tres índices y la matriz antisimétrica que genera las rotaciones infinitesimales tiene dos índices, no habrá confusión aunque se use la misma letra.

nótese que en la expresión  $\mathbf{Z}_S$  hemos puesto un rótulo para denotar el sistema de referencia en donde se observa a  $\mathbf{Z}$ , situación diferente a la que se vé en la Ec. (13.70), en la cual no se especifica el sistema que observa a  $\mathbf{V}$ . La diferencia consiste en que  $\mathbf{V}$  es un vector “primario” que no depende del sistema de referencia, en tanto que  $\mathbf{Z}$  ya es una razón de cambio temporal, que en consecuencia depende explícitamente del sistema elegido.

Por otro lado, usando (13.70) con la nueva notación resulta

$$\mathbf{Z}_S = \mathbf{Z}_r + \boldsymbol{\omega} \times \mathbf{V} \quad (13.84)$$

reemplazando (13.84) en (13.83) resulta

$$\mathbf{W}_S = \left[ \frac{d}{dt} (\mathbf{Z}_r + \boldsymbol{\omega} \times \mathbf{V}) \right]_r + \boldsymbol{\omega} \times \mathbf{Z}_r + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{V}) \quad (13.85)$$

es claro que la segunda derivada de  $\mathbf{V}$ , vista por  $S^{rot}$  sería

$$\mathbf{W}_r = \left[ \frac{d}{dt} \left( \frac{d\mathbf{V}}{dt} \right)_r \right]_r = \left[ \frac{d\mathbf{Z}_r}{dt} \right]_r$$

de modo que

$$\mathbf{W}_S = \mathbf{W}_r + \left[ \frac{d}{dt} (\boldsymbol{\omega} \times \mathbf{V}) \right]_r + \boldsymbol{\omega} \times \mathbf{Z}_r + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{V}) \quad (13.86)$$

tomaremos el caso de  $\boldsymbol{\omega}$  constante i.e. eje de rotación y rapidez angular constantes

$$\mathbf{W}_S = \mathbf{W}_r + \boldsymbol{\omega} \times \left[ \frac{d\mathbf{V}}{dt} \right]_r + \boldsymbol{\omega} \times \mathbf{Z}_r + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{V})$$

usando (13.81) queda

$$\mathbf{W}_S = \mathbf{W}_r + 2\boldsymbol{\omega} \times \mathbf{Z}_r + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{V}) \quad (13.87)$$

ahora retornamos a escribir estas segundas derivadas en términos del vector original  $\mathbf{V}$ , para ello usamos (13.81, 13.82) y resulta

$$\left[ \frac{d}{dt} \left( \frac{d\mathbf{V}}{dt} \right)_S \right]_S = \left[ \frac{d}{dt} \left( \frac{d\mathbf{V}}{dt} \right)_r \right]_r + 2\boldsymbol{\omega} \times \left( \frac{d\mathbf{V}}{dt} \right)_r + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{V})$$

que nos induce a escribir la ecuación de operadores equivalente

$$\left[ \frac{d}{dt} \left( \frac{d(\dots)}{dt} \right)_S \right]_S = \left[ \frac{d}{dt} \left( \frac{d(\dots)}{dt} \right)_r \right]_r + 2\boldsymbol{\omega} \times \left( \frac{d(\dots)}{dt} \right)_r + \boldsymbol{\omega} \times [\boldsymbol{\omega} \times (\dots)] \quad (13.88)$$

Una anotación metodológica importante, para la validez de las ecuaciones (13.80, 13.88) no es estrictamente necesario suponer que ninguno de los sistemas de referencia  $S_0$  o  $S'$  sea inercial o que alguno de los dos esté relacionado con un sistema inercial a través de una traslación pura. Ambos sistemas podrían ser rotantes con respecto a un sistema inercial, las relaciones aquí obtenidas son puramente cinemáticas y parten simplemente de dos premisas:  $S_0$  y  $S'$  tienen origen común y establecemos a  $S_0$  como el “sistema no rotante” y a  $S'$  como el “sistema rotante”. Lo que hacemos entonces es comparar los vectores y sus derivadas vistas por ambos sistemas,  $\boldsymbol{\omega}$  es entonces la rotación de  $S'$  vista por  $S_0$ . Cuando entremos al estudio de la dinámica, será importante que uno de los dos sistemas sea inercial (o de lo contrario introducimos un tercero que sí lo sea) a fin de establecer un marco de referencia en donde se puedan emplear las leyes de Newton. Pero las Ecs. (13.80, 13.88) mantendrán su validez incluso cuando ambos sistemas sean no inerciales.

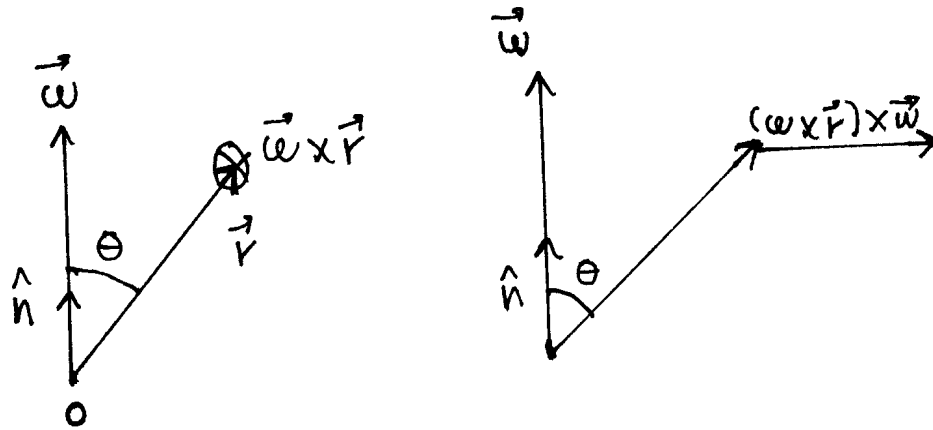


Figura 13.7: Ilustración de la aceleración centrífuga. Para una posición dada de la partícula, la aceleración centrífuga va en la dirección perpendicular al eje y alejándose de él.

### 13.12. Sistemas no inerciales rotantes

Las Ecs. (13.80, 13.88) son las leyes cinemáticas básicas sobre las cuales se escribirán las ecuaciones de movimiento del cuerpo rígido. Sin embargo, su validez es mucho más general dado que son aplicables a un vector arbitrario.

Las Ecs. (13.80, 13.88) serán de particular utilidad en el estudio del movimiento de una partícula o sistemas de partículas relativo a un sistema coordenado rotante. Describiremos brevemente el problema de una partícula que se mueve relativo a un sistema rotante. Para ello tomaremos como vector base el vector posición de la partícula (que coincide en ambos sistemas  $S_0$  y  $S'$ ), medido desde el origen común de  $S_0$  y  $S'$ . Dado que la base de la descripción dinámica de la partícula es la segunda ley de Newton, debemos calcular la primera y segunda derivada de este vector con respecto a ambos sistemas de referencia. Partamos de la segunda ley de Newton en un sistema inercial

$$\mathbf{F} = m\mathbf{a}_S$$

por simplicidad asumiremos que  $S_0$  es un sistema inercial y podemos entonces tomarlo coincidente con  $S$ . Debemos entonces calcular la relación entre la aceleración vista por  $S_0$  denotada por  $\mathbf{a}_S$  y la vista por  $S'$  denotada como  $\mathbf{a}_r$ . Primero aplicamos el operador (13.80)

$$\begin{aligned} \left[ \frac{d\mathbf{r}}{dt} \right]_S &= \left[ \frac{d\mathbf{r}}{dt} \right]_r + \boldsymbol{\omega} \times \mathbf{r} \\ \mathbf{v}_S &= \mathbf{v}_r + \boldsymbol{\omega} \times \mathbf{r} \end{aligned} \tag{13.89}$$

y la aceleración se obtiene aplicando el operador (13.88) al vector  $\mathbf{r}$

$$\begin{aligned} \left[ \frac{d}{dt} \left( \frac{d\mathbf{r}}{dt} \right) \right]_S &= \left[ \frac{d}{dt} \left( \frac{d\mathbf{r}}{dt} \right) \right]_r + 2\boldsymbol{\omega} \times \left( \frac{d\mathbf{r}}{dt} \right)_r + \boldsymbol{\omega} \times [\boldsymbol{\omega} \times \mathbf{r}] \\ \mathbf{a}_S &= \mathbf{a}_r + 2(\boldsymbol{\omega} \times \mathbf{v}_r) + \boldsymbol{\omega} \times [\boldsymbol{\omega} \times \mathbf{r}] \end{aligned} \tag{13.90}$$

despejando  $\mathbf{v}_r$  y  $\mathbf{a}_r$  en la Ecs. (13.89, 13.90) se obtiene

$$\mathbf{v}_r = \mathbf{v}_S - \boldsymbol{\omega} \times \mathbf{r} \tag{13.91}$$

$$\mathbf{a}_r = \mathbf{a}_S + 2(\mathbf{v}_r \times \boldsymbol{\omega}) + [(\boldsymbol{\omega} \times \mathbf{r}) \times \boldsymbol{\omega}] \tag{13.92}$$

analicemos los términos que aparecen a la derecha de la expresión (13.92). El término  $[(\boldsymbol{\omega} \times \mathbf{r}) \times \boldsymbol{\omega}]$  se ilustra en la Fig. 13.7, en tal figura se asume que  $\boldsymbol{\omega}$  y  $\mathbf{r}$  están en el plano del papel de modo que  $\boldsymbol{\omega} \times \mathbf{r}$  es perpendicular al

papel hacia adentro. Colocando la cola del vector  $[(\boldsymbol{\omega} \times \mathbf{r}) \times \boldsymbol{\omega}]$  en el punto donde se ubica la partícula, vemos que tal vector es perpendicular a  $\boldsymbol{\omega}$  y apunta en la dirección que se aleja de  $\boldsymbol{\omega}$  y por tanto del eje de rotación, razón por la cual se le conoce como aceleración centrífuga. Su magnitud es  $\omega^2 r \sin \theta$  siendo  $\theta$  el ángulo entre  $\boldsymbol{\omega}$  y  $\mathbf{r}$ . El término  $2(\mathbf{v}_r \times \boldsymbol{\omega})$  se conoce como aceleración de coriolis y es claramente perpendicular al eje de rotación y a la velocidad de la partícula vista por el sistema rotante.

Multiplicando (13.92) por la masa de la partícula vemos que un observador en el sistema rotante observa el movimiento de la partícula como si ésta se moviera bajo la influencia de una fuerza efectiva:

$$m\mathbf{a}_r = m\mathbf{a}_S + 2m(\mathbf{v}_r \times \boldsymbol{\omega}) + m[(\boldsymbol{\omega} \times \mathbf{r}) \times \boldsymbol{\omega}] \quad (13.93)$$

$$\mathbf{F}_{fict} = \mathbf{F} + 2m(\mathbf{v}_r \times \boldsymbol{\omega}) + m[(\boldsymbol{\omega} \times \mathbf{r}) \times \boldsymbol{\omega}] \quad (13.94)$$

Naturalmente, los términos correspondientes a las aceleraciones centrífuga y de coriolis se denominan efecto centrífugo y efecto de coriolis. Es más adecuado usar los términos “efecto de coriolis” o “efecto centrífugo” que los términos “fuerza de Coriolis” o “fuerza centrífuga”. Ya que la primera terminología nos recuerda que estos efectos surgen de emplear un observador no inercial y no corresponde a interacciones entre sistemas físicos. Esto se puede ver claramente del hecho de que cada uno de estos términos está relacionado con la masa de la partícula por una aceleración que no está relacionada con la partícula sino con el sistema rotante. En la literatura también se usa el término “fuerza ficticia” para indicar este tipo de efectos.

Nótese que el efecto de coriolis solo aparece si la partícula tiene una velocidad no nula con respecto al sistema rotante y produce una aceleración normal a la velocidad instantánea, esto implica a su vez que este efecto (tomado aisladamente) solo produce cambio de dirección pero no de rapidez en la partícula. De otra parte el efecto centrífugo desaparece si el vector posición de la partícula es paralelo al vector velocidad angular i.e. al eje de rotación.

### 13.12.1. La tierra como sistema rotante

Un sistema rotante de particular interés en la Física es el planeta tierra. Aunque un sistema de referencia fijo a la tierra puede tomarse en buena aproximación como un sistema inercial, para ciertas medidas de alta precisión es necesario considerar el hecho de que este sistema rota con respecto a un sistema mucho más cercano al ideal inercial (sistema de referencia de las estrellas fijas). En tal caso, debe estudiarse los efectos centrífugo y de coriolis para un sistema de referencia fijo a la tierra. Estimaremos entonces el orden de magnitud de éstos efectos en la tierra. La tierra rota en dirección antihoraria con respecto al polo norte, con una velocidad angular relativa a las estrellas fijas dada por

$$\omega = \left( \frac{2\pi}{(24 \times 3600) s} \right) \left( \frac{366,25}{365,25} \right) = 7,292 \times 10^{-5} s^{-1}$$

el primer término entre paréntesis se refiere a la velocidad angular relativa al radio vector del sol. La cantidad en el segundo paréntesis se refiere a el cociente entre el número de días siderales en un año, dividido por el número de días en un año solar, es decir es el factor de corrección para obtener la velocidad angular relativa a las estrellas fijas. Tomado este valor de  $\omega$  junto con el radio de la tierra en el ecuador, la máxima aceleración centrípeta estará dada por

$$\omega^2 r = 3,38 cm/s^2$$

que corresponde a un 0,34% de la aceleración de la gravedad. Este efecto es pequeño pero no despreciable. En realidad, el efecto medido de la gravedad representa el efecto combinado del campo gravitacional de la distribución de masa de la tierra y los efectos de la aceleración centrípeta. En general se denota la suma de los dos efectos como el campo de gravedad, para distinguirlo del campo gravitacional.

Un efecto interesante del término de coriolis es el de convertir un movimiento horizontal en línea recta sobre una esfera rotante (e.g. la tierra), en movimiento circular. Supongamos que  $\mathbf{v}_r$  es la velocidad (con respecto a la esfera) de un objeto que se mueve horizontalmente, es decir tangencial a la esfera. Veremos que la componente horizontal de la fuerza de coriolis es perpendicular a  $\mathbf{v}_r$  y su magnitud es independiente de la dirección de  $\mathbf{v}_r$ .

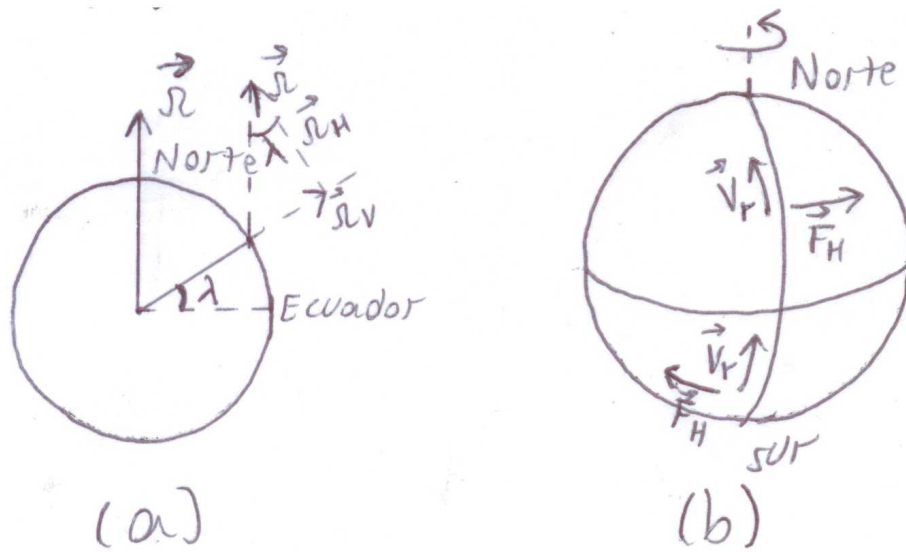


Figura 13.8: Efecto coriolis sobre una masa ubicada sobre una esfera rotante. (a) Descomposición de la velocidad angular  $\Omega$  en sus componentes vertical  $\Omega_V$  y horizontal  $\Omega_H$ . Siendo  $\lambda$  la latitud de la partícula. (b) Dirección de la componente horizontal  $F_H$  de la fuerza de coriolis en los hemisferios norte y sur, para una partícula que viaja a velocidad horizontal  $v_r$  con respecto a la tierra en dirección norte. El hemisferio norte se define con  $\Omega$ , utilizando la regla de la mano derecha.

Para verlo consideremos una masa  $m$  que se mueve con velocidad  $\mathbf{v}_r$  (con respecto a la esfera) a una latitud  $\lambda$  como se ilustra en la Fig. 13.8a. sobre la superficie de una esfera que rota a velocidad angular uniforme  $\Omega$  (donde la latitud se mide con respecto al ecuador definido por el eje de rotación). Descomponiendo a  $\Omega$  en sus componentes vertical y horizontal, la fuerza de coriolis se escribe

$$\mathbf{F} = -2m\Omega \times \mathbf{v}_r = -2m\Omega_H \times \mathbf{v}_r - 2m\Omega_V \times \mathbf{v}_r$$

puesto que  $\Omega_H$  y  $\mathbf{v}_r$  son horizontales, entonces  $\Omega_H \times \mathbf{v}_r$  es vertical. Por tanto la componente horizontal de la fuerza de coriolis es

$$\mathbf{F}_H = -2m\Omega_V \times \mathbf{v}_r = -2m\Omega \sin \lambda (\mathbf{u}_\rho \times \mathbf{v}_r)$$

donde  $\mathbf{u}_\rho$  es el vector unitario radial de las coordenadas esféricas. Puesto que  $\mathbf{u}_\rho$  y  $\mathbf{v}_r$  son perpendiculares, vemos que la magnitud de  $\mathbf{F}_H$  está dada por

$$\|\mathbf{F}_H\| = 2m\Omega |\sin \lambda| \|\mathbf{u}_\rho \times \mathbf{v}_r\| = 2mv_r \Omega |\sin \lambda|$$

además  $\mathbf{F}_H$  es siempre perpendicular a  $\mathbf{v}_r$ , de modo que en ausencia de otras fuerzas horizontales, producirá movimiento circular horizontal (con una posible componente vertical). Puesto que  $\mathbf{F}_H$  depende de  $\sin \lambda$ , es claro que invierte su sentido cuando se invierte la latitud como se observa en la Fig. 13.8b. El movimiento será en el sentido horario en el hemisferio norte y antihorario en el hemisferio sur. El flujo del viento sobre la tierra está fuertemente influenciado por la fuerza de coriolis formando patrones de movimiento circular estables. Sin embargo, un mejor entendimiento de la dinámica de los sistemas atmosféricos requiere introducir otras fuerzas.

### 13.12.2. Superficie de un líquido rotante

Supongamos que un balde rota con velocidad angular constante  $\omega$ . Nos preguntamos por el perfil que forma la superficie del agua en el balde (ver Fig 13.9a). Nótese que en un sistema no inercial rotante con velocidad angular  $\omega$ , el problema es totalmente estático y por tanto más simple. Consideremos un pequeño volumen

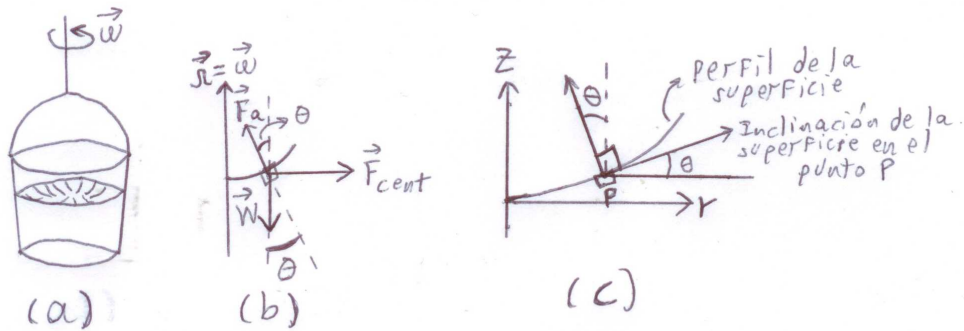


Figura 13.9: (a) Superficie de un líquido que rota uniformemente en un balde. (b) Diagrama de fuerzas aparentes sobre un pequeño elemento de volumen en la superficie del agua, visto por un sistema de referencia que rota con el balde. (c) Este diagrama muestra claramente que el ángulo  $\theta$  mide la inclinación o pendiente de la superficie del agua en el punto donde se ubica el elemento de volumen.

de agua de masa  $m$  sobre la superficie del líquido. Las fuerzas aparentes que actúan sobre este elemento de volumen (ver Fig. 13.9b) son: la fuerza de contacto  $\mathbf{F}_a$  que es la resultante de la interacción del elemento de volumen con otros elementos vecinos, el peso  $\mathbf{W}$  y la fuerza ficticia centrífuga (puesto que el elemento está en reposo con respecto al sistema rotante, no hay fuerza de coriolis). El elemento está en reposo con respecto al sistema rotante de modo que

$$\mathbf{F}_a + \mathbf{W} + \mathbf{F}_{cent} = \mathbf{0}$$

en coordenadas cilíndricas, las fuerzas vienen dadas por

$$\begin{aligned} \mathbf{F}_{cent} &= -m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = -m\omega^2 r \mathbf{u}_z \times (\mathbf{u}_z \times \mathbf{u}_r) = -m\omega^2 r \mathbf{u}_z \times \mathbf{u}_\theta = m\omega^2 r \mathbf{u}_r \\ \mathbf{F}_a &= F_a \cos \theta \mathbf{u}_z - F_a \sin \theta \mathbf{u}_r \quad ; \quad \mathbf{W} = -mg \mathbf{u}_z \end{aligned}$$

y las ecuaciones de movimiento quedan

$$\begin{aligned} F_a \cos \theta - mg &= 0 \quad ; \quad -F_a \sin \theta + m\omega^2 r = 0 \quad \Rightarrow \\ F_a \cos \theta &= mg \quad ; \quad F_a \sin \theta = m\omega^2 r \end{aligned}$$

dividiendo las ecuaciones se obtiene

$$\tan \theta = \frac{\omega^2 r}{g}$$

a diferencia de los sólidos, los líquidos no ejercen una fuerza estática tangencial a la superficie (no hay un análogo a la fuerza de fricción tipo Coulomb). Por tanto,  $\mathbf{F}_a$  es perpendicular a la superficie del líquido (ver Fig. 13.9c), de modo que la pendiente de la superficie en cualquier punto es entonces

$$\frac{dz}{dr} = \tan \theta = \frac{\omega^2 r}{g}$$

es fácil integrar esta ecuación para encontrar la ecuación de la superficie. Tomando  $z = 0$  sobre el eje en la superficie del líquido tenemos

$$\int_0^z dz' = \int_0^r \frac{\omega^2 r'}{g} dr' \quad \Rightarrow \quad z = \frac{\omega^2}{2g} r^2$$

la superficie es un paraboloides de revolución.



13.12.3. Cuenta sobre un alambre (revisado)

En la sección 2.8.3, Pág. 29, resolvimos el problema de una cuenta que se desliza sin rozamiento por un alambre que rota uniformemente alrededor de un eje fijo perpendicular al alambre, ilustrado en la Fig. 2.2b. Debido a la ausencia de fricción, la fuerza  $\mathbf{N}$  de la cuenta sobre el alambre es perpendicular al alambre. En la sección 2.8.3 se utilizó la formulación Lagrangiana. Encontraremos por el método de fuerzas ficticias la solución para el movimiento de la cuenta y el valor de  $\mathbf{N}$ . Para ello nos montamos en un sistema rotante con la velocidad angular  $\boldsymbol{\omega}$  del alambre. Es claro que en este sistema rotante la cuenta tiene un movimiento puramente radial i.e.  $\mathbf{v}_r = \dot{r}\mathbf{u}_r$ . Como la única fuerza aplicada sobre la cuenta es  $\mathbf{N}$ , la suma de fuerzas en el sistema rotante es  $\mathbf{N}$  más las fuerzas ficticias

$$\mathbf{N} + \mathbf{F}_{cent} + \mathbf{F}_{cor} = m\ddot{\mathbf{r}}\mathbf{u}_r$$

$$\begin{aligned} \mathbf{N} &= N\mathbf{u}_\theta \quad ; \quad \mathbf{F}_{cor} = -2m\boldsymbol{\omega} \times \mathbf{v}_r = -2m\omega\dot{r}(\mathbf{u}_z \times \mathbf{u}_r) = -2m\omega\dot{r}\mathbf{u}_\theta \\ \mathbf{F}_{cent} &= -m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = -m\omega^2 r\mathbf{u}_z \times (\mathbf{u}_z \times \mathbf{u}_r) = m\omega^2 r\mathbf{u}_r \end{aligned}$$

y las ecuaciones de la fuerza quedan

$$N - 2m\omega\dot{r} = 0 \quad ; \quad m\omega^2 r = m\ddot{r} \tag{13.95}$$

la segunda ecuación coincide con la Ec. (2.45), cuya solución general es la Ec. (2.46) y la solución para  $\dot{r}(0) = 0$  es la Ec. (2.47). Reemplazando la solución (2.47) en la primera de las Ecs. (13.95) se encuentra

$$N = F_{cor} = 2m\omega\dot{r} = 2m\omega^2 C \cosh \omega t$$

que coincide con (2.48), Pág. 29.

13.12.4. Deflexión de una masa en caída libre

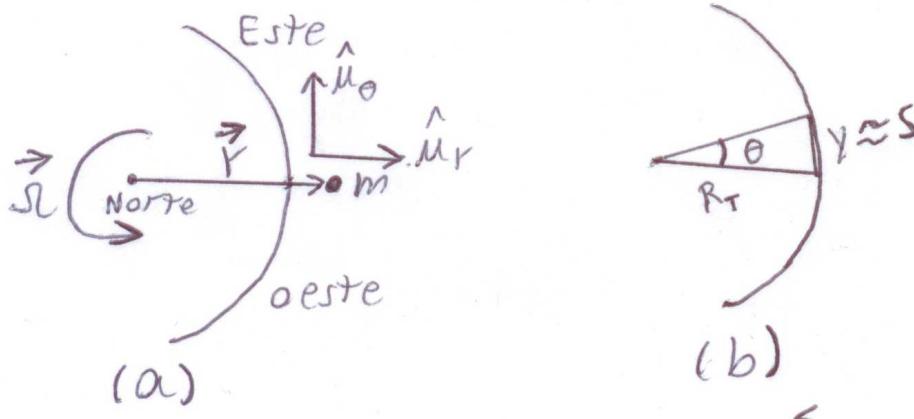


Figura 13.10: (a) Sistema coordenado cilíndrico en el cual los vectores unitarios  $\mathbf{u}_r$  y  $\mathbf{u}_\theta$  están sobre el plano ecuatorial y la velocidad angular de la tierra va a lo largo de  $\mathbf{u}_z$ . (b) El valor de la longitud de la cuerda  $\gamma$  (que mide la deflexión) es aproximadamente igual al valor del arco subtendido por  $\theta$ , ya que dicho ángulo es muy pequeño.

Otro efecto interesante de la fuerza de coriolis es la deflexión con respecto a la vertical hacia el este, de una masa  $m$  en caída libre. Calcularemos esta deflexión para una masa  $m$  que se suelta desde una altura  $h$  en el ecuador.

Tomaremos un sistema de coordenadas cilíndricas con el plano  $XY$  sobre el plano ecuatorial, donde  $\mathbf{u}_\theta$  apunta hacia el este como se vé en la Fig. 13.10a. La fuerza aparente sobre  $m$  estará dada por

$$\mathbf{F} = -mg\mathbf{u}_r - 2m\boldsymbol{\Omega} \times \mathbf{v}_r - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = -mg\mathbf{u}_r - 2m\Omega \mathbf{u}_z \times \mathbf{v}_r + m\Omega^2 r\mathbf{u}_r$$

la fuerza centrífuga va en la dirección radial de modo que solo cambia la gravedad aparente pero no produce deflexión con respecto al desplazamiento vertical. Si la partícula parte del reposo, la fuerza de coriolis estará en el plano ecuatorial. Por tanto  $m$  se mueve sobre el plano ecuatorial  $XY$ . En consecuencia, la velocidad de la partícula con respecto a la tierra es

$$\mathbf{v}_r = \dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta$$

de modo que la fuerza aparente es

$$\begin{aligned}\mathbf{F} &= (m\Omega^2 r - mg)\mathbf{u}_r - 2m\Omega\mathbf{u}_z \times (\dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta) \\ \mathbf{F} &= m(\Omega^2 r + 2r\dot{\theta}\Omega - g)\mathbf{u}_r - 2m\dot{r}\Omega\mathbf{u}_\theta\end{aligned}$$

usando las expresiones para la aceleración radial y transversal, obtenemos las ecuaciones

$$\ddot{r} - r\dot{\theta}^2 = \Omega^2 r + 2r\dot{\theta}\Omega - g \quad ; \quad r\ddot{\theta} + 2\dot{r}\dot{\theta} = -2\dot{r}\Omega \quad (13.96)$$

es una buena aproximación considerar que  $m$  cae en forma casi vertical y  $\dot{\theta} \ll \Omega$ , con lo cual omitiremos todos los términos que contengan  $\dot{\theta}$ , y se obtiene

$$\ddot{r} \simeq \Omega^2 r - g \quad ; \quad r\ddot{\theta} \simeq -2\dot{r}\Omega \quad (13.97)$$

la variación de  $r$  es muy pequeña puesto que oscila entre  $h + R_T$  y  $R_T$  siendo  $R_T$  el radio de la tierra y considerando  $h \ll R_T$ . Por esta razón tomaremos  $g$  constante y  $r \simeq R_T$  de modo que la primera de las Ecs. (13.97) se puede escribir como

$$\ddot{r} \simeq -g + \Omega^2 R_T \equiv g'$$

donde  $g'$  es la gravedad aparente generada por la corrección de la fuerza centrífuga. La solución radial es entonces

$$\dot{r} \simeq -g't \quad ; \quad r \simeq r_0 - \frac{1}{2}g't^2 \quad (13.98)$$

reemplazando la primera de las Ecs. (13.98) en la segunda de las Ecs. (13.97) y usando  $r \simeq R_T$  obtenemos

$$\begin{aligned}R_T\ddot{\theta} &\simeq 2g't\Omega \Rightarrow \ddot{\theta} \simeq \frac{2g'\Omega}{R_T}t \Rightarrow \\ \dot{\theta} &\simeq \frac{g'\Omega}{R_T}t^2 \quad ; \quad \theta \simeq \frac{g'\Omega}{3R_T}t^3\end{aligned} \quad (13.99)$$

donde hemos usado las condiciones iniciales  $\dot{\theta}_0 = \theta_0 = 0$ . Como se ilustra en la Fig. 13.10b, la deflexión horizontal está dada aproximadamente por  $y \simeq R_T\theta$  y de la Ec. (13.99) resulta

$$y \simeq \frac{g'\Omega}{3}t^3 \quad (13.100)$$

el tiempo  $t_h$  para caer una distancia  $h$  está dado por

$$r - r_0 \simeq -h = -\frac{1}{2}g't_h^2 \Rightarrow t_h = \sqrt{\frac{2h}{g'}}$$

reemplazando este tiempo en (13.100), resulta

$$y \simeq \frac{g'\Omega}{3} \left(\frac{2h}{g'}\right)^{3/2} = \frac{\Omega}{3\sqrt{g'}}(2h)^{3/2}$$

para una altura  $h = 50\text{mts}$  se obtiene

$$y \simeq 0,77\text{cm}$$

que es una cantidad muy pequeña pero medible. Puesto que  $\theta$  es positivo, la deflexión de la partícula es hacia el este.

### 13.13. Ejercicios

1. Realice explícitamente el producto  $\mathbf{A} \equiv \mathbf{BCD}$  de las tres matrices dadas en las Ecs. (13.15, 13.16, 13.17), y demuestre que se obtiene la matriz (13.19).
2. Demuestre que la matriz  $\mathbf{A}$  definida en (13.19) es ortogonal real propia.
3. Relacione adecuadamente la expresión (5.9) de la Pág. 74, válida para rotaciones infinitesimales, con la Ec. (13.39) de la Pág. 352, válida para rotaciones finitas. Tenga en cuenta las convenciones quirales usadas en cada caso.
4. Demuestre que la matriz  $\mathbf{A}$  dada en la Ec. (13.45), Pág. 353, es ortogonal real propia.
5. Demuestre que a primer orden,  $\det \mathbf{A} = +1$  para la matriz (13.49), que describe una rotación infinitesimal.
6. Construya la matriz de rotación para los ángulos de Euler en la  $y$ -convención.
7. Encuentre la diferencia entre la gravedad aparente en el ecuador y en los polos, asumiendo que la tierra es esférica. Encuentre el valor de la gravedad aparente en función de la latitud. Es de anotar sin embargo, que la suposición de tierra esférica no es buena, ya que el efecto del achatamiento de los polos sobre la gravedad aparente, es comparable con el efecto de la latitud.

## Ecuaciones de movimiento del cuerpo rígido

En el capítulo 13, hemos desarrollado las herramientas matemáticas para el estudio sistemático de la cinemática del cuerpo rígido. Los ángulos de Euler proveen un conjunto conveniente (aunque asimétrico) de 3 coordenadas generalizadas independientes para determinar la orientación relativa de un cuerpo rígido. En particular, el método de transformaciones ortogonales propias provee una elegante técnica para generar las orientaciones relativas. Finalmente, las transformaciones ortogonales infinitesimales proveen un marco para generar rotaciones finitas por transformaciones sucesivas infinitesimales (tal como ya se vió en el capítulo de transformaciones canónicas). En este capítulo se pondrá en acción el formalismo ya descrito para encontrar las ecuaciones dinámicas de movimiento de Euler para el cuerpo rígido en una forma conveniente. Veremos también las soluciones para algunas configuraciones simples pero muy importantes.

### 14.1. Momento angular y energía cinética de movimiento alrededor de un punto

El teorema de Chasles establece que el movimiento general de un cuerpo rígido consta de una traslación mas una rotación. Esto permite separar el movimiento en una fase traslacional y una rotacional. En particular si el cuerpo rígido posee un punto fijo, es claro que alrededor de dicho punto solo tenemos rotación y la componente traslacional estará ausente. Incluso cuando no hay un punto fijo, los seis grados de libertad para determinar la configuración del cuerpo rígido se pueden separar en tres grados de libertad que determinan la posición de un punto de referencia y otros tres grados de libertad (e.g. los ángulos de Euler) que determinan la orientación del cuerpo rígido. Si adicionalmente tomamos como punto de referencia el centro de masa del cuerpo rígido, entonces una división similar ocurre para el momento angular según se vé en la Ec. (1.25), en donde se observa que el momento angular visto por el laboratorio (sistema  $S$ ) se escribe como la suma del momento angular con respecto al centro de masa (sistema  $S_0$  con origen en el centro de masa) mas el momento angular debido a la traslación del centro de masa. El primer término involucra solo variables angulares, en tanto que el segundo incluye solo a las coordenadas cartesianas del centro de masa. Adicionalmente, la Ec. (1.32) nos muestra que se puede hacer una división análoga para la energía cinética del cuerpo rígido, en una notación adecuada para nuestras variables independientes escribimos

$$T = \frac{1}{2}Mv_{CM}^2 + T_{CM}(\phi, \theta, \psi)$$

para el potencial ocurre con frecuencia que una división similar es también posible, si bien no hay un teorema que lo garantice. El potencial puede usualmente escribirse en dos términos donde cada uno de ellos involucra solo uno de los conjuntos coordenados traslacional o rotacional. Como ejemplos sencillos se puede ver que la energía potencial en un campo gravitacional uniforme dependerá solo de la coordenada vertical del centro de gravedad (que para campo uniforme coincide con el centro de masa). Por otro lado, si la fuerza sobre un cuerpo es debida a un campo magnético uniforme  $\mathbf{B}$ , actuando sobre su momento dipolar magnético  $\mathbf{M}$  el potencial

será proporcional a  $\mathbf{M} \cdot \mathbf{B}$  que involucra solo la orientación del cuerpo. En realidad casi todos los problemas solubles en la práctica permiten este tipo de división de modo que el Lagrangiano,  $L = T - V$  se podrá dividir en dos partes una que involucra solo coordenadas traslacionales y otra que solo involucra variables rotacionales (coordenadas angulares). Los dos grupos de coordenadas se separarán completamente y por tanto el problema traslacional y el rotacional se resolverán también por separado.

Es por tanto de gran importancia encontrar expresiones para el momento angular y la energía cinética de movimiento con respecto a algún punto fijo en el cuerpo. Para ello emplearemos la relación (13.80) que liga a las derivadas temporales entre sistemas rotantes y no rotantes. No obstante, es necesario demostrar previamente la siguiente aseveración: **el ángulo de rotación asociado al desplazamiento de un cuerpo rígido, así como el vector velocidad angular de rotación, son independientes de la escogencia de origen del sistema coordenado fijo al cuerpo.**

Para probarlo tomemos a partir del origen de  $S$  dos vectores  $\mathbf{R}_1, \mathbf{R}_2$  asociados a dos posibles orígenes de sistemas de ejes fijos al cuerpo  $S'_1$  y  $S'_2$  respectivamente. Sea  $\mathbf{R}$  el vector relativo que va desde el origen de  $S'_1$  hasta el origen de  $S'_2$ . Claramente resulta que

$$\mathbf{R}_2 = \mathbf{R}_1 + \mathbf{R}$$

si el origen de  $S'_2$  se considera como un medida relativa a  $S'_1$ , entonces la derivada temporal de  $\mathbf{R}_2$  relativa a  $S$  se escribe

$$\left(\frac{d\mathbf{R}_2}{dt}\right)_S = \left(\frac{d\mathbf{R}_1}{dt}\right)_S + \left(\frac{d\mathbf{R}}{dt}\right)_S \quad (14.1)$$

calculemos el segundo miembro de la derecha usando (13.80)

$$\left[\frac{d\mathbf{R}}{dt}\right]_S = \left[\frac{d\mathbf{R}}{dt}\right]_{r_1} + \boldsymbol{\omega}_1 \times \mathbf{R}$$

donde  $\boldsymbol{\omega}_1$  es el vector velocidad angular apropiado para el sistema  $S'_1$ . Teniendo en cuenta además que las derivadas de  $\mathbf{R}$  relativas a todo sistema de ejes fijos al cuerpo deben ser nulas se obtiene:

$$\left[\frac{d\mathbf{R}}{dt}\right]_S = \boldsymbol{\omega}_1 \times \mathbf{R}$$

con lo cual la Ec. (14.1) queda

$$\left(\frac{d\mathbf{R}_2}{dt}\right)_S = \left(\frac{d\mathbf{R}_1}{dt}\right)_S + \boldsymbol{\omega}_1 \times \mathbf{R} \quad (14.2)$$

por otro lado, podemos también considerar el origen de  $S'_1$  como una medida relativa a  $S'_2$ . El origen del primer sistema coordenado se considera fijo en el segundo sistema con vector posición  $-\mathbf{R}$ . En este caso la derivada del vector posición  $\mathbf{R}_1$  con respecto a los ejes de  $S$  se escribe

$$\begin{aligned} \left(\frac{d\mathbf{R}_1}{dt}\right)_S &= \left(\frac{d\mathbf{R}_2}{dt}\right)_S - \left(\frac{d\mathbf{R}}{dt}\right)_S = \left(\frac{d\mathbf{R}_2}{dt}\right)_S - \boldsymbol{\omega}_2 \times \mathbf{R} \\ \left(\frac{d\mathbf{R}_2}{dt}\right)_S &= \left(\frac{d\mathbf{R}_1}{dt}\right)_S + \boldsymbol{\omega}_2 \times \mathbf{R} \end{aligned} \quad (14.3)$$

restando las Ecs. (14.2, 14.3) se obtiene la relación

$$(\boldsymbol{\omega}_1 - \boldsymbol{\omega}_2) \times \mathbf{R} = 0$$

es decir cualquier diferencia en los vectores velocidad angular en dos puntos arbitrarios debe ser paralela a la línea que une los dos puntos. Asumiendo que  $\boldsymbol{\omega}$  es un campo vectorial continuo, la única solución posible para todos los pares de puntos (i.e. para todas las orientaciones de  $\mathbf{R}$ ) es que los dos vectores velocidad angular sean iguales

$$\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2$$

de modo que el vector velocidad angular es el mismo para todo sistema coordenado fijo al cuerpo rígido. Esto se puede vislumbrar intuitivamente teniendo en cuenta que las partículas del cuerpo rígido se mueven y rotan como un todo.

### 14.1.1. Momento angular, velocidad angular y tensor de inercia

Hemos garantizado que el vector velocidad angular es independiente del origen elegido para el sistema coordenado fijo al cuerpo  $S'$ . Ahora bien, dado un sistema  $S_0$  no rotante<sup>1</sup> con origen en cierto punto  $O$ , el momento angular del cuerpo rígido alrededor de este origen será

$$\mathbf{L} = m_i (\mathbf{r}_i \times \mathbf{v}_i) \quad (14.4)$$

siendo  $\mathbf{r}_i$ ,  $\mathbf{v}_i$  el radio vector y la velocidad respectivamente de la  $i$ -ésima partícula relativa al origen escogido. Es conveniente colocar el origen de  $S_0$  en un punto por donde pase el eje de rotación. Cuando hay un punto fijo del cuerpo rígido respecto a un sistema inercial, este punto es un buen origen ya que el eje de rotación pasa por él (teorema de Euler), y el sistema  $S_0$  será inercial. Si no hay punto fijo un origen conveniente será el centro de masa, ya que en este caso el movimiento se podrá separar en una traslación pura del centro de masa y una rotación pura alrededor de un eje que pasa por el centro de masa.

Como el problema de la traslación pura es muy simple, nos concentraremos en la rotación pura. Supondremos por simplicidad que existe un punto fijo con respecto a un sistema inercial de modo que  $S_0$  coincide con nuestro sistema inercial  $S$ . Su origen estará en el punto fijo y el eje de rotación pasa por este punto (de lo contrario el punto no podría ser fijo). Es claro que cada partícula  $i$  del cuerpo rígido realizará un movimiento circular **con la misma frecuencia angular** alrededor del eje de rotación<sup>2</sup>, donde el radio del círculo será  $r_i \sin \theta_i$  siendo  $\mathbf{r}_i$  la posición de la partícula y  $\theta_i$  el ángulo entre el eje de rotación y  $\mathbf{r}_i$ . La rapidez instantánea de la partícula está dada por

$$v_i = \omega R_i = \omega r_i \sin \theta_i$$

y su desplazamiento es tangente al círculo de modo que en dirección y magnitud tenemos que

$$\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i \quad (14.5)$$

como corresponde a todo movimiento circular. Es importante enfatizar que para que esta expresión y otras que vienen de aquí en adelante sean válidas,  $\mathbf{r}_i$  debe medirse con respecto a un origen tal que el eje de rotación pase por él<sup>3</sup>. Otra forma de ver esta relación es usando la Ec. (13.80) aplicada al vector  $\mathbf{r}_i$  en donde el sistema rotante  $S'$  se define como un sistema con ejes fijos al cuerpo

$$\left( \frac{d\mathbf{r}_i}{dt} \right)_S = \left( \frac{d\mathbf{r}_i}{dt} \right)_{S'} + \boldsymbol{\omega} \times \mathbf{r}_i$$

el primer término a la derecha es nulo porque en  $S'$  la partícula debe estar en reposo, ya que el sistema es fijo al cuerpo. Por tanto se reproduce la Ec. (14.5). Usando (14.5) en la Ec. (14.4) se obtiene

$$\mathbf{L} = m_i [\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i)]$$

expandiendo el triple producto cruz resulta

$$\mathbf{L} = m_i [\boldsymbol{\omega} r_i^2 - \mathbf{r}_i (\mathbf{r}_i \cdot \boldsymbol{\omega})] \quad (14.6)$$

escribiendo esta ecuación por componentes se tiene

$$L_n = m_i \left\{ \omega_n r_i^2 - (\mathbf{r}_i)_n \left[ (\mathbf{r}_i)_j \omega_j \right] \right\}$$

<sup>1</sup>Con esto queremos decir que existe un sistema inercial  $S$  con ejes paralelos a los de  $S_0$  todo el tiempo. Es decir entre  $S$  y  $S_0$  solo hay traslación pura.

<sup>2</sup>El eje de rotación puede variar de dirección, pero para nuestros fines, es posible tomar el eje de rotación instantáneo definido en un tiempo  $t$ . En consecuencia, el movimiento circular tendrá un plano fijo solo en un intervalo de tiempo infinitesimal, pero esto no limita nuestra demostración.

<sup>3</sup>Hay un aspecto delicado en el sentido de que  $\mathbf{n}$  como vector no define completamente al eje de rotación, ya que los vectores no tienen su origen localizado, puesto que una traslación paralela los deja idénticos. En contraste, el eje de rotación está bien localizado en el espacio y no se puede trasladar paralelamente sin alterar el problema físico.

donde los índices  $j, n$  denotan componentes en tanto que el índice  $i$  rotula a una partícula, hay suma sobre los índices  $i, j$ . Expandiendo la componente  $x$  del momento angular resulta

$$\begin{aligned} L_x &= m_i \left\{ \omega_x r_i^2 - (\mathbf{r}_i)_x \left[ (\mathbf{r}_i)_j \omega_j \right] \right\} = m_i \left\{ \omega_x r_i^2 - x_i \left[ (\mathbf{r}_i)_x \omega_x + (\mathbf{r}_i)_y \omega_y + (\mathbf{r}_i)_z \omega_z \right] \right\} \\ L_x &= m_i \left\{ \omega_x r_i^2 - x_i [x_i \omega_x + y_i \omega_y + z_i \omega_z] \right\} \\ L_x &= m_i (r_i^2 - x_i^2) \omega_x - m_i x_i y_i \omega_y - m_i x_i z_i \omega_z \end{aligned}$$

y ecuaciones similares para las otras componentes de  $\mathbf{L}$ . Vemos que  $\mathbf{L}$  está relacionado con  $\boldsymbol{\omega}$  a través de una transformación lineal. Parametrizaremos las transformaciones lineales de las componentes de  $\mathbf{L}$  en la forma

$$\begin{aligned} L_x &= I_{xx} \omega_x + I_{xy} \omega_y + I_{xz} \omega_z \\ L_y &= I_{yx} \omega_x + I_{yy} \omega_y + I_{yz} \omega_z \\ L_z &= I_{zx} \omega_x + I_{zy} \omega_y + I_{zz} \omega_z \end{aligned} \quad (14.7)$$

Comparando las expresiones de  $L_x$  se tiene

$$I_{xx} = m_i (r_i^2 - x_i^2) \quad ; \quad I_{xy} = -m_i x_i y_i \quad ; \quad I_{xz} = -m_i x_i z_i$$

y procediendo de la misma manera con las componentes  $L_y$  y  $L_z$ , los coeficientes toman la forma

$$I_{qw} = m_i (r_i^2 \delta_{qw} - q_i w_i) \quad ; \quad q, w \equiv x, y, z \quad (14.8)$$

donde a los elementos diagonales se les conoce como **coeficientes de momento de inercia**, en tanto que los elementos no diagonales se denominan **productos de inercia**. Las Ecs. (14.8) son adecuadas si el sistema de partículas es discreto. Para cuerpos rígidos contínuos la suma sobre  $i$  debe reemplazarse por una integral volumétrica, la masa se convierte en un diferencial  $dm$ , de modo que los elementos de matriz se escriben como

$$\begin{aligned} I_{ww} &= \int (r^2 - w^2) dm = \int_V \rho(\mathbf{r}) (r^2 - w^2) dV \quad , \quad w \equiv x, y, z \\ I_{qw} &= - \int_V \rho(\mathbf{r}) q w dV \quad , \quad q, w \equiv x, y, z \end{aligned} \quad (14.9)$$

donde  $\rho(\mathbf{r})$  denota la densidad del sólido en el punto  $\mathbf{r}$  y  $dV$  es un diferencial de volumen alrededor de  $\mathbf{r}$ . Dado que en el contínuo desaparece el índice de partícula, podemos condensar la notación para los elementos diagonales y no diagonales si denotamos los ejes por  $X_i$  con  $i = 1, 2, 3$  y el elemento de matriz  $I_{jk}$  se escribirá como

$$I_{jk} = \int_V \rho(\mathbf{r}) (r^2 \delta_{jk} - x_j x_k) dV \quad (14.10)$$

De la Ec. (14.7) podemos escribir la ecuación que relaciona al momento angular total con la velocidad angular de rotación en la forma

$$L_i = I_{ij} \omega_j \quad \Rightarrow \quad \mathbf{L}_S = \mathbf{I}_S \boldsymbol{\omega}_S \quad (14.11)$$

donde los elementos matriciales  $I_{ij}$  están dados por (14.10). El subíndice  $S$  enfatiza que esta ecuación y todos los observables se miden en el sistema inercial  $S$  y con la base de vectores unitarios definida por  $S$ . En la práctica sin embargo, resulta poco útil esta relación ya que los elementos de la matriz  $\mathbf{I}_S$  dependen de la dinámica del cuerpo y solo se podrá calcular cuando esté resuelto el problema. Por tanto reescribiremos a los vectores  $\mathbf{L}$  y  $\boldsymbol{\omega}$  y al operador  $\mathbf{I}$  en una base de vectores unitarios de un sistema  $S'$  de ejes fijos al cuerpo. Para ello usamos la matriz pasiva de cambio de base  $\mathbf{B}$  que nos lleva de la base en  $S$  a la base en  $S'$ , esta matriz es ortogonal real propia ya que pasa una base ortonormal a otra base ortonormal, multiplicando por  $\mathbf{B}$  a ambos lados de (14.11)

$$\begin{aligned} \mathbf{B} \mathbf{L}_S &= \mathbf{B} \mathbf{I}_S \mathbf{B}^{-1} \mathbf{B} \boldsymbol{\omega}_S \quad \Rightarrow \quad (\mathbf{B} \mathbf{L}_S) = (\mathbf{B} \mathbf{I}_S \mathbf{B}^{-1}) (\mathbf{B} \boldsymbol{\omega}_S) \\ &\Rightarrow \quad \mathbf{L}_{S'} = \mathbf{I}_{S'} \boldsymbol{\omega}_{S'} \end{aligned} \quad (14.12)$$

recordando que estas transformaciones son de tipo pasivo tenemos que  $\mathbf{L}_{S'}$ ,  $\boldsymbol{\omega}_{S'}$  representa a los mismos vectores que  $\mathbf{L}_S$ ,  $\boldsymbol{\omega}_S$ . Así mismo,  $\mathbf{I}_{S'}$  representa al mismo operador que  $\mathbf{I}_S$ . Simplemente, se están escribiendo en bases diferentes. Por ejemplo  $\mathbf{L}_{S'}$  **NO** representa el momento angular medido por el sistema  $S'$  (el cual obviamente es cero), lo que representa es al vector momento angular que mide el sistema  $S$  pero escrito en la base de vectores unitarios  $\mathbf{u}'_i$  fija al sistema  $S'$ .

¿Qué se logra con esta transformación pasiva? basta notar el hecho de que si la matriz que representa al operador  $\mathbf{I}$  en la nueva base conserva su forma, es decir si  $\mathbf{I}_{S'}$  viene dada por las Ecs. (14.10) con el simple reemplazo  $x_k \rightarrow x'_k$  esta matriz será independiente del tiempo, ya que las coordenadas  $x'_i, y'_i, z'_i$  son constantes en el tiempo para cada partícula  $i$ -ésima cuando  $S'$  está fijo al cuerpo. En consecuencia  $\mathbf{I}_{S'}$  solo dependerá de factores geométricos. Los elementos de matriz serán constantes en el tiempo características del cuerpo involucrado y dependientes del origen y orientación del sistema coordenado particular fijo al cuerpo.

Sin embargo, para poder afirmar que  $\mathbf{I}_{S'}$  es constante en el tiempo, es necesario demostrar que efectivamente la representación matricial de  $\mathbf{I}$  conserva su forma ante una transformación ortogonal propia (cambio de base), esto se vé simplemente del hecho de que  $\mathbf{I}$  es un tensor euclidiano de segundo rango, según se puede demostrar a través de la regla del cociente expresada en la Ec. (12.186). Para ello notamos que se cumplen todas las condiciones exigidas para la validez de esta regla del cociente: la ecuación (14.11) es válida en cualquier sistema cartesiano rotado según se puede ver en el procedimiento que nos llevó a la Ec. (14.12) pues aunque elegimos el sistema primado como un sistema de ejes del cuerpo, esta condición no es necesaria para llegar a (14.12), el vector  $\boldsymbol{\omega}$  es arbitrario pues en principio un cuerpo rígido se puede poner a rotar en cualquier dirección y con cualquier velocidad angular y finalmente las cantidades  $\mathbf{L}$  y  $\boldsymbol{\omega}$  son tensores bien definidos (vectores axiales). En consecuencia, la regla del cociente nos lleva a que  $\mathbf{I}$  es un tensor polar cartesiano de segundo rango, conocido como **tensor de inercia**.

Con base en la discusión anterior, entenderemos que de ahora en adelante, la ecuación

$$\mathbf{L} = \mathbf{I}\boldsymbol{\omega} \quad (14.13)$$

se escribirá en componentes usando la base de vectores unitarios definida por algún sistema coordenado fijo al cuerpo rígido  $S'$ <sup>4</sup>. Pero no debemos perder de vista que los observables involucrados en esta ecuación son los medidos con respecto al sistema de referencia  $S$ . El operador  $\mathbf{I}$  está representado por una matriz  $3 \times 3$  cuyos elementos son los coeficientes de momento de inercia y los productos de inercia.  $\mathbf{L}$  y  $\boldsymbol{\omega}$  son vectores columna. Es claro que el operador  $\mathbf{I}$  actúa desde un punto de vista activo sobre  $\boldsymbol{\omega}$ . Esto se puede ver de varias maneras: los vectores  $\mathbf{L}$  y  $\boldsymbol{\omega}$  son vectores físicamente diferentes que incluso tienen diferentes dimensiones, y no simplemente el mismo vector expresado en sistemas coordenados distintos. A diferencia del operador de rotación que es adimensional, el operador  $\mathbf{I}$  tiene dimensiones de masa por longitud al cuadrado, y no está restringido por ninguna condición de ortogonalidad, puesto que no está “obligado” a conservar la norma del vector original.

## 14.2. Tensor de inercia y momento de inercia

Considerado como un operador lineal que transforma a  $\boldsymbol{\omega}$  en  $\mathbf{L}$ , la matriz  $\mathbf{I}$  posee las propiedades de transformación de un tensor de segundo rango. Por esta razón se conoce como **tensor de inercia**. A partir de las Ecs. (14.8, 14.10) se puede ver claramente que este tensor se puede escribir en forma de afinor (en sus versiones discreta y continua) de la forma

$$\mathbf{I} = m_i (r_i^2 \mathbf{1} - \mathbf{r}_i \mathbf{r}_i) \quad (14.14)$$

$$\mathbf{I} = \int \rho(\mathbf{r}) (r^2 \mathbf{1} - \mathbf{r} \mathbf{r}) dV \quad (14.15)$$

donde el índice  $i$  es un rótulo para partículas. Por simplicidad continuaremos utilizando la versión discreta, el procedimiento para el continuo es análogo. Multiplicando (14.14) por el vector  $\boldsymbol{\omega}$  y usando la Ec. (14.13),

<sup>4</sup>Naturalmente esta base de vectores unitarios es diferente en cada instante de tiempo, pero esto no compromete la validez de la Ec. (14.13), la cual se evalúa en cada instante de tiempo.



resulta

$$\mathbf{I} \cdot \boldsymbol{\omega} = m_i [r_i^2 \boldsymbol{\omega} - \mathbf{r}_i (\mathbf{r}_i \cdot \boldsymbol{\omega})] = \mathbf{L}$$

que coincide con la Ec. (14.6) y nos muestra la ventaja de usar el algoritmo algebraico de afinores. La energía cinética del cuerpo rígido se escribe como

$$T = \frac{1}{2} m_i v_i^2$$

donde  $\mathbf{v}_i$  es la velocidad de la partícula  $i$ -ésima relativa al punto fijo en el sistema  $S_0$ . A partir de la Ec. (14.5), esta energía cinética se puede escribir de la forma

$$T = \frac{1}{2} m_i \mathbf{v}_i \cdot (\boldsymbol{\omega} \times \mathbf{r}_i)$$

utilizando la identidad vectorial  $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ , tenemos que

$$T = \frac{\boldsymbol{\omega}}{2} \cdot [m_i (\mathbf{r}_i \times \mathbf{v}_i)]$$

la cantidad sumada sobre  $i$  en paréntesis cuadrados, se reconoce como el momento angular total del cuerpo alrededor del origen de  $S_0$  y en consecuencia la energía cinética se puede escribir en la forma

$$T = \frac{\boldsymbol{\omega} \cdot \mathbf{L}}{2} = \frac{\boldsymbol{\omega} \cdot \mathbf{I} \cdot \boldsymbol{\omega}}{2} \quad (14.16)$$

definamos ahora

$$\boldsymbol{\omega} \equiv \omega \mathbf{n} \quad (14.17)$$

siendo  $\mathbf{n}$  la dirección del eje de rotación, reescribimos la expresión como

$$T = \frac{\omega^2}{2} \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} = \frac{1}{2} I \omega^2 \quad ; \quad I \equiv \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} \quad (14.18)$$

usando (14.14) y la definición de  $I$  en (14.18) vemos que esta cantidad está dada por

$$I \equiv \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} = m_i [r_i^2 - (\mathbf{r}_i \cdot \mathbf{n})^2] \quad (14.19)$$

puesto que  $\mathbf{n}$  es un vector axial e  $\mathbf{I}$  es un tensor polar de segundo rango, la definición de  $I$  en (14.18) nos dice que  $I$  es un tensor polar de rango cero (escalar) y se conoce como el **momento de inercia** con respecto al eje de rotación definido por  $\mathbf{n}$  y que pasa por el origen de  $S_0$ . Podemos escribir esta cantidad de la forma:

$$I \equiv \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} = m_i [r_i^2 - r_i^2 \cos^2 \theta_i] = m_i r_i^2 \sin^2 \theta_i$$

La figura 14.1 muestra que la distancia entre el eje de rotación y la partícula  $i$ -ésima está dada por  $R_i = r_i \sin \theta_i$ . Por lo tanto, el momento de inercia se puede escribir como

$$I = m_i R_i^2 \quad (14.20)$$

siendo  $R_i$  la distancia entre la partícula  $i$ -ésima y el eje de rotación. De esta expresión es claro que el valor del momento de inercia depende de la dirección y localización del eje de rotación<sup>5</sup>. Dado que  $\omega$  cambia en general su dirección con respecto al cuerpo a medida que el tiempo transcurre, el momento de inercia será también en general función del tiempo<sup>6</sup>. No obstante, si el cuerpo está restringido a rotar sobre un eje fijo, el momento de inercia será una constante. En tal caso, la energía cinética (14.16) está casi en la forma requerida para construir ecuaciones de movimiento de Lagrange o Hamilton, solo se requiere escribir  $\omega$  como la derivada temporal de algún ángulo lo cual está sugerido en la Ec. (13.69), y no presenta mayor dificultad.

<sup>5</sup>Es decir para calcular el momento de inercia debemos ubicar la dirección del eje y un punto por donde pasa. Por otro lado, recordemos que el origen se debe tomar de tal forma que pase por el eje de rotación, en ese sentido se dice que el momento de inercia (y el tensor de inercia) depende de la orientación del eje y del origen.

<sup>6</sup>De la expresión (14.19) vemos que  $I$  depende del tiempo solo a través de  $\mathbf{n}$ , puesto que el tensor  $\mathbf{I}$  es independiente del tiempo, al menos si usamos una base fija al cuerpo.

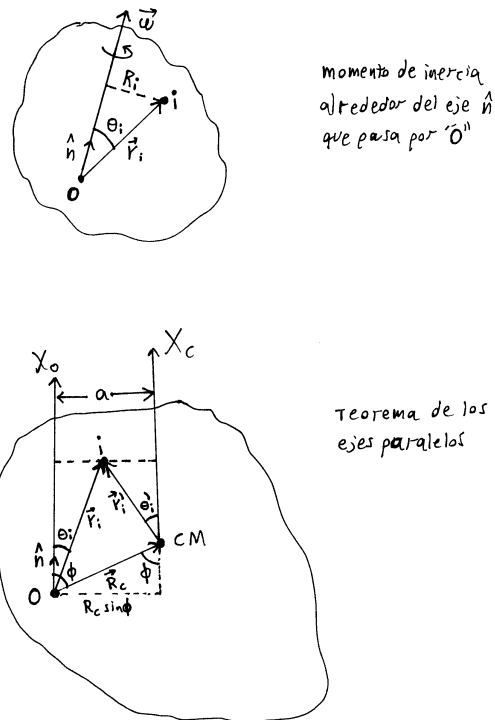


Figura 14.1: Ilustración del momento de inercia con respecto al eje definido por  $\mathbf{n}$  y que pasa por el punto  $O$  que se elige como origen. La partícula  $i$ -ésima del cuerpo rígido está a una distancia  $R_i$  de dicho eje.

#### 14.2.1. Teorema de los ejes paralelos y de los ejes perpendiculares

Tanto el tensor de inercia como el momento de inercia dependen del origen elegido para los ejes fijos al cuerpo. Si tenemos en cuenta que además el momento de inercia depende del eje de rotación, vemos que esta cantidad debe recalcularse cada vez que cambiamos de ejes. Sin embargo, existe un teorema que relaciona a los momentos de inercia entre dos ejes paralelos cuando uno de ellos pasa por el centro de masa del cuerpo rígido.

**Theorem 23** *Teorema de Steiner: si tenemos dos ejes paralelos y uno de ellos pasa por el centro de masa del cuerpo rígido, los momentos de inercia vistos por ambos ejes están relacionados por*

$$I = I_C + Ma^2$$

siendo  $M$  la masa total del cuerpo y siendo "a" la distancia entre los ejes.

Definamos un eje  $X_O$  que pasa por un cierto origen  $O$ , y sea  $X_C$  un eje que pasa por el centro de masa del cuerpo y que es paralelo a  $X_O$ . Denotemos por  $\mathbf{R}_C$  a la posición del centro de masa (CM) con respecto al origen  $O$ . Recurriendo a la definición del momento de inercia (14.20) y a la Fig. 14.1, escribiremos el momento de inercia con respecto a  $X_O$

$$I_O = m_i R_i^2 = m_i r_i^2 \sin^2 \theta_i = m_i (\mathbf{r}_i \times \mathbf{n})^2 \quad (14.21)$$

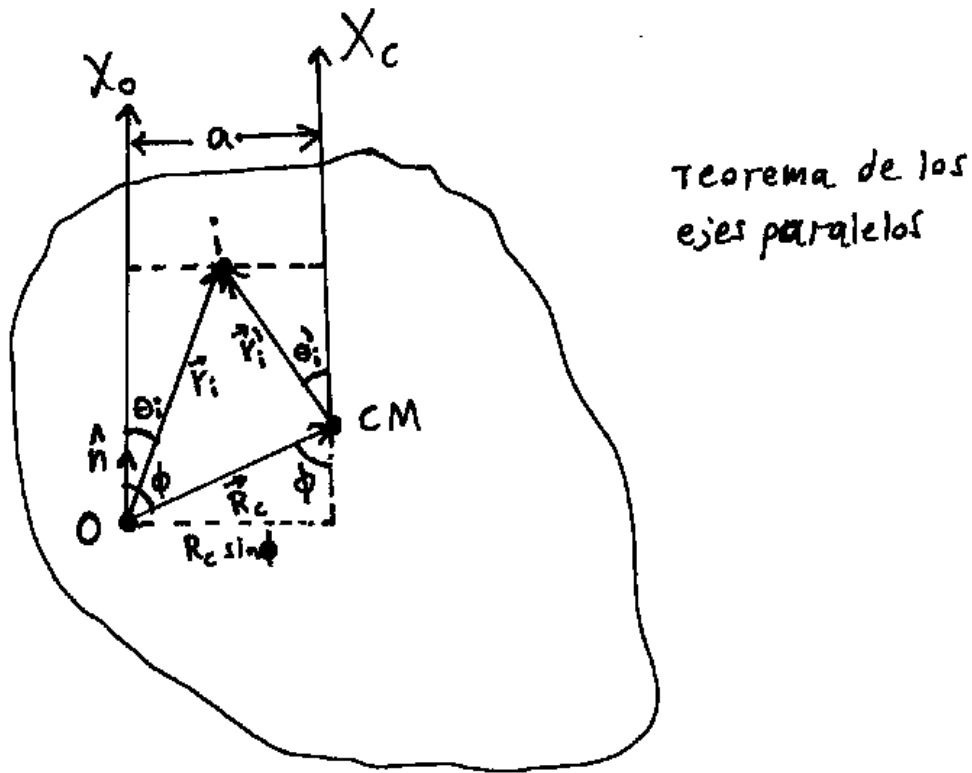


Figura 14.2: Ilustración del teorema de los ejes paralelos o teorema de Steiner para el momento de inercia.

donde la posición de la partícula  $i$ -ésima con respecto al origen  $O$  se define por  $\mathbf{r}_i$ , y su posición con respecto al CM se denota por  $\mathbf{r}'_i$ , la Fig. 14.2 muestra que

$$\mathbf{r}_i = \mathbf{R}_C + \mathbf{r}'_i \tag{14.22}$$

es claro además de la Fig. 14.2, que la distancia entre la partícula  $i$ -ésima y el eje  $X_C$  viene dada por  $R_{iC}^2 = r_i'^2 \sin^2 \theta'_i$ . Reemplazando (14.22) en (14.21), el momento de inercia se escribe (escribiremos la suma explícita)

$$\begin{aligned} I_O &= \sum_i m_i [(\mathbf{r}'_i + \mathbf{R}_C) \times \mathbf{n}]^2 = \sum_i m_i [(\mathbf{r}'_i \times \mathbf{n})^2 + (\mathbf{R}_C \times \mathbf{n})^2 + 2(\mathbf{r}'_i \times \mathbf{n}) \cdot (\mathbf{R}_C \times \mathbf{n})] \\ I_O &= \sum_i \left[ m_i r_i'^2 \sin^2 \theta'_i + m_i (\mathbf{R}_C \times \mathbf{n})^2 + 2(m_i \mathbf{r}'_i \times \mathbf{n}) \cdot (\mathbf{R}_C \times \mathbf{n}) \right] \\ I_O &= \sum_i m_i R_{iC}^2 + \left( \sum_i m_i \right) (\mathbf{R}_C \times \mathbf{n})^2 + 2M \left[ \left( \frac{1}{M} \sum_i m_i \mathbf{r}'_i \right) \times \mathbf{n} \right] \cdot (\mathbf{R}_C \times \mathbf{n}) \end{aligned}$$

el término  $\sum_i m_i \mathbf{r}'_i / M$  se refiere a la posición del centro de masa vista por el propio centro de masa de modo que se anula. Por otro lado, el primer término a la derecha es claramente el momento de inercia visto por el eje  $X_C$ .

$$I_O = I_C + M (\mathbf{R}_C \times \mathbf{n})^2 = I_C + M (R_C^2 \sin^2 \phi)$$

y nuevamente la Fig. 14.2 muestra que  $R_C \sin \phi$  es la distancia entre los dos ejes, de modo que (omitiendo el subíndice  $O$ )

$$I = I_C + Ma^2 \tag{14.23}$$

lo cual demuestra el teorema de Steiner. El segundo término de la derecha equivale al momento de inercia que tendría el cuerpo rígido con respecto al eje  $X_O$  si toda la masa del cuerpo estuviera condensada en un punto

que está a una distancia  $a$  de tal eje (en particular, un punto que pasa por el eje  $X_C$ ). Podemos ver como corolario que dada una orientación específica, de todos los ejes con esta orientación el que pasa por el centro de masa es el que da el mínimo valor posible del momento de inercia. Vemos además que el teorema de Steiner nos enfatiza que para calcular el momento de inercia, no solo necesitamos la orientación del eje de rotación, sino también un punto por donde pasa.

Para el tensor de inercia existe una propiedad análoga. Reemplacemos (14.22) en (14.14) y escribamos la suma de forma explícita

$$\begin{aligned} \mathbf{I} &= \sum_i m_i \left[ (\mathbf{R}_C + \mathbf{r}'_i)^2 \mathbf{1} - (\mathbf{R}_C + \mathbf{r}'_i) (\mathbf{R}_C + \mathbf{r}'_i) \right] \\ \mathbf{I} &= \sum_i m_i \left\{ [R_C^2 \mathbf{1} + r_i'^2 \mathbf{1} + 2 (\mathbf{R}_C \cdot \mathbf{r}'_i) \mathbf{1}] - [\mathbf{R}_C \mathbf{R}_C + \mathbf{R}_C \mathbf{r}'_i + \mathbf{r}'_i \mathbf{R}_C + \mathbf{r}'_i \mathbf{r}'_i] \right\} \end{aligned}$$

el tensor de inercia medido en el origen  $O$  se puede escribir en forma de afinor de la siguiente forma<sup>7</sup>

$$\begin{aligned} \mathbf{I} &= MR_C^2 \mathbf{1} + \left\{ \sum_i m_i [(r_i'^2 + 2\mathbf{R}_C \cdot \mathbf{r}'_i) \mathbf{1} - (\mathbf{r}'_i \mathbf{r}'_i + \mathbf{R}_C \mathbf{r}'_i + \mathbf{r}'_i \mathbf{R}_C)] \right\} - M\mathbf{R}_C \mathbf{R}_C \\ \mathbf{I} &= \left\{ \sum_i m_i [r_i'^2 \mathbf{1} - \mathbf{r}'_i \mathbf{r}'_i + 2\mathbf{R}_C \cdot \mathbf{r}'_i \mathbf{1} - \mathbf{R}_C \mathbf{r}'_i - \mathbf{r}'_i \mathbf{R}_C] \right\} + MR_C^2 \mathbf{1} - M\mathbf{R}_C \mathbf{R}_C \\ \mathbf{I} &= \left\{ \sum_i m_i [r_i'^2 \mathbf{1} - \mathbf{r}'_i \mathbf{r}'_i] \right\} + 2\mathbf{R}_C \cdot \left( \sum_i m_i \mathbf{r}'_i \right) \mathbf{1} - \mathbf{R}_C \sum_i m_i \mathbf{r}'_i - \left\{ \sum_i m_i \mathbf{r}'_i \right\} \mathbf{R}_C + MR_C^2 \mathbf{1} - M\mathbf{R}_C \mathbf{R}_C \end{aligned}$$

todas las sumas de la forma  $m_i \mathbf{r}'_i$  se anulan ya que están asociadas a la posición del CM vista por el propio CM. La expresión se reduce a

$$\mathbf{I} = \left[ \sum_i m_i (r_i'^2 \mathbf{1} - \mathbf{r}'_i \mathbf{r}'_i) \right] + M (R_C^2 \mathbf{1} - \mathbf{R}_C \mathbf{R}_C) \quad (14.24)$$

$$\mathbf{I} = \mathbf{I}_C + M (R_C^2 \mathbf{1} - \mathbf{R}_C \mathbf{R}_C) \quad (14.25)$$

de la Ec. (14.14) vemos que el segundo término a la derecha de (14.25) correspondería al tensor de inercia con respecto a  $O$  asociado a una partícula puntual de masa  $M$  situada en el centro de masa. En tal sentido, tanto el momento de inercia como el tensor de inercia admiten una descomposición relativa al CM muy semejante a la que existe para la energía cinética, el torque y el momento angular.

Finalmente, para distribuciones de masa bidimensionales, es fácil relacionar el momento de inercia de un sistema de tres ejes perpendiculares con origen común, si colocamos el sólido rígido plano sobre el plano  $XY$  es claro que sus momentos de inercia con respecto a cada eje vendrán dados por

$$I_x = m_i y_i^2 \quad ; \quad I_y = m_i x_i^2 \quad ; \quad I_z = m_i (x_i^2 + y_i^2)$$

de modo que se obtiene la identidad

$$I_x + I_y = I_z$$

que se conoce como **teorema de los ejes perpendiculares**, válido solo para sólidos bidimensionales continuos o discretos.

### 14.3. Compendio de propiedades del tensor de inercia

A partir de la Ec. (14.8) o de la Ec. (14.10), se observan algunas propiedades inmediatas

<sup>7</sup>Nótese que en estas expresiones tenemos sumas de la forma  $\mathbf{R}_C \mathbf{r}'_i + \mathbf{r}'_i \mathbf{R}_C$ . Estas sumas corresponden a diadas (o productos tensoriales entre dos vectores) que no son en general conmutativas.

1. La representación matricial del tensor de inercia depende de la base de vectores cartesianos elegida pero también del origen de coordenadas elegido, ya que depende del vector posición y de sus coordenadas.
2.  $I_{ij} = I_{ji}$  el tensor de inercia es simétrico y real, lo cual corresponde a una matriz hermítica real. Recordemos que la propiedad de simetría es invariante ante un cambio de base **ortogonal**. Por tanto, esta propiedad es válida en cualquier sistema ortogonal de ejes<sup>8</sup>. El carácter simétrico y real a su vez implica que el tensor tiene seis grados de libertad.
3.  $I_{ii} \geq 0$ , en cualquier base de modo que los elementos diagonales son no negativos. La ecuación (14.8) nos muestra que  $I_{xx}$  se puede escribir como

$$I_{xx} = m_i (y_i^2 + z_i^2)$$

que solo es cero si todos los  $y_i$  y  $z_i$  son nulos. Esto indicaría que la masa del cuerpo rígido está toda distribuída sobre el eje  $X$ . Es decir, un elemento diagonal puede ser nulo solo si la distribución de masa es unidimensional o si consta de una sola masa puntual. Si la distribución es unidimensional solo un elemento diagonal puede ser cero (y solo si se elige la base de modo que la distribución de masa esté a lo largo de un eje). Si la distribución es una masa puntual, todos los elementos del tensor de inercia son cero (una masa puntual no puede rotar alrededor de un eje que pase por ella).

4. De lo anterior se deduce que si el cuerpo rígido es plano o volumétrico tendremos que  $I_{ii} > 0$ , para los tres ejes, y para cualquier base y origen.

Dado que la representación matricial de este tensor es hermítica (simétrica y real) para cualquier base ortogonal y origen coordinado, sus valores propios son reales (ver sección 12.11). Más aún dado que sus elementos diagonales son no negativos en cualquier base, lo serán en particular en la base que diagonaliza tal tensor, y dado que los elementos en la diagonal de la matriz diagonalizada son los valores propios, se deduce que los valores propios de  $\mathbf{I}$  son no negativos, y de hecho serán estrictamente positivos para distribuciones de masa superficiales o volumétricas. Apelando al teorema 15, Pág. 315, encontramos que  $\mathbf{I}$  es un operador lineal definido positivo si la distribución de masa es superficial o volumétrica, y singular positivo si la distribución es lineal o consta de una sola masa puntual.

Es claro que estas consideraciones han sido independientes del origen elegido. Adicionalmente, una transformación de similaridad con una matriz ortogonal real propia, preserva el carácter definido positivo (o singular positivo) de una matriz. Por tanto, el carácter definido positivo o singular positivo de  $\mathbf{I}$  es intrínseco, i.e. independiente de la base y el origen utilizados.

Otra forma de ver la positividad del tensor de inercia es a través de la definición del momento de inercia (14.18) combinada con la Ec. (14.20) para el momento de inercia con lo cual

$$\mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} \equiv I = m_i R_i^2 \geq 0 \quad (14.26)$$

siendo  $\mathbf{n}$  un vector unitario real arbitrario y  $R_i$  la distancia de la masa  $m_i$  al eje de rotación. Puesto que  $I$  es positivo y  $\mathbf{n}$  es un vector unitario real pero por lo demás arbitrario, la Ec. (14.26) corresponde a la definición de una matriz positiva<sup>9</sup>. En esta expresión se vé que la matriz es definida positiva a menos que todos los  $R_i$  sean nulos, lo cual solo es posible si la distribución de masa es lineal o si consta de una sola masa puntual. Se requiere además que el eje pase por todas las masas de la distribución.

Sintetizamos esta discusión enumerando las siguientes propiedades

1. El tensor de inercia es simétrico real. Así mismo, es definido positivo para distribuciones superficiales y volumétricas, y singular positivo para una distribución lineal de masas o para una sola masa puntual<sup>10</sup>.

<sup>8</sup>En tensores, es de suma importancia saber si cierta propiedad es independiente de la base, o es válida sólo en una base particular.

<sup>9</sup>El teorema 16, Pág. 315, nos dice que una matriz simétrica y real  $\mathbf{I}$  es positiva si y solo si  $\mathbf{v} \cdot \mathbf{I} \cdot \mathbf{v} \geq 0$ , para  $\mathbf{v}$  un vector real arbitrario. Escribiendo  $\mathbf{v} = v\mathbf{n} \neq \mathbf{0}$ , esto es equivalente a escribir  $v^2 \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} \geq 0$ , que a su vez es equivalente a  $\mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} \geq 0$ , para  $\mathbf{n}$  unitario y real pero por lo demás arbitrario.

<sup>10</sup>Un conjunto discreto de varias masas puntuales no colineales se considera una distribución superficial o volumétrica para nuestros propósitos.

2. Cuando  $\mathbf{I}$  es definido positivo, todos sus valores propios son estrictamente positivos. Cuando es singular positivo, al menos un valor propio es cero y los demás son positivos (ver ejercicio 3, Pág. 414).
3. Siempre es posible **para un origen dado** hacer una rotación a ejes principales **reales** donde este tensor sea diagonal, incluso en presencia de degeneración de los valores propios<sup>11</sup>. Esta característica es muy importante, ya que garantiza que la diagonalización se puede realizar dentro del espacio vectorial real  $\mathbb{R}^3$ .
4. Cuando hay degeneración simple ( $\lambda_1 = \lambda_2 \neq \lambda_3$ ), y encontramos vectores propios  $\mathbf{V}_1, \mathbf{V}_2$  linealmente independientes asociados al valor propio degenerado, cualquier conjunto de ejes en el plano generado por ellos son ejes principales. Esto es un indicativo de simetría axial, como ocurre con los sólidos de revolución cuando uno de los ejes principales es el eje de simetría<sup>12</sup>. Además un vector ortogonal a  $\mathbf{V}_1$  y  $\mathbf{V}_2$  también será vector propio y por tanto también define un eje principal (ver sección 12.11.2). Similarmente, si  $\mathbf{V}_1$  es vector propio asociado a  $\lambda_1 = \lambda_2$  y  $\mathbf{V}_3$  es vector propio asociado a  $\lambda_3$  entonces  $\mathbf{V}_2 = \mathbf{V}_1 \times \mathbf{V}_3$  también es vector propio asociado a  $\lambda_1 = \lambda_2$  (sección 12.11.2).
5. La degeneración completa indica simetría esférica como es el caso de esferas, cascarones y anillos esféricos etc. (con densidad uniforme).
6. La rotación que nos lleva a ejes principales tiene una interesante analogía con la rotación que lleva a la ecuación de un elipsoide a su modo normal (ver sección 12.11.3). En tal caso hablamos del elipsoide de inercia. Los momentos principales de inercia determinan las longitudes de los ejes del elipsoide de inercia.
7. Recordemos que la degeneración simple tiene como análogo geométrico a un elipsoide de revolución y la degeneración total corresponde a una esfera.

Vimos además otras propiedades como son:

1. La cantidad  $\mathbf{I}$  es un tensor cartesiano polar de segundo rango, lo cual justifica el término tensor de inercia. Esto además implica que esta cantidad conserva su forma (definida por 14.10) en cualquier base cartesiana. En notación diádica,  $\mathbf{I}$  se escribe en la forma de la Ec. (14.14)

$$\mathbf{I} = m_i (r_i^2 \mathbf{1} - \mathbf{r}_i \mathbf{r}_i) \quad (14.27)$$

2. Tanto el momento de inercia como el tensor de inercia poseen un teorema de ejes paralelos, estos teoremas se escriben:

$$I = I_C + Ma^2 \quad (14.28)$$

$$\mathbf{I} = \mathbf{I}_C + M (R_C^2 \mathbf{1} - \mathbf{R}_C \mathbf{R}_C) \quad (14.29)$$

La Ec. (14.28) nos dice que si tenemos dos ejes paralelos uno de los cuales pasa por el centro de masa del cuerpo, siendo  $a$  la distancia entre los ejes y  $M$  la masa del cuerpo rígido, el momento de inercia en los dos ejes está relacionado por dicha ecuación. La Ec. (14.29) nos dice que si tenemos dos sistemas de ejes  $S \equiv X_1 X_2 X_3$  y  $S_C \equiv X_{1C} X_{2C} X_{3C}$  paralelos entre sí donde el origen de  $S_C$  yace sobre el centro de masa del cuerpo, los tensores de inercia medidos por ambos sistemas coordinados están relacionados por esta ecuación, siendo  $\mathbf{R}_C$  el vector posición del origen de  $S_C$  medido con respecto al origen de  $S$ .

3. A partir de la definición de momento de inercia (14.19), se vé que esta cantidad depende de la orientación y localización del eje con que se mide. El vector unitario  $\mathbf{n}$  define la dirección, pero como  $\mathbf{I}$  depende del origen (el cual se debe elegir de modo que el eje pase por él), tenemos la dirección del eje y un punto por donde pasa, lo cual define al eje unívocamente.

<sup>11</sup>Esta característica solo depende del carácter simétrico y real y no de su carácter positivo (ver sección 12.11.1).

<sup>12</sup>Si el origen que elegimos no pasa por el eje de simetría, el eje de simetría no será en general eje principal con respecto a dicho origen.

4. En el caso de cuerpos rígidos planos, tenemos un teorema de ejes perpendiculares: si la distribución de masa del cuerpo yace sobre el plano  $XY$  entonces los momentos de inercia con respecto a éstos ejes están relacionados por

$$I_x + I_y = I_z$$

Ahora bien, dado que el tensor de inercia es simétrico y real, se garantiza la existencia de ejes principales en los cuales su representación matricial es diagonal. Veremos que aspecto adquieren las ecuaciones antes vistas, cuando se expresan en esta base de ejes principales. Cuando calculamos el tensor sobre los ejes principales obtendremos una diada de la forma

$$\mathbf{I} = I_{kk} \mathbf{u}_k \mathbf{u}_k \equiv I_k \mathbf{u}_k \mathbf{u}_k \quad ; \quad k = 1, 2, 3$$

con un origen dado, la relación entre el momento angular y la frecuencia angular Ec. (14.13) se simplifica en la forma

$$L_k = I_k \omega_k \quad ; \quad k = 1, 2, 3 \quad (\text{no hay suma}) \quad (14.30)$$

de la misma manera la expresión para la energía cinética será

$$\begin{aligned} T &= \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{I} \cdot \boldsymbol{\omega} = \frac{1}{2} (\omega_m \mathbf{u}_m) \cdot (I_k \mathbf{u}_k \mathbf{u}_k) \cdot (\omega_n \mathbf{u}_n) \\ T &= \frac{1}{2} (\omega_m I_k \omega_n) (\mathbf{u}_m \cdot \mathbf{u}_k) (\mathbf{u}_k \cdot \mathbf{u}_n) = \frac{1}{2} (\omega_m I_k \omega_n) \delta_{mk} \delta_{kn} \\ T &= \frac{1}{2} I_k \omega_k^2 \end{aligned} \quad (14.31)$$

se observa además, que si el eje de rotación coincide con alguno de los ejes principales digamos  $X_3$ , entonces  $\omega_k = \omega \delta_{k3}$  entonces la Ec. (14.30) nos da

$$L_k = I_k \omega \delta_{k3} = (I_3 \omega) \delta_{k3}$$

es decir que solo sobrevive la componente 3, de modo que cuando la rotación es alrededor de alguno de los ejes principales, el momento angular es paralelo a la velocidad angular. Recordemos que la dirección de los ejes principales coincide con la dirección de los vectores propios (conociendo los vectores propios, solo hay que elegir un origen para determinar los ejes, pero este origen debe pasar por el eje de rotación), por lo tanto el efecto de  $\mathbf{I}$  sobre cualquier vector paralelo a un eje principal nos da otro vector también paralelo al eje principal, sea  $\mathbf{v}$  un vector paralelo a  $\mathbf{u}_j$

$$\begin{aligned} \mathbf{I} \cdot \mathbf{v} &= (I_k \mathbf{u}_k \mathbf{u}_k) \cdot (v \mathbf{u}_j) = I_k v \mathbf{u}_k (\mathbf{u}_k \cdot \mathbf{u}_j) = I_k v \mathbf{u}_k \delta_{jk} \\ \mathbf{I} \cdot \mathbf{v} &= I_j v \mathbf{u}_j \quad (\text{no suma}) \end{aligned}$$

lo cual demuestra la anterior afirmación. Este vector  $\mathbf{v}$  también es naturalmente vector propio de  $\mathbf{I}$  ya que la acción de un operador sobre uno de sus vectores propios es también un vector propio<sup>13</sup>.

Un concepto muy relacionado con el momento de inercia es el factor geométrico denominado **radio de giro** definido por

$$I \equiv MR_0^2 \quad (14.32)$$

es decir dado un eje  $X_O$ , el radio de giro es la distancia a este eje a la cual habría que colocar una masa puntual con la masa de todo el sólido, para que tuviese el mismo momento de inercia (con respecto a  $X_O$ ) que el sólido completo. Recordando la definición original del momento de inercia Ec. (14.19)

$$I = \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n}$$

y comparando con la Ec. (12.103) podemos retomar el análisis hecho en la sección 12.11.3. La Ec. (12.104) en este contexto nos queda

$$\rho \equiv \frac{\mathbf{n}}{\sqrt{I}} = \frac{\mathbf{n}}{R_0 \sqrt{M}} \quad (14.33)$$

<sup>13</sup>Existen en todo caso posibles diferencias relativas a las unidades.

El radio vector a un punto sobre el elipsoide de inercia es entonces inversamente proporcional a  $R_0$  con respecto a la dirección del vector. Las ecuaciones análogas a (12.106) y a (12.108) nos dan

$$I_{11}\rho_1^2 + I_{22}\rho_2^2 + I_{33}\rho_3^2 + 2I_{12}\rho_1\rho_2 + 2I_{13}\rho_1\rho_3 + 2I_{23}\rho_2\rho_3 = 1 \quad (14.34)$$

$$I_1\rho_1'^2 + I_2\rho_2'^2 + I_3\rho_3'^2 = 1 \quad (14.35)$$

siendo la última la ecuación escrita en la base que diagonaliza al tensor de inercia, en la cual el **elipsoide de inercia** adquiere su forma normal, lo cual geoméricamente significa que los ejes de la base (ejes principales) van a lo largo de los ejes del elipsoide.  $I_i$  son los valores propios del tensor de inercia (en cualquier base), en presencia de degeneración simple el elipsoide de inercia corresponderá a un elipsoide de revolución, en tanto que en el caso de degeneración completa el elipsoide se convierte en esfera.

El hecho de que el tensor de inercia y todas las cantidades asociadas a él tales como los ejes principales, momentos principales, elipsoide de inercia etc., dependan del punto particular fijo en el cuerpo (origen elegido), tiene como consecuencia que si cambiamos de punto en el cuerpo todas estas cantidades en general se modifican. La Ec. (14.29) nos indica la manera en que el tensor de inercia se modifica cuando pasamos del punto de referencia centro de masa a otro punto. En particular, la transformación que diagonaliza a  $\mathbf{I}'$  cuando el punto fijo es el centro de masa no diagonalizará en general al término diferencia  $M(R_C^2\mathbf{1} - \mathbf{R}_C\mathbf{R}_C)$  de modo que en general no definirá ejes principales en el nuevo punto. El tensor diferencia solo será también diagonal en el caso en el cual  $\mathbf{R}_C$  vaya a lo largo de uno de los ejes principales relativos al centro de masa. En tal caso, el nuevo tensor de inercia  $\mathbf{I}$  posee los mismos ejes principales que en el centro de masa, pero los momentos principales de inercia cambian, excepto en lo que corresponde al desplazamiento de ejes, donde es claro que el elemento diagonal del tensor diferencia será nulo. El teorema de los ejes paralelos para la forma diagonalizada del tensor de inercia tiene entonces una forma algo especial y restringida.

El análogo geométrico del elipsoide de inercia resulta muy cómodo para resolver al menos cualitativamente, algunos aspectos de la dinámica del cuerpo rígido, para esto resulta más conveniente parametrizar al vector  $\rho$  usando las Ecs. (14.18) y (14.17) en la Ec. (14.33)

$$\begin{aligned} \rho &= \frac{\mathbf{n}}{\sqrt{I}} = \frac{\boldsymbol{\omega}}{\omega\sqrt{I}} = \frac{\boldsymbol{\omega}}{\omega\sqrt{2T/\omega^2}} = \frac{\boldsymbol{\omega}}{\sqrt{2T}} \\ \rho &= \frac{\boldsymbol{\omega}}{\sqrt{2T}} \end{aligned} \quad (14.36)$$

en el sistema coordenado de los ejes principales del cuerpo, definimos la función

$$F(\rho) = \rho \cdot \mathbf{I} \cdot \rho = \rho_i^2 I_i \quad (14.37)$$

donde las superficies con  $F$  constante son elipsoides. De acuerdo con la Ec. (14.35), la superficie asociada a  $F = 1$  define el elipsoide de inercia (cuyos semiejes nos dan los momentos principales de inercia). La Ec. (14.36), nos muestra que a medida que el eje de rotación cambia en el tiempo, el vector  $\rho$  se mueve con él, y su extremo siempre está en un punto de la superficie del elipsoide<sup>14</sup>. El gradiente de  $F$  evaluado en ese punto nos da la dirección de la normal al elipsoide en ese punto. De acuerdo con (14.37) tal gradiente con respecto a  $\rho$  viene dado por

$$\begin{aligned} (\nabla_\rho F)_m &= \partial_m (\rho \cdot \mathbf{I} \cdot \rho) = \partial_m (\rho_k I_{kj} \rho_j) = (\partial_m \rho_k) I_{kj} \rho_j + \rho_k I_{kj} (\partial_m \rho_j) \\ (\nabla_\rho F)_m &= \delta_{mk} I_{kj} \rho_j + \rho_k I_{kj} \delta_{mj} = I_{mj} \rho_j + \rho_k I_{km} \end{aligned}$$

usando la simetría del tensor de inercia

$$(\nabla_\rho F)_m = I_{mj} \rho_j + I_{mk} \rho_k = 2I_{mk} \rho_k$$

<sup>14</sup>La Ec. (14.36), también muestra que si la energía cinética no es constante, el vector  $\rho$  no dictamina unívocamente la magnitud de  $\boldsymbol{\omega}$ , pero sigue emulando la dinámica de su orientación.



usando la Ec. (14.36), el gradiente se escribe (en el lenguaje de diadas) como

$$\begin{aligned}\nabla_{\rho} F &= 2\mathbf{I} \cdot \rho = \frac{2}{\sqrt{2T}} (\mathbf{I} \cdot \boldsymbol{\omega}) \\ \nabla_{\rho} F &= \sqrt{\frac{2}{T}} \mathbf{L}\end{aligned}\tag{14.38}$$

con lo cual se observa que  $\boldsymbol{\omega}$  se moverá de tal forma que la normal<sup>15</sup> al elipsoide de inercia (en el punto donde el extremo de  $\rho$  toca al elipsoide) va en la dirección del momento angular.

Es claro geoméricamente que en un punto de intersección entre el elipsoide de inercia y un eje principal, la normal a la superficie elipsoidal va a lo largo del eje principal. En consecuencia, cuando  $\rho$  va a lo largo de un eje principal, es paralelo a  $\nabla_{\rho} F$ , que a su vez implica que los vectores  $\boldsymbol{\omega}$  y  $\mathbf{L}$  son paralelos entre sí, en concordancia con lo que se encontró para rotación alrededor de un eje principal.

## 14.4. Ecuaciones de Euler para el movimiento de un cuerpo rígido

Con las herramientas anteriores estamos en capacidad de plantear ecuaciones de movimiento para un cuerpo rígido. Si existen ligaduras no holónomas tales efectos deben manejarse para cada problema particular. Quizás el caso más simple de ligadura no holónoma que se emplea en la mecánica del cuerpo rígido, ocurre en el caso de las **ligaduras por rodadura**, las cuales deben introducirse en la ecuaciones de movimiento por el método de los multiplicadores indeterminados de Lagrange.

La estrategia general es buscar un punto de referencia que permita separar el movimiento en una parte traslacional y otra rotacional. Cuando el cuerpo rígido posee un punto fijo respecto a un sistema inercial, este punto será una buena elección ya que solo tendremos que tratar con el movimiento rotacional alrededor de tal punto. Cuando no hay puntos fijos, casi siempre el mejor punto de referencia será el centro de masa del cuerpo, en virtud del teorema de Chasles y del hecho de que la energía cinética y el momento angular se pueden escribir como un término relativo a la traslación del centro de masa mas uno relativo a la rotación alrededor del centro de masa. La energía cinética por ejemplo tendrá la forma

$$T = \frac{1}{2} M V_C^2 + \frac{1}{2} I_C \omega^2$$

además ocurre que una separación similar es posible para el potencial en muchos problemas de aplicación. Esto implica que las ecuaciones de movimiento se podrán separar en una parte traslacional y otra rotacional que se resuelven por separado. Por las ecuaciones de movimiento de Newton, esto es posible utilizando las Ecs. (1.15) para la traslación del CM y las Ecs. (1.28) para la rotación alrededor del CM. Por otro lado, si el sistema es holónomo y conservativo la formulación de Lagrange o Hamilton puede ser apropiada. El Lagrangiano tendrá la forma

$$L(q, \dot{q}) = L_C(q_C, \dot{q}_C) + L_r(q_r, \dot{q}_r)$$

siempre que el potencial admita una separación similar. El primer término condensa los efectos de traslación y tendrá coordenadas y velocidades generalizadas del centro de masa. El segundo término condensa los efectos de rotación alrededor del CM y depende de variables angulares, que definen la orientación del cuerpo en cada instante. Esta estructura permite de nuevo separar los problemas traslacional y rotacional. Bien sea con el formalismo Newtoniano, Lagrangiano o Hamiltoniano, es muy ventajoso trabajar en el sistema de ejes principales del punto de referencia, que como ya vimos simplifica mucho la expresión para la energía cinética y el momento angular.

Para trabajar el movimiento rotacional casi siempre se usan los ángulos de Euler como coordenadas generalizadas. Sin embargo, cuando el movimiento rotacional es bidimensional, podemos definir un solo parámetro independiente de rotación y se simplifica el problema.

<sup>15</sup>Es claro que esta normal va hacia afuera del elipsoide en este punto, ya que el gradiente va en la dirección en la cual aumenta  $F$ .

Bien sea que nuestro punto de referencia sea un punto fijo o el centro de masa, podemos obtener las ecuaciones de movimiento de un cuerpo rígido empleando las Ecs. (1.15, 1.20) del formalismo Newtoniano. Consideramos entonces un sistema inercial con origen en el punto fijo, o un sistema de ejes no rotante (con respecto a un sistema inercial) con origen en el CM. En ambos casos es válida la Ec. (1.20) como se puede verificar con la Ec. (1.28), estos sistemas los llamaremos simplemente sistemas no rotantes  $S_0$ , de modo que se puede escribir

$$\left(\frac{d\mathbf{L}}{dt}\right)_{S_0} = \mathbf{N} \quad (14.39)$$

donde cambiamos la notación para el torque por  $\mathbf{N}$ . Podemos relacionar esta razón de cambio con aquella que mide un sistema fijo al cuerpo usando (13.80)

$$\left(\frac{d\mathbf{L}}{dt}\right)_{S_0} = \left(\frac{d\mathbf{L}}{dt}\right)_{S'} + \boldsymbol{\omega} \times \mathbf{L} \quad (14.40)$$

combinando (14.39) con (14.40) se obtiene

$$\left(\frac{d\mathbf{L}}{dt}\right)_{S'} + \boldsymbol{\omega} \times \mathbf{L} = \mathbf{N} \quad (14.41)$$

nótese que esta ecuación es estrictamente de carácter mixto en el sentido de que la razón de cambio de  $\mathbf{L}$  es medida por el sistema rotante fijo al cuerpo en tanto que los vectores  $\mathbf{N}$ ,  $\mathbf{L}$  y  $\boldsymbol{\omega}$  están medidos en el sistema no rotante, sin embargo las coordenadas de estos vectores se escribirán en la base de vectores unitarios asociados a  $S'$ . A pesar de este carácter mixto, esta ecuación resulta ser la más conveniente para extraer la dinámica del cuerpo rígido. Borraremos ahora el subíndice " $S'$ " sobreentendiendo que se usa la base de un sistema coordinado fijo al cuerpo y escribiremos esta ecuación en componentes<sup>16</sup>

$$\frac{dL_k}{dt} + \varepsilon_{kmn}\omega_m L_n = N_k$$

si tomamos ejes principales con respecto al punto de referencia, podemos emplear (14.30) para escribir

$$I_k \frac{d\omega_k}{dt} + \varepsilon_{kmn}\omega_m \omega_n I_n = N_k \quad (\text{no suma sobre } k) \quad (14.42)$$

no hay suma sobre  $k$  pero sí sobre los otros índices. Hemos usado el hecho de que los momentos principales de inercia son independientes del tiempo, lo cual a su vez está ligado al hecho de escribir los vectores en la base generada por  $S'$ . Vale decir sin embargo, que aunque todos los observables en esta ecuación están escritos en la base de ejes principales,  $\omega$ ,  $\mathbf{L}$  y  $\mathbf{N}$  como vectores están medidos en el sistema  $S_0$ . En forma expandida estas ecuaciones se escriben

$$\begin{aligned} I_1 \dot{\omega}_1 - \omega_2 \omega_3 (I_2 - I_3) &= N_1 \\ I_2 \dot{\omega}_2 - \omega_3 \omega_1 (I_3 - I_1) &= N_2 \\ I_3 \dot{\omega}_3 - \omega_1 \omega_2 (I_1 - I_2) &= N_3 \end{aligned} \quad (14.43)$$

Las Ecs. (14.42) o (14.43) se denominan **ecuaciones de Euler de movimiento para un cuerpo rígido** con un punto fijo. Estas ecuaciones definen un sistema acoplado, sin embargo en presencia de degeneración simple (e.g.  $I_1 = I_2 \neq I_3$  simetría cilíndrica) se observa un desacople parcial de la tercera ecuación ya que las componentes  $N_1$  y  $N_2$  del torque solo afectan a  $\omega_1$  y  $\omega_2$ . Cuando la degeneración es total (simetría esférica) las ecuaciones se desacoplan completamente, esto nos muestra de nuevo que la degeneración está asociada con simetrías del sistema.

Es necesario hacer notar que las ecuaciones de Euler solo contienen la descripción del movimiento rotacional, y deben ser complementadas con las ecuaciones de la dinámica traslacional del centro de masa o del punto

<sup>16</sup>Es necesario que todos los vectores estén escritos en la misma base para que sea válida la igualación componente a componente.

de referencia. Finalmente, es necesario señalar otro inconveniente con las Ecs. (14.43), las componentes de los vectores  $\boldsymbol{\omega}$  y  $\mathbf{N}$  que allí aparecen, están evaluadas en la base fija al cuerpo, de modo que para obtener el valor de cada uno de estos vectores necesitamos conocer la orientación relativa de los vectores unitarios base con respecto a los vectores base del sistema no rotante en el instante de tiempo en que evaluemos a estos vectores. Por lo tanto, necesitaremos en general conocer la dinámica del sistema de ejes rotante para determinar completamente a estos vectores en cualquier instante.

#### 14.4.1. Ecuaciones de Euler con el formalismo de Lagrange

Podemos también deducir estas ecuaciones a partir del formalismo de Lagrange partiendo de la Ec. (2.17)

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j$$

y dado que las coordenadas generalizadas serán angulares (e.g. los ángulos de Euler) las fuerzas generalizadas serán los torques asociados a tales ángulos. Sin embargo, solo uno de los ángulos de Euler tiene su torque asociado dirigido a lo largo de uno de los ejes del cuerpo y las otras dos ecuaciones deberán obtenerse por permutación cíclica. Comenzamos por escribir la energía cinética con base en los momentos principales

$$T = \frac{1}{2} I_k \omega_k^2 \quad (14.44)$$

como la base más conveniente para escribir las ecuaciones es la asociada a  $S'$ , emplearemos las relaciones (13.64) que nos dan las componentes de la velocidad angular en términos de los ángulos de Euler y en la base de los vectores fijos al cuerpo (ver sección 13.10)

$$\omega_1 = \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \quad ; \quad \omega_2 = \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \quad ; \quad \omega_3 = \dot{\phi} \cos \theta + \dot{\psi} \quad (14.45)$$

las coordenadas generalizadas son  $q_1 = \theta$ ,  $q_2 = \phi$ ,  $q_3 = \psi$ , los ángulos de Euler. Escribamos la ecuación de movimiento para  $\psi$

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\psi}} \right) - \frac{\partial T}{\partial \psi} &= N_\psi \\ \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{\psi}} \left( \frac{1}{2} I_k \omega_k^2 \right) \right] - \frac{\partial}{\partial \psi} \left( \frac{1}{2} I_k \omega_k^2 \right) &= N_\psi \\ \frac{d}{dt} \left( I_k \omega_k \frac{\partial \omega_k}{\partial \dot{\psi}} \right) - I_k \omega_k \frac{\partial \omega_k}{\partial \psi} &= N_\psi \end{aligned} \quad (14.46)$$

debemos calcular las cantidades  $\frac{\partial \omega_k}{\partial \dot{\psi}}$  y  $\frac{\partial \omega_k}{\partial \psi}$  para lo cual usamos (14.45)

$$\begin{aligned} \frac{\partial \omega_1}{\partial \dot{\psi}} &= \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi = \omega_2 \quad ; \quad \frac{\partial \omega_2}{\partial \dot{\psi}} = -\dot{\phi} \sin \theta \sin \psi - \dot{\theta} \cos \psi = -\omega_1 \\ \frac{\partial \omega_3}{\partial \dot{\psi}} &= \frac{\partial \omega_1}{\partial \dot{\psi}} = \frac{\partial \omega_2}{\partial \dot{\psi}} = 0 \quad ; \quad \frac{\partial \omega_3}{\partial \dot{\psi}} = 1 \end{aligned} \quad (14.47)$$

y reemplazando (14.47) en (14.46) resulta

$$\begin{aligned} \frac{d}{dt} \left( I_3 \omega_3 \frac{\partial \omega_3}{\partial \dot{\psi}} \right) - I_1 \omega_1 \frac{\partial \omega_1}{\partial \psi} - I_2 \omega_2 \frac{\partial \omega_2}{\partial \psi} &= N_\psi \\ I_3 \dot{\omega}_3 - I_1 \omega_1 (\omega_2) - I_2 \omega_2 (-\omega_1) &= N_\psi \\ I_3 \dot{\omega}_3 - \omega_1 \omega_2 (I_1 - I_2) &= N_3 \end{aligned}$$

que nos reproduce la tercera de las Ecs. (14.43), las otras ecuaciones se obtienen por permutación cíclica de ésta

$$\begin{aligned} I_2 \dot{\omega}_2 - \omega_3 \omega_1 (I_3 - I_1) &= N_2 \\ I_1 \dot{\omega}_1 - \omega_2 \omega_3 (I_2 - I_3) &= N_1 \end{aligned}$$

## 14.5. Precesión libre

Un problema muy sencillo pero con muchas aplicaciones consiste en la dinámica de un cuerpo rígido sin fuerzas ni torques netos. En este caso el centro de masa define un sistema inercial ya que tiene velocidad uniforme con respecto al laboratorio, por supuesto también es posible en este caso tomar cualquier punto fijo del cuerpo rígido con respecto al laboratorio como punto de referencia, siempre que no haya torque neto con respecto a este punto (recordemos que el torque depende en general del origen elegido). En este caso, con el origen colocado en el centro de masa, el movimiento será de rotación pura de modo que las ecuaciones de Euler me determinarían la dinámica completa del cuerpo rígido, y el momento angular solo se debe a la rotación alrededor del centro de masa. En ausencia de torques netos, las ecuaciones de Euler (14.43), se escriben en la forma

$$I_1\dot{\omega}_1 = \omega_2\omega_3(I_2 - I_3) \quad ; \quad I_2\dot{\omega}_2 = \omega_3\omega_1(I_3 - I_1) \quad ; \quad I_3\dot{\omega}_3 = \omega_1\omega_2(I_1 - I_2) \quad (14.48)$$

Ahora bien, ya hemos visto que para un sistema de partículas exento de torque neto externo, el momento angular es una constante de movimiento. Por otro lado, dado que la fuerza neta externa es cero, la energía potencial externa se puede tomar como cero y por tanto se debe conservar la energía propia (que al ser medida por el centro de masa, corresponde a la energía interna, ver sección 1.5.5 Pág. 11), la cual consta de la energía cinética mas la energía potencial interna, pero esta última se conserva en virtud de que las distancias relativas de las partículas permanecen constantes (que es precisamente la condición de cuerpo rígido), en consecuencia la energía cinética del sistema debe conservarse. Se concluye entonces que para un cuerpo rígido exento de torques y fuerzas netas externas, tenemos dos integrales primeras de movimiento correspondientes a la conservación del momento angular y la energía cinética. Con estas dos integrales es posible en principio integrar completamente las Ecs. (14.48) en términos de funciones elípticas. No obstante, será más ilustrativo usar el análogo geométrico del elipsoide (conocido como construcción de Poincot) que nos dará muchas características del movimiento sin resolver completamente el problema. Más adelante, abordaremos el problema algebraicamente para sólidos con simetría axial.

### 14.5.1. Construcción de Poincot para la precesión libre

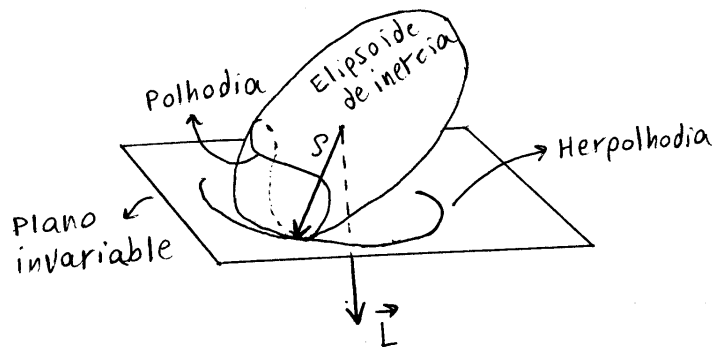


Figura 14.3: Ilustración de la construcción de Poincot para el estudio de la precesión libre. El elipsoide rueda sin deslizar sobre el plano invariable definido por el momento angular. La Polhodia y Herpolodia son las curvas trazadas sobre la superficie del elipsoide y el plano invariable respectivamente.

Ya hemos visto que en el sistema coordenado de los ejes principales del sólido, la superficie que forma el elipsoide de inercia nos da una relación entre  $\boldsymbol{\omega}$  y  $\mathbf{L}$ . En particular cuando la energía cinética es constante, el vector  $\boldsymbol{\rho}$  dado por (14.36) nos define completamente la dinámica del vector  $\boldsymbol{\omega}$ , y ya vimos que el extremo de  $\boldsymbol{\rho}$  define un punto sobre la superficie del elipsoide, y que la normal a la superficie del elipsoide en este punto determina precisamente la dirección del momento angular como se puede apreciar en (14.38).

En el caso particular de sólido libre, el momento angular es fijo en el sistema inercial y por tanto define la dirección de un plano, y dado que precisamente  $\mathbf{L}$  es perpendicular a la superficie del elipsoide en el punto donde está el extremo de  $\boldsymbol{\rho}$ , podemos definir unívocamente un plano normal a  $\mathbf{L}$  y tangente al elipsoide en el punto donde está ubicado el extremo de  $\boldsymbol{\rho}$  (que denominaremos plano invariable). A medida que va cambiando el valor de  $\boldsymbol{\omega}$  (i.e. de  $\boldsymbol{\rho}$ ), el elipsoide (que está fijo al cuerpo) se debe mover entonces sobre este plano tangente a fin de preservar la conexión entre  $\boldsymbol{\omega}$  y  $\mathbf{L}$ . La fig 14.3 muestra la configuración geométrica resultante. Adicionalmente, se puede mostrar que la distancia entre el origen del elipsoide y el plano tangente a éste en el extremo de  $\boldsymbol{\rho}$  debe ser constante en el tiempo. Esta distancia es la proyección de  $\boldsymbol{\rho}$  sobre la dirección de  $\mathbf{L}$

$$\boldsymbol{\rho} \cdot \hat{\mathbf{L}} = \frac{\boldsymbol{\rho} \cdot \mathbf{L}}{L} = \frac{\boldsymbol{\omega} \cdot \mathbf{L}}{L\sqrt{2T}} = \frac{\sqrt{2T}}{L}$$

donde hemos usado las Ecs. (14.16, 14.36). Es claro que esta cantidad es constante en nuestro caso ya que  $T$  y  $L$  son constantes. En consecuencia, el centro del elipsoide está a una “altura” constante con respecto al plano invariable. Adicionalmente, el centro del elipsoide coincide con el origen de todos los sistemas de referencia que estamos usando ya que este origen se eligió para que no exista traslación respecto al sistema inercial del centro de masa (o punto fijo), por tanto dicho punto es estacionario con respecto al plano invariable. De lo anterior se puede ver que el movimiento libre de fuerzas de un cuerpo rígido se puede representar considerando que es el del elipsoide de inercia rodando sin deslizar sobre el plano invariable, con el centro del elipsoide fijo y a una altura constante en el plano. La rodadura tiene lugar sin deslizamiento puesto que el punto de contacto está definido por el extremo de  $\boldsymbol{\rho}$ , el cual por estar dirigido según el eje instantáneo de rotación, es la dirección del cuerpo que está momentáneamente en reposo. La curva trazada por el punto de contacto sobre la superficie del elipsoide de inercia se denomina **polhodia**. En tanto que la curva que traza el punto de contacto sobre el plano invariable se denomina **herpolhodia**. Ocurre entonces que la polhodia rueda sin deslizar sobre la herpolhodia que está en el plano invariable.

La construcción de Poinsot es adecuada para describir completamente el movimiento de un cuerpo rígido libre de fuerzas, ya que la dirección del plano invariable y la altura del elipsoide de inercia se determinan con los valores de  $T$  y  $\mathbf{L}$ , las cuales están entre las condiciones iniciales del problema. El problema se reduce entonces al problema geométrico de trazar la polhodia y la herpolhodia. Es notable el hecho de que la herpolhodia es siempre cóncava hacia el origen de aquí su nombre que significa “en forma de serpiente”. La dirección y magnitud de la velocidad angular están dictaminadas por el vector  $\boldsymbol{\rho}$  como se puede ver de (14.36) teniendo en cuenta que en nuestro caso  $T$  es constante. Por otro lado, la orientación instantánea del cuerpo está determinada por la orientación del elipsoide de inercia, ya que éste está fijo al cuerpo. Muchas descripciones del movimiento libre de un cuerpo rígido basadas en estos argumentos se pueden encontrar en la literatura.

Es claro sin embargo, que este análisis no permite encontrar la dependencia paramétrica de la polhodia o la herpolhodia con el tiempo. Además, tampoco garantiza que todos los puntos de dichas curvas sean trazados en el movimiento real, ya que es posible que solo una parte de dichas curvas se trace con un conjunto dado de condiciones iniciales. En tal sentido, este análisis tiene limitaciones similares a las discutidas en las secciones 10.3 y 10.4, concernientes al análisis de curvas de potencial efectivo.

### Precesión libre de cuerpos con simetría axial

En el caso especial de simetría axial del cuerpo (que se representa por la degeneración simple de los valores propios del tensor de inercia), el elipsoide de inercia se convierte en un elipsoide de revolución, de modo que claramente la polhodia sobre el elipsoide es un círculo alrededor del eje de simetría. De la misma manera, la herpolhodia será también un círculo sobre el plano invariable. Un observador fijo al cuerpo verá que  $\boldsymbol{\omega}$  se mueve sobre la generatriz de un cono (denominado el **cono del cuerpo**) cuya intersección con el elipsoide de inercia forma la polhodia. De manera análoga, un observador fijo en los ejes inerciales ve que  $\boldsymbol{\omega}$  se mueve sobre la superficie de un **cono del espacio** cuya intersección con el plano invariable genera la herpolhodia. Por esta razón es usual describir el movimiento libre de un sólido con simetría axial como el generado por el rodamiento del cono del cuerpo sobre el cono del espacio.

Adicionalmente, si el momento de inercia alrededor del eje de simetría es menor que aquél correspondiente a los otros dos ejes principales, se vé a través de la Ec. (12.108) que el elipsoide sería **prolato** i.e. la longitud a lo largo del eje de simetría sería mayor a la longitud de la sección de revolución (alargado como un balón de rugby). Si por el contrario, el momento de inercia alrededor del eje de simetría es mayor, el elipsoide es **oblato** (achatado).

En el caso de elipsoides prolatos, el cono del cuerpo es exterior al cono del espacio, para elipsoides oblatos el cono del cuerpo rueda sobre el interior del cono del espacio. No obstante, en ambos casos se observa que la dirección de  $\omega$  precesa en el tiempo alrededor del eje de simetría del cuerpo. Por esta razón hablamos del fenómeno de **precesión libre**.

Puede verse además, que si la rotación es alrededor de un eje principal, un eje del elipsoide coincide con la dirección de  $\rho$  (y por tanto de  $\omega$ ) y es perpendicular al plano invariable. En consecuencia  $\omega$  es paralelo a  $\mathbf{L}$  como ya se vió con otros métodos, y la precesión de  $\omega$  desaparece ya que dicho vector será constante.

### 14.5.2. Elipsoide de Binet y el momento angular

La construcción de Poinsot nos dice como se mueve  $\omega$  con respecto al sistema inercial y con respecto al sistema de ejes del cuerpo. Sin embargo, esta construcción no nos describe la dinámica del vector  $\mathbf{L}$  vista por el sistema de ejes del cuerpo<sup>17</sup>. Para describir el movimiento de  $\mathbf{L}$  medido por un observador que se mueve con los ejes principales del cuerpo, recurrimos a otro artificio geométrico. Combinando las Ecs. (14.30) y (14.31), vemos que en los ejes principales, la energía cinética se puede escribir como

$$T = \frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3} \quad (14.49)$$

puesto que  $T$  es constante, esta relación define un elipsoide, referido como el **elipsoide de Binet**, que también está fijo en los ejes del cuerpo pero que NO coincide con el elipsoide de inercia<sup>18</sup>. Si adoptamos la convención  $I_3 \leq I_2 \leq I_1$  y escribimos la ecuación para el elipsoide en forma estándar

$$\frac{L_1^2}{2TI_1} + \frac{L_2^2}{2TI_2} + \frac{L_3^2}{2TI_3} = 1 \quad (14.50)$$

el elipsoide tendrá semiejes que en orden decreciente corresponden a  $\sqrt{2TI_1}$ ,  $\sqrt{2TI_2}$ ,  $\sqrt{2TI_3}$ . Por otro lado, dado que se conserva la magnitud del momento angular podemos escribir una ecuación de la forma

$$\frac{L_1^2 + L_2^2 + L_3^2}{L^2} = 1 \quad (14.51)$$

que corresponde a una esfera de radio  $L$  en el espacio  $L_1, L_2, L_3$ . El vector  $\mathbf{L}$  se mueve de tal manera que describe un camino tanto sobre el elipsoide (14.50) como sobre la esfera (14.51). Es decir, el camino de  $\mathbf{L}$  visto por un observador fijo al cuerpo, es la intersección del elipsoide y la esfera<sup>19</sup>. Naturalmente, para que esta imagen sea consistente es necesario garantizar que dicha intersección existe, y es claro que esto equivale a demostrar que el valor del radio  $L$  de la esfera está entre el valor del semieje menor y el semieje mayor. Cada semieje del elipsoide está dado por la cantidad

$$a_i \equiv \sqrt{2TI_i}$$

ahora usando convenio de suma de índices, reescribimos (14.50) en la forma

$$\frac{L_i L_i}{a_i a_i} = 1$$

<sup>17</sup>Es claro que con respecto al sistema de ejes inerciales,  $\mathbf{L}$  es constante.

<sup>18</sup>Comparando las Ecs. (14.35) y (14.50), es inmediato ver que las longitudes de los semiejes del elipsoide de Binet son diferentes a las del elipsoide de Poinsot. Incluso las "longitudes" de los ejes tienen dimensiones diferentes.

<sup>19</sup>El elipsoide (14.50) es una manifestación de la conservación de  $T$ , en tanto que la esfera (14.51) es una manifestación de la conservación de  $\|\mathbf{L}\|$ .

multiplicando la ecuación por  $a_j^2$  siendo  $j$  un índice fijo  $j = 1, 2, 3$  se obtiene

$$a_j^2 = L_i L_i \frac{a_j^2}{a_i a_i}$$

si  $a_j$  es el menor de los semiejes principales (por convención  $a_3$ ) tendremos que  $a_j/a_i \leq 1$  y por tanto

$$a_3^2 = L_1^2 \frac{a_3^2}{a_1^2} + L_2^2 \frac{a_3^2}{a_2^2} + L_3^2 \frac{a_3^2}{a_3^2} \leq L_1^2 + L_2^2 + L_3^2 = L^2$$

similarmente para el mayor de los semiejes se puede probar

$$a_1^2 \geq L^2$$

por lo tanto

$$\sqrt{2TI_3} \leq L \leq \sqrt{2TI_1} \tag{14.52}$$

que es lo que se quería demostrar. Si  $a_3 \neq a_1$ , la esfera está fuera del elipsoide sobre el eje  $L_3$  y dentro del elipsoide en el eje  $L_1$ <sup>20</sup>. Las Figs. 14.4 muestran las curvas que resultan de la intersección del elipsoide con la esfera para diferentes valores de  $L$  y desde diferentes perspectivas. Las curvas que aparecen como líneas rectas en la Fig. 14.4 corresponden a  $L = \sqrt{2TI_2}$ .

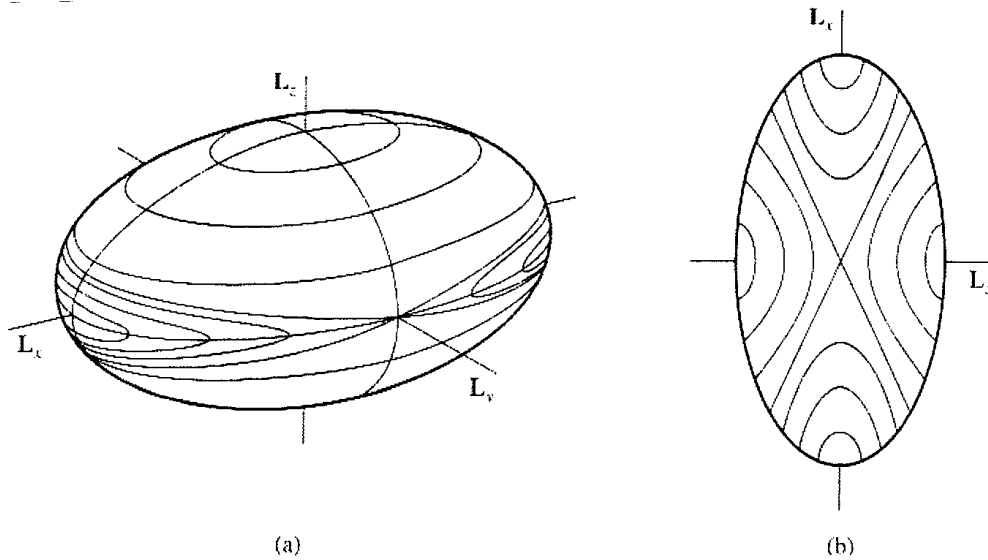


Figura 14.4: Ilustración de las trayectorias posibles del vector  $\mathbf{L}$  sobre la superficie del elipsoide de Binet fijo al cuerpo, para el estudio de la precesión libre.

### 14.5.3. Elipsoide de Binet, rotación estacionaria y condiciones para la rotación estable

Las ecuaciones de Euler para sólido libre (14.48) muestran claramente que la rotación estacionaria (i.e. con todas las componentes de  $\boldsymbol{\omega}$  constantes) solo es posible para un sólido libre de fuerzas si se cumple la condición

$$\omega_1 \omega_2 (I_1 - I_2) = \omega_2 \omega_3 (I_2 - I_3) = \omega_3 \omega_1 (I_3 - I_1) = 0$$

<sup>20</sup>Si  $a_3 = a_1$  entonces por definición todos los momentos de inercia serán iguales y el elipsoide coincide con la esfera.

Recordemos que  $\omega_i$  son componentes medidas en la base de los ejes principales del cuerpo, por tanto la estacionaridad se está midiendo usando una base de ejes fijos al cuerpo. Si en general suponemos que todos los momentos principales de inercia son diferentes como es de esperarse con un sólido irregular, vemos que el movimiento estacionario para dicho sólido solo es posible si al menos dos de las componentes de  $\omega$  son cero, es decir  $\omega$  va a lo largo de un solo eje principal<sup>21</sup>. Si existe degeneración simple por ejemplo  $I_1 = I_2$  solo existe la condición de que  $\omega_3 = 0$  de modo que  $\omega$  puede estar en cualquier dirección dentro del plano  $X_1X_2$  lo cual es de esperarse ya que cualquier eje sobre ese plano es eje principal. Si la degeneración es total  $\omega$  puede estar en cualquier dirección ya que cualquier eje es principal en virtud de la simetría esférica.

Retornemos al caso no degenerado en donde los ejes principales forman un conjunto con una orientación bien definida. En este caso como ya vimos el movimiento solo puede ser estacionario si hay rotación alrededor de alguno de los ejes principales. Una pregunta natural que surge en toda condición estacionaria es su estabilidad, es decir la desviación del comportamiento estacionario ante pequeñas perturbaciones, para lo cual examinaremos lo que ocurre en cada eje. Supongamos que como condición inicial la rotación comienza alrededor del tercer eje principal (recordemos que por convención  $I_3$  es el menor de los tres momentos de inercia), ocurrirá movimiento estable alrededor de  $L_3$  si comenzando con  $\mathbf{L}$  en la dirección  $L_3$  garantizamos que no hay otras orientaciones posibles en la vecindad del eje  $L_3$ ; geoméricamente esto significa que la esfera y el elipsoide de Binet solo se deben intersectar en un punto en una vecindad de  $L_3$  de modo que el radio de la esfera debe coincidir con el valor del semieje menor, es decir  $L^2 = 2TI_3$ . Cuando hay una pequeña desviación de este valor, el radio de la esfera del momento angular es solo un poco mayor de este valor<sup>22</sup>, y la intersección con el elipsoide de la energía cinética será una pequeña figura cerrada alrededor del eje  $L_3$ . El movimiento es entonces estable ya que una pequeña perturbación produce una desviación ligera de  $\mathbf{L}$  alrededor del eje original.

Una situación similar ocurre en el otro extremo en el cual el movimiento se hace alrededor del eje de mayor momento de inercia. El movimiento estable ocurre cuando  $L^2 = 2TI_1$  y cuando este movimiento es perturbado, el radio de la esfera del momento angular es solo ligeramente menor que este valor. La intersección es de nuevo un pequeño círculo alrededor del eje principal y el movimiento es estable. En contraste, para el eje intermedio  $L_2$ , la energía cinética tiene dos órbitas que circundan al elipsoide y se cruzan entre sí donde  $\pm L_2$  pasa a través del elipsoide (ver Figs. 14.4). En consecuencia, hay dos órbitas diferentes con valores ligeramente menores que  $\sqrt{2TI_2}$  y otras dos órbitas diferentes con valores ligeramente mayores que  $\sqrt{2TI_2}$ <sup>23</sup>, todas las cuatro órbitas tienen caminos largos y apartados sobre la superficie del elipsoide (Figs. 14.4). Este es un fenómeno típico de bifurcación que por supuesto implica inestabilidad.

Lo anterior se puede entender teniendo en cuenta que para el eje con momento intermedio el radio de curvatura del elipsoide en una dirección es mas grande que el de la esfera de contacto en tanto que es menor en la dirección perpendicular. En los otros extremos en cambio los radios de curvatura son o bien mas grandes o bien más pequeños que el radio de la esfera en todas las direcciones. Estos criterios de estabilidad son de gran utilidad en aplicaciones concernientes a la estabilidad de sistemas rotantes tales como naves espaciales. No obstante, cuando hay fuerzas disipativas presentes, estos argumentos de estabilidad deben ser modificados. Por ejemplo, para un cuerpo con  $L$  constante pero con  $T$  suavemente decreciente, la única rotación estable es alrededor del eje principal con mayor momento de inercia. Esto se puede ver teniendo en cuenta que la energía cinética de rotación alrededor de un eje principal  $i$ -ésimo viene dada por  $T = L^2/2I_i$  la cual tiene un valor mínimo para el eje con mayor momento de inercia, y por tanto conduce a un movimiento estable alrededor de tal eje. Cuando se pone a rotar al cuerpo alrededor de otro de los ejes principales el efecto de decrecimiento ligero de la energía cinética es el de cambiar gradualmente el eje de rotación hasta que termina rotando alrededor del eje con máximo valor de momento de inercia.

<sup>21</sup>En tal caso  $\mathbf{L}$  es paralelo a  $\omega$ , y como  $\mathbf{L}$  es constante, también lo será  $\omega$ .

<sup>22</sup>La Ec. (14.52) nos garantiza que el radio de la esfera no puede ser menor que el semieje menor del elipsoide de Binet. Similarmente tampoco puede ser mayor que el semieje mayor. Estrictamente, tanto la esfera como el elipsoide pueden cambiar, por ejemplo si se introduce una pequeña fuerza que produce un pequeño torque por un breve instante, puede ocurrir en general un ligero cambio en el momento angular y también en la energía cinética y ambas figuras deben cambiar. En cualquier caso al volver a ser libre el cuerpo rígido se tiene que mantener la relación (14.52) de modo que no cambia nuestra discusión.

<sup>23</sup>Nótese que en este caso el radio de la esfera puede aumentar o disminuir dentro de cierto intervalo sin violar la condición (14.52).



#### 14.5.4. Solución algebraica para la precesión libre con simetría axial

Para un cuerpo simétrico libre de fuerzas, no es difícil obtener la solución analítica de las ecuaciones de Euler. Tomemos el eje de simetría como el eje principal  $L_3$  de modo que  $I_1 = I_2$ . Las ecuaciones de Euler (14.48) se reducen entonces a

$$I_1\dot{\omega}_1 = \omega_2\omega_3(I_1 - I_3) \quad ; \quad I_1\dot{\omega}_2 = \omega_3\omega_1(I_3 - I_1) \quad ; \quad I_3\dot{\omega}_3 = 0 \quad (14.53)$$

la última de estas ecuaciones establece que  $\omega_3$  es constante, y se puede tratar como una condición inicial conocida del problema. Las dos ecuaciones siguientes se escriben como

$$\dot{\omega}_1 = -\Omega\omega_2 \quad ; \quad \dot{\omega}_2 = \Omega\omega_1 \quad ; \quad \Omega \equiv \frac{I_3 - I_1}{I_1}\omega_3 \quad (14.54)$$

derivando la primera de estas ecuaciones y teniendo en cuenta la segunda, se obtiene

$$\ddot{\omega}_1 = -\Omega\dot{\omega}_2 = -\Omega^2\omega_1 \quad \Rightarrow \quad \ddot{\omega}_1 = -\Omega^2\omega_1$$

esta ecuación tiene la típica solución armónica

$$\omega_1 = A \cos(\Omega t + \delta) \quad (14.55)$$

y sustituyendo esta expresión en la segunda de las Ecs. (14.54) resulta

$$\dot{\omega}_2 = \Omega A \cos(\Omega t + \delta) \quad \Rightarrow \quad \omega_2 = A \sin(\Omega t + \delta) \quad (14.56)$$

estas soluciones muestran que  $\omega_3\mathbf{u}_3$  es constante en tanto que  $\omega_1\mathbf{u}_1 + \omega_2\mathbf{u}_2$  es constante en magnitud y rota uniformemente alrededor del eje de simetría  $X_3$  del cuerpo con frecuencia angular  $\Omega$ . Esto indica entonces que el vector  $\boldsymbol{\omega}$  completo es constante en magnitud y precesa alrededor de  $X_3$  con la misma frecuencia, tal como lo predijo la construcción de Poinsot.

Otra forma interesante de ver la precesión es construyendo un vector  $\boldsymbol{\Omega}$  a lo largo de  $X_3$  y con magnitud descrita por la tercera de las Ecs. (14.54). En tal caso, las Ecs. (14.53) se pueden sintetizar en una ecuación vectorial

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\Omega} \times \boldsymbol{\omega}$$

la cual muestra inmediatamente la precesión de  $\boldsymbol{\omega}$  con frecuencia  $\boldsymbol{\Omega}$ .

Es importante tener en cuenta que la precesión aquí descrita es vista por los ejes del cuerpo, que a su vez rotan en el espacio con una frecuencia  $\omega$  que es usualmente mayor que  $\Omega$ . De otra parte, la tercera de las Ecs. (14.54) nos dice que entre más cerca esté  $I_1$  de  $I_3$  más lenta será la precesión i.e. menor será  $\Omega$  con respecto a  $\omega$ . Las constantes  $A$  (amplitud de precesión) y  $\omega_3$  se pueden evaluar en términos de las constantes de movimiento más usuales como son  $T$  y  $L^2$ , es fácil ver que estas cantidades se pueden escribir en términos de  $A$  y  $\omega_3$  de la siguiente forma. Empleemos primero las Ecs. (14.30, 14.31) para ejes principales con  $I_1 = I_2$  y tendremos en cuenta las Ecs. (14.55, 14.56)

$$\begin{aligned} T &= \frac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \frac{1}{2}I_3\omega_3^2 = \frac{1}{2}I_1[A^2 \cos^2(\Omega t + \delta) + A^2 \sin^2(\Omega t + \delta)] + \frac{1}{2}I_3\omega_3^2 \\ T &= \frac{1}{2}I_1A^2 + \frac{1}{2}I_3\omega_3^2 \\ L^2 &= (\mathbf{I} \cdot \boldsymbol{\omega}) \cdot (\mathbf{I} \cdot \boldsymbol{\omega}) = [(I_k \mathbf{u}_k \mathbf{u}_k) \cdot (\omega_m \mathbf{u}_m)] \cdot [(I_j \mathbf{u}_j \mathbf{u}_j) \cdot (\omega_n \mathbf{u}_n)] \\ &= [I_k \omega_m \mathbf{u}_k (\mathbf{u}_k \cdot \mathbf{u}_m)] \cdot [I_j \omega_n \mathbf{u}_j (\mathbf{u}_j \cdot \mathbf{u}_n)] = [I_k \omega_m \mathbf{u}_k \delta_{km}] \cdot [I_j \omega_n \mathbf{u}_j \delta_{jn}] \\ &= [I_k \omega_k \mathbf{u}_k] \cdot [I_j \omega_j \mathbf{u}_j] = I_k I_j \omega_k \omega_j (\mathbf{u}_k \cdot \mathbf{u}_j) = I_k I_j \omega_k \omega_j \delta_{kj} = I_k^2 \omega_k^2 \\ &= I_1^2(\omega_1^2 + \omega_2^2) + I_3^2 \omega_3^2 = I_1^2 A^2 + I_3^2 \omega_3^2 \end{aligned}$$

quedando finalmente

$$T = \frac{1}{2}I_1A^2 + \frac{1}{2}I_3\omega_3^2 \quad ; \quad L^2 = I_1^2A^2 + I_3^2\omega_3^2 \quad (14.57)$$

y estas relaciones se pueden invertir para obtener  $A$  y  $\omega_3$  en términos de  $T$  y  $L^2$ .

Por ejemplo, es de esperarse que la tierra sufra una precesión aproximadamente libre ya que los torques externos son muy débiles. La tierra es aproximadamente simétrica alrededor del eje polar  $X_3$  y oblata en los polos, de modo que  $I_1 \simeq I_2 < I_3$ . Numéricamente, el radio de estos momentos es tal que

$$\frac{I_3 - I_1}{I_1} = 0,00327$$

y la magnitud de la frecuencia angular de la precesión debe ser

$$\Omega = \frac{\omega_3}{305,81039} \cong \frac{\omega_3}{306}$$

y dado que  $\omega_3$  es casi igual en magnitud que  $\omega$  este resultado predice un periodo de precesión de unos 306 días es decir unos 10 meses. Si alguna perturbación ocurre sobre el eje de figura de la tierra, esperaríamos que el eje de rotación precese alrededor del eje de figura (alrededor del polo norte) una vez cada diez meses. En realidad se ha medido una precesión alrededor del polo con una amplitud del orden de unas pocas décimas de segundo de latitud, aunque modelos más realistas deben incluir otros efectos tales como fuerzas de fricción que se producen por la no rigidez del planeta.

#### 14.5.5. Estabilidad de sólidos irregulares con precesión libre por métodos algebraicos

Retornemos al caso de sólidos irregulares para los cuales los tres momentos de inercia a lo largo de ejes principales son diferentes. Veremos el problema de la estabilidad directamente con las ecuaciones de Euler y sus soluciones. Asumiremos que en  $t = 0$  el movimiento comienza con  $\omega_2 = \omega_3 = 0$  y  $\omega_1 \neq 0$ . Y que inmediatamente después de una breve perturbación,  $\omega_2$  y  $\omega_3$  adquieren valores no nulos pero mucho menores que  $\omega_1$ . Cuando la perturbación termina, el movimiento vuelve a ser libre de fuerzas y torques de modo que las ecuaciones de Euler quedan

$$I_1 \dot{\omega}_1 = \omega_2 \omega_3 (I_2 - I_3) \quad (14.58)$$

$$I_2 \dot{\omega}_2 = \omega_3 \omega_1 (I_3 - I_1) \quad (14.59)$$

$$I_3 \dot{\omega}_3 = \omega_1 \omega_2 (I_1 - I_2) \quad (14.60)$$

dado que  $\omega_2$  y  $\omega_3$  son pequeños con respecto a  $\omega_1$  el término a la derecha de la ecuación (14.58) es de segundo orden y se puede despreciar, por tanto tenemos que a primer orden  $I_1 \dot{\omega}_1 = 0$  y  $\omega_1$  se puede considerar constante. Derivando (14.59)

$$I_2 \ddot{\omega}_2 = (\dot{\omega}_3 \omega_1 + \omega_3 \dot{\omega}_1) (I_3 - I_1) \approx \dot{\omega}_3 \omega_1 (I_3 - I_1) \quad (14.61)$$

y sustituyendo (14.60) en (14.61) resulta

$$\begin{aligned} I_2 \ddot{\omega}_2 &= \left[ \omega_1 \omega_2 \frac{(I_1 - I_2)}{I_3} \right] \omega_1 (I_3 - I_1) = \left[ \frac{(I_1 - I_2)(I_3 - I_1)}{I_3} \right] \omega_1^2 \omega_2 \Rightarrow \\ \ddot{\omega}_2 &= A \omega_2 \quad ; \quad A \equiv \left[ \frac{(I_1 - I_2)(I_3 - I_1)}{I_2 I_3} \right] \omega_1^2 \end{aligned} \quad (14.62)$$

si  $I_1$  es el mayor o el menor de los tres momentos de inercia  $A$  será negativo y la Ec. (14.62) corresponde a la ecuación de un oscilador armónico simple. En tal caso,  $\omega_2$  oscila con frecuencia  $\sqrt{A}$  con amplitud acotada, y es fácil demostrar que  $\omega_3$  también realiza movimiento armónico simple, en tanto que  $\omega_1$  es constante a primer orden. Esto implica que el vector  $\omega$  solo sufre una ligera desviación de su valor original. En contraste, si  $I_1$  es el eje intermedio,  $A$  será positivo y las componentes  $\omega_2$  y  $\omega_3$  tenderán a crecer exponencialmente con el tiempo, dando un movimiento inestable. Esta descripción coincide con el análisis cualitativo de la sección 14.5.3, basado en el elipsoide de Binet.

## 14.6. La peonza simétrica pesada con un punto fijo

Consideraremos ahora un ejemplo un tanto más complejo pero muy ilustrativo y de amplias aplicaciones: la peonza simétrica pesada en un campo gravitacional uniforme con un punto fijo sobre el eje de simetría del cuerpo. Este ejemplo permite estudiar diversos sistemas que van desde los trompos infantiles hasta los instrumentos giroscópicos de navegación. Puesto que la mayor parte del tiempo utilizaremos los ejes fijos al cuerpo, denotaremos  $X_1X_2X_3$  a dichos ejes, en tanto que los ejes inerciales se denotarán por  $X'_1X'_2X'_3$  contrario a lo tradicional. La configuración se muestra en la Fig. 14.5

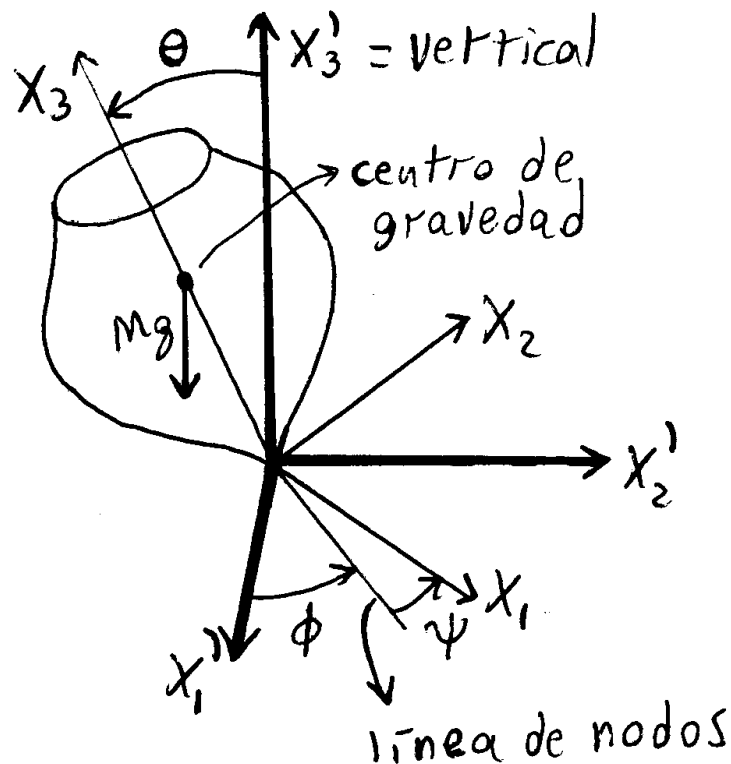


Figura 14.5: Peonza simétrica pesada con un punto fijo que no está en el centro de masa. El origen de ambos sistemas coordenados se ubica en el punto fijo de la peonza. El campo gravitacional va en la dirección  $-X'_3$ .

### 14.6.1. Planteamiento del Lagrangiano

El eje de simetría es claramente un eje principal del cuerpo. Dado que un punto es estacionario, la configuración de la peonza está completamente especificada por los tres ángulos de Euler:  $\theta$  nos da la inclinación del eje  $X_3$  (eje de simetría fijo al cuerpo) con respecto a la vertical i.e. con respecto al eje  $X'_3$  asociado al sistema inercial;  $\phi$  mide el azimuth de la peonza alrededor de la vertical<sup>24</sup>, en tanto que  $\psi$  mide el ángulo de rotación de la peonza alrededor del eje de simetría  $X_3$ , (ver Fig. 14.5). La distancia entre el centro de masa y

<sup>24</sup>Recordemos que  $\phi$  como ángulo de Euler mide el ángulo entre el eje inercial  $X'_1$  y la línea de nodos (intersección del plano  $X'_1X'_2$  con el plano  $X_1X_2$ ). Este ángulo es diferente al ángulo azimutal  $\varphi$  que se usaría en coordenadas esféricas si midiéramos el vector posición del centro de masa de la peonza. Sin embargo es claro que  $\dot{\phi} = \dot{\varphi}$ , de modo que ambos ángulos medirían la misma precesión.

el punto fijo, ambos sobre el eje de simetría está dada por  $l$ . La razón de cambio de cada ángulo nos da las características del movimiento de la peonza

1.  $\dot{\psi}$  mide la rotación (o spinning) de la peonza alrededor de su eje de figura  $X_3$ .
2.  $\dot{\phi}$  mide la precesión o rotación del eje de figura  $X_3$  alrededor del eje vertical  $X'_3$ .
3.  $\dot{\theta}$  mide la nutación, o cabeceo arriba y abajo del eje de figura  $X_3$  relativo al eje vertical  $X'_3$ .

En la mayoría de casos de interés estas razones de cambio están en la siguiente jerarquía:  $\dot{\psi} \gg \dot{\theta} \gg \dot{\phi}$ . Puesto que  $I_1 = I_2 \neq I_3$  las ecuaciones de Euler (14.43) quedan

$$\begin{aligned} I_1 \dot{\omega}_1 + \omega_2 \omega_3 (I_3 - I_1) &= N_1 \\ I_2 \dot{\omega}_2 + \omega_1 \omega_3 (I_1 - I_3) &= N_2 \\ I_3 \dot{\omega}_3 &= N_3 \end{aligned} \quad (14.63)$$

consideremos el caso en el cual inicialmente  $N_3 = N_2 = 0$ ,  $N_1 \neq 0$  y  $\omega_1 = \omega_2 = 0$ ,  $\omega_3 \neq 0$ , con lo cual  $\omega_3$  será constante. El torque  $N_1$  hace que  $\omega_1$  cambie ya que debido a este torque  $\dot{\omega}_1 \neq 0$  incluso al principio, como se puede ver de la primera ecuación (14.63) teniendo en cuenta que  $\omega_2$  es nulo al principio. Dado que para tiempos posteriores a  $t = 0$  se tiene que  $\omega_1$  ya no es nulo, la segunda ecuación requiere que  $\omega_2$  también comience a cambiar. Sin embargo, no es fácil descifrar el significado de esta dinámica puesto que lo que se mide en el laboratorio son los cambios en los ángulos de Euler  $\psi$ ,  $\phi$ ,  $\theta$  y sus ángulos asociados en el sistema  $X'_1 X'_2 X'_3$  del laboratorio. No se observa directamente a  $\dot{\omega}_1$ ,  $\dot{\omega}_2$ ,  $\dot{\omega}_3$  y sus ángulos asociados en los ejes principales. Esto sugiere que las ecuaciones de Euler no nos dan la descripción más útil del movimiento del sistema. En consecuencia, utilizaremos el formalismo Lagrangiano en lugar de las ecuaciones de Euler. El origen tanto del sistema inercial como del fijo al cuerpo está en el punto fijo de la peonza, de modo que para el sistema inercial con este origen hay una rotación pura alrededor de algún eje que pasa por este origen y que nos lleva al sistema de ejes fijo al cuerpo (teorema de Euler). La energía cinética para cuerpos simétricos en rotación pura se escribe como

$$T = \frac{1}{2} I_1 (\omega_1^2 + \omega_2^2) + \frac{1}{2} I_3 \omega_3^2 \quad (14.64)$$

como la idea es escribirla en términos de los ángulos de Euler, usaremos las Ecs. (13.64)

$$\begin{aligned} \omega_1^2 + \omega_2^2 &= \left( \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \right)^2 + \left( \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \right)^2 \\ &= \dot{\phi}^2 \sin^2 \theta \sin^2 \psi + \dot{\theta}^2 \cos^2 \psi + 2\dot{\theta}\dot{\phi} \sin \theta \sin \psi \cos \psi + \dot{\phi}^2 \sin^2 \theta \cos^2 \psi + \dot{\theta}^2 \sin^2 \psi - 2\dot{\theta}\dot{\phi} \sin \theta \cos \psi \sin \psi \\ &= \dot{\phi}^2 \sin^2 \theta (\sin^2 \psi + \cos^2 \psi) + \dot{\theta}^2 (\cos^2 \psi + \sin^2 \psi) \end{aligned}$$

$$\omega_1^2 + \omega_2^2 = \dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2 \quad (14.65)$$

$$\omega_3^2 = \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 \quad (14.66)$$

con lo cual la energía cinética (14.64) en términos de ángulos de Euler queda

$$T = \frac{1}{2} I_1 \left( \dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2 \right) + \frac{I_3}{2} \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 \quad (14.67)$$

por otro lado, la energía potencial gravitacional del cuerpo rígido es la suma de las energías potenciales individuales de cada partícula

$$\begin{aligned} V &= -\mathbf{F}_i \cdot \mathbf{r}_i = -m_i \mathbf{g} \cdot \mathbf{r}_i = -M \left( \frac{m_i \mathbf{r}_i}{M} \right) \cdot \mathbf{g} \\ V &= -M \mathbf{R}_C \cdot \mathbf{g} \end{aligned}$$

siendo  $\mathbf{g}$  el campo gravitacional uniforme en la superficie de la tierra y  $\mathbf{R}_C$  la posición del centro de gravedad de la peonza (que coincide con el centro de masa para campo gravitacional uniforme). En consecuencia, la energía potencial es la misma que si todo el cuerpo estuviera concentrado en su centro de masa. A partir de la geometría de la Fig. 14.5 se ve que

$$\begin{aligned} V &= -M\mathbf{R}_C \cdot \mathbf{g} = -M \|\mathbf{R}_C\| \|\mathbf{g}\| \cos(\pi - \theta) = -Mgl \cos(\pi - \theta) \\ V &= Mgl \cos \theta \end{aligned} \quad (14.68)$$

dada la energía cinética (14.67) y la energía potencial (14.68), el Lagrangiano para la peonza simétrica pesada con punto fijo queda en la forma

$$L = \frac{1}{2}I_1 (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{I_3}{2} (\dot{\phi} \cos \theta + \dot{\psi})^2 - Mgl \cos \theta \quad (14.69)$$

### 14.6.2. Reducción del problema a cuadraturas

Los ángulos  $\phi$  y  $\psi$  son coordenadas cíclicas en el Lagrangiano (14.69). En consecuencia, sus momentos conjugados serán constantes de movimiento. Hemos visto que el momento conjugado a una coordenada angular de rotación corresponde a la componente del momento angular total a lo largo del eje de rotación (ver sección 5.1.2, Pág. 73). Para el ángulo  $\phi$  el eje de rotación es el eje vertical  $X'_3$ , para  $\psi$  es el eje de simetría  $X_3$  del cuerpo, y para  $\theta$  es la línea de nodos. De lo anterior, se puede ver en forma directa porqué los momentos  $p_\phi$  y  $p_\psi$  son constantes, ya que el torque generado por la gravedad va a lo largo de la línea de nodos, de modo que no hay componente del torque a lo largo del eje vertical ni del eje de simetría, los cuales son por definición perpendiculares a la línea de nodos. La ausencia de torques en estos ejes conduce entonces a que las componentes del momento angular total a lo largo de ellos sean constantes en el tiempo.

Aparecen entonces dos integrales primeras de movimiento

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3 (\dot{\psi} + \dot{\phi} \cos \theta) = I_3 \omega_3 \equiv I_1 a \quad (14.70)$$

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = (I_1 \sin^2 \theta + I_3 \cos^2 \theta) \dot{\phi} + I_3 \dot{\psi} \cos \theta \equiv I_1 b \quad (14.71)$$

donde hemos usado la tercera de las Ecs. (13.64). Además hemos expresado estas constantes de movimiento en términos de nuevas constantes  $a$  y  $b$ . Es inmediato ver que existe una constante adicional de movimiento dado que el sistema es conservativo: la energía total del sistema (que coincide con la función energía y numéricamente con el Hamiltoniano)

$$E = T + V = \frac{1}{2}I_1 (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{I_3}{2}\omega_3^2 + Mgl \cos \theta \quad (14.72)$$

donde hemos usado la Ec. (14.67) y la tercera de las Ecs. (13.64). Puesto que tenemos tres coordenadas generalizadas, tres cuadraturas son suficientes para resolver el problema. Éstas se pueden obtener de las integrales primeras de movimiento sin recurrir explícitamente a las ecuaciones de Euler-Lagrange. Despejando  $\dot{\psi}$  en (14.70) se obtiene

$$I_3 \dot{\psi} = I_1 a - I_3 \dot{\phi} \cos \theta \quad (14.73)$$

que a su vez se puede sustituir en (14.71) para eliminar  $\psi$  en dicha ecuación

$$\begin{aligned} I_1 \dot{\phi} \sin^2 \theta + I_3 \dot{\phi} \cos^2 \theta + (I_1 a - I_3 \dot{\phi} \cos \theta) \cos \theta &\equiv I_1 b \\ I_1 \dot{\phi} \sin^2 \theta + I_1 a \cos \theta &\equiv I_1 b \end{aligned}$$

y despejando  $\dot{\phi}$

$$\dot{\phi} = \frac{b - a \cos \theta}{\sin^2 \theta} \quad (14.74)$$

esta ecuación sería integrable si conociéramos a  $\theta$  en función del tiempo. Sustituimos (14.74) de nuevo en (14.73) de lo cual resulta una ecuación análoga para  $\dot{\psi}$

$$\begin{aligned} I_3 \dot{\psi} &= I_1 a - I_3 \left( \frac{b - a \cos \theta}{\sin^2 \theta} \right) \cos \theta \\ \dot{\psi} &= \frac{I_1 a}{I_3} - \cos \theta \left( \frac{b - a \cos \theta}{\sin^2 \theta} \right) \end{aligned} \quad (14.75)$$

que de nuevo es integrable siempre que conozcamos a  $\theta$  en función del tiempo. Finalmente, se observa que las Ecs. (14.74, 14.75) se pueden usar para eliminar  $\dot{\phi}$  y  $\dot{\psi}$  en la ecuación (14.72) de la energía<sup>25</sup>, con lo cual se obtendría una ecuación diferencial ordinaria para  $\theta$ . Antes de hacer esto, notemos que la Ec. (14.70) nos dice que  $\omega_3$  es una constante en el tiempo y por tanto también lo será la expresión  $I_3 \omega_3^2/2$ . En consecuencia la expresión

$$E' = E - \frac{I_3 \omega_3^2}{2}$$

es también una constante de movimiento, esto permite eliminar esta constante de la Ec. (14.72) ya que un corrimiento constante en la energía no tiene ninguna consecuencia física. Reemplazando (14.74) en (14.72) y redefiniendo  $E'$  se tiene

$$\begin{aligned} E' &= \frac{1}{2} I_1 \left[ \left( \frac{b - a \cos \theta}{\sin^2 \theta} \right)^2 \sin^2 \theta + \dot{\theta}^2 \right] + Mgl \cos \theta \\ E' &= \frac{1}{2} I_1 \dot{\theta}^2 + \frac{1}{2} I_1 \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + Mgl \cos \theta \end{aligned} \quad (14.76)$$

la Ec. (14.76) tiene la forma de un problema unidimensional equivalente si definimos un potencial efectivo de la forma

$$V_{ef}(\theta) = Mgl \cos \theta + \frac{1}{2} I_1 \frac{(b - a \cos \theta)^2}{\sin^2 \theta}$$

que será muy útil para describir el movimiento en  $\theta$ , al igual que en el caso de fuerzas centrales en el cual la construcción de un potencial efectivo nos sirvió para analizar el movimiento radial<sup>26</sup>. Es conveniente hacer el cambio de variable

$$u = \cos \theta \quad ; \quad du = -\sin \theta d\theta$$

ya que en tal caso la ecuación (14.76) queda de la forma

$$E' = \frac{1}{2} I_1 \dot{\theta}^2 + \frac{1}{2} I_1 \frac{(b - au)^2}{1 - u^2} + Mglu$$

y como

$$\frac{du}{dt} = -\sin \theta \frac{d\theta}{dt} \quad \Rightarrow \quad \dot{u}^2 = \dot{\theta}^2 \sin^2 \theta = \dot{\theta}^2 (1 - u^2) \quad \Rightarrow \quad \frac{\dot{u}^2}{(1 - u^2)} = \dot{\theta}^2$$

la expresión para  $E'$  queda

$$\begin{aligned} E' &= \frac{1}{2} I_1 \frac{\dot{u}^2}{(1 - u^2)} + \frac{1}{2} I_1 \frac{(b - au)^2}{(1 - u^2)} + Mglu \\ E' (1 - u^2) &= \frac{1}{2} I_1 \dot{u}^2 + \frac{1}{2} I_1 (b - au)^2 + Mglu (1 - u^2) \end{aligned}$$

despejando  $\dot{u}^2$

$$\dot{u}^2 = \frac{2E'}{I_1} (1 - u^2) - (b - au)^2 - \frac{2Mgl}{I_1} u (1 - u^2) \quad (14.77)$$

<sup>25</sup>Estrictamente  $\dot{\psi}$  ya ha sido eliminada en (14.72) a través del factor  $\omega_3$ .

<sup>26</sup>Así como en el escenario de fuerzas centrales, los términos adicionales en el potencial efectivo corresponden a las componentes de la energía cinética asociadas a las coordenadas cíclicas que en este caso son dos. Sin embargo, en la definición de  $E'$  hemos sustraído las componentes que dependen de  $\dot{\psi}$ , las cuales están contenidas en  $\omega_3$ .

es conveniente introducir las siguientes constantes de movimiento

$$\alpha = \frac{2E'}{I_1} = \frac{2E - I_3\omega_3^2}{I_1} ; \quad \beta = \frac{2Mgl}{I_1} ; \quad a = \frac{p_\psi}{I_1} = \frac{I_3\omega_3}{I_1} ; \quad b = \frac{p_\phi}{I_1} \quad (14.78)$$

donde  $a$  y  $b$  provienen de las Ecs. (14.70, 14.71). Estas constantes están asociadas a los momentos conservados, la energía total y la energía potencial. Por ejemplo,  $\alpha$  es proporcional a la energía del sistema excluyendo la energía cinética angular asociada a  $X_3$  (es decir excluyendo el spinning asociado a  $\omega_3$ ). Usando (14.78) en (14.77) se obtiene

$$\begin{aligned} \dot{u}^2 &= \alpha(1-u^2) - (b-au)^2 - \beta u(1-u^2) \\ \dot{u}^2 &= (1-u^2)(\alpha - \beta u) - (b-au)^2 \end{aligned} \quad (14.79)$$

que se puede reducir de manera inmediata a cuadraturas

$$t = \int_{u(0)}^{u(t)} \frac{du}{\sqrt{(1-u^2)(\alpha - \beta u) - (b-au)^2}} \quad (14.80)$$

integrando esta ecuación obtenemos en principio  $\theta = \theta(t)$  que al ser sustituido en (14.74) y (14.75) nos reduce a cuadraturas la dinámica de  $\psi$  y  $\phi$ . En cuanto a la solución de (14.80), el polinomio del radical es cúbico de manera que las soluciones son integrales elípticas. De nuevo la forma analítica de las soluciones puede ser muy compleja y poco ilustrativa. Por lo tanto es más provechoso para los propósitos actuales realizar consideraciones generales, con las cuales se pueden extraer aspectos de la naturaleza general del movimiento sin integrar las ecuaciones de movimiento.

Antes de analizar las soluciones posibles de (14.80), es importante analizar las constantes definidas en (14.78). La Fig. 14.5 muestra un caso en el cual el punto fijo no yace en el centro de masa. Si la peonza está rotando sobre una superficie horizontal, entonces tanto  $\alpha$  como  $\beta$  son mayores que cero. Si en cambio la peonza posee un soporte que le permite estar por debajo de la horizontal (de modo que  $0 \leq \theta \leq \pi$ , entonces  $\beta$  sigue siendo mayor que cero pero  $\alpha$  puede ser positivo o negativo<sup>27</sup>. En otras aplicaciones como algunos giróscopos, el punto fijo puede estar en el centro de masa, en tal caso se tiene que  $\beta = l = 0$  y  $\alpha \geq 0$ . Enfocaremos nuestra atención al caso en el cual la energía cinética rotacional alrededor de  $X_3$  es mucho mayor que la energía cinética asociada a los otros dos ejes.

### 14.6.3. Análisis cualitativo del movimiento

Definamos una función  $f(u)$  usando (14.79), y ordenando las potencias de  $u$

$$\dot{u}^2 \equiv f(u) = \beta u^3 - (\alpha + a^2)u^2 + (2ab - \beta)u + (\alpha - b^2) \quad (14.81)$$

esta ecuación nos muestra por un lado que la función  $f(u)$  es un polinomio de grado 3 en  $u$ , y por otro lado que debe ser no negativa ya que  $f(u) = \dot{u}^2$ . Es decir valores de  $u$  (y por tanto de  $\theta$ ) que conduzcan a valores negativos de  $f(u)$  estarán físicamente descartados. Esto a su vez nos llevará a que las raíces de este polinomio de tercer grado nos dan los valores de  $\theta$  en los cuales  $\dot{\theta}$  cambia de signo, es decir los **ángulos de retorno** en  $\theta$ . El conocimiento de estos ángulos nos puede dar información cualitativa del movimiento del sistema<sup>28</sup>.

Para el giróscopo  $f(u)$  es cuadrática puesto que  $\beta = 0$ , pero para la peonza contribuye en general todo el polinomio cúbico. Dado que en muchas aplicaciones los giróscopos usan montajes libres de torque, la precesión y la nutación estarán suprimidos de modo que el movimiento del giróscopo es trivial. Para entender el movimiento general de un cuerpo rotante, consideraremos solo el caso en el cual  $\beta > 0$ .

<sup>27</sup>Esto se puede ver teniendo en cuenta que  $E'$  en la Ec. (14.76), puede ser menor que cero para algunos valores de  $\theta$  en donde  $\cos \theta < 0$ . Es decir para algunos valores en los cuales  $\pi/2 < \theta \leq \pi$ , con lo cual el centro de masa está por debajo de la horizontal y por tanto la energía potencial es negativa.

<sup>28</sup>Puesto que  $0 \leq \theta \leq \pi$ , y  $u \equiv \cos \theta$ ; tenemos que  $u$  es una función monótonamente decreciente de  $\theta$  en el intervalo físico de éste.

Las ecuaciones cúbicas con coeficientes reales poseen tres raíces que pueden venir en cuatro tipos de combinaciones<sup>29</sup>

1. Una raíz real y un par conjugado de raíces complejas
2. Tres raíces reales, dos de las cuales son iguales
3. Tres raíces reales todas distintas
4. Tres raíces reales todas idénticas

Estas posibilidades dependen de los signos relativos y magnitudes de las cuatro constantes definidas en (14.78). Para nuestro problema específico, hay también una restricción física y es que  $-1 \leq u \leq 1$  para que  $\theta$  sea real. Consideraremos solo el caso en el cual  $\cos \theta \equiv u > 0$  que corresponde a que el centro de masa permanezca por encima de la horizontal. Recordemos sin embargo que muchos giróscopos poseen un soporte puntual que permite que el centro de masa se sitúe por debajo del plano horizontal.

Es claro que  $f(\pm 1) = -(b \mp a)^2$ . En consecuencia,  $f(u)$  es negativa en ambos extremos del intervalo  $[-1, 1]$ , excepto en el caso inusual en el que al menos uno de los valores  $u = \pm 1$  sea una raíz. Esto último corresponde a tener una peonza vertical, caso que analizaremos más adelante. Por tanto, en el presente análisis asumiremos que el polinomio no tiene raíces en  $u = \pm 1$  y por tanto que  $f(\pm 1) < 0$ .

Para  $u$  grande, el término dominante en  $f(u)$  es  $\beta u^3$ . Dado que  $\beta$  es una constante siempre positiva según se vé en la Ec. (14.78),  $f(u)$  será positiva para valores grandes positivos de  $u$  y negativa para valores grandes negativos de  $u$ . Ahora bien, ya que  $f(\infty) > 0$  y  $f(1) < 0$ , el polinomio debe cortar al eje  $u$  en algún punto del intervalo  $(1, +\infty)$ . En conclusión,  $f(u)$  tiene una raíz real en el intervalo  $(1, \pm\infty)$ . Ahora bien puesto que un valor de  $u$  en este intervalo está físicamente excluido, vemos que a lo más existen dos raíces reales de  $u$  que correspondan a una situación físicamente posible. Dicho algebraicamente, hay a lo más dos raíces reales en el intervalo  $-1 < u < 1$ .

Por otro lado, la ausencia de raíces de  $f(u)$  en el intervalo  $-1 < u < 1$ , indicaría que no hay ningún valor posible para  $\theta$  lo cual sería físicamente inconsistente. Esto indica que tiene que haber al menos una raíz real en tal intervalo. Vemos entonces que  $f(u)$  tiene al menos dos raíces reales distintas: una en  $(-1, 1)$  y la otra en  $(1, +\infty)$ . Esto descarta los casos 1 y 4.

Concluimos entonces que solo es posible que haya una o dos raíces en el intervalo  $(-1, 1)$ . Además,  $f(u)$  es negativa en los extremos del intervalo  $(-1, 1)$ . Analicemos las posibles situaciones

- Si existe un  $u_0 \in (-1, 1)$  tal que  $f(u_0) > 0$ , deben existir dos raíces reales diferentes en  $(-1, 1)$ . Una a la izquierda y otra a la derecha de  $u_0$ .
- El caso  $f(u) < 0$  para todo  $u \in (-1, 1)$ , es físicamente inconsistente, puesto que no habría raíces en el intervalo.
- Si  $f(u) < 0$  para todo  $u \in (-1, 1)$ , excepto para un punto  $u = u_1$  en el cual  $f(u_1) = 0$ , es claro que  $f(u_1)$  es un máximo local y  $u_1$  sería una raíz. Puede verse fácilmente que esta raíz debe tener degeneración simple. Para verlo basta con escribir el polinomio y su derivada en términos de las tres raíces

$$f(u) = \beta(u - u_1)(u - u_2)(u - u_3) \quad ; \quad f'(u) = \beta[(u - u_2)(u - u_3) + (u - u_1)(u - u_3) + (u - u_1)(u - u_2)]$$

evaluando la derivada en  $u = u_1$  y usando la condición de extremo (máximo local) en este punto, se tiene

$$f'(u_1) = \beta(u_1 - u_2)(u_1 - u_3) = 0$$

lo cual solo es posible si  $u_1 = u_2$  y/o  $u_1 = u_3$ . La degeneración no puede ser triple, ya que una de las raíces está en el intervalo  $(1, +\infty)$ .

<sup>29</sup>Puede verse que para cualquier polinomio de grado  $n$  con coeficientes reales, el número de soluciones complejas (no reales) debe ser siempre par, pues es inmediato ver que si  $u_a$  es una raíz, entonces  $u_a^*$  también será una raíz.



- El anterior análisis descarta la posibilidad de que existan dos raíces  $u_1 \neq u_2$  en  $(-1, 1)$ , en donde una o ambas correspondan a máximos locales de  $f(u)$ , ya que al contar la degeneración de éstas, tendríamos más raíces que el grado del polinomio. Esto contradice al teorema fundamental del álgebra.

Lo anterior nos deja entonces con solo dos escenarios posibles para nuestro polinomio cúbico: En el **escenario (a)**, todas las tres raíces son reales y diferentes, estando dos de ellas en el intervalo  $(-1, 1)$ . En el **escenario (b)**, hay una raíz con degeneración simple en el intervalo  $(-1, 1)$ ; que corresponde a un máximo local, y la otra está fuera de este intervalo.

En el escenario (a) con dos raíces distintas  $u_1 < u_2$  en el intervalo  $(-1, 1)$ , y teniendo en cuenta que  $f(\pm 1) < 0$ , puede verse que<sup>30</sup>

$$\begin{aligned} f(u) &< 0 & \text{si } u \in [-1, u_1) \cup (u_2, 1] \\ f(u) &> 0 & \text{si } u \in (u_1, u_2) \end{aligned}$$

En el escenario (b), solo hay una raíz  $u_1$  en el intervalo  $(-1, 1)$ , que coincide con un máximo local (degeneración). Como  $f(\pm 1) < 0$  se concluye que  $f(u) < 0$  en todo  $[-1, 1]$  excepto en la raíz  $u_1$  donde  $f(u_1) = 0$ .

Analicemos físicamente las dos soluciones algebraicas posibles. Cuando hay dos raíces distintas  $u_1, u_2$  en  $(-1, 1)$ , hemos visto que  $f(u) \geq 0$  en  $[u_1, u_2]$ , lo cual significa que todo el intervalo es una región accesible físicamente. Cuando hay una sola raíz en  $(-1, 1)$  hay un solo valor posible físicamente para  $u$  ya que en el resto del intervalo  $f(u) < 0$ .

En resumen, el movimiento físico corresponde a casos en que  $f(u) = \dot{u}^2$  es un valor no negativo estando  $u$  en el intervalo  $-1 \leq u \leq 1$ , lo cual corresponde a que  $\theta$  esté entre 0 y  $\pi$ . Se llega a que para cualquier peonza física no vertical,  $f(u)$  debe tener dos raíces  $u_1$  y  $u_2$  (posiblemente degeneradas) en el intervalo  $(-1, 1)$  y la peonza se mueve de manera que el  $\cos \theta$  permanezca entre estas raíces. Si no hay degeneración (ver Fig. 14.7a, Pag. 407) la peonza **cabecea** o **nuta** dentro del intervalo de  $\theta$  permitido. Si hay degeneración (ver Fig. 14.7b, Pag. 407) solo hay un valor de  $\theta$  permitido y no se produce nutación. Es importante por tanto, localizar estas dos raíces para extraer información cualitativa del comportamiento de  $\dot{\phi}$  y  $\dot{\theta}$  para los valores de  $\theta$  permitidos.

Es usual describir el movimiento de la peonza trazando la curva que se produce con la intersección entre el eje de figura y una esfera de radio unidad centrada en el punto fijo. Esta curva es conocida como el **locus** del eje de figura. La coordenada de un punto sobre el locus se especifica con los ángulos de Euler  $\theta$  y  $\phi$  definidos por los ejes del cuerpo<sup>31</sup>. Las raíces  $u_1 \leq u_2$  definen el intervalo permitido para  $\theta$

$$\theta_2 \leq \theta \leq \theta_1 \quad ; \quad \theta_{1,2} = \arccos u_{1,2} \quad (14.82)$$

El locus yace entonces entre dos círculos de colatitud  $\theta_1$  y  $\theta_2$  trazados sobre la esfera unidad (ver Figs. 14.6). El valor de  $\dot{\theta}$  es nulo en ambos círculos y por tanto son puntos de retorno en  $\theta$ .

Es natural ahora preguntarse si hay puntos de retorno posibles para  $\phi$ , esto corresponde a hacer  $\dot{\phi} = 0$ , y de la Ec. (14.74) vemos que esto equivale a la condición

$$b - au = 0$$

La forma del locus está determinada en gran medida por el valor de la raíz de esta ecuación, la cual denotamos por  $u'$

$$u' \equiv \frac{b}{a} \quad (14.83)$$

<sup>30</sup>Se deja al lector la demostración de que  $f(u) < 0$  en el intervalo  $[-1, u_1)$ . Probemos que  $f(u) > 0$  para todo  $u \in (u_1, u_2)$ . Si  $f(u) > 0$  para algún punto en  $(u_1, u_2)$  entonces para que  $f(u) \leq 0$  en otros puntos de este intervalo tendría que haber más raíces del polinomio, lo cual no es posible. Por otro lado, si  $f(u) < 0$  en todo el intervalo  $(u_1, u_2)$ , entonces  $f(u)$  posee un máximo local en  $u_1$ , ya que  $f(u)$  también es negativa en una vecindad a la izquierda de  $u_1$ . Pero un máximo local implica que la raíz es degenerada, lo cual nos lleva a que el polinomio tendría cuatro raíces. De nuevo, esto no es posible. La demostración de que  $f(u) < 0$  en el intervalo  $(u_2, 1]$  se sigue fácilmente.

<sup>31</sup>Es decir, el locus describe la precesión y nutación pero no dice nada sobre el spinning. No obstante, en ausencia de fricción esto no implica ninguna pérdida de información ya que el spinning está descrito por  $\omega_3$  que es constante de movimiento.

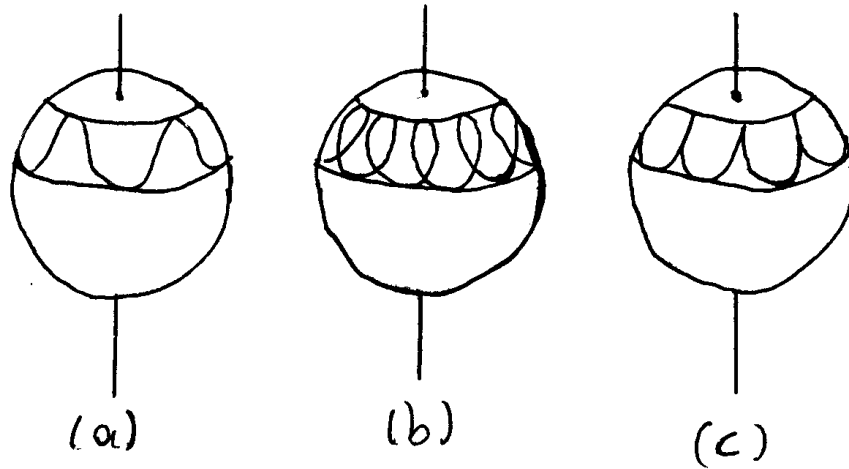


Figura 14.6: Ilustración de la precesión y nutación de una peonza simétrica pesada con un punto fijo. La nutación se realiza entre dos círculos.

supongamos por ejemplo que las condiciones iniciales son tales que  $u'$  es mayor que  $u_2$ . En tal caso la Ec. (14.74) nos dice que  $\dot{\phi}$  siempre tendrá el mismo signo para todos los valores permitidos de  $\theta$  entre  $\theta_2$  y  $\theta_1$  (si  $u_1 < u_2 \Rightarrow \theta_2 < \theta_1$ ), puesto que los puntos de retorno de  $\phi$  ocurren en regiones no accesibles para  $\theta$ . Por lo tanto, el locus del eje de figura debe ser tangente a los círculos límite de tal forma que  $\dot{\phi}$  está en la misma dirección tanto en  $\theta_2$  como en  $\theta_1$ . En otras palabras, la precesión del eje de figura alrededor del eje  $X'_3$  nunca revierte su sentido. Sin embargo, a diferencia de la precesión libre, en este caso el eje de figura también produce un cabeceo o nutación (variación de  $\theta$ ) hacia arriba y hacia abajo dentro del intervalo definido por  $\theta_2$  y  $\theta_1$ . Hay entonces una nutación durante la precesión de la forma ilustrada en la Fig. 14.6a. Por otro lado, vale la pena aclarar que la precesión de la que hablamos aquí es del eje de simetría del cuerpo y no de la velocidad angular o del momento angular con respecto a ejes del cuerpo que es la precesión que se discute en el caso libre.

Si ahora asumimos que  $u'$  está entre  $u_1$  y  $u_2$ , tenemos que existe un punto de retorno para  $\phi$  dentro del intervalo permitido para  $\theta$  i.e.  $\dot{\phi} = 0$ , para algún valor de  $\theta$  entre  $\theta_2$  y  $\theta_1$ . En otras palabras, la dirección de la precesión será diferente en los dos círculos límite, de manera que el locus del eje de figura exhibe lazos cerrados como muestra la figura 14.6b. No obstante, el promedio de  $\dot{\phi}$  no es nulo y hay un sentido neto de avance de la precesión.

Otro caso interesante ocurre cuando  $u'$  coincide con una de las raíces de  $f(u)$ . Por ejemplo, si  $u' = u_2$ , tanto  $\dot{\phi}$  como  $\dot{\theta}$  deben anularse en el círculo superior lo cual requiere que el locus tenga cúspides tocando este círculo, como muestra la figura 14.6c. Este caso no es tan excepcional como podría pensarse a priori, pues se cumple cuando asumimos que la peonza simétrica está rotando inicialmente alrededor del eje de figura solamente. En otras palabras, asumamos que en  $t = 0$  solo hay rotación alrededor del eje de figura y en ese instante se suelta dicho eje. El movimiento subsecuente viene descrito por las condiciones iniciales  $\theta_0, \phi_0$  (orientación inicial del eje de figura) y  $\dot{\phi}(0) = \dot{\theta}(0) = 0$ <sup>32</sup>. La cantidad  $u_0 = \cos \theta_0$  tiene que ser entonces una de las raíces de  $f(u)$ , ya que  $\theta(0) = 0$  significa que el punto inicial es de retorno en  $\theta$ . De hecho, corresponde a la raíz del círculo límite superior (correspondiente a  $\theta_2$ )

$$\cos \theta_0 = u_0 = u_2 = u' = \frac{b}{a} \quad (14.84)$$

para probarlo, notemos que con estas condiciones iniciales  $E'$  en  $t = 0$ , coincide con  $Mgl \cos \theta_0$  y que los términos en  $E'$  provenientes de la energía cinética no pueden ser negativos. Por tanto, tan pronto como

<sup>32</sup>Por supuesto las condiciones iniciales  $\psi(0) = \psi_0$  y  $\dot{\psi}(0) = \omega_3$  también son necesarias. La primera se puede tomar  $\psi_0 = 0$  sin pérdida de generalidad y la segunda es constante de movimiento.

$\dot{\phi}$  y  $\dot{\theta}$  comienzan a ser no nulos, la energía solo se puede conservar si hay un decrecimiento del término potencial  $Mgl \cos \theta$  i.e. aumentando  $\theta$  <sup>33</sup>. El valor inicial de  $\theta$  corresponde por tanto a  $\theta_2$ , el mínimo valor posible para  $\theta$  que corresponde al círculo superior. Es decir, tan pronto como el eje de figura se suelta la peonza comienza a caer hasta que llega al valor máximo de  $\theta$  que es  $\theta_1$  para retornar de nuevo hacia  $\theta_2$  y así sucesivamente. Adicionalmente, el eje también está precesando mientras se cumple el ciclo de nutación ver Fig. 14.6c. Finalmente, dado que en este caso  $u' = u_2$ , el locus presenta cúspides en el círculo superior.

Desde el punto de vista de la energía (omitiendo la energía constante de spinning) podemos decir que cuando  $\theta = \theta_2$  las energías cinéticas de nutación y precesión son nulas y la potencial toma el valor máximo, cuando la peonza cae hasta el punto más bajo, la energía potencial alcanza su valor mínimo y la energía cinética de nutación vuelve a ser nula ya que  $\dot{\theta} = 0$ , en tanto que la energía cinética de precesión alcanza su valor máximo.

#### 14.6.4. Análisis cuantitativo aproximado de la peonza rápida

Tomemos de nuevo las condiciones iniciales  $\dot{\phi}(0) = \dot{\theta}(0) = 0$ , y supongamos adicionalmente que la energía cinética de rotación alrededor de  $X_3$  es mucho mayor que el máximo valor posible para el cambio en la energía potencial,

$$\frac{1}{2}I_3\omega_3^2 \gg 2Mgl \quad (14.85)$$

los efectos de los torques gravitacionales, es decir, la precesión y la nutación que los acompañan, serán solo pequeñas perturbaciones de la rotación dominante (spinning) alrededor del eje de figura. Esta condición define el régimen de **peonza rápida**. Bajo estas suposiciones es posible obtener expresiones analíticas aproximadas para la amplitud y frecuencia de nutación, así como para la frecuencia promedio de precesión.

La extensión de la nutación para las condiciones iniciales dadas se obtiene a través de  $u_1 - u_0$ , siendo  $u_0$  la raíz de  $f(u)$  dada en (14.84) y siendo  $u_1$  la otra raíz física de  $f(u)$ . Las condiciones iniciales nos conducen a

$$E = \frac{1}{2}I_3\omega_3^2 + Mgl \cos \theta_0 \quad \Rightarrow \quad E' = Mgl \cos \theta_0$$

que se puede reescribir como

$$\frac{2E'}{I_1} = \frac{2Mgl}{I_1} \cos \theta_0$$

y teniendo en cuenta (14.78, 14.84) tales condiciones equivalen a la igualdad

$$\alpha = \beta u_0$$

esta relación nos permite reescribir  $f(u)$  en la Ec. (14.81) de la forma

$$\begin{aligned} f(u) &= \beta u^3 - (\beta u_0 + a^2) u^2 + (2ab - \beta) u + (\beta u_0 - b^2) \\ f(u) &= \beta (u^3 - u_0 u^2 - u + u_0) - a^2 u^2 + 2abu - b^2 \\ f(u) &= \beta (u_0 - u) (1 - u^2) - [au - b]^2 \end{aligned}$$

por otro lado la Ec. (14.84) nos dice que

$$b = au_0 \quad (14.86)$$

de modo que

$$f(u) = \beta (u_0 - u) (1 - u^2) - [au - au_0]^2 = \beta (u_0 - u) (1 - u^2) - a^2 (u - u_0)^2$$

<sup>33</sup>En otras palabras, existe una energía cinética asociada al spinning, otra asociada a la precesión, y otra a la nutación. La primera es una constante en el tiempo puesto que está asociada a  $\omega_3$ , y las dos siguientes son nulas en  $t = 0$ . En consecuencia, la energía cinética adquiere el valor mínimo posible en  $t = 0$ . Por lo tanto, tan pronto como la peonza adquiere energía cinética de nutación y precesión, debe compensar con una disminución de la energía potencial.

quedando finalmente

$$f(u) = (u_0 - u) [\beta(1 - u^2) - a^2(u_0 - u)] \quad (14.87)$$

esta parametrización permite encontrar las otras raíces diferentes de  $u_0$  por medio de la expresión cuadrática en los brackets. La raíz buscada  $u_1$  debe en consecuencia satisfacer tal ecuación cuadrática que la podemos expresar en la forma

$$(1 - u_1^2) - \frac{a^2}{\beta}(u_0 - u_1) = 0 \quad (14.88)$$

de acuerdo con la discusión de la sección 14.6.3, una de las raíces de  $f(u)$  es real con  $u > 1$ , es decir no es físicamente accesible. En este caso debe corresponder a una de las raíces de la ecuación cuadrática (14.88), y por tanto tal ecuación solo nos dará una solución física. Definiendo la cantidad

$$x_1 \equiv u_0 - u_1 \quad (14.89)$$

de acuerdo con la sección 14.6.3, vemos que si  $u_1$  es una solución física, entonces  $x_1 > 0$ . Solo las soluciones positivas de  $x_1$  son físicamente accesibles. Podemos reescribir (14.88) como

$$\begin{aligned} \left[1 - (u_0 - x_1)^2\right] - \frac{a^2}{\beta}x_1 &= 0 \quad \Rightarrow \quad 1 - u_0^2 + 2u_0x_1 - x_1^2 - \frac{a^2}{\beta}x_1 = 0 \\ \Rightarrow \quad x_1^2 + \left(\frac{a^2}{\beta} - 2u_0\right)x_1 + u_0^2 - 1 &= 0 \end{aligned}$$

teniendo en cuenta (14.84) la ecuación cuadrática queda de la forma

$$x_1^2 + px_1 - q = 0 \quad ; \quad p \equiv \frac{a^2}{\beta} - 2 \cos \theta_0 \quad ; \quad q = \sin^2 \theta_0 \quad (14.90)$$

### Nutación en el límite de peonza rápida

Hasta el momento no se ha realizado ninguna aproximación, aunque si se han usado las condiciones iniciales  $\dot{\theta}(0) = \dot{\phi}(0) = 0$ , ya que de ellas depende la validez de las Ecs. (14.84, 14.86). La condición de peonza rápida (14.85) implica que  $p$  es mucho mayor que  $q$ , lo cual se puede ver calculando el radio  $a^2/\beta$  y empleando (14.78)

$$\begin{aligned} \frac{a^2}{\beta} &= \left(\frac{I_3\omega_3}{I_1}\right)^2 \frac{I_1}{2Mgl} \\ \frac{a^2}{\beta} &= \left(\frac{I_3}{I_1}\right) \frac{I_3\omega_3^2}{2Mgl} \end{aligned} \quad (14.91)$$

claramente bajo la condición (14.85) este cociente es mucho mayor que la unidad excepto en el caso en el cual  $I_3 \ll I_1$  (que correspondería a una peonza tipo cigarro). De esto se deduce entonces que  $p \gg q$ . Calculemos las raíces de (14.90), usando la aproximación  $\sqrt{1 + 2x} \simeq 1 + x$  para  $|x| \ll 1$

$$\begin{aligned} x_1 &= \frac{-p \pm \sqrt{p^2 + 4q}}{2} = \frac{p}{2} \left(-1 \pm \sqrt{1 + \frac{4q}{p^2}}\right) \cong \frac{p}{2} \left[-1 \pm \left(1 + \frac{2q}{p^2}\right)\right] \\ x_1^a &\cong \frac{q}{p} \quad ; \quad x_1^b \cong -\left[p + \frac{q}{p}\right] \end{aligned}$$

y dado que  $q$  y  $p$  son ambos positivos en esta aproximación, solo hay una raíz positiva y por tanto solo una solución física como se anticipó. Por lo tanto, la única raíz físicamente realizable de (14.90), está dada a primer orden en  $q/p$  por

$$x_1 \simeq \frac{q}{p}$$

usando la segunda de las Ecs. (14.90) y despreciando  $2 \cos \theta_0$  comparado con  $a^2/\beta$  este resultado se escribe como

$$x_1 \simeq \frac{\beta \sin^2 \theta_0}{a^2} = \frac{I_1 2Mgl}{I_3 I_3 \omega_3^2} \sin^2 \theta_0 \quad (14.92)$$

donde hemos usado (14.78). En consecuencia, la extensión de la nutación medida por  $x_1 = u_0 - u_1$  disminuye en la misma forma que  $1/\omega_3^2$ , de modo que a mayor spinning de la peonza menor es la extensión de la nutación. También vemos que la extensión de la nutación disminuye cuando disminuye  $\theta_0$  (al menos en el régimen en el cual estamos interesados i.e.  $0 < \theta_0 \leq \pi/2$ ), es decir a medida que nos acercamos a la vertical en la condición inicial. Debemos recordar sin embargo que la condición  $\theta_0 = 0$  de peonza vertical ha sido excluida del análisis hasta el momento.

La frecuencia de la nutación para la peonza rápida se puede encontrar de manera aproximada teniendo en cuenta que en (14.87) podemos hacer la aproximación

$$(1 - u^2) \cong (1 - u_0^2) = \sin^2 \theta_0 \quad (14.93)$$

en virtud de lo pequeño de la extensión de la nutación. Teniendo en cuenta esta aproximación y definiendo

$$x \equiv u_0 - u \quad (14.94)$$

tenemos que la Ec. (14.87) queda

$$f(u) = x [\beta \sin^2 \theta_0 - a^2 x] = a^2 x \left[ \frac{\beta \sin^2 \theta_0}{a^2} - x \right]$$

de las Ecs. (14.79, 14.81) vemos que  $f(u) = \dot{u}^2 = \dot{x}^2$ . Adicionalmente, usando la Ec. (14.92), se tiene que

$$f(u) = \dot{x}^2 \simeq a^2 x (x_1 - x) \quad ; \quad x \equiv u_0 - u$$

si corremos el origen de  $x$  hasta el punto medio de su intervalo permitido por medio del cambio de variable

$$y = x - \frac{x_1}{2} \Rightarrow \dot{y} = \dot{x}$$

la ecuación diferencial queda en la forma

$$\begin{aligned} \dot{y}^2 &= a^2 \left( y + \frac{x_1}{2} \right) \left[ x_1 - \left( y + \frac{x_1}{2} \right) \right] \Rightarrow \dot{y}^2 = a^2 \left( y + \frac{x_1}{2} \right) \left[ \frac{x_1}{2} - y \right] \\ \dot{y}^2 &= a^2 \left( \frac{x_1^2}{4} - y^2 \right) \end{aligned}$$

al derivar a ambos lados se obtiene

$$2y\dot{y} = -2a^2 y \dot{y}$$

que se simplifica a la ecuación familiar de un oscilador armónico

$$\begin{aligned} \ddot{y} &= -a^2 y \Rightarrow y = A \cos(at + \delta) \Rightarrow x - \frac{x_1}{2} = A \cos(at + \delta) \Rightarrow \\ x &= \frac{x_1}{2} + A \cos(at + \delta) \end{aligned}$$

puesto que  $u(0) = u_0$  se tiene que  $x(0) = 0$ , y la condición inicial  $\dot{\theta}(0) = 0$  implica  $\dot{x}(0) = -\dot{u}(0) = 0$ . Con estas condiciones iniciales la amplitud  $A$  y la fase  $\delta$  vienen dadas por

$$A = -\frac{x_1}{2} \quad ; \quad \delta = 0$$

la solución completa es

$$x = \frac{x_1}{2} (1 - \cos at) \quad (14.95)$$

donde  $x_1$  viene dada por (14.92). La frecuencia angular de la nutación del eje de figura entre  $\theta_0$  y  $\theta_1$  es entonces

$$a = \frac{I_3}{I_1} \omega_3 \quad (14.96)$$

donde se ha tenido en cuenta (14.78). Vemos que esta frecuencia angular de nutación se incrementa cuando aumenta el spinning inicial de la peonza.

### Precesión en el límite de peonza rápida

La velocidad angular de precesión se obtiene a través de (14.74)

$$\dot{\phi} = \frac{b - a \cos \theta}{\sin^2 \theta} = \frac{a \left( \frac{b}{a} - \cos \theta \right)}{\sin^2 \theta} = \frac{a(u_0 - u)}{\sin^2 \theta} \cong \frac{ax}{\sin^2 \theta_0} \quad (14.97)$$

donde hemos tenido en cuenta las definiciones (14.84) y (14.94). Sustituyendo (14.95) en (14.97) se obtiene

$$\dot{\phi} \cong \frac{a}{\sin^2 \theta_0} \frac{x_1}{2} (1 - \cos at) \quad (14.98)$$

La Ec. (14.98), nos muestra que la velocidad de precesión es directamente proporcional a la extensión de la nutación  $x_1$ . Usando (14.92)

$$\begin{aligned} \dot{\phi} &\cong \frac{a}{2 \sin^2 \theta_0} \left( \frac{I_1}{I_3} \frac{2Mgl}{I_3 \omega_3^2} \right) \sin^2 \theta_0 (1 - \cos at) \\ \dot{\phi} &\cong \frac{a}{2} \left( \frac{I_1^2}{I_3^2 \omega_3^2} \right) \frac{2Mgl}{I_1} (1 - \cos at) \end{aligned}$$

finalmente, aplicando las definiciones (14.78)

$$\begin{aligned} \dot{\phi} &\cong \frac{a}{2} \frac{1}{a^2} \beta (1 - \cos at) \\ \dot{\phi} &\cong \frac{\beta}{2a} (1 - \cos at) \end{aligned} \quad (14.99)$$

con lo cual se vé que la rata de precesión no es constante sino que varía armónicamente con el tiempo con la misma frecuencia que la nutación. No obstante, debe tenerse en cuenta que para este caso no hay inversión del sentido de precesión ya que  $\dot{\phi}$  nunca invierte su signo según se vé en (14.99), esto se debe a que  $\dot{\phi} = 0$  ocurre al mismo tiempo que  $\dot{\theta} = 0$  en un extremo del intervalo permitido. La frecuencia **promedio** de precesión es sin embargo

$$\begin{aligned} \overline{\dot{\phi}} &\cong \frac{1}{2\pi/a} \int_0^{2\pi/a} \frac{\beta}{2a} (1 - \cos at) dt = \frac{a}{2\pi} \frac{\beta}{2a} \int_0^{2\pi/a} dt \\ \overline{\dot{\phi}} &\cong \frac{\beta}{2a} = \frac{Mgl}{I_3 \omega_3} \end{aligned} \quad (14.100)$$

que indica que la rata de precesión promedio disminuye a medida que aumenta la velocidad rotacional inicial de la peonza.

### Características generales del movimiento de la peonza rápida

Tenemos ahora un cuadro completo del movimiento de la peonza simétrica pesada con punto fijo, cuando el eje de figura está inicialmente en reposo. Una vez se suelta dicho eje, la peonza comienza a caer bajo la influencia de la gravedad. Pero a medida que cae, el torque resultante alrededor del eje de caída (línea de nodos) causa que la peonza adquiera una velocidad de precesión, directamente proporcional a la extensión de su caída, el eje de figura empieza entonces a precesar alrededor de la vertical. La caída inicial resulta en una nutación periódica del eje de figura que se combina con la precesión. A medida que incrementamos el valor inicial de la velocidad angular alrededor del eje de figura la extensión de la nutación decrece rápidamente aunque la frecuencia de la nutación aumenta, y la precesión alrededor de la vertical también disminuye. En la práctica, para peonzas suficientemente rápidas, la nutación se amortigua muy rápidamente debido a la fricción en el pivote y resulta muy difícil de observar. En consecuencia, la peonza parece precesar uniformemente alrededor de la vertical. Puesto que la peonza parece precesar uniformemente solo en apariencia, se habla frecuentemente de precesión **pseudoregular**.

Nótese sin embargo que aunque en la práctica la nutación suele ser inobservable para peonzas rápidas, ignorar completamente la nutación nos lleva a la conclusión inconsistente de que una vez que se suelta el eje de figura, la peonza comienza inmediatamente a precesar uniformemente. Es decir, que a partir de la condición inicial  $\dot{\phi}(t=0) = 0$ , se alcanza una velocidad  $\dot{\phi}$  finita en un tiempo infinitesimal, implicando una aceleración angular  $\ddot{\phi}$  infinita. Nuestro tratamiento de precesión pseudoregular resuelve la inconsistencia, mostrando que la precesión crece continuamente a partir del reposo sin ninguna aceleración infinita y la tendencia inicial de la peonza es moverse en la dirección de la fuerza de gravedad.

14.6.5. Peonza con precesión regular

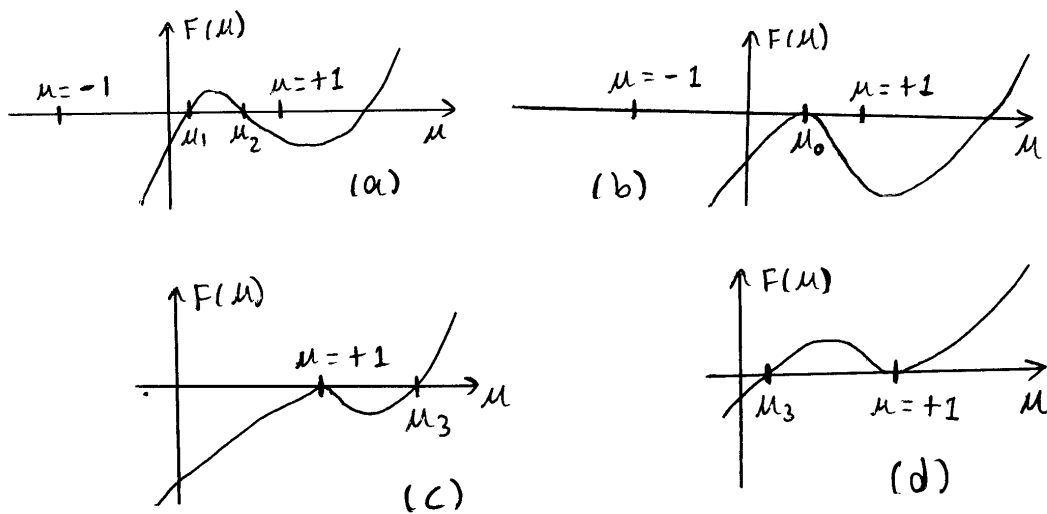


Figura 14.7: Comportamiento de las raíces del polinomio  $f(u)$  dado en la Ec. (14.81). (a) Hay dos raíces en el intervalo  $(-1, 1)$  y el sistema exhibe nutación. (b) Hay solo una raíz física  $u_0$  en el intervalo  $(-1, 1)$  y por tanto un solo valor permitido de  $\theta$ , en este caso no hay nutación (precesión regular). (c) Corresponde a un máximo local con una única raíz física en  $u = 1$  que corresponde a peonza vertical sin nutación. Finalmente, en (d) hay dos raíces físicas donde una de ellas corresponde a  $u = 1$ , esto describe una peonza inicialmente vertical con nutación.

En vista de lo anterior, resulta de interés saber qué condiciones iniciales conducen a una precesión verdaderamente regular. En tal caso  $\theta$  permanece constante con su valor inicial  $\theta_0$ , lo cual significa que  $\theta_1 = \theta_2 = \theta_0$ . Esto implica que  $f(u)$  debe tener una degeneración simple en las raíces que yacen en el intervalo  $[-1, 1]$  i.e.  $u_1 = u_2 = u_0$  (ver Fig. 14.7b). Recordando que en este caso  $f(u)$  era no positivo en todo el intervalo  $[-1, 1]$ , la única manera de obtener una sola raíz (o más bien dos raíces degeneradas) en el interior de tal intervalo es considerar el caso en el cual  $f(u)$  posee un máximo local en la raíz, como se aprecia en la Fig. 14.7b, es decir

$$f(u_0) = \dot{u}_0^2 = 0 \quad ; \quad \left. \frac{df}{du} \right|_{u=u_0} = 0 \tag{14.101}$$

de la Ec. (14.74), es claro que la condición  $\theta(t) = \theta_0$  (ausencia de nutación) nos conduce a  $\dot{\phi} = cte = \dot{\phi}_0$ . Es decir, que la ausencia de nutación implica automáticamente el límite de precesión regular. Utilizando la Ec. (14.81) en la forma dada por la Ec. (14.79), la primera de las condiciones en (14.101) con  $\dot{u}_0^2 = 0$  conduce a

$$(\alpha - \beta u_0) = \frac{(b - a u_0)^2}{1 - u_0^2} \tag{14.102}$$

utilizando (14.79), la segunda de las condiciones en (14.101) se traduce en

$$\begin{aligned} \left. \frac{df(u)}{du} \right|_{u_0} &= (1 - u_0^2)(-\beta) - 2u_0(\alpha - \beta u_0) - 2(b - au_0)(-a) = 0 \Rightarrow \\ \beta(1 - u_0^2) &= -2u_0(\alpha - \beta u_0) + 2a(b - au_0) \\ \frac{\beta}{2} &= \frac{a(b - au_0)}{1 - u_0^2} - u_0 \frac{(\alpha - \beta u_0)}{1 - u_0^2} \end{aligned} \quad (14.103)$$

sustituyendo (14.102) en (14.103) y teniendo en cuenta (14.74), se llega a una ecuación cuadrática para  $\dot{\phi}$

$$\begin{aligned} \frac{\beta}{2} &= \frac{a(b - au_0)}{1 - u_0^2} - u_0 \frac{1}{1 - u_0^2} \frac{(b - au_0)^2}{1 - u_0^2} \\ \frac{\beta}{2} &= a \frac{(b - a \cos \theta_0)}{\sin^2 \theta_0} - \cos \theta_0 \left( \frac{b - a \cos \theta_0}{\sin^2 \theta_0} \right)^2 \\ \frac{\beta}{2} &= a\dot{\phi} - \dot{\phi}^2 \cos \theta_0 \end{aligned} \quad (14.104)$$

y teniendo en cuenta las definiciones de (14.78) para  $\beta$  y  $a$  podemos escribir estos términos en dos formas alternativas, dependiendo de si  $a$  se expresa en términos de  $\omega_3$  o de los valores constantes de  $\dot{\phi}$  y  $\dot{\psi}$  (ver Ecs. 14.78 y 14.70)

$$a = \frac{I_3}{I_1} (\dot{\psi} + \dot{\phi} \cos \theta) = \frac{I_3}{I_1} \omega_3 \quad ; \quad \beta = \frac{2Mgl}{I_1} \quad (14.105)$$

reemplazando (14.105) en (14.104) usando las dos formas de “ $a$ ”

$$\begin{aligned} \frac{Mgl}{I_1} &= \frac{I_3}{I_1} \omega_3 \dot{\phi} - \dot{\phi}^2 \cos \theta_0 \\ \frac{Mgl}{I_1} &= \frac{I_3}{I_1} (\dot{\psi} + \dot{\phi} \cos \theta_0) \dot{\phi} - \dot{\phi}^2 \cos \theta_0 \end{aligned}$$

donde hemos tenido en cuenta que  $\theta = \theta_0$  para todo tiempo. Estas ecuaciones se pueden escribir como

$$\begin{aligned} Mgl &= \dot{\phi} (I_3 \omega_3 - I_1 \dot{\phi} \cos \theta_0) \\ Mgl &= \dot{\phi} [I_3 \dot{\psi} - (I_1 - I_3) \dot{\phi} \cos \theta_0] \end{aligned} \quad (14.106)$$

las condiciones iniciales para el problema de la peonza pesada simétrica con punto fijo, requieren la especificación de  $(\theta, \phi, \psi, \dot{\theta}, \dot{\phi})$  y  $\dot{\psi}$  ú  $\omega_3$  en el tiempo  $t = 0$ . Dado que las variables  $\phi$  y  $\psi$  son cíclicas sus valores iniciales no son muy relevantes, y en general podemos elegir cualquier valor que deseemos para cada una de las otras cuatro constantes. Pero si hacemos la exigencia adicional de que el eje de figura precese uniformemente sin nutación, nuestra escogencia de estas cuatro condiciones iniciales ya no es completamente libre, ya que debe satisfacer cualquiera de las Ecs. (14.106), que en general nos dan una relación entre las cantidades  $\theta_0$ ,  $\dot{\phi}_0$  y  $\omega_3$  ó entre las cantidades  $\theta_0$ ,  $\dot{\phi}_0$  y  $\dot{\psi}_0$ . Para  $\dot{\theta} = 0$  podemos aún elegir valores iniciales de  $\theta$  y  $\omega_3$  **casi** arbitrariamente, pero el valor de  $\dot{\phi}$  queda entonces determinado a través de las Ecs. (14.74, 14.78). La frase casi arbitrariamente, se refiere a que al ser las Ecs. (14.106) cuadráticas, y teniendo en cuenta que  $\dot{\phi}$  debe ser real, entonces el discriminante de (14.106) debe ser no negativo. Aplicando esta condición a la primera de las Ecs. (14.106), se obtiene

$$I_3^2 \omega_3^2 \geq 4MglI_1 \cos \theta_0 \quad (14.107)$$

para  $\pi/2 \leq \theta_0 < \pi$  es decir cuando el centro de masa esta al nivel o por debajo del punto fijo (excluyendo por ahora peonza vertical con  $\theta = \pi$ ), el miembro derecho de esta expresión es no positivo, de modo que cualquier



valor de  $\omega_3$  puede conducir a precesión uniforme. Pero para  $0 < \theta_0 < \pi/2$ , el valor escogido para  $\omega_3$  tiene que ser mayor a un valor mínimo  $\omega'_3$  que se obtiene igualando los miembros en (14.107)

$$\omega_3 \geq \omega'_3 = \frac{2}{I_3} \sqrt{Mgl I_1 \cos \theta_0} \quad (14.108)$$

para lograr la precesión uniforme. Condiciones similares se pueden obtener a partir de la segunda de las Ecs. (14.106) para los valores permitidos de  $\dot{\psi}$ .

Nótese que las Ecs. (14.106) no se pueden satisfacer con  $\dot{\phi}(0) = 0$  para valores finitos de  $\dot{\psi}$  o de  $\omega_3$ . Es decir que para obtener precesión uniforme, debemos darle siempre a la peonza un impulso que origine la precesión, i.e. tal que  $\dot{\phi}_0 \neq 0$ . Sin esta velocidad precesional adecuadamente generada, podemos en el mejor de los casos obtener solo una precesión pseudoregular. Esto es lógico ya que sin un torque adicional, no es posible generar la precesión sin una caída previa del centro de masa (que generaría la nutación), como ya vimos anteriormente.

### Límite de precesión regular lenta

Como un resultado de la naturaleza cuadrática de (14.106), habrá en general dos soluciones para  $\dot{\phi}$  conocidas como precesión “rápida” y “lenta”. Usando la primera de las Ecs. (14.106), las soluciones para  $\dot{\phi}$  son

$$\dot{\phi}_1 = \frac{I_3 \omega_3 - \sqrt{I_3^2 \omega_3^2 - 4Mgl I_1 \cos \theta_0}}{2I_1 \cos \theta_0} ; \quad \dot{\phi}_2 = \frac{I_3 \omega_3 + \sqrt{I_3^2 \omega_3^2 - 4Mgl I_1 \cos \theta_0}}{2I_1 \cos \theta_0} \quad (14.109)$$

donde  $\dot{\phi}_1$  corresponderá a precesión lenta y  $\dot{\phi}_2$  a precesión rápida. Que se obtenga una u otra, depende de las condiciones iniciales con las cuales se prepara el sistema.

Vamos a analizar el límite de precesión lenta por dos caminos. Cuando la precesión es lenta, es razonable despreciar a  $\dot{\phi} \cos \theta_0$  con respecto a  $\omega_3$ . De la Ec. (14.105) vemos que esto es equivalente a despreciar a  $\dot{\phi} \cos \theta_0$  comparado con “ $a$ ” siempre que  $I_3/I_1$  no sea demasiado pequeño. Aplicando esta aproximación en la Ec. (14.104) tenemos

$$\frac{\beta}{2} = (a - \dot{\phi} \cos \theta_0) \dot{\phi} \simeq a \dot{\phi}$$

vemos entonces que una solución aproximada para  $\dot{\phi}$  viene dada por

$$\dot{\phi} \simeq \frac{\beta}{2a} = \frac{Mgl}{I_3 \omega_3} \quad \text{precesión lenta} \quad (14.110)$$

expresión que coincide con la correspondiente para el promedio de la precesión pseudoregular de una peonza rápida Ec. (14.100). Este resultado es lógico ya que si la precesión es lenta no hay mucha diferencia entre aplicar un ligero impulso al giróscopo sin nutación, o tener una muy pequeña nutación sin imprimirle impulso en lo absoluto. Analicemos la aproximación realizada en términos de parámetros más físicos

$$\begin{aligned} a &>> \dot{\phi} \cos \theta_0 \Rightarrow \frac{I_3}{I_1} \omega_3 \gg \frac{Mgl}{I_3 \omega_3} \cos \theta_0 \Rightarrow \omega_3^2 \gg \frac{Mgl I_1}{I_3^2} \cos \theta_0 \\ \omega_3 &>> \frac{\sqrt{Mgl I_1 \cos \theta_0}}{I_3} \end{aligned}$$

Al comparar con la Ec. (14.108), se puede ver que con el valor de  $\dot{\phi}$  dado en (14.110), despreciar a  $\dot{\phi} \cos \theta_0$  con respecto a “ $a$ ” es equivalente a requerir que  $\omega_3$  tenga un valor mucho mayor a su cota inferior permitida.

Veamos otra forma de obtener el límite de precesión lenta. Comenzaremos reescribiendo  $\dot{\phi}_1$  de la Ec. (14.109) en la forma

$$\begin{aligned} \dot{\phi}_1 &= \frac{I_3 \omega_3 - I_3 \omega_3 \sqrt{1 - \frac{4Mgl I_1 \cos \theta_0}{I_3^2 \omega_3^2}}}{2I_1 \cos \theta_0} = \frac{I_3 \omega_3 \left[ 1 - \sqrt{1 - \left( \frac{2Mgl}{\frac{1}{2} I_3 \omega_3^2} \right) \frac{I_1 \cos \theta_0}{I_3}} \right]}{2I_1 \cos \theta_0} \\ \dot{\phi}_1 &= \frac{I_3 \omega_3 [1 - \sqrt{1 - x}]}{2I_1 \cos \theta_0} ; \quad x \equiv \frac{I_1 \cos \theta_0}{I_3} y, \quad y \equiv \frac{2Mgl}{\frac{1}{2} I_3 \omega_3^2} \end{aligned} \quad (14.111)$$

en este punto recordamos que en el límite de peonza rápida, la velocidad angular de precesión era proporcional a la extensión de la nutación, la cual a su vez era proporcional a  $\omega_3^{-2}$ , de modo que la velocidad angular de precesión disminuye con el aumento de la velocidad angular de spinning. Este resultado no es válido aquí, ya que en este caso no tenemos nutación. Sin embargo, es de esperarse que el resultado no sea demasiado diferente al caso en el que  $\omega_3$  sea muy grande de modo que la extensión de la nutación es muy pequeña y la precesión muy lenta. Tomaremos en consecuencia el límite de peonza rápida Ec. (14.85), que al compararse con (14.111) nos lleva a que  $y \ll 1$ , y si  $I_1 \cos \theta_0$  no es mucho mayor que  $I_3$  nos conduce a que  $x \ll 1$ . Usando la aproximación  $\sqrt{1-x} \simeq 1 - \frac{1}{2}x$  para  $|x| \ll 1$  en la Ec. (14.111), se obtiene

$$\begin{aligned}\dot{\phi}_1 &\simeq \frac{I_3 \omega_3 \left[1 - \left(1 - \frac{1}{2}x\right)\right]}{2I_1 \cos \theta_0} = \frac{I_3 \omega_3}{4I_1 \cos \theta_0} x = \frac{I_3 \omega_3}{4I_1 \cos \theta_0} \left(\frac{2Mgl}{\frac{1}{2}I_3 \omega_3^2}\right) \frac{I_1 \cos \theta_0}{I_3} \\ \dot{\phi}_1 &\simeq \left(\frac{Mgl}{I_3 \omega_3}\right)\end{aligned}$$

que coincide con el límite de precesión lenta (14.110) obtenido con la aproximación  $a \gg \dot{\phi} \cos \theta_0$ . Por tanto, el límite de peonza rápida (14.85) implica el límite de precesión lenta, siempre que  $I_1 \cos \theta_0$  no sea mucho mayor que  $I_3$ . Esto explica de una manera más natural por qué este límite coincide con la precesión pseudoregular (14.100), ya que ésta fué obtenida en el límite de peonza rápida. No obstante, la obtención del límite de precesión lenta solo dependió de que  $|x| \ll 1$ , esto también se puede lograr si

$$\frac{I_1 \cos \theta_0}{I_3} \ll 1$$

siempre que  $y$  no sea mucho mayor que uno. Por ejemplo, tomemos  $I_1 \ll I_3$  y/o  $\theta_0 \sim \pi/2$ , con  $y \sim 1$ . En este caso obtenemos límite de precesión lenta, aunque claramente no tenemos límite de peonza rápida. De hecho, al tomar  $|x| \ll 1$  en (14.111) se obtiene

$$\begin{aligned}x &\equiv \frac{I_1 \cos \theta_0}{I_3} y = \frac{I_1 \cos \theta_0}{I_3} \left(\frac{2Mgl}{\frac{1}{2}I_3 \omega_3^2}\right) \ll 1 \\ \Rightarrow \omega_3 &\gg \frac{2\sqrt{Mgl I_1 \cos \theta_0}}{I_3}\end{aligned}$$

que al comparar con la Ec. (14.108), nos muestra de nuevo que el límite de precesión lenta, equivale a requerir valores de  $\omega_3$  mucho mayores a su cota inferior permitida. Esta es entonces la condición más general de precesión lenta.

### Límite de precesión regular rápida

Existe un límite de precesión rápida (no confundir con el límite de peonza rápida cuyo límite se refiere al spinning y no a la precesión), que se obtiene cuando  $\dot{\phi}$  es grande de tal forma que  $Mgl$  es pequeño comparado con los otros términos en las Ecs. (14.106)

$$\dot{\phi} \simeq \frac{I_3 \omega_3}{I_1 \cos \theta_0} \quad \text{precesión rápida}$$

vemos entonces que la precesión rápida es independiente de los torques gravitacionales, ya que precisamente éstos se han despreciado. Es natural entonces suponer que dicho límite está más relacionado con la precesión libre, dado que no hay más torques en el sistema.

Alternativamente, este límite se puede obtener de la solución  $\dot{\phi}_2$  en la Ec. (14.109), despreciando el término  $Mgl I_1 \cos \theta_0$  (que está relacionado con el torque gravitacional) con respecto a  $I_3^2 \omega_3^2$

$$\dot{\phi}_2 = \frac{I_3 \omega_3 + I_3 \omega_3 \sqrt{1 - \frac{4Mgl I_1 \cos \theta_0}{I_3^2 \omega_3^2}}}{2I_1 \cos \theta_0} \simeq \frac{I_3 \omega_3 + I_3 \omega_3}{2I_1 \cos \theta_0} = \frac{I_3 \omega_3}{I_1 \cos \theta_0}$$

Debemos observar sin embargo, que las soluciones (14.109) conducen a una solución de precesión lenta  $\dot{\phi}_1$  y otra de precesión rápida  $\dot{\phi}_2$ , independientemente de que se cumplan los límites anteriores. Lo que hemos analizado entonces son valores extremos de  $\dot{\phi}_1$  y de  $\dot{\phi}_2$ .

### 14.6.6. Peonza inicialmente vertical

En la sección 14.6.3 se descartó del análisis la posibilidad de que una de las raíces de  $f(u)$  estuviera en  $u = \pm 1$ . Un límite interesante es el caso en que  $u = 1$  es una de las raíces de  $f(u)$ . Este caso corresponde a una peonza inicialmente vertical con el centro de masa por encima del origen, y se debe analizar por aparte ya que en todo el tratamiento que se ha hecho hasta ahora, el  $\sin^2 \theta_0$  ha aparecido a menudo en el denominador.

Trabajemos entonces bajo la condición de que la peonza esté inicialmente vertical con una velocidad angular inicial de spinning  $\omega_3$ . En  $t = 0$  el eje de figura coincide con el eje vertical, y recordando que  $p_\phi$  ( $p_\psi$ ) es la componente total del momento angular alrededor del eje vertical (de figura), se concluye que  $p_\phi = p_\psi$  en  $t = 0$  y en todo tiempo ya que son constantes de movimiento. De las Ecs. (14.70, 14.71) vemos entonces que lo anterior conduce a que  $a = b$ . Por otro lado, dado que la energía cinética inicial es solo de spinning se obtiene que

$$E' = E - \frac{1}{2}I_3\omega_3^2 = Mgl$$

de esta expresión y las relaciones (14.78) se obtiene que  $\alpha = \beta$ . Con estas condiciones, la Ec. (14.79) queda en la forma

$$\begin{aligned} f(u) &= \dot{u}^2 = (1 - u^2)\beta(1 - u) - a^2(1 - u)^2 \\ f(u) &= \dot{u}^2 = (1 - u)(1 + u)\beta(1 - u) - a^2(1 - u)^2 \\ f(u) &= \dot{u}^2 = (1 - u)^2[\beta(1 + u) - a^2] \end{aligned}$$

lo cual nos indica que  $u = 1$  es una raíz doble siempre que sea raíz de  $f(u)$ , donde la tercera raíz viene dada por

$$u_3 = \frac{a^2}{\beta} - 1 \quad (14.112)$$

Si  $a^2/\beta > 2$ , lo cual se cumple por ejemplo en el límite de peonza rápida (siempre que  $I_3/I_1$  no sea mucho menor que uno), se tiene que  $u_3 > 1$ , de modo que solo la raíz  $u = 1$  corresponde a una situación física. Esto significa que la peonza continúa con su eje de figura vertical, con energía cinética de spinning únicamente. En tal caso la raíz degenerada  $u = 1$  corresponde a un máximo local de  $f(u)$  con la tercera raíz en algún punto fuera del intervalo  $[-1, 1]$ , como se vé en la Fig. 14.7c.

Por otro lado, si  $a^2/\beta < 2$  se tiene que la tercera raíz es menor que 1, tomaremos solo el caso en que  $u_3$  está en el intervalo  $(0, 1)$ . En este escenario, la raíz  $u = 1$  corresponde a un mínimo local y  $f(u)$  se comporta como muestra la Fig. 14.7d, la peonza nutará entre 0 y  $\theta_3$  que serán los puntos de retorno en  $\theta$ . Estas desigualdades muestran en consecuencia que existe un valor crítico  $a^2/\beta = 2$  que define la transición entre un movimiento perfectamente vertical y uno con nutación y precesión<sup>34</sup>. Aplicando (14.91), este valor límite nos da

$$\frac{a^2}{\beta} = \left(\frac{I_3}{I_1}\right) \frac{I_3\omega'^2}{2Mgl} = 2$$

la única variable que se puede modular en este término (para una peonza dada) es la velocidad angular. Por tanto, existe una velocidad angular crítica por encima de la cual solo el movimiento vertical es posible.

$$\omega'^2 = 4\frac{MglI_1}{I_3^2} \quad (14.113)$$

<sup>34</sup>Vemos que cuando  $u = 1$  es una raíz, es posible que las tres raíces de  $f(u)$  sean todas físicas, aunque solo hay a lo más dos raíces físicas diferentes. Es también posible que las tres raíces sean degeneradas. Estas características difieren del análisis realizado en la sección 14.6.3, cuando descartábamos la posibilidad de que  $u = \pm 1$  fueran raíces de  $f(u)$ .

expresión que es idéntica al valor expresado en la Ec. (14.108) de la frecuencia mínima para precesión constante, y que ahora se aplica para  $\theta_0 = 0$ .

En la práctica, si la peonza gira inicialmente con eje vertical y  $\omega_3$  es superior al valor crítico, ella rotará un tiempo solo sobre el eje vertical (trompo dormido) pero su spinning va disminuyendo por efecto de la fricción, cuando  $\omega_3$  pasa el valor crítico, comienza a bambolearse en cantidades cada vez mayores a medida que disminuye el spinning.

#### 14.6.7. Efectos de fricción, torques adicionales y aplicaciones de la peonza simétrica pesada

Los efectos de fricción (que no se pueden introducir en forma directa en el formalismo Lagrangiano) producen ciertos fenómenos inesperados. Un ejemplo notable es el “tippie-top” que consiste básicamente en algo mas que media esfera con un soporte unido a la parte plana. Cuando se comienza a rotar con la superficie esférica hacia abajo sobre una superficie dura y lisa, el artefacto comienza a deslizarse y nutar hasta que eventualmente se gira quedando con la superficie esférica hacia arriba, pivoteando sobre el soporte, comportándose como un trompo dormido. La inversión completa del vector momento angular es el resultado del torque friccional que ocurre cuando el trompo se desliza sobre su superficie esférica.

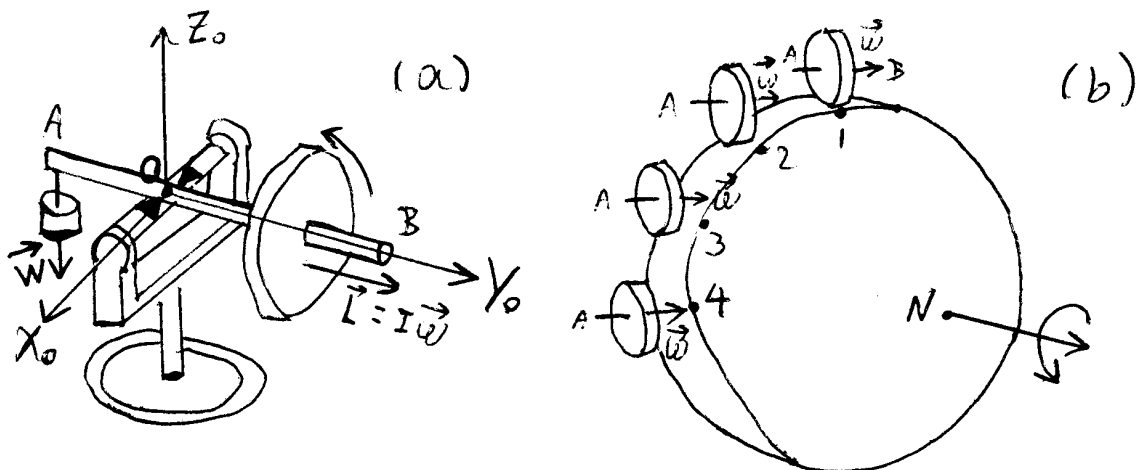


Figura 14.8: (a) Giróscopo al cual se le coloca un contrapeso  $W$  para compensar el torque debido al peso del disco, de modo que el centro de masa de la configuración coincida con el punto fijo. (b) Fenómeno de inercia giroscópica: después de 6 horas, un observador en la tierra ve un giro de  $\pi/2$  en el eje del giróscopo, debido a que dicho eje permanece invariable con respecto a un sistema inercial.

Hay una amplia gama de aplicaciones concernientes al uso de peonzas rápidas. Una de ellas es el giróscopo que consiste en una peonza simétrica que se monta sobre un soporte que permite la reorientación del eje de figura en cualquier dirección, el eje de figura gira muy rápidamente debido a la aplicación de ciertos torques externos. El centro de masa se hace coincidir con el punto fijo a través de algún contrapeso colocado al otro extremo de la peonza (ver Fig. 14.8a). Una vez se libera el eje de figura con una gran velocidad angular (con  $\dot{\phi} = \dot{\theta} = 0$ ), se observa que la dirección del eje de figura permanece constante (medida por un observador inercial, por ejemplo las estrellas fijas) sin importar como reorientemos el montaje, fenómeno conocido como inercia giroscópica, el cual se debe a la conservación del momento angular en ausencia de torques netos externos. En particular, si el montaje se deja por seis horas (asumiendo muy bajo amortiguamiento) se puede observar una rotación del eje de figura de  $\pi/2$  que en realidad se debe justamente a que el eje de figura permanece constante en dirección con respecto a un sistema inercial, en tanto que la tierra ha tenido una rotación de  $\pi/2$ , como se aprecia en la Fig. 14.8b. Esta característica hace del giróscopo un excelente instrumento para mantener direcciones específicas, por ejemplo para proveer plataformas estables.

Si se le aplican torques externos adecuados al giróscopo, puede generar precesión y nutación. En la mayoría de ocasiones se puede utilizar el límite de peonza rápida para estos dispositivos y como vimos antes, en la práctica la nutación se atenúa rápidamente por efectos de fricción y pronto se vuelve inobservable. En realidad, en muchos giróscopos el montaje se realiza para atenuar deliberadamente el efecto de nutación. En tal caso, el estado estacionario del sistema solo presenta precesión, lo cual simplifica fuertemente el tratamiento matemático. Podemos entonces generalizar el tratamiento hecho para la peonza simétrica de una manera simple.

Si  $\mathbf{R}$  es el radio vector que va desde el punto fijo hasta el centro de masa a lo largo del eje de figura, el torque gravitacional que se ejerce sobre la peonza viene dado por

$$\mathbf{N} = \mathbf{R} \times M\mathbf{g} \quad (14.114)$$

donde  $\mathbf{g}$  define al campo gravitacional uniforme en la vecindad de la superficie terrestre. Definamos  $\mathbf{L}_3$  como la componente del momento angular de la peonza a lo largo del eje de figura, el cual describe el momento angular de espín del sólido, y definamos  $\boldsymbol{\omega}_p$  como el vector de precesión, el cual está alineado con el eje vertical y posee magnitud igual a la velocidad angular promedio de precesión  $\bar{\dot{\phi}}$  dada aproximadamente por (14.100). De las definiciones es claro que

$$\boldsymbol{\omega}_p \equiv \left| \bar{\dot{\phi}} \right| \mathbf{u}'_3 \quad ; \quad \mathbf{L}_3 \equiv I_3 \omega_3 \mathbf{u}_3$$

La dirección de la precesión pseudoregular está dada por

$$\mathbf{u} \equiv \frac{\mathbf{u}'_3 \times \mathbf{u}_3}{\|\mathbf{u}'_3 \times \mathbf{u}_3\|}$$

que es la dirección de la línea de nodos (ver Fig. 14.5). Por tanto,  $\boldsymbol{\omega}_p \times \mathbf{L}_3$  está dado por

$$\boldsymbol{\omega}_p \times \mathbf{L}_3 = \left| \bar{\dot{\phi}} \right| \mathbf{u}'_3 \times I_3 \omega_3 \mathbf{u}_3 = \left| \bar{\dot{\phi}} \right| I_3 \omega_3 (\mathbf{u}'_3 \times \mathbf{u}_3) = \left| \bar{\dot{\phi}} \right| I_3 \omega_3 \sin \theta \mathbf{u}$$

y aplicando (14.100), se obtiene

$$\begin{aligned} \boldsymbol{\omega}_p \times \mathbf{L}_3 &= \left( \frac{Mgl}{I_3 \omega_3} \right) I_3 \omega_3 \sin \theta \mathbf{u} = Mgl \sin \theta \mathbf{u} = |M\mathbf{g}| |\mathbf{R}| \sin \theta \mathbf{u} = |\mathbf{R} \times M\mathbf{g}| \mathbf{u} \\ \boldsymbol{\omega}_p \times \mathbf{L}_3 &= \mathbf{R} \times M\mathbf{g} \end{aligned}$$

que coincide con el hecho ya discutido de que el torque va a lo largo de la línea de nodos  $\mathbf{u}$ . Nos queda finalmente

$$\boldsymbol{\omega}_p \times \mathbf{L}_3 = \mathbf{N} \quad (14.115)$$

Puesto que el torque alrededor del punto fijo o del centro de masa se puede colocar en la forma  $\mathbf{R} \times \mathbf{F}$ , en forma similar a (14.114), la rata de precesión promedio para una peonza rápida se puede derivar siempre de (14.115), donde la dirección de la fuerza  $\mathbf{F}$  define al eje de precesión. Casi todas las aplicaciones prácticas de los giróscopos involucran el comportamiento del equilibrio (despreciando transientes) que se puede derivar de (14.115).

En virtud de la ecuación (14.115) se puede ver que un giróscopo originalmente libre de torques puede servir para medir ratas de rotación angular y torques aplicados por medio del fenómeno de precesión. Pues dicha ecuación nos dice que la rata de precesión es proporcional al torque, en contraste con el caso de un cuerpo sin rotación de espín en el cual es la aceleración angular la que es proporcional al torque. Una vez que el torque se remueve, el cuerpo que no tiene espín continúa moviéndose; bajo condiciones similares el giróscopo continúa con su espín sin precesar.

El caso del girocompás es más complejo, ya que en este caso lo que se pretende es mantener el objeto con dirección fija con respecto a un sistema de referencia no inercial (la tierra). Para ello se le adiciona al dispositivo una precesión aplicada a una rata tal que contrarreste la rata de rotación terrestre. Una vez ajustado en la dirección de rotación de la tierra i.e. la dirección norte, el girocompas preserva esta dirección, por lo menos en vehículos que se muevan lentamente.

Dos ejemplos interesantes como fenómenos en sí y para la aplicación de las técnicas aquí descritas, son: la precesión de los equinoccios, y la precesión del movimiento de cargas en un campo magnético [1]. De hecho, las aplicaciones del giróscopo constituyen un campo muy activo de investigación [14, 15].

### 14.7. Ejercicios

1. Sea un sólido de revolución como el de la Fig. 3.2, Pág. 51, con densidad  $\rho$  constante. Demuestre que los momentos de inercia con respecto a los ejes  $X$  e  $Y$  vienen dados por (ver problema 4 pág 52, y Ref. [8])

$$I_X = \frac{\pi\rho}{2} \int_{x_0}^{x_f} f(x)^4 dx \quad ; \quad I_Y = \frac{I_X}{2} + \pi\rho \int_{x_0}^{x_f} x^2 f(x)^2 dx \quad (14.116)$$

2. Demuestre que para un sistema ortogonal de ejes  $XYZ$ , se cumplen las siguientes propiedades

$$I_X + I_Y + I_Z = 2 \sum_i m_i r_i^2 \quad (14.117)$$

$$I_X \leq I_Y + I_Z \quad ; \quad I_Y \leq I_X + I_Z \quad ; \quad I_Z \leq I_X + I_Y \quad (14.118)$$

donde  $\mathbf{r}_i$  es la posición de la  $i$ -ésima partícula, con respecto al origen del sistema coordenado. Las Ecs. (14.118) nos dicen que los momentos de inercia  $I_X$ ,  $I_Y$ ,  $I_Z$  cumplen una desigualdad triangular (Ver Ref. [12], y Sec. 2.3 de la Ref. [13]).

3. Demuestre que para una distribución lineal de masa el tensor  $\mathbf{I}$  tiene un valor propio nulo no-degenerado y un valor propio no-nulo con degeneración 2. Una masa puntual tiene en cambio un valor propio nulo totalmente degenerado. ¿Es esto una manifestación de la simetría del sistema?.
4. Demuestre que el momento de inercia se puede escribir en la forma

$$I = m_i (\mathbf{r}_i \times \mathbf{n}) \cdot (\mathbf{r}_i \times \mathbf{n}) \quad (14.119)$$

y que el tensor de inercia se puede escribir en la forma

$$\mathbf{I} = -m_i (\mathbf{R}_i)^2 \quad ; \quad (\mathbf{R}_i)_{mn} = \varepsilon_{mnk} x_k^{(i)} \quad (14.120)$$

donde los índices  $m, n, k$  denotan componentes y el índice  $i$  rotula partículas.  $\mathbf{R}_i$  es claramente una matriz antisimétrica asociada al punto material  $i$ -ésimo del sistema (discreto) de partículas. ¿Cuál es el análogo en el continuo de las Ecs. (14.119, 14.120)?.

5. Demostrar que el movimiento general de un cuerpo rígido con un punto fijo cumple la siguiente condición entre la energía cinética  $T$ , el torque externo neto  $\mathbf{N}$  y el vector velocidad angular  $\boldsymbol{\omega}$

$$\frac{dT}{dt} = \boldsymbol{\omega} \cdot \mathbf{N}$$

6. Dos esferas tienen el mismo diámetro y la misma masa. Una de ellas es sólida y la otra es hueca (una capa esférica). Describa un experimento no destructivo para determinar cuál es la esfera sólida y cual es la hueca.
7. Un cubo homogéneo de aristas de longitud  $l$ , tiene una de sus aristas en contacto con una superficie horizontal, y esta inicialmente en una posición de equilibrio inestable. Se le da entonces un pequeño empujón para que comience a caer. Muestre que la velocidad angular del cubo cuando una de las caras golpea la superficie viene dado por

$$\omega^2 = \frac{Ag}{l} (\sqrt{2} - 1)$$

donde  $A$  es una constante tal que  $A = 3/2$  si la arista no desliza sobre la superficie y  $A = 12/5$  cuando hay deslizamiento sin fricción.

8. Considere una lámina delgada homogénea que yace en el plano  $X_1X_2$ . Muestre que el tensor de inercia toma la forma

$$\mathbf{I} = \begin{pmatrix} A & -C & 0 \\ -C & B & 0 \\ 0 & 0 & A+B \end{pmatrix}$$

y que si hacemos una rotación en un ángulo  $\theta$  alrededor del eje  $X_3$ , el tensor de inercia resultante posee la misma estructura con los reemplazos  $A \rightarrow A'$ ,  $B \rightarrow B'$ ,  $C \rightarrow C'$  dados por

$$A' = A \cos^2 \theta - C \sin 2\theta + B \sin^2 \theta ; \quad B' = A \sin^2 \theta + C \sin 2\theta + B \cos^2 \theta ; \quad C' = C \cos 2\theta - \frac{1}{2}(B - A) \sin 2\theta$$

y que por tanto los ejes  $X'_1$  y  $X'_2$  serán ejes principales si la rotación se hace en un ángulo

$$\theta = \frac{1}{2} \tan^{-1} \left( \frac{2C}{B - A} \right)$$

9. Aplique el teorema anterior para calcular el tensor de inercia en los siguientes casos (a) una lámina elíptica cuyo centro coincide con el origen de coordenadas. (b) Una lámina elíptica en donde uno de los focos coincide con el origen de coordenadas. Utilice valores arbitrarios de la excentricidad  $\varepsilon$  y el semieje mayor  $a$ .
10. Encuentre la matriz  $\mathbf{B}$  que diagonaliza al tensor  $\mathbf{I}$  mediante una transformación de similaridad, utilizando los siguientes métodos (a) Directamente de los resultados del problema 8. (b) Calculando los valores y vectores propios de  $\mathbf{I}$  y construyendo la matriz diagonalizante con los vectores propios (ver sección 12.3). Verifique que la matriz diagonalizada tiene los valores propios en la diagonal y dé una interpretación física a tales valores propios.

# Capítulo 15

## Oscilaciones

Un tipo de sistema clásico de muy amplia utilidad en diversas ramas de la Física es aquél que consiste de un conjunto de elementos que oscilan alrededor de sus posiciones de equilibrio. Si la desviación del equilibrio estable de cada elemento del sistema es suficientemente pequeña, entonces el movimiento de éste se puede simular como un conjunto de osciladores armónicos acoplados. Estudiaremos en este capítulo el comportamiento de un sistema discreto de osciladores. Las técnicas matemáticas a emplear serán muy similares a las que se usan para estudiar la mecánica del cuerpo rígido, aunque los sistemas mecánicos considerados aquí, no necesariamente deben involucrar cuerpos rígidos.

El hecho de que existan agregados de partículas que sean estables, indica la presencia de fuerzas restauradoras que en límite de pequeñas oscilaciones alrededor de sus posiciones de equilibrio, pueden aproximarse a osciladores armónicos. De esta manera podemos estudiar sistemas tan diversos como partículas acopladas por resortes, vibraciones de moléculas o de la red cristalina de un sólido por efectos térmicos, etc. Adicionalmente, se puede utilizar el límite apropiado para que una red discreta de osciladores se convierta en un sistema continuo (tales como cuerdas, superficies elásticas etc.), como veremos más adelante cuando estudiemos la mecánica de medios continuos, o teoría clásica de campos.

### 15.1. Pequeñas oscilaciones y equilibrio estable

Consideraremos sistemas conservativos en donde la energía potencial solo es función de la posición. Adicionalmente, asumiremos que las ecuaciones de transformación que definen a las coordenadas generalizadas del sistema  $q_1, \dots, q_n$  no dependen explícitamente del tiempo. En consecuencia, se excluye la posibilidad de ligaduras dependientes del tiempo explícitamente. Decimos que el sistema está en equilibrio cuando cada una de las fuerzas generalizadas (2.21) asociadas a las coordenadas  $q_i$  se anulan

$$Q_i = - \left( \frac{\partial V}{\partial q_i} \right)_0 = 0 \quad (15.1)$$

la energía potencial posee por lo tanto un extremo en la configuración de equilibrio del sistema, que describimos por  $q_{01}, \dots, q_{0n}$  y que denotamos con el subíndice “0” en la Ec. (15.1). Si el sistema está inicialmente en equilibrio con velocidades generalizadas nulas i.e.  $\dot{q}_i(0) = 0$ , el sistema continuará indefinidamente en el equilibrio.

Intuitivamente tenemos la idea de lo que es el equilibrio estable e inestable. Por ejemplo, un lápiz que se suspende verticalmente con su punta como punto de apoyo, abandona fácilmente su configuración de equilibrio ante la mas mínima perturbación. En cambio un péndulo simple vertical no tiende a alejarse de su posición de equilibrio cuando ejercemos una pequeña perturbación, más bien el sistema oscila ligeramente alrededor del equilibrio gracias a la presencia de fuerzas restauradoras que apuntan siempre hacia el punto de equilibrio.

Clasificamos como posición de equilibrio estable a aquella en la cual una pequeña perturbación del sistema resulta solo en un movimiento acotado en cierta región alrededor del equilibrio, con la condición adicional de que



esta región de oscilación se puede hacer arbitrariamente pequeña modulando adecuadamente la perturbación<sup>1</sup>. El equilibrio se dice inestable si una perturbación arbitrariamente pequeña se traduce en movimiento acotado en una región finita o en movimiento no acotado. Vemos pues que el péndulo vertical y el lápiz descritos arriba están en equilibrio estable e inestable respectivamente de acuerdo con esta definición.

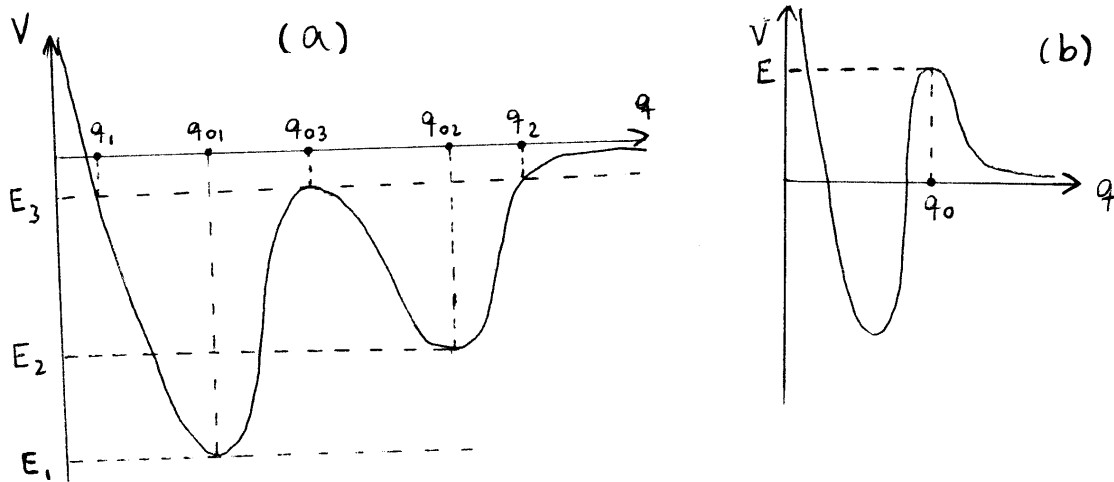


Figura 15.1: Curvas de energía potencial ilustrando el equilibrio estable e inestable en un sistema con un grado de libertad. (a) Si la partícula está en  $q_{01}$  en  $t = 0$  y la energía es  $E_1 + \Delta E$ , con  $\Delta E$  suficientemente pequeño, el movimiento es acotado, y la región permitida disminuye arbitrariamente a medida que disminuye  $\Delta E$ , hablamos entonces de equilibrio estable. Si la partícula está en  $q_{03}$  para  $t = 0$ , y su energía es  $E_3 + \Delta E$ , el movimiento es acotado, pero el tamaño mínimo del intervalo permitido es  $[q_1, q_2]$  aún para  $\Delta E$  arbitrariamente pequeño. El equilibrio es inestable. (b) Si la partícula está en  $q_0$ , el movimiento será no acotado para energías  $E + \Delta E$  independientemente del tamaño de  $\Delta E$ , este punto es de equilibrio inestable.

Podemos ver que si el extremo en el potencial es un mínimo local entonces el equilibrio tiene que ser estable. Para demostrarlo tendremos en cuenta que en cierta vecindad alrededor del punto de equilibrio, el potencial se incrementa en cualquier dirección. Supongamos que el sistema tiene inicialmente la energía total  $V(q_{0i})$  i.e. está en reposo en la configuración de equilibrio. Cuando hacemos un incremento de energía  $\Delta E$  (perturbación) el sistema adquiere energía cinética cuyo valor en  $t = 0$  es justamente  $\Delta E$ , como en cualquier dirección la energía potencial aumenta, entonces la energía cinética debe disminuir, podemos hacer un incremento de energía (perturbación) suficientemente pequeño para que

$$\Delta E \leq V_{\text{máx}} - V(q_{0i}) \quad (15.2)$$

donde  $V_{\text{máx}}$  es el máximo valor que toma el potencial dentro de una región alrededor del equilibrio en la cual el potencial aumenta en todas direcciones (denotamos a esta región por  $R_{\text{máx}}$ ). Si se cumple la condición (15.2) entonces existe un valor del potencial  $V_c$  dentro de la región en cuestión para la cual  $\Delta E = V_c - V(q_{0i})$ . El conjunto de puntos en donde  $V(q_{0i}) \leq V \leq V_c$  define una región que denotamos por  $R_c$ , y claramente  $R_c \subseteq R_{\text{máx}}$ . En los puntos de  $R_c$  en donde  $\Delta E = V_c - V(q_{0i})$  la energía cinética es cero (punto de retorno) y solo puede volver a aumentar si el sistema se mueve hacia donde disminuye el potencial. Esto impide al sistema salir de la región  $R_c$  ya que en los puntos inmediatamente fuera de tal región el potencial es superior al de cualquier punto en  $R_c$ . Finalmente para una función bien comportada, la región  $R_c$  disminuye arbitrariamente a medida que disminuimos a  $\Delta E$ .

<sup>1</sup>El equilibrio inestable también puede producir movimiento acotado como se vé en la Fig. 15.1a. Sin embargo, vemos que el intervalo permitido mínimo para el movimiento, no se puede hacer arbitrariamente pequeño, incluso si la perturbación se hace arbitrariamente pequeña. En otros casos con el de la figura 15.1b, cualquier incremento en la energía conduce a movimiento no acotado.

En un punto en el cual hay un máximo local del potencial, existe una vecindad alrededor del punto en la cual la energía cinética aumenta a medida que nos alejamos del equilibrio dado que el potencial disminuye en cualquier dirección. Cuando hay un máximo local de una función bien comportada existe una región finita en la cual el potencial decrece monótonamente y se puede ver que dentro de esta región no hay puntos de retorno. En consecuencia si  $\Delta E > 0$ , el sistema se mueve mínimo en esta región finita accesible aún cuando la perturbación sea arbitrariamente pequeña. Gráficamente, estas características son fáciles de visualizar para sistemas unidimensionales como se puede apreciar en las figuras 15.1a,b.

Estaremos interesados en el movimiento del sistema dentro de una pequeña región alrededor de un punto de equilibrio estable. Dado que las desviaciones del equilibrio son pequeñas, podemos hacer una expansión de Taylor para las funciones involucradas alrededor del punto de equilibrio y retener solo los términos de más bajo orden. Denotaremos las desviaciones de las coordenadas generalizadas con respecto al equilibrio por  $\eta_i$

$$q_i = q_{0i} + \eta_i \quad (15.3)$$

estas se tomarán como las nuevas coordenadas generalizadas del movimiento. Una expansión de Taylor del potencial alrededor del punto de equilibrio nos da

$$\begin{aligned} V(q_1, \dots, q_n) &= V(q_{01}, \dots, q_{0n}) + \left( \frac{\partial V}{\partial q_i} \right)_0 (q_i - q_{0i}) + \frac{1}{2!} \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 (q_i - q_{0i})(q_j - q_{0j}) + \dots \\ V(q_1, \dots, q_n) &= V(q_{01}, \dots, q_{0n}) + \left( \frac{\partial V}{\partial q_i} \right)_0 \eta_i + \frac{1}{2} \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 \eta_i \eta_j + \dots \end{aligned} \quad (15.4)$$

se ha usado convención de suma sobre índices repetidos. Los términos lineales en  $\eta_i$  se anulan debido a la condición de equilibrio (15.1). El primer término es una constante que se puede remover redefiniendo el cero de potencial. Por tanto, los primeros términos no triviales son los términos cuadráticos en  $\eta_i$ . Usaremos entonces la aproximación

$$V(q_1, \dots, q_n) \simeq \frac{1}{2} V_{ij} \eta_i \eta_j \quad ; \quad V_{ij} \equiv \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 \quad (15.5)$$

Los coeficientes  $V_{ij}$  son constantes que dependen de los valores de las  $q_i$  en el punto de equilibrio. De la definición es claro que los  $V_{ij}$  son simétricos i.e.  $V_{ij} = V_{ji}$  (siempre que las segundas derivadas parciales sean continuas). Es posible que bajo diversas circunstancias, algunos de estos coeficientes se anulen. Por ejemplo, si el potencial es independiente de una coordenada  $q_k$  el equilibrio ocurrirá para un valor arbitrario de esta coordenada y hablaremos de equilibrio neutro o indiferente con respecto a dicha coordenada. También puede ocurrir que el potencial se comporte como cuadrático (sin algunos términos cruzados) en tal punto, de modo que uno o más de los  $V_{ij}$  serán nulos. Cualquiera de estos casos debe ser tratado por aparte con respecto a la formulación que desarrollaremos.

Una expansión en serie es también posible para la energía cinética. Como las transformaciones a las  $q_i$  no dependen explícitamente del tiempo (y por lo tanto, tampoco las transformaciones a las  $\eta_i$ ) la energía cinética resulta ser una función homogénea de grado dos en las velocidades como se puede verificar a partir de (2.24)

$$T = \frac{1}{2} m_{ij} \dot{q}_i \dot{q}_j = \frac{1}{2} m_{ij} \dot{\eta}_i \dot{\eta}_j \quad (15.6)$$

la Ec. (2.24) muestra que los coeficientes  $m_{ij}$  son funciones de las coordenadas y que  $m_{ij} = m_{ji}$ . Por tanto también pueden ser expandidos alrededor de la configuración de equilibrio.

$$m_{ij}(q_1, \dots, q_n) = m_{ij}(q_{01}, \dots, q_{0n}) + \left( \frac{\partial m_{ij}}{\partial q_k} \right)_0 \eta_k + \dots \quad (15.7)$$

pero dado que la ecuación (15.6) ya es cuadrática en los  $\dot{\eta}_i$ , el término más bajo que no se anula (primer término no trivial) en  $T$ , se obtiene del primer término en la expansión (15.7). Definimos entonces

$$T_{ij} \equiv m_{ij}(q_{01}, \dots, q_{0n}) \quad (15.8)$$

y escribimos la energía cinética a orden cuadrático en  $\dot{\eta}$  como

$$T \simeq \frac{1}{2} T_{ij} \dot{\eta}_i \dot{\eta}_j \quad (15.9)$$

es claro de (15.8), que los coeficientes  $T_{ij}$  también son simétricos, de modo que la expresión total no se afecta con un intercambio de índices. El Lagrangiano se obtiene entonces a partir de las Ecs. (15.5) y (15.9)

$$L = \frac{1}{2} (T_{ij} \dot{\eta}_i \dot{\eta}_j - V_{ij} \eta_i \eta_j) \quad (15.10)$$

tomando las  $\eta$ 's como coordenadas generalizadas, podemos construir las ecuaciones de Lagrange

$$\begin{aligned} \frac{\partial L}{\partial \dot{\eta}_k} &= \frac{T_{ij}}{2} \frac{\partial \dot{\eta}_i}{\partial \dot{\eta}_k} \dot{\eta}_j + \frac{T_{ij}}{2} \frac{\partial \dot{\eta}_j}{\partial \dot{\eta}_k} \dot{\eta}_i = \frac{T_{ij}}{2} \delta_{ik} \dot{\eta}_j + \frac{T_{ij}}{2} \delta_{jk} \dot{\eta}_i \\ \frac{\partial L}{\partial \eta_k} &= \frac{T_{kj}}{2} \dot{\eta}_j + \frac{T_{ik}}{2} \dot{\eta}_i = \frac{T_{ki}}{2} \dot{\eta}_i + \frac{T_{ki}}{2} \dot{\eta}_i \end{aligned}$$

donde hemos usado la simetría de los  $T_{ij}$  y el hecho de que los índices de suma son mudos, tenemos entonces que

$$\frac{\partial L}{\partial \dot{\eta}_k} = T_{ki} \dot{\eta}_i \quad \Rightarrow \quad \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\eta}_k} \right) = T_{ki} \ddot{\eta}_i$$

similarmente, usando la simetría de las  $V_{ij}$

$$\begin{aligned} \frac{\partial L}{\partial \eta_k} &= -\frac{V_{ij}}{2} \left( \frac{\partial \eta_i}{\partial \eta_k} \eta_j + \frac{\partial \eta_j}{\partial \eta_k} \eta_i \right) = -\frac{V_{ij}}{2} (\delta_{ik} \eta_j + \delta_{jk} \eta_i) \\ &= -\frac{V_{kj}}{2} \eta_j - \frac{V_{ik}}{2} \eta_i = -V_{ki} \eta_i \end{aligned}$$

las ecuaciones de Lagrange quedan

$$T_{ki} \ddot{\eta}_i + V_{ki} \eta_i = 0 \quad (15.11)$$

cada una de las Ecs. (15.11) involucra en general a todos los  $\eta_i$  y forman un conjunto acoplado de ecuaciones diferenciales simultáneas que debemos resolver para describir el movimiento del sistema alrededor del equilibrio.

En la mayoría de aplicaciones, la energía cinética se puede escribir de modo que no posea productos cruzados, con lo cual el Lagrangiano (15.10) se escribe como

$$L = \frac{1}{2} (T_i \dot{\eta}_i^2 - V_{ij} \eta_i \eta_j) \quad (15.12)$$

y las ecuaciones de movimiento (15.11) quedan

$$T_k \ddot{\eta}_k + V_{ki} \eta_i = 0 \quad (\text{no suma sobre } k) \quad (15.13)$$

## 15.2. Solución de las ecuaciones de movimiento como problema de valores propios

Las Ecs. (15.11) son ecuaciones diferenciales lineales con coeficientes constantes. Como se trata de osciladores acoplados es natural usar una función de ensayo de naturaleza armónica, por lo cual realizaremos el siguiente ansatz

$$\eta_i = C a_i e^{-i\omega t} \quad (15.14)$$

donde  $C a_i$  es una amplitud compleja de la oscilación para cada coordenada  $\eta_i$ , el factor  $C$  es un factor de escala que se introduce por conveniencia y que es igual para todas las coordenadas. Por supuesto se sobreentiende

que es la parte real de (15.14) la que corresponde a la solución física del problema. Sustituyendo (15.14) en (15.11) se obtiene

$$\begin{aligned} -C\omega^2 a_i T_{ki} e^{-i\omega t} + C a_i V_{ki} e^{-i\omega t} &= 0 \\ (V_{ki} - \omega^2 T_{ki}) a_i &= 0 \end{aligned} \quad (15.15)$$

El sistema (15.15), consiste en un conjunto de  $n$  ecuaciones homogéneas para las  $a_i$ , y por tanto solo tendrán solución no trivial si el determinante de los coeficientes se anula

$$\det(\mathbf{V} - \omega^2 \mathbf{T}) = \begin{vmatrix} V_{11} - \omega^2 T_{11} & V_{12} - \omega^2 T_{12} & V_{13} - \omega^2 T_{13} & \cdots \\ V_{21} - \omega^2 T_{21} & V_{22} - \omega^2 T_{22} & V_{23} - \omega^2 T_{23} & \cdots \\ V_{31} - \omega^2 T_{31} & V_{32} - \omega^2 T_{32} & V_{33} - \omega^2 T_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} = 0 \quad (15.16)$$

esta condición secular, que consiste en encontrar las  $n$  raíces  $\omega_i^2$  del polinomio característico, nos provee de los valores de las frecuencias  $\omega$  para los cuales (15.14) es solución de (15.11). Para cada valor de  $\omega^2$  las Ecs. (15.15) pueden ser resueltas para las amplitudes  $a_i$  o más exactamente para los cocientes de  $n-1$  amplitudes con la amplitud restante.

Matricialmente podemos reescribir las Ecs. (15.15) sustituyendo  $\omega^2 \rightarrow \lambda$  en la forma

$$\mathbf{V}\mathbf{a} = \lambda \mathbf{T}\mathbf{a} \quad (15.17)$$

que coincide con la ecuación de valores propios modificada que se estudió en la sección 12.14 como se puede ver de la Ec. (12.117)<sup>2</sup>. Tenemos además que tanto  $\mathbf{T}$  como  $\mathbf{V}$  son matrices simétricas y reales. Por otro lado, la energía cinética dada por (15.9) se puede reescribir en forma matricial

$$T = \frac{1}{2} \tilde{\eta} \mathbf{T} \eta$$

pero la energía cinética es positiva y solo es cero cuando todas las velocidades generalizadas son cero, esto nos lleva a concluir que  $\mathbf{T}$  es una matriz definida positiva ya que un término de la forma  $\mathbf{z} \mathbf{T} \mathbf{z}$  corresponde a dos veces la energía cinética cuando  $\mathbf{z}$  se interpreta como arreglo vectorial de velocidades generalizadas  $\dot{\eta}$ . Finalmente, la condición de que el potencial se esté evaluando en un mínimo local nos lleva a la condición

$$\eta_i V_{ij} \eta_j \geq 0 \quad \Leftrightarrow \quad \tilde{\eta} \mathbf{V} \eta \geq 0$$

de modo que la matriz  $\mathbf{V}$  es positiva. Como consecuencia se cumplen todas las hipótesis establecidas en la sección 12.14 es decir tenemos una ecuación de valores propios modificada de la forma (15.17) donde  $\mathbf{T}$  y  $\mathbf{V}$  son matrices reales y simétricas, siendo  $\mathbf{T}$  definida positiva y  $\mathbf{V}$  positiva. En dicha sección se concluyó que bajo estas hipótesis

1. Los valores propios  $\lambda$  en (15.17) son reales no negativos: en este contexto, significa que  $\omega^2$  es no negativo y por lo tanto las frecuencias  $\omega$  son reales. De la Ec. (12.123) se puede ver claramente que si  $\mathbf{V}$  no fuera positiva (i.e. no estuviéramos en un mínimo local del potencial) entonces es posible que  $\lambda$  sea negativo y por lo tanto  $\omega$  sería imaginario. Por otro lado, se puede ver que introducir un  $\omega$  imaginario en el ansatz (15.14) produciría movimiento no acotado en  $\eta_i$  con el tiempo. Este movimiento sería obviamente inestable con lo cual se vé desde otro punto de vista, que la estabilidad requiere un mínimo local de potencial.
2. Las componentes de los autovectores  $\mathbf{a}$  son reales, excepto por una posible fase global: En nuestro contexto, vemos que una fase compleja global en la amplitud se puede absorber en el factor  $C$  de la Ec. (15.14), de modo que la posibilidad de  $\mathbf{a}$  complejo resulta aún más irrelevante.

<sup>2</sup>Similarmente, (15.16), corresponde a la ecuación secular modificada (12.131).

3. Los autovectores  $\mathbf{a}$  correspondientes a valores propios diferentes son ortogonales bajo la métrica definida por  $\mathbf{T}$ . En este caso, si interpretamos a  $\mathbf{z}$  como un vector cuyas componentes son velocidades generalizadas, podemos decir que la norma al cuadrado inducida para  $\mathbf{z}$  viene dada por dos veces la energía cinética

$$\begin{aligned}\|\mathbf{z}\|^2 &\equiv \tilde{\mathbf{z}}\mathbf{T}\mathbf{z} \equiv \tilde{\eta}\mathbf{T}\dot{\eta} = 2T \\ \|\mathbf{z}\| &= \sqrt{2T}\end{aligned}$$

la ortogonalidad con la métrica  $\mathbf{T}$  está definida por la propiedad (12.125). Si queremos normalizar los vectores propios debemos imponer una condición de normalización de la forma (12.126), recordando que ambas ecuaciones se sintetizan en la Ec. (12.127)

$$\tilde{\mathbf{A}}\mathbf{T}\mathbf{A} = \mathbf{1} \quad (15.18)$$

siendo  $\mathbf{A}$  la matriz de vectores propios ortonormales bajo la métrica  $\mathbf{T}$ .

4. La matriz de autovectores ortonormales  $\mathbf{A}$  diagonaliza tanto a  $\mathbf{T}$  como a  $\mathbf{V}$  a través de una transformación de congruencia.  $\mathbf{T}$  se diagonaliza a la identidad Ec. (15.18) y  $\mathbf{V}$  se diagonaliza a una matriz diagonal en donde los valores propios yacen en la diagonal.

$$\tilde{\mathbf{A}}\mathbf{V}\mathbf{A} = \lambda \quad ; \quad \lambda \equiv \text{diag}(\omega_1^2, \omega_2^2, \dots, \omega_n^2) \quad (15.19)$$

retomando la Ec. (12.123) en nuestro contexto tenemos que

$$\omega_k^2 = \frac{\tilde{\mathbf{a}}_{(k)}\mathbf{V}\mathbf{a}_{(k)}}{\tilde{\mathbf{a}}_{(k)}\mathbf{T}\mathbf{a}_{(k)}} = \frac{2V(\mathbf{a}_{(k)})}{2T(\mathbf{a}_{(k)})}$$

el denominador es dos veces la energía cinética con velocidades generalizadas  $a_{i(k)}$ , en tanto que el numerador es dos veces la energía potencial asociada a coordenadas  $a_{i(k)}$ . Insistimos nuevamente en que la positividad del numerador está asociada con el hecho de que la energía potencial está evaluada en un mínimo local.

Finalmente recordemos que la transformación de congruencia que diagonaliza a  $\mathbf{T}$  y  $\mathbf{V}$  se puede interpretar como el paso desde un sistema de ejes inclinados no cartesianos a un conjunto de ejes principales cartesianos. El proceso total de obtención de las frecuencias fundamentales de pequeñas oscilaciones es un tipo particular de transformaciones a ejes principales (que diagonaliza a  $\mathbf{V}$ ) que también involucran una transformación del tensor métrico (desde  $\mathbf{T}$  hasta la identidad que es el tensor métrico cartesiano<sup>3</sup>).

### 15.2.1. Un ejemplo con dos grados de libertad

A manera de ejemplo, pensemos en una partícula de masa  $m$  con dos grados de libertad  $(x_1, x_2)$  para la cual el Lagrangiano viene dado por (15.10)

$$L = \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) - \frac{1}{2}V_{ij}x_ix_j$$

donde las  $V_{ij}$  son constantes. Como las coordenadas son cartesianas podemos normalizar al tensor  $\mathbf{T}$  como  $T_{ij} = \delta_{ij}$ . De modo que la ecuación de valores propios modificada (12.117), se convierte en la ecuación de valores propios ordinaria Ec. (12.35). La transformación de congruencia (12.130) tiene soluciones solo cuando se satisface la relación (12.132), que en dos dimensiones se escribe

$$\begin{vmatrix} V_{11} - \lambda & V_{12} \\ V_{21} & V_{22} - \lambda \end{vmatrix} = 0 \quad \Leftrightarrow \quad (V_{11} - \lambda)(V_{22} - \lambda) - V_{12}V_{21} = 0$$

<sup>3</sup>Estrictamente hay que definir  $x'_i = x_i\sqrt{m_i}$  para que el tensor métrico (tensor cinético) sea de la forma  $T_{ij} = \delta_{ij}$  (ver sección 12.15, Pág. 320).

que resulta en una ecuación cuadrática para  $\lambda$  con soluciones

$$\lambda_{1,2} = \frac{1}{2} \left( V_{11} + V_{22} \pm \sqrt{(V_{11} - V_{22})^2 + 4V_{12}V_{21}} \right) \quad (15.20)$$

asociado a los valores propios  $\lambda_i$  hay asociados vectores propios  $\mathbf{a}_{(i)}$ . Denotamos la  $j$ -ésima componente del  $i$ -ésimo vector en la forma  $a_{ji}$ . El vector propio  $\mathbf{a}_{(i)}$  se define con

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix} = \lambda_i \begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix} \Rightarrow \begin{pmatrix} V_{11}a_{1i} + V_{12}a_{2i} \\ V_{21}a_{1i} + V_{22}a_{2i} \end{pmatrix} = \lambda_i \begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix}$$

$$\begin{pmatrix} V_{11}a_{1i} + V_{12}a_{2i} - \lambda_i a_{1i} \\ V_{21}a_{1i} + V_{22}a_{2i} - \lambda_i a_{2i} \end{pmatrix} = 0$$

para  $i = 1$  queda

$$V_{11}a_{11} + V_{12}a_{21} - \lambda_1 a_{11} = 0 \quad (15.21)$$

$$V_{21}a_{11} + V_{22}a_{21} - \lambda_1 a_{21} = 0 \quad (15.22)$$

similarmente se pueden escribir las expresiones para  $i = 2$

$$V_{11}a_{12} + V_{12}a_{22} - \lambda_2 a_{12} = 0 \quad (15.23)$$

$$V_{21}a_{12} + V_{22}a_{22} - \lambda_2 a_{22} = 0 \quad (15.24)$$

solo una del par de ecuaciones para cada  $i$  es linealmente independiente, tomemos las Ecs. (15.21) y (15.24) y asumiendo suma sobre  $j$  pero no sobre  $i$  reescribimos estas ecuaciones en la forma

$$V_{ij}a_{ji} - \lambda_i a_{ii} = 0 \quad ; \quad \text{suma sobre } j \text{ no sobre } i \quad (15.25)$$

reescribiendo la Ec. (15.25) y usando la condición de normalización de los autovectores<sup>4</sup> escribimos

$$a_{ji}(V_{ij} - \lambda_i \delta_{ij}) = 0 \quad ; \quad a_{1i}^2 + a_{2i}^2 = 1 \quad ; \quad i = 1, 2 \quad ; \quad \text{suma sobre } j \text{ no sobre } i \quad (15.26)$$

este problema se puede resolver en forma general. Sin embargo, nos restringiremos aquí a estudiar dos casos límites

**Primer caso:** Asumimos  $V_{11} > V_{22} > 0$  y  $0 \neq |V_{21}| = |V_{12}| \ll (V_{11} - V_{22})$ . Con estas condiciones podemos definir un parámetro pequeño de la forma

$$\delta \equiv \frac{V_{12}}{(V_{11} - V_{22})}$$

escribiendo los autovalores (15.20) en términos de  $\delta$

$$\lambda_{1,2} = \frac{1}{2} \left( V_{11} + V_{22} \pm (V_{11} - V_{22}) \sqrt{1 + \frac{4V_{12}^2}{(V_{11} - V_{22})^2}} \right) = \frac{1}{2} \left( V_{11} + V_{22} \pm (V_{11} - V_{22}) \sqrt{1 + 4\delta^2} \right)$$

$$\lambda_{1,2} \cong \frac{1}{2} [V_{11} + V_{22} \pm (V_{11} - V_{22}) (1 + 2\delta^2)] = \frac{1}{2} [V_{11} + V_{22} \pm (V_{11} - V_{22}) \pm 2\delta^2 (V_{11} - V_{22})]$$

$$\lambda_{1,2} \cong \frac{1}{2} \left[ V_{11} + V_{22} \pm (V_{11} - V_{22}) \pm 2\delta^2 V_{12} \frac{(V_{11} - V_{22})}{V_{12}} \right] = \frac{1}{2} \left[ V_{11} + V_{22} \pm (V_{11} - V_{22}) \pm 2\delta^2 V_{12} \frac{1}{\delta} \right]$$

$$\lambda_{1,2} \cong \frac{1}{2} [V_{11} + V_{22} \pm (V_{11} - V_{22}) \pm 2V_{12}\delta]$$

los valores propios quedan finalmente

$$\lambda_1 \cong V_{11} + V_{12}\delta \quad ; \quad \lambda_2 \cong V_{22} - V_{12}\delta \quad (15.27)$$

<sup>4</sup>Recordemos que la normalización se hace de acuerdo con la métrica impuesta por  $\mathbf{T}$ . Pero al ser las coordenadas cartesianas, ésta se reduce a la métrica trivial Euclídea.

veamos como quedan los autovectores a primer orden en  $\delta$ . Para ello reemplacemos (15.27) en la Ec. (15.21)

$$\begin{aligned} V_{11}a_{11} + V_{12}a_{21} - (V_{11} + V_{12}\delta)a_{11} &= 0 \\ V_{12}a_{21} - V_{12}a_{11}\delta &= 0 \\ a_{21} &= a_{11}\delta \end{aligned} \quad (15.28)$$

junto con la condición de normalización

$$\begin{aligned} a_{11}^2 + a_{21}^2 &= 1 \Rightarrow a_{11}^2 + (a_{11}\delta)^2 = 1 \Rightarrow a_{11}^2(1 + \delta^2) = 1 \\ a_{11} &= \sqrt{\frac{1}{1 + \delta^2}} \cong 1 - \frac{1}{2}\delta^2 = 1 + \mathcal{O}(\delta^2) \end{aligned} \quad (15.29)$$

donde hemos tomado la raíz positiva, recuérdese que aún con la condición de normalización, hay una ambigüedad en el vector propio concerniente al sentido de dicho vector. El vector propio normalizado asociado a  $\lambda_1$  se obtiene entonces de (15.29) y (15.28) y se escribe como

$$\mathbf{a}_{(1)} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \simeq \begin{pmatrix} 1 - \frac{1}{2}\delta^2 \\ \delta - \frac{1}{2}\delta^3 \end{pmatrix} = \begin{pmatrix} 1 \\ \delta \end{pmatrix} + \mathcal{O}(\delta^2)$$

reemplazando (15.27) en (15.24) y con un procedimiento análogo se obtiene el vector propio asociado a  $\lambda_2$

$$\mathbf{a}_{(2)} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \simeq \begin{pmatrix} -\delta + \frac{1}{2}\delta^3 \\ 1 - \frac{1}{2}\delta^2 \end{pmatrix} = \begin{pmatrix} -\delta \\ 1 \end{pmatrix} + \mathcal{O}(\delta^2)$$

la matriz asociada a los vectores propios se escribe como

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & -\delta \\ \delta & 1 \end{pmatrix} + \mathcal{O}(\delta^2) \quad (15.30)$$

que nos da las relaciones (válidas a primer orden en  $\delta$ )

$$a_{11} = a_{22} \quad ; \quad a_{12} = -a_{21} \quad (15.31)$$

**Segundo caso:** Examinemos el caso en el cual  $V_{12} > V_{22} > 0$  y  $|V_{11} - V_{22}| \ll V_{12} = V_{21}$ . En este caso un parámetro conveniente es  $\varepsilon = (V_{11} - V_{22}) / (8V_{12})$  que es una cantidad pequeña. A primer orden en  $\varepsilon$ , los autovalores y autovectores están descritos por

$$\begin{aligned} \lambda_1 &\simeq \frac{1}{2}(V_{11} + V_{22}) + V_{12} + (V_{11} - V_{22})\varepsilon \\ \lambda_2 &\simeq \frac{1}{2}(V_{11} + V_{22}) - V_{12} - (V_{11} - V_{22})\varepsilon \\ \mathbf{A} &\simeq \begin{pmatrix} \frac{1}{\sqrt{2}}(1 + 2\varepsilon) & -\frac{1}{\sqrt{2}}(1 - 2\varepsilon) \\ \frac{1}{\sqrt{2}}(1 - 2\varepsilon) & \frac{1}{\sqrt{2}}(1 + 2\varepsilon) \end{pmatrix} \end{aligned} \quad (15.32)$$

en este caso también se cumplen las relaciones (15.31), aunque los valores numéricos difieren fuertemente. En la primera matriz los valores fuera de la diagonal son de orden  $\delta$  en tanto que los diagonales son de orden 1. En cambio, los elementos de la matriz en el segundo caso son todos del orden de  $1/\sqrt{2}$ .

Es fácil ver que las matrices (15.30, 15.32) son ortogonales propias a primer orden en  $\delta$  y  $\varepsilon$ . Por tanto, la transformación de congruencia que diagonaliza a  $\mathbf{V}$  es en ambos casos una transformación de similaridad (ver ejercicio 1, Pág. 450). En realidad, esto es cierto incluso para el problema exacto, en virtud de que la matriz de la energía cinética ya es trivial (ver sección 12.15).

Con esta formulación, podemos resolver problemas que involucren a dos osciladores acoplados, cada uno con un grado de libertad. Se deja como ejercicio al lector, el planteamiento de la solución de pequeñas oscilaciones para dichos sistemas.

### 15.3. Problema de valores propios con degeneración

Ya mencionamos que en el caso en el cual existe degeneración hay más grados de libertad disponibles para crear los vectores propios. Los vectores propios posibles escogidos no es de esperarse que sean ortogonales entre sí. No obstante, siempre es posible construir un conjunto de vectores propios ortogonales entre sí, con base en los vectores previamente elegidos. Estos vectores ortogonales se pueden entonces usar para construir una matriz ortogonal  $\mathbf{A}$ .

Vamos a considerar por simplicidad el caso de una doble raíz. Sean  $\mathbf{a}'_{(k)}$  y  $\mathbf{a}'_{(l)}$  un cierto par de vectores propios asociados a una raíz doble  $\lambda$ . Asumiremos que estos vectores ya han sido normalizados según (12.126) i.e. normalizados según la métrica  $\mathbf{T}$ . Cualquier combinación lineal entre  $\mathbf{a}'_k$  y  $\mathbf{a}'_l$  también será un autovector correspondiente a la raíz  $\lambda$ . Construimos entonces un vector  $\mathbf{a}_l$  en la forma

$$\mathbf{a}_l = c_1 \mathbf{a}'_k + c_2 \mathbf{a}'_l \quad (15.33)$$

donde  $c_1, c_2$  son constantes elegidas de tal forma que  $\mathbf{a}_l$  sea ortogonal a  $\mathbf{a}'_k$ . Multiplicando ambos miembros de (15.33) por  $\mathbf{T}\mathbf{a}'_k$  y usando las condiciones de ortonormalidad (12.125) y (12.126) se obtiene

$$\begin{aligned} \tilde{\mathbf{a}}_l \mathbf{T}\mathbf{a}'_k &= c_1 \tilde{\mathbf{a}}'_k \mathbf{T}\mathbf{a}'_k + c_2 \tilde{\mathbf{a}}'_l \mathbf{T}\mathbf{a}'_k = 0 \\ \tilde{\mathbf{a}}_l \mathbf{T}\mathbf{a}'_k &= c_1 + c_2 \tilde{\mathbf{a}}'_l \mathbf{T}\mathbf{a}'_k = 0 \end{aligned}$$

con lo cual se obtiene el cociente entre las constantes

$$\frac{c_1}{c_2} = -\tilde{\mathbf{a}}'_l \mathbf{T}\mathbf{a}'_k \equiv \tau_l \quad (15.34)$$

requiriendo además que  $\mathbf{a}_l$  esté normalizado según (12.126), tenemos una condición adicional sobre tales constantes

$$\tilde{\mathbf{a}}_l \mathbf{T}\mathbf{a}_l = 1 = c_1^2 + c_2^2 + 2c_1 c_2 \tau_l \quad (15.35)$$

las condiciones (15.34) y (15.35) fijan a  $c_1$  y  $c_2$  en forma casi unívoca. Hemos formado entonces el conjunto ortonormal  $\mathbf{a}_l$  y  $\mathbf{a}_k = \mathbf{a}'_k$ . Estos además son ortogonales a los otros vectores propios correspondientes a valores propios diferentes, lo cual ya se ha demostrado. De esta forma ya tenemos un conjunto de  $n$  autovectores  $\mathbf{a}_j$  cuyas componentes pueden formar la matriz  $\mathbf{A}$  que diagonaliza simultáneamente a  $\mathbf{T}$  y a  $\mathbf{V}$ .

Este procedimiento constituye un caso particular del llamado proceso de ortogonalización de Gram-Schmidt (ver ejercicio 8, Pág. 336). Cuando tenemos una raíz de multiplicidad  $m$ , los vectores ortonormalizados se forman por combinaciones lineales de un conjunto original de vectores propios linealmente independientes  $\mathbf{a}'_1, \dots, \mathbf{a}'_m$ . El conjunto ortonormal  $\mathbf{a}_1, \dots, \mathbf{a}_m$  se forma de la siguiente forma: el primero se elige igual a un múltiplo (normalizado) de  $\mathbf{a}'_1$ ;  $\mathbf{a}_2$  se toma como una combinación lineal de  $\mathbf{a}'_1$  y  $\mathbf{a}'_2$  y así sucesivamente. Con este algoritmo, el número de constantes a determinar equivale a la suma de los primeros  $m$  números i.e.  $\frac{1}{2}m(m+1)$ . Las condiciones de normalización nos dan  $m$  condiciones, junto con  $\frac{1}{2}m(m-1)$  condiciones de ortogonalidad, de esta forma los vectores quedan definidos casi unívocamente excepto por sus sentidos pues si se cambia  $\mathbf{a}_k$  por  $-\mathbf{a}_k$  el nuevo vector sigue siendo ortogonal a los otros y normal, esto está relacionado con el hecho de que las condiciones de normalización son cuadráticas e invariantes ante un cambio de signo en todas las componentes de un vector dado<sup>5</sup>. Finalmente, cabe destacar que este conjunto ortogonal no es en general único.

Recordemos además que en general la degeneración está ligada con alguna simetría, por ejemplo en el caso de matrices 3x3 (como el tensor de inercia), la indeterminación en las componentes de los autovectores para una raíz doble indican que todos los vectores en un cierto plano son vectores propios, con lo cual basta elegir dos vectores perpendiculares en dicho plano para formar ejes principales, y los vectores en  $\mathbf{A}$  son vectores normalizados que van a lo largo de tales ejes.

<sup>5</sup>De hecho se le puede colocar una fase global  $e^{i\varphi_k}$  a cada vector propio  $\mathbf{a}_k$ , sin alterar la ortonormalidad ni la condición de vector propio. En particular, un cambio de signo en un vector se puede interpretar como una fase global de la forma  $e^{i\pi}$ .



### 15.3.1. Un ejemplo bidimensional con degeneración

Un caso muy simple que ilustra las características enunciadas para el problema degenerado, consiste en el ejemplo bidimensional dado en la sección (15.2.1). Este ejemplo presenta degeneración cuando la matriz del potencial cumple las condiciones

$$V_{11} = V_{22} \equiv V_0, \quad V_{12} = V_{21} = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = V_0$$

como un paso al límite podemos analizar el problema de dos formas

a) Usando el primer caso en la sección (15.2.1), tomando el límite  $V_{12} \rightarrow 0$ , para luego tomar el límite  $V_{11} \rightarrow V_{22}$  pero de tal forma que  $\delta \equiv V_{12}/(V_{11} - V_{22})$  tienda a cero. En este caso los autovectores en (15.30) quedan de la forma

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (15.36)$$

b) Usando el segundo caso en la sección (15.2.1), tomando el límite  $V_{11} \rightarrow V_{22}$  para luego tomar el límite  $V_{12} \rightarrow 0$ , pero de modo que  $\varepsilon \equiv (V_{11} - V_{22})/(8V_{12})$  tienda a cero. En este caso la Ec. (15.32) nos da los vectores propios

$$\mathbf{b}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} ; \quad \mathbf{b}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad (15.37)$$

cada par de autovectores es combinación lineal del otro i.e.

$$\mathbf{b}_1 = \frac{1}{\sqrt{2}}(\mathbf{a}_1 + \mathbf{a}_2) \quad , \quad \mathbf{b}_2 = \frac{1}{\sqrt{2}}(\mathbf{a}_2 - \mathbf{a}_1) \quad ; \quad \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix}$$

y la matriz de transformación es no singular (por tanto invertible) de manera que se conserva la independencia lineal. Nótese que ambos pares de vectores (15.36) y (15.37) poseen la estructura genérica

$$\mathbf{c}_1 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} ; \quad \mathbf{c}_2 = \begin{pmatrix} -c_2 \\ c_1 \end{pmatrix} \quad (15.38)$$

de modo que los anteriores resultados se generalizan para pares de vectores propios con la estructura (15.38). Por tanto, los vectores propios están determinados por cualquier par de números  $c_1$  y  $c_2$  que formen la estructura (15.38) y cumplan la condición de normalización definida en (12.126). Esto nos muestra que hay un conjunto infinito de pares de vectores propios normalizados y linealmente independientes en el caso en que hay presencia de degeneración.

Existe otra forma de ver los resultados anteriores. Nótese que los autovectores aproximados dados por (15.30) corresponden al caso en el cual los términos principales de la energía potencial  $V_{11}$  y  $V_{22}$ , están sobre la diagonal en tanto  $V_{12}$  está en las posiciones fuera de la diagonal. Si tomamos los autovectores de las Ecs. (15.32) en el límite  $\varepsilon \rightarrow 0$  y hacemos que los autovectores (15.32) transformen a  $V$  según la prescripción  $\mathbf{V}' = \mathbf{A}\mathbf{V}\mathbf{A}$ , obtenemos el tensor de energía potencial transformado

$$\mathbf{V}' = \begin{pmatrix} \frac{1}{2}(V_{11} + V_{22}) + V_{12} & \frac{1}{2}(V_{11} - V_{22}) \\ \frac{1}{2}(V_{11} - V_{22}) & \frac{1}{2}(V_{11} + V_{22}) - V_{12} \end{pmatrix}$$

en el cual el término diferencia  $(V_{11} - V_{22})$  está fuera de la diagonal. En consecuencia, el conjunto de autovectores dados por las Ecs. (15.32) corresponden a la situación física en la cual el pequeño término de energía  $(V_{11} - V_{22})$  está fuera de la diagonal.

## 15.4. Frecuencias de vibración libre y coordenadas normales

Hemos visto que con los valores adecuados de las constantes, la expresión (15.14) representa una solución consistente para las ecuaciones de movimiento (15.11) asociadas a pequeñas oscilaciones. Sin embargo, esta no

es necesariamente la solución más general del problema ya que no ha tenido en cuenta las condiciones iniciales. Por otro lado, en virtud de la linealidad de la Ec. (15.11) se observa inmediatamente que una superposición de soluciones de tal ecuación es también una solución. Combinando este hecho con el teorema matemático que nos dice que las funciones  $e^{i\omega_n t}$  son completas en un intervalo acotado (teorema de Fourier) podemos deducir que la solución más general a las Ecs. (15.11) consiste en una superposición de soluciones de la forma (15.14). Por tanto, si el sistema se desvía ligeramente del equilibrio, es posible que realice oscilaciones alrededor de la configuración de equilibrio con alguna de las frecuencias permitidas  $\omega_1, \dots, \omega_n$ . Por esta razón a las soluciones de la ecuación secular se les conoce usualmente como *frecuencias de vibración libre* o *frecuencias de resonancia* del sistema.

Escribamos entonces la solución general como superposición de soluciones del tipo (15.14), representada como una suma sobre el índice  $k$

$$\eta_i = C_k a_{ik} e^{-i\omega_k t} \quad (15.39)$$

donde  $C_k$  es un factor complejo asociado a cada frecuencia de resonancia  $\omega_k$ . Podría objetarse que para cada solución  $\lambda_k$  de la ecuación secular tanto  $\omega_k$  como  $-\omega_k$  son frecuencias de resonancia ya que ambas están asociadas al mismo vector propio  $\mathbf{a}_k$  pero los factores de escala para cada frecuencia denotados por  $C_k^+$  y  $C_k^-$  podrían ser en general diferentes. Con esta argumentación la solución debe escribirse como

$$\eta_i = a_{ik} (C_k^+ e^{-i\omega_k t} + C_k^- e^{+i\omega_k t}) \quad (15.40)$$

recordemos sin embargo que el movimiento físico está descrito por la parte real de la solución compleja planteada, y la parte real de cualquiera de las dos ecuaciones (15.39) o (15.40) se puede escribir en la forma

$$\eta_i = f_k a_{ik} \cos(\omega_k t + \delta_k) \quad (15.41)$$

donde las amplitudes  $f_k$  y fases  $\delta_k$  se determinan a partir de las condiciones iniciales. Por tanto cualquiera de las soluciones (15.39, 15.40, 15.41) puede representar al movimiento físico. Tomaremos entonces la más conveniente para nuestros propósitos Ec. (15.39). Nótese que en la Ec. (15.39), la fase del movimiento real  $\delta_k$  está contenida en la fase compleja de  $C_k$ , en tanto que la magnitud de  $C_k$  estaría relacionada con la amplitud asociada a cada frecuencia  $\omega_k$ .

Los factores de escala  $C_k$  se pueden determinar fácilmente gracias a las propiedades de ortonormalidad de  $\mathbf{A}$  Ec. (12.127)<sup>6</sup>. En el tiempo  $t = 0$  la parte real de (15.39) se reduce a

$$\eta_i(0) = a_{ik} \operatorname{Re} C_k \quad (15.42)$$

donde hemos tenido en cuenta que los elementos de  $\mathbf{A}$  son reales. Para las velocidades generalizadas  $\dot{\eta}_i$  se obtiene una expresión análoga. Derivando (15.39) y evaluando la parte real de la solución en  $t = 0$

$$\begin{aligned} \dot{\eta}_i &= -i\omega_k C_k a_{ik} e^{-i\omega_k t} \quad ; \quad \operatorname{Re} \dot{\eta}_i(0) = -\omega_k a_{ik} \operatorname{Re}(iC_k) = -\omega_k a_{ik} \operatorname{Re}[i(\operatorname{Re} C_k + i \operatorname{Im} C_k)] = -\omega_k a_{ik} \operatorname{Re}[i \operatorname{Re} C_k - \operatorname{Im} C_k] \\ \dot{\eta}_i(0) &= \omega_k a_{ik} \operatorname{Im} C_k \end{aligned} \quad (15.43)$$

las  $2n$  ecuaciones (15.42) y (15.43) nos permiten evaluar las  $n$  partes reales y  $n$  partes imaginarias a partir de las condiciones iniciales<sup>7</sup>. Para resolver (15.42) es más conveniente escribir dicha ecuación en notación matricial

$$\boldsymbol{\eta}(0) = \mathbf{A} \operatorname{Re} \mathbf{C}$$

multiplicando a la izquierda por el factor  $\tilde{\mathbf{A}} \mathbf{T}$  y usando (12.127) se obtiene la solución para  $\operatorname{Re} \mathbf{C}$

$$\operatorname{Re} \mathbf{C} = \tilde{\mathbf{A}} \mathbf{T} \boldsymbol{\eta}(0) \quad \Leftrightarrow \quad \operatorname{Re} C_l = a_{jl} T_{jk} \eta_k(0) \quad (15.44)$$

<sup>6</sup>Es necesario recordar que la condición de ortonormalidad dada en la Ec. (12.127), no coincide con la condición usual de ortogonalidad, de modo que en general  $\mathbf{A}^{-1} \neq \tilde{\mathbf{A}}$ , como se discutió en la sección 12.15.

<sup>7</sup>Estrictamente, las Ecs. (15.42) y (15.43) deberían contener  $\Re \eta_i(0)$  y  $\Im \eta_i(0)$  a la izquierda de estas expresiones. Sin embargo, asumimos que las condiciones iniciales son reales y por ello dicha notación no es necesaria.

similarmente, de la Ec. (15.43) y usando (12.127) en componentes, se obtiene la parte imaginaria de  $\mathbf{C}$

$$\begin{aligned} \dot{\eta}_k(0) &= \omega_p a_{kp} \operatorname{Im} C_p \Rightarrow \tilde{a}_{lj} T_{jk} \dot{\eta}_k(0) = \omega_p (\tilde{a}_{lj} T_{jk} a_{kp}) \operatorname{Im} C_p \Rightarrow a_{jl} T_{jk} \dot{\eta}_k(0) = \omega_p \delta_{lp} \operatorname{Im} C_p \Rightarrow \\ a_{jl} T_{jk} \dot{\eta}_k(0) &= \omega_l \operatorname{Im} C_l \quad \text{no suma sobre } l \end{aligned}$$

la parte imaginaria de  $\mathbf{C}$  en términos de las condiciones iniciales queda entonces

$$\operatorname{Im} C_l = \frac{1}{\omega_l} a_{jl} T_{jk} \dot{\eta}_k(0) \quad \text{no suma sobre } l \quad (15.45)$$

Las Ecs. (15.44) (15.45), permiten calcular directamente los factores complejos (y por tanto las amplitudes y fases) en términos de las condiciones iniciales  $\boldsymbol{\eta}(0)$ ,  $\dot{\boldsymbol{\eta}}(0)$  y las matrices  $\mathbf{A}$  y  $\mathbf{T}$ .

La solución para cada coordenada dada por (15.39) es en general una suma de oscilaciones armónicas simples en todas las frecuencias  $\omega_k$  que satisfacen la ecuación secular. A menos que todas las frecuencias sean conmensurables, el conjunto de todas las  $\eta_i$  no retornan en general a sus valores iniciales en ningún instante posterior, esto significa que la configuración de las  $\eta_i$  no es una función periódica simple del tiempo<sup>8</sup>. No obstante, la condición de ortonormalidad de  $\mathbf{A}$  nos permite encontrar a partir de las  $\eta_i$  un conjunto de coordenadas generalizadas  $\zeta_i$  que son todas funciones periódicas en el tiempo. A este conjunto de coordenadas generalizadas las llamaremos **coordenadas normales**. Partiendo de (15.39) y multiplicando por  $a_{ni}^{-1}$  a ambos lados

$$\begin{aligned} a_{ni}^{-1} \eta_i &= a_{ni}^{-1} a_{ik} C_k e^{-i\omega_k t} = \delta_{nk} C_k e^{-i\omega_k t} \\ a_{ni}^{-1} \dot{\eta}_i &= C_n e^{-i\omega_n t} \end{aligned}$$

nótese que el resultado de la derecha en la última ecuación depende solo de la frecuencia  $n$  y de su correspondiente factor de escala  $C_n$ ; es por tanto natural definir las nuevas coordenadas generalizadas en la forma

$$a_{ni}^{-1} \eta_i \equiv \zeta_n \quad \Rightarrow \quad \zeta_n = C_n e^{-i\omega_n t} \quad (15.46)$$

la relación inversa será

$$\begin{aligned} a_{mn} a_{ni}^{-1} \eta_i &= a_{mn} \zeta_n \quad \Rightarrow \quad \delta_{mi} \eta_i = a_{mn} \zeta_n \\ \eta_m &= a_{mn} \zeta_n \end{aligned} \quad (15.47)$$

matricialmente se escribe

$$\boldsymbol{\eta} = \mathbf{A} \boldsymbol{\zeta} \quad ; \quad \mathbf{A}^{-1} \boldsymbol{\eta} = \boldsymbol{\zeta} \quad (15.48)$$

comparando con el procedimiento de la sección 12.14.1, Ecs. (12.134), (12.136), vemos que las coordenadas  $\zeta_i$  son precisamente las que diagonalizan la forma cuadrática positiva, que en nuestro contexto está asociada con el potencial

$$V = \frac{1}{2} \tilde{\boldsymbol{\eta}} \mathbf{V} \boldsymbol{\eta} \quad (15.49)$$

por tanto podemos tomar el resultado (12.137) para escribir

$$V = \frac{1}{2} \tilde{\boldsymbol{\zeta}} \boldsymbol{\lambda} \boldsymbol{\zeta} = \frac{1}{2} \omega_k^2 \zeta_k \zeta_k \quad (15.50)$$

la energía cinética es una forma cuadrática definida positiva que se escribe como

$$T = \frac{1}{2} \tilde{\dot{\boldsymbol{\eta}}} \mathbf{T} \dot{\boldsymbol{\eta}} \quad (15.51)$$

<sup>8</sup>Si las coordenadas generalizadas  $\eta_i$  son todas funciones periódicas simples en el tiempo, esto implicaría que la trayectoria en el espacio de fase es cerrada como vimos en el movimiento periódico de libración en el capítulo de Hamilton Jacobi, sección 9.10.

comparando con la Ec. (12.134) vemos que  $\chi \rightarrow \dot{\eta}$  y la coordenada normal asociada es simplemente  $\psi \rightarrow \dot{\zeta}$ . Es decir que en este caso las coordenadas normales asociadas a la matriz definida positiva se obtienen simplemente derivando temporalmente (15.48)<sup>9</sup>

$$\dot{\eta} = \mathbf{A}\dot{\zeta} \quad ; \quad \mathbf{A}^{-1}\dot{\eta} = \dot{\zeta} \quad (15.52)$$

y la forma cuadrática para la energía cinética se obtiene de (12.138)

$$T = \frac{1}{2}\dot{\zeta}\dot{\zeta} = \frac{1}{2}\dot{\zeta}_i\dot{\zeta}_i \quad (15.53)$$

Usando (15.50) y (15.53) podemos construir el Lagrangiano y las Ecuaciones de movimiento en coordenadas normales

$$L = \frac{1}{2}(\dot{\zeta}_k\dot{\zeta}_k - \omega_k^2\zeta_k^2) \quad (15.54)$$

$$\ddot{\zeta}_k + \omega_k^2\zeta_k = 0 \quad (15.55)$$

las soluciones de (15.55) son inmediatas

$$\zeta_k = C_k e^{-i\omega_k t} \quad (15.56)$$

en concordancia con (15.46). Cada una de estas nuevas coordenadas es entonces una función periódica simple del tiempo que involucra a una sola frecuencia de resonancia.

En el lenguaje de la formulación Lagrangiana, el espacio de configuraciones  $\eta \equiv (\eta_1, \dots, \eta_n)$ , es un espacio cartesiano  $n$ -dimensional en el cual cada  $\eta_i$  tiene un eje asociado. El cambio a los ejes asociados a las coordenadas  $\zeta$ , se puede considerar un cambio a **ejes principales**, ya que cuando el espacio de configuraciones se describe con  $\zeta$ , la energía cinética y potencial quedan en su forma canónica (diagonalizada), al igual que el Lagrangiano y las ecuaciones de movimiento. No obstante, esta analogía no es completa, ya que la transformación  $\eta \rightarrow \zeta$  no es unitaria en el caso más general, sino que requiere un reescalamiento de los ejes (ver sección 12.15), lo cual se manifiesta en el hecho de que la matriz  $\mathbf{A}$  que diagonaliza a  $\mathbf{T}$  y  $\mathbf{V}$  no es en general ortogonal. Vale enfatizar que de acuerdo con la discusión en la sección 12.15, la necesidad o no del reescalamiento depende solo de la estructura de  $\mathbf{T}$  y no de  $\mathbf{V}$ . Cuando el reescalamiento no es necesario, la analogía a ejes principales es completa.

Cada conjunto de  $n$  coordenadas normales  $\zeta_k \equiv (\zeta_{k1}, \dots, \zeta_{kn})$ , corresponde a una vibración del sistema con una sola frecuencia y estas oscilaciones componentes suelen denominarse *modos normales de vibración*<sup>10</sup>. En un modo normal dado, todas las partículas vibran con la misma frecuencia  $\omega_k$ , amplitud  $|C_k|$  y fase  $\delta_k$ .

Debemos recordar sin embargo, que los modos normales se deben superponer para formar la solución física, cada modo contribuye con un “peso” es decir con cierta amplitud relativa y fase. La Ec. (15.39) nos muestra que los “pesos” o amplitudes con que cada modo normal contribuye dependen de los coeficientes  $C_k$  asociados con cada modo, pero también de los elementos de matriz  $a_{ij}$ . Las fases en cambio, solo dependen de los  $\omega_k$  y  $C_k$  ya que la matriz  $a_{ij}$  es real.

Si el problema de valores propios está asociado a un número finito de grados de libertad, el número de valores propios es finito. Por tanto, existe un valor propio  $\omega_0^2$  menor que todos los demás, a  $\omega_0$  se le suele denominar **frecuencia angular fundamental**. Ordenando las frecuencias angulares en la forma  $\omega_0 < \omega_1 < \omega_2 < \dots$ , decimos  $\omega_1$  es la primer frecuencia angular excitada,  $\omega_2$  la segunda frecuencia angular excitada etc. Correspondientemente, hablamos de un modo fundamental, el primer modo excitado, etc. Es notable el hecho de que cuando el número de grados de libertad se extiende al infinito numerable, ocurre para la mayor parte de problemas de interés físico, que los valores propios siguen siendo discretos (aunque infinitos) y existe un valor propio menor que todos los demás. Por tal razón, en la extensión infinito numerable se puede extender para casi todos los casos el lenguaje de modo fundamental, primer modo excitado etc.

<sup>9</sup>Recordemos que en el caso más general, el conjunto de coordenadas normales asociadas a la matriz positiva es totalmente independiente del conjunto de coordenadas normales asociadas a la matriz definida positiva, como se aprecia en la sección 12.14.1.

<sup>10</sup>Aunque los modos normales de oscilación no son en general la solución Física del problema, éstas pueden corresponder a soluciones Físicas si las condiciones iniciales son las adecuadas.

Nótese que el punto de partida fué la estipulación de que la oscilación del movimiento es pequeña. Esto fué lo que permitió representar el potencial como una forma cuadrática, la cual es característica del movimiento armónico simple. La transformación a coordenadas normales enfatiza este hecho puesto que el Lagrangiano en tales coordenadas se puede escribir como la suma de Lagrangianos desacoplados correspondientes a osciladores armónicos de frecuencias  $\omega_k$  (ver Ec. 15.54). Por lo tanto, podemos considerar que el movimiento completo para pequeñas oscilaciones se puede obtener excitando los osciladores armónicos con diferentes intensidades y fases. Este punto de vista es muy útil en mecánica cuántica, por ejemplo para la cuantización del campo electromagnético. Las frecuencias de los osciladores armónicos se identifican con las frecuencias de los fotones y las amplitudes de excitación se convierten en los números de ocupación discretos y cuantizados, i.e. el número de fotones para cada frecuencia. Caso similar se presenta en el estudio de otras excitaciones colectivas que se pueden desacoplar en sumas de osciladores armónicos tales como las redes de fonones y magnones en la Física del estado sólido.

En este punto surge otra analogía interesante, el Lagrangiano (15.10) describe un conjunto acoplado (interactuante) de partículas reales. La transformación a coordenadas normales  $\zeta$  nos lleva a un Lagrangiano (15.54) que puede pensarse como un conjunto de partículas (imaginarias), desacopladas (no interactuantes) i.e. un conjunto independiente de osciladores armónicos cada uno con su frecuencia característica<sup>11</sup>. Para resolver el problema real con base en el problema desacoplado (más simple), basta con aplicar la transformación inversa que nos regresa a las coordenadas que describen a las partículas reales. La estrategia fue pasar desde un problema acoplado o interactuante, hacia un problema desacoplado o no interactuante, para resolver el problema desacoplado y luego regresarse al problema real acoplado. Ya usamos esta estrategia en la sección 10.1, en la cual un problema de dos cuerpos acoplados de coordenadas  $\mathbf{r}_1$  y  $\mathbf{r}_2$ , se convirtió en un problema de dos cuerpos desacoplados (imaginarios), uno asociado a la coordenada relativa  $\mathbf{r}$  y otro a la coordenada del centro de masa  $\mathbf{R}$ . Podemos decir en este nuevo lenguaje, que las componentes de  $\mathbf{r}$  y  $\mathbf{R}$  son las coordenadas normales del problema de dos cuerpos. Es interesante ver como la Física recurre una y otra vez a las mismas estrategias generales, aunque difiera en los detalles.

## 15.5. Vibraciones libres de una molécula triatómica

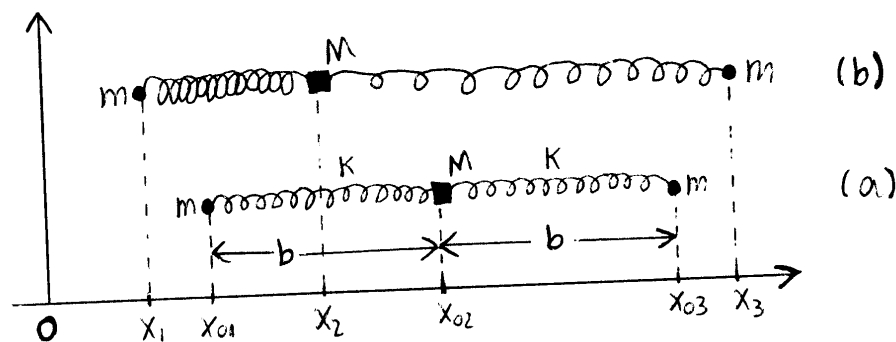


Figura 15.2: Molécula triatómica lineal. La masa de cada átomo de los extremos es  $m$ , en tanto que el átomo central tiene masa  $M$ . (a) Los puntos  $x_{0i}$  señalan las posiciones de equilibrio y (b) los puntos  $x_i$  indican los valores de las posiciones de los átomos en un instante dado.

Ilustraremos la obtención de los modos normales de oscilación considerando el caso de una molécula triatómica lineal, la cual en estado de equilibrio consiste de dos masas  $m$  en los extremos y una masa  $M$  en el centro, la distancia entre cada masa  $m$  y la masa  $M$  se denota por  $b$ . Todos los átomos están en línea recta como lo indica la figura 15.2. Por simplicidad asumiremos que solo hay vibraciones longitudinales y el

<sup>11</sup>Esta es sin embargo, solo una imagen conveniente, ya que los Lagrangianos (15.10) y (15.54) describen exactamente el mismo problema Físico real, aunque en coordenadas distintas.

potencial interatómico se aproximará a aquél equivalente a dos resortes acoplados idénticos que unen a la masa  $M$  con las dos masas  $m$ . Esta aproximación está justificada siempre que haya pequeñas oscilaciones alrededor del equilibrio. No obstante, este modelo hace implícitamente una aproximación adicional, ya que se desprecia la interacción entre las masas  $m$ . De acuerdo con la figura 15.2, la energía potencial se puede escribir en la forma

$$V = \frac{k}{2}(x_2 - x_1 - b)^2 + \frac{k}{2}(x_3 - x_2 - b)^2$$

introducimos entonces las coordenadas relativas a la posición de equilibrio

$$\eta_i = x_i - x_{0i}$$

la figura 15.2 muestra que

$$x_{02} - x_{01} = b = x_{03} - x_{02}$$

La energía potencial en estas coordenadas se escribe como

$$V = \frac{k}{2}(\eta_2 - \eta_1)^2 + \frac{k}{2}(\eta_3 - \eta_2)^2 \quad (15.57)$$

$$V = \frac{k}{2}(\eta_1^2 + 2\eta_2^2 + \eta_3^2 - 2\eta_1\eta_2 - 2\eta_2\eta_3) \quad (15.58)$$

para construir la matriz asociada a  $V$  definimos

$$V \equiv \frac{1}{2}\tilde{\eta}_i V_{ij} \eta_j = \frac{1}{2} \begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \end{pmatrix} \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} \quad (15.59)$$

comparando (15.58) con (15.59) y escribiendo los términos de tal modo que  $V_{ij}$  sea simétrico, el tensor  $\mathbf{V}$  se escribe de la forma

$$\mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \quad (15.60)$$

la energía cinética en estas coordenadas tiene la forma

$$T = \frac{m}{2}(\dot{\eta}_1^2 + \dot{\eta}_3^2) + \frac{M}{2}\dot{\eta}_2^2 = \frac{1}{2} \begin{pmatrix} \dot{\eta}_1 & \dot{\eta}_2 & \dot{\eta}_3 \end{pmatrix} \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} \begin{pmatrix} \dot{\eta}_1 \\ \dot{\eta}_2 \\ \dot{\eta}_3 \end{pmatrix} \quad (15.61)$$

de modo que el tensor  $\mathbf{T}$  es diagonal

$$\mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} \quad (15.62)$$

con base en los tensores (15.60) y (15.62), la ecuación secular (15.16) queda

$$\det(\mathbf{V} - \omega^2 \mathbf{T}) = \begin{vmatrix} k - \omega^2 m & -k & 0 \\ -k & 2k - \omega^2 M & -k \\ 0 & -k & k - \omega^2 m \end{vmatrix} = 0$$

que nos da el polinomio cúbico característico para  $\omega^2$

$$-\omega^2 (k - \omega^2 m) [k(M + 2m) - \omega^2 Mm] = 0$$

cuyas soluciones son inmediatas

$$\omega_1 = 0, \quad \omega_2 = \sqrt{\frac{k}{m}}, \quad \omega_3 = \sqrt{\frac{k}{m} \left(1 + \frac{2m}{M}\right)} \quad (15.63)$$

### 15.5.1. Modos normales de frecuencia cero: traslaciones y rotaciones rígidas

Introduciendo el valor propio  $\omega_1 = 0$  en (15.56) se obtiene

$$\zeta_1 = C_1$$

este resultado es a priori sorprendente ya que no corresponde a un movimiento oscilatorio, sino a una traslación de la molécula como un todo a lo largo del eje de la molécula (dado que solo estamos permitiendo movimiento longitudinal). Pero precisamente hay que tener en cuenta que la traslación rígida de la molécula a lo largo de su eje no produce ningún cambio en la energía potencial<sup>12</sup>, este es un ejemplo del equilibrio indiferente que ya hemos comentado. Este modo de frecuencia nula corresponde entonces a una traslación rígida de la molécula, pues no hay fuerza restauradora asociada a este movimiento. Hemos hecho la suposición de que la molécula posee tres grados de libertad para el movimiento vibracional, pero en realidad uno de ellos es un grado de libertad asociado a este movimiento rígido. Otro aspecto interesante de las frecuencias de resonancia nulas se puede ver a través de la Ec. (12.123), la cual muestra que un valor nulo de  $\omega$  solo se puede obtener si la energía potencial es positiva singular<sup>13</sup>. La Ec. (15.57) muestra que con  $k \neq 0$ , la condición necesaria y suficiente para que  $V$  sea nulo es que  $\eta_1 = \eta_2 = \eta_3$ . Es decir existe un vector propio no nulo de la forma

$$\boldsymbol{\eta} = (\eta, \eta, \eta)^T \quad (15.64)$$

con valor propio nulo, que efectivamente corresponde a traslación rígida de la molécula. Equivalentemente, el vector (15.64) es un vector no nulo que hace que la forma bilineal (15.59) se anule cuando  $V_{ij}$  viene dado por (15.60).

Por otro lado, si relajamos la condición de movimiento longitudinal (que solo nos da un grado de libertad traslacional) y permitimos vibraciones en todas las tres direcciones, el número de grados de libertad de traslación rígida se incrementa a 6 ya que la molécula podrá moverse a lo largo de los tres ejes, o realizar rotaciones uniformes alrededor de tales ejes. Por lo tanto, en un sistema general de  $n$  grados de libertad habrá 6 frecuencias nulas (3 traslacionales y 3 rotacionales) y solo  $n - 6$  frecuencias verdaderas de vibración.

Además de la traslación o rotación rígida, existe una condición adicional bajo la cual ocurren frecuencias nulas y es el caso en el cual tanto la primera como la segunda derivada del potencial se anulan en la posición de equilibrio. En este escenario  $k = 0$  (no hay término cuadrático en la expansión de Taylor del potencial) y aún son posibles pequeñas oscilaciones si la cuarta derivada es no nula (la tercera debe ser nula para que haya equilibrio estable). No obstante, en este caso la oscilación no será de tipo armónico pues el primer término no nulo en la expansión del potencial será de orden cuártico, el caso más simple es un potencial de la forma  $V = ax^4$  con  $a$  positivo, siendo  $x = 0$  el punto de equilibrio. Este es un caso de pequeñas oscilaciones diferente a la aproximación armónica usual, pero no se presenta con mucha frecuencia.

De la discusión anterior se vé entonces que las frecuencias nulas de resonancia no tienen ninguna incidencia en el estudio de los modos vibracionales. Una forma de remover los modos traslacionales (aunque no los rotacionales) es imponer la condición de que el centro de masa permanezca estacionario en el origen. Es decir, ubicar el sistema de referencia inercial en el CM

$$m(x_1 + x_3) + Mx_2 = 0 \quad (15.65)$$

esta ecuación se puede usar para eliminar coordenadas en  $V$  y  $T$ .

<sup>12</sup>Es importante mencionar que esta traslación no es un movimiento que se realiza en el tiempo, ya que de ser así, tendría que incluirse la energía cinética del centro de masa en el Lagrangiano. Por esta razón,  $\zeta_1$  es constante y no lineal en el tiempo. Desde un punto de vista pasivo, también se puede ver como un corrimiento del origen. Recordemos que la invarianza del sistema ante estas traslaciones nos lleva a la conservación del momento lineal total del sistema y es una manifestación de la homogeneidad del espacio. Si la molécula interactúa con agentes externos, este modo traslacional adquiere relevancia.

<sup>13</sup>Si la energía potencial fuera definida positiva entonces  $\omega = 0$  solo podría corresponder a la situación trivial de reposo de todas las partículas en sus posiciones de equilibrio.

### 15.5.2. Vectores propios de la ecuación secular

Volvamos ahora al examen de los modos vibracionales reales. La frecuencia de resonancia  $\omega_2$  descrita en (15.63), se reconoce como la frecuencia de oscilación de una masa  $m$  suspendida por un resorte de constante  $k$  atado a un soporte fijo. En consecuencia, es de esperarse que los átomos de los extremos sean los que estén vibrando en tanto que el átomo del centro está en reposo, con el fin de que sirva como soporte fijo para los resortes. Solo en el tercer modo de vibración  $\omega_3$  la masa central  $M$  participa de la vibración en el movimiento oscilatorio. Estas predicciones se pueden verificar encontrando los vectores propios asociados a cada modo de vibración y utilizando las condiciones iniciales del problema.

Las componentes  $a_{ij}$  de la matriz  $\mathbf{A}$  de los vectores propios se obtienen para cada frecuencia a través de las ecuaciones secular y de normalización

$$(\mathbf{V} - \omega_j^2 \mathbf{T}) \mathbf{a}_{(j)} = 0 \quad ; \quad \tilde{\mathbf{a}}_{(j)} \mathbf{T} \mathbf{a}_{(j)} = 1$$

utilizando (15.60, 15.62), estas ecuaciones se escriben explícitamente como

$$\begin{aligned} (k - \omega_j^2 m) a_{1j} & -ka_{2j} & & = 0 \\ -ka_{1j} & + (2k - \omega_j^2 M) a_{2j} & -ka_{3j} & = 0 \\ & -ka_{2j} & + (k - \omega_j^2 m) a_{3j} & = 0 \end{aligned} \quad (15.66)$$

$$m(a_{1j}^2 + a_{3j}^2) + Ma_{2j}^2 = 1 \quad ; \quad j = 1, 2, 3 \quad (15.67)$$

las Ecs. (15.66, 15.67) determinan completamente a los vectores propios **reales** excepto por un signo global<sup>14</sup>. Con  $\omega_1 = 0$  se obtiene de la primera y tercera de las Ecs. (15.66) que todos los coeficientes  $a_{k1}$  son iguales. Este hecho junto con la condición de normalización (15.67) nos fija el valor de  $a_{k1}$

$$a_{11} = a_{21} = a_{31} = \frac{1}{\sqrt{2m + M}} \quad (15.68)$$

que concuerda con (15.64).

Para el segundo modo, tomando el  $\omega_2$  de (15.63) vemos que  $(k - \omega_2^2 m)$  se anula y las ecuaciones (15.66) muestran que  $a_{22} = 0$  y  $a_{12} = -a_{32}$ . Nuevamente estas relaciones junto con la condición de normalización (15.67) nos determinan el segundo vector propio

$$a_{12} = \frac{1}{\sqrt{2m}} \quad ; \quad a_{22} = 0 \quad ; \quad a_{32} = -\frac{1}{\sqrt{2m}} \quad (15.69)$$

Finalmente, para el valor de  $\omega_3$  en (15.63), se obtiene que  $a_{13} = a_{33}$ . Los vectores propios normalizados son

$$a_{13} = a_{33} = \frac{1}{\sqrt{2m(1 + \frac{2m}{M})}} \quad ; \quad a_{23} = -\frac{2}{\sqrt{2M(2 + \frac{M}{m})}} \quad (15.70)$$

### 15.5.3. Modos normales y modos reales de la molécula triatómica

Con base en las Ecs. (15.68, 15.69, 15.70) se construye la matriz  $\mathbf{A}$  de vectores propios

$$\mathbf{A} = \begin{pmatrix} \frac{1}{\sqrt{2m+M}} & \frac{1}{\sqrt{2m}} & \frac{1}{\sqrt{2m(1+\frac{2m}{M})}} \\ \frac{1}{\sqrt{2m+M}} & 0 & -\frac{2}{\sqrt{2M(2+\frac{M}{m})}} \\ \frac{1}{\sqrt{2m+M}} & -\frac{1}{\sqrt{2m}} & \frac{1}{\sqrt{2m(1+\frac{2m}{M})}} \end{pmatrix} \equiv \begin{pmatrix} a & b & c \\ a & 0 & -\frac{2m}{M}c \\ a & -b & c \end{pmatrix} \quad (15.71)$$

<sup>14</sup>Nótese que (15.67) no es una normalización usual. Esto se debe a que el tensor cinético (15.62) que es el análogo al tensor métrico, no es trivial, ya que si bien es diagonal, no es proporcional a la identidad.



la inversa y el determinante están dados por

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{m}{a(M+2m)} & \frac{M}{a(M+2m)} & \frac{m}{a(M+2m)} \\ \frac{1}{2b} & 0 & -\frac{1}{2b} \\ \frac{M}{2c(M+2m)} & -\frac{M}{c(M+2m)} & \frac{M}{2c(M+2m)} \end{pmatrix} ; \quad \det \mathbf{A} = -2abc \left(1 + \frac{2m}{M}\right) \quad (15.72)$$

el lector puede chequear que esta matriz es invertible siempre que las masas sean no nulas. Adicionalmente, se observa que el determinante no es la unidad, y  $\mathbf{A}^{-1} \neq \tilde{\mathbf{A}}$ , de modo que esta matriz no es ortogonal. Como ya se mencionó, esto se debe al carácter no trivial del tensor cinético (15.62). Esto implica entonces que la diagonalización simultánea de  $\mathbf{T}$  y  $\mathbf{V}$  (en la cual  $\mathbf{T}$  se convierte en la identidad) requiere de una transformación de contracción o dilatación (ver sección 12.15). Nótese en particular que si  $m = M$  se obtiene  $\mathbf{A}^{-1} = m\tilde{\mathbf{A}}$ , con lo cual  $\mathbf{A}$  es una matriz “cuasi-ortogonal”, esto se debe a que en este límite el tensor cinético (15.62) se vuelve proporcional a la identidad, con lo cual todos los ejes requieren el mismo factor de contracción o dilatación para trivializar a  $\mathbf{T}$ . En contraste, cuando  $m \neq M$ , se requieren dos factores diferentes de contracción para tal trivialización.

Por simplicidad utilizaremos las condiciones iniciales  $\eta_i(0) = A_i$ ,  $\dot{\eta}_i(0) = 0$  para  $i = 1, 2, 3$ ; que corresponden a colocar las partículas en  $t = 0$  fuera del equilibrio con velocidades iniciales nulas. La Ec. (15.45) nos muestra que para estas condiciones iniciales la parte imaginaria de los  $C_l$  es nula<sup>15</sup>. Por tanto los  $C_l$  son reales y vienen dados por la Ec. (15.44)

$$\begin{aligned} \text{Re} \mathbf{C} &= \mathbf{C} = \mathbf{A}^{-1} \boldsymbol{\eta}(0) = \begin{pmatrix} \frac{m}{a(M+2m)} & \frac{M}{a(M+2m)} & \frac{m}{a(M+2m)} \\ \frac{1}{2b} & 0 & -\frac{1}{2b} \\ \frac{M}{2c(M+2m)} & -\frac{M}{c(M+2m)} & \frac{M}{2c(M+2m)} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \\ \mathbf{C} &= \begin{pmatrix} \frac{MA_2 + mA_1 + mA_3}{a(M+2m)} \\ \frac{A_1 - A_3}{2b} \\ \frac{M(A_1 - 2A_2 + A_3)}{2c(M+2m)} \end{pmatrix} \end{aligned} \quad (15.73)$$

Aplicando (15.63, 15.73), obtenemos las ecuaciones asociadas a los modos normales de oscilación

$$\begin{aligned} \zeta_1 &= C_1 e^{-i\omega_1 t} = \frac{MA_2 + mA_1 + mA_3}{a(M+2m)} ; \quad \omega_1 = 0 \\ \zeta_2 &= C_2 e^{-i\omega_2 t} = \frac{A_1 - A_3}{2b} e^{-i\omega_2 t} ; \quad \omega_2 \equiv \sqrt{\frac{k}{m}} \\ \zeta_3 &= C_3 e^{-i\omega_3 t} = \frac{M(A_1 - 2A_2 + A_3)}{2c(M+2m)} e^{-i\omega_3 t} ; \quad \omega_3 \equiv \sqrt{\frac{k}{m} \left(1 + \frac{2m}{M}\right)} \end{aligned} \quad (15.74)$$

donde las constantes  $a, b, c$  están definidas en (15.71). Las amplitudes físicas  $\eta_i(t)$  se obtienen aplicando (15.71, 15.63) en las Ecs. (15.39) y teniendo en cuenta que  $\omega_1 = 0$

$$\begin{aligned} \eta_1(t) &= C_1 a_{11} e^{-i\omega_1 t} + C_2 a_{12} e^{-i\omega_2 t} + C_3 a_{13} e^{-i\omega_3 t} = a C_1 + b C_2 e^{-i\omega_2 t} + c C_3 e^{-i\omega_3 t} \\ \eta_2(t) &= C_1 a_{21} e^{-i\omega_1 t} + C_2 a_{22} e^{-i\omega_2 t} + C_3 a_{23} e^{-i\omega_3 t} = a C_1 - \frac{2m}{M} c C_3 e^{-i\omega_3 t} \\ \eta_3(t) &= C_1 a_{31} e^{-i\omega_1 t} + C_2 a_{32} e^{-i\omega_2 t} + C_3 a_{33} e^{-i\omega_3 t} = a C_1 - b C_2 e^{-i\omega_2 t} + c C_3 e^{-i\omega_3 t} \end{aligned} \quad (15.75)$$

<sup>15</sup>Para  $\omega_l = 0$ , no podemos utilizar (15.45). La Ec. (15.43) nos muestra que la parte imaginaria de  $C_k$  asociada a un  $\omega_k = 0$ , puede elegirse arbitrariamente para reproducir adecuadamente las condiciones iniciales. Por tanto, elegiremos  $\text{Im} C_k = 0$  también en este caso.

y sustituyendo (15.73) en (15.75), se obtiene

$$\begin{aligned}\eta_1(t) &= \frac{MA_2 + mA_1 + mA_3}{(M + 2m)} + \frac{(A_1 - A_3)}{2}e^{-i\omega_2 t} + \frac{M(A_1 - 2A_2 + A_3)}{2(M + 2m)}e^{-i\omega_3 t} \\ \eta_2(t) &= \frac{MA_2 + mA_1 + mA_3}{(M + 2m)} - \frac{m(A_1 - 2A_2 + A_3)}{(M + 2m)}e^{-i\omega_3 t} \\ \eta_3(t) &= \frac{MA_2 + mA_1 + mA_3}{(M + 2m)} - \frac{(A_1 - A_3)}{2}e^{-i\omega_2 t} + \frac{M(A_1 - 2A_2 + A_3)}{2(M + 2m)}e^{-i\omega_3 t}\end{aligned}\quad (15.76)$$

El primer término a la derecha de estas ecuaciones es constante e igual para todas las coordenadas reales. Por tanto, representa una traslación rígida de la molécula<sup>16</sup>. Como ya se mencionó, este término se puede anular exigiendo que la coordenada del centro de masa permanezca en el origen para todo tiempo (es decir montándose en el sistema de referencia del CM). Esto nos lleva a exigir

$$X_{CM} = \frac{m(x_1 + x_3) + Mx_2}{2m + M} = 0$$

es claro que solo el numerador es relevante en esta ligadura

$$m(x_1 + x_3) + Mx_2 = 0 = m(x_1 - x_{01} + x_3 - x_{03}) + M(x_2 - x_{02}) + m(x_{01} + x_{03}) + Mx_{02}$$

en particular  $X_{CM}(t = 0) = 0$ , de modo que  $m(x_{01} + x_{03}) + Mx_{02} = 0$ , y se obtiene

$$m(x_1 + x_3) + Mx_2 = m(\eta_1 + \eta_3) + M\eta_2 = 0$$

aplicando esta ecuación en  $t = 0$  se obtiene

$$m(A_1 + A_3) + MA_2 = 0 \Leftrightarrow A_2 = -\frac{m}{M}(A_1 + A_3) \quad (15.77)$$

con esta ligadura, se anula el primer término a la derecha en las Ecs. (15.76) y nos quedamos solo con los modos vibracionales. Sin embargo, debe tenerse en cuenta que las  $A_i$  ya no son todas independientes. Utilizando la ligadura (15.77) en cada una de sus formas se obtiene

$$\begin{aligned}\eta_1(t) &= \frac{(A_1 - A_3)}{2}e^{-i\omega_2 t} + \frac{1}{2}(A_1 + A_3)e^{-i\omega_3 t} ; \quad \eta_2(t) = -\frac{m(A_1 + A_3)}{M}e^{-i\omega_3 t} \\ \eta_3(t) &= -\frac{(A_1 - A_3)}{2}e^{-i\omega_2 t} + \frac{1}{2}(A_1 + A_3)e^{-i\omega_3 t}\end{aligned}\quad (15.78)$$

aplicando la ligadura (15.77) a los modos normales (15.74) se obtiene

$$\begin{aligned}\zeta_1 &= 0 ; \quad \zeta_2 = \frac{A_1 - A_3}{2b}e^{-i\omega_2 t} ; \quad \zeta_3 = \frac{1}{2c}(A_1 + A_3)e^{-i\omega_3 t} \\ \zeta_1 &= 0 ; \quad \zeta_2 = \frac{\sqrt{2m}(A_1 - A_3)}{2}e^{-i\omega_2 t} ; \quad \zeta_3 = \frac{\sqrt{2m(1 + \frac{2m}{M})}}{2}(A_1 + A_3)e^{-i\omega_3 t}\end{aligned}\quad (15.79)$$

donde hemos usado las definiciones de las constantes  $a, b, c$  Ec. (15.71). Recordemos que los modos normales son soluciones que se pueden interpretar como un conjunto de osciladores (imaginarios) desacoplados. Pero las soluciones (15.78) sí son soluciones reales cuando se interpretan como elongaciones de las partículas.

<sup>16</sup>La igualdad de estos términos se debe a la igualdad de las componentes del vector  $(a, a, a)$  asociado a  $\omega_1 = 0$ , como se vé en la primera columna de la matriz (15.71). Otra manera de ver que este vector está asociado a un modo traslacional, es observando que la energía potencial no cambia al pasar desde el vector cero (equilibrio) al vector (15.68), puesto que en ambos casos la energía potencial es cero. La única manera de que la energía potencial permanezca intacta es que todas las partículas conserven sus distancias relativas, y dado que solo se permiten modos longitudinales, esto solo puede asociarse a una traslación rígida del sistema.

Las soluciones para las elongaciones reales  $\eta_i$  son superposiciones de los modos normales, mostrando que las partículas reales están acopladas. Es claro que si hacemos  $A_1 = -A_3$  en las Ecs. (15.78, 15.79), se eliminan las contribuciones asociadas a  $\omega_3$  tanto para los modos normales como para los reales

$$\eta_1(t) = A_1 e^{-i\omega_2 t} ; \eta_2(t) = 0 ; \eta_3(t) = -A_1 e^{-i\omega_2 t} = A_1 e^{-i(\omega_2 t + \pi)} \quad (15.80)$$

$$\zeta_1 = 0 ; \zeta_2 = \sqrt{2m} A_1 e^{-i\omega_2 t} ; \zeta_3 = 0 \quad (15.81)$$

Recordando que hemos considerado  $\dot{\eta}_i = 0$ , Físicamente estas condiciones iniciales significan que soltamos las partículas desde el reposo de modo que las partículas de los extremos se estiran ambas o se contraen ambas en la misma magnitud (movimiento en antifase i.e.  $A_1 = e^{i\pi} A_3$ ), dejando sin elongar a la partícula de la mitad<sup>17</sup>.

Examinemos las soluciones para todo tiempo con estas condiciones iniciales. En este caso, la Ec. (15.81) muestra que solo sobrevive el modo normal asociado a  $\omega_2$ , de modo que estamos en el segundo “modo normal puro”. En cuanto al movimiento Físico, las Ecs. (15.80) muestran que éste corresponde a que la partícula  $M$  está en reposo, en tanto que las partículas de los extremos oscilan armónicamente siempre en antifase. Estas soluciones son las esperadas por argumentos de simetría<sup>18</sup>. Además, ya que la masa  $M$  actúa como extremo fijo, es lógico que el movimiento armónico de cada masa  $m$ , esté dado por el valor que se obtendría si cada una estuviera sola, atada a un resorte de constante  $k$  y fija a una pared. Efectivamente la frecuencia real de cada masa  $m$  es  $\omega_2 = \sqrt{k/m}$ .

Es fácil ver de las Ecs. (15.79) que la obtención del “tercer modo normal puro” se obtiene haciendo  $A_1 = A_3$ . Sustituyendo estas condiciones en las Ecs. (15.78, 15.79), las coordenadas reales y normales quedan

$$\eta_1(t) = A_1 e^{-i\omega_3 t} ; \eta_2(t) = -\frac{2A_1 m}{M} e^{-i\omega_3 t} = \frac{2A_1 m}{M} e^{-i(\omega_3 t + \pi)} ; \eta_3(t) = A_1 e^{-i\omega_3 t}$$

$$\zeta_1 = 0 ; \zeta_2 = 0 ; \zeta_3 = \sqrt{2m \left(1 + \frac{2m}{M}\right)} A_1 e^{-i\omega_3 t}$$

de nuevo examinemos físicamente las condiciones iniciales. La condición  $A_1 = A_3$  indica que las dos partículas de los extremos se elongan en la misma dirección (en fase) y con la misma amplitud. A fin de mantener el CM fijo, la ligadura (15.77) nos da  $A_2 = -2mA_1/M$ , de modo que la partícula  $M$  se elonga en dirección contraria a las otras (antifase con las masas de los extremos), con una amplitud que depende de la amplitud de las masas de los extremos y del cociente entre las masas (por ejemplo si  $m = M$ , la amplitud de  $M$  debe ser el doble de la de las partículas de los extremos).

Las soluciones físicas  $\eta_i(t)$  muestran que en el tercer modo normal puro, las masas de los extremos oscilan siempre en fase y la del centro siempre en antifase con respecto a las anteriores. Todas las partículas oscilan con la misma frecuencia, que es la asociada al tercer modo normal puro.

En general, cualquier vibración longitudinal de la molécula que no involucre traslación rígida será una combinación lineal de los modos asociados a  $\omega_2$  y  $\omega_3$ . Las amplitudes de los modos normales y reales, así como las fases relativas en los modos reales, estarán determinadas por las condiciones iniciales.

#### 15.5.4. Análisis cualitativo de vibraciones transversales y longitudinales

Hasta el momento nos hemos limitado al estudio de las vibraciones longitudinales. El caso más general en el cual se consideran vibraciones transversales es más complicado matemáticamente ya que el movimiento general involucra 9 grados de libertad, por tanto solo haremos un análisis cualitativo de dicho movimiento. En primer lugar, el número de modos con frecuencia cero nos va a dar los grados de libertad asociados al movimiento como cuerpo rígido, en consecuencia habrá 6 grados de libertad asociados a tales frecuencias sin importar el número de átomos de la molécula, es decir  $3n - 6$  modos vibracionales para una molécula de  $n$  átomos.

<sup>17</sup>Esto se vé al utilizar la condición inicial  $A_1 = -A_3$  en la ligadura (15.77), de lo cual se obtiene  $A_2 = 0$ .

<sup>18</sup>De hecho, al estar el átomo central en reposo, los átomos de los extremos deben vibrar en antifase (o contrafase), para que el momento lineal total sea nulo todo el tiempo (en virtud de que estamos en el sistema de referencia CM).

En el caso de una molécula lineal habrá tres grados de libertad de traslación pero solo dos grados de libertad de rotación rígida, puesto que una rotación alrededor del eje de la molécula no es relevante y no aparecerá como un modo de rotación rígida. Esto implica que tendremos cuatro grados de libertad asociados a la vibración: 2 de las vibraciones longitudinales como ya vimos, y otros dos asociados a vibraciones perpendiculares al eje (transversales). No obstante, la simetría axial de la molécula nos dice que los dos modos vibracionales transversales deben ser degenerados. No hay nada que distinga a una vibración a lo largo de cualquier eje perpendicular a la línea de la molécula, por lo cual las frecuencias asociadas deben ser idénticas.

La indeterminación de los autovectores inherentes a los modos degenerados se refleja en el hecho de que todas las direcciones perpendiculares al eje de la molécula son equivalentes, cualquier par ortogonal de ejes que yace en el plano perpendicular a tal eje se pueden escoger como las direcciones de los modos normales de vibración. El movimiento transversal depende de las amplitudes y fases de éstos modos degenerados. Por ejemplo, si ambos modos están presentes y en fase los átomos se moverán en línea recta pasando a través de la configuración de equilibrio. Pero si están en antifase, el movimiento compuesto producirá una elipse que corresponde a una figura de Lissajous, tal como ocurre con el oscilador armónico bidimensional isotrópico. Esto significa que los dos modos representan una rotación más que una vibración.

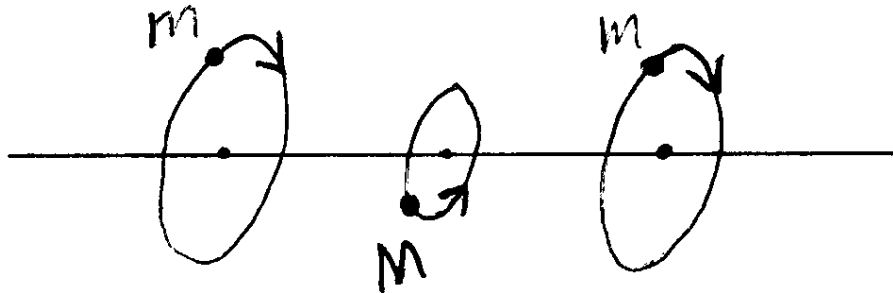


Figura 15.3: *Movimiento de los átomos de una molécula triatómica lineal, cuando los modos degenerados están en desfase de  $\pi/2$ . Obsérvese que la masa  $M$  gira en dirección contraria a las masas  $m$  de los extremos. Los puntos a lo largo de la recta son las posiciones de equilibrio de las masas.*

La simetría también nos lleva a que las amplitudes asociadas a modos normales de los átomos de los extremos, sean idénticas en magnitud. El cálculo detallado muestra además que los átomos de los extremos viajan en la misma dirección que las figuras de Lissajous. Por tanto el átomo del centro viaja en dirección opuesta con el fin de mantener el centro de masa en reposo. La figura 15.3 muestra el movimiento de los átomos en el caso en el cual los modos degenerados están en desfase  $\pi/2$ .

Cuando la complejidad de la molécula crece, la complejidad de la ecuación secular aumenta significativamente. No obstante, en muchos casos las simetrías de la molécula han jugado un papel simplificador fundamental a través de la teoría de grupos, gracias a la cual es a veces posible factorizar la ecuación secular en bloques más pequeños que se pueden diagonalizar por separado para encontrar las amplitudes y las frecuencias normales del sistema. Estos métodos encuentran importantes aplicaciones en espectroscopía infrarroja y en espectroscopía Raman.

## 15.6. Modos normales puros y soluciones físicas asociadas

De la discusión en la sección 15.5.3 concerniente a una molécula triatómica lineal, observamos algunas características de los modos normales puros y las soluciones asociadas para las elongaciones reales de las partículas. Para frecuencias nulas, podemos eliminar el correspondiente grado de libertad (normal y real) exigiendo que el centro de masa esté en reposo en el origen, al menos para los grados de libertad traslacionales. Puesto que los modos normales de frecuencia nula ya fueron discutidos en la sección 15.5.1, nos restringiremos

a analizar los modos vibracionales.

En la sección 15.5.3, vimos que era posible obtener un modo normal puro ajustando las condiciones iniciales del sistema real apropiadamente. Dicho modo normal está asociado a un movimiento real en el cual las partículas se mueven todas con la misma frecuencia (o no se mueven en lo absoluto como  $\eta_2(t)$  en la Ec. 15.80) y se mueven en fase o en antifase. Veremos que estas características son válidas en un marco más general.

Las Ecs. (15.44, 15.45) nos muestran como obtener los factores complejos  $C_l$  asociados con cada modo normal, a partir de las condiciones iniciales  $\eta_i(0)$  y  $\dot{\eta}_i(0)$ . De las Ecs. (15.56) es claro que el  $m$ -ésimo modo normal puro se obtiene cuando los factores  $C_k$  están dados por

$$C_k = \delta_{km} C_m \quad k = 1, \dots, n \quad ; \quad m \text{ es fijo} \quad (15.82)$$

$$\zeta_k = \delta_{km} C_k e^{-i\omega_k t} \quad ; \quad m \text{ es fijo} \quad (15.83)$$

por tanto, deben ajustarse las condiciones iniciales  $\eta_i(0)$  y  $\dot{\eta}_i(0)$  de tal modo que las Ecs. (15.44, 15.45) me reproduzcan (15.82). Una vez realizado dicho ajuste, vemos de las Ecs. (15.39) que las elongaciones reales de las partículas están dadas por

$$\begin{aligned} \eta_i(t) &= C_k a_{ik} e^{-i\omega_k t} = \delta_{km} C_m a_{ik} e^{-i\omega_k t} \\ \eta_i(t) &= C_m a_{im} e^{-i\omega_m t} \quad ; \quad i = 1, \dots, n \quad ; \quad \text{no suma sobre } m \end{aligned} \quad (15.84)$$

de (15.84) se vé que la solución Física para cada partícula  $i$ , consta de un solo término también armónico y de la misma frecuencia que el modo normal, a menos que  $a_{im}$  sea nulo en cuyo caso la partícula estará en reposo. Este fué el caso de  $\eta_2(t)$  para el segundo modo normal en la Ec. (15.80), ya que  $a_{22} = 0$  como se puede ver en (15.71). Es decir, en el movimiento real asociado al  $m$ -ésimo modo normal puro, cada partícula tiene movimiento armónico simple con la frecuencia  $\omega_m$  del correspondiente modo, o está en reposo. Por otro lado, teniendo en cuenta que los elementos  $a_{ik}$  son reales, al comparar la solución real (15.84) con el modo normal (15.83) vemos que si  $a_{im} > 0$ , entonces  $\eta_i(t)$  tiene la misma fase que  $\zeta_m(t)$

$$\begin{aligned} \zeta_m(t) &= C_m e^{-i\omega_m t} = |C_m| e^{i\delta_m} e^{-i\omega_m t} = |C_m| e^{-i(\omega_m t + \delta_m)} \quad ; \quad \text{no suma sobre } m \\ \eta_i(t) &= C_m a_{im} e^{-i\omega_m t} = |C_m| e^{i\delta_m} |a_{im}| e^{-i\omega_m t} = |C_m a_{im}| e^{-i(\omega_m t + \delta_m)} \quad ; \quad \text{no suma sobre } m \end{aligned}$$

Pero si  $a_{im} < 0$ , entonces  $\eta_i(t)$  está en antifase con  $\zeta_m(t)$

$$\eta_i(t) = C_m a_{im} e^{-i\omega_m t} = |C_m| e^{i\delta_m} |a_{im}| e^{i\pi} e^{-i\omega_m t} = |C_m a_{im}| e^{-i(\omega_m t + \delta_m + \pi)}$$

Por tanto, en el  $m$ -ésimo modo normal puro, la  $i$ -ésima partícula vibrante está en fase (desfase) con tal modo normal si  $a_{im}$  es positivo (negativo); si  $a_{im} = 0$ , la  $i$ -ésima partícula estará en reposo.

Insistimos en este punto en que los modos normales no son estrictamente “imaginarios” pues de hecho son una solución totalmente equivalente a la que se encuentra con las coordenadas de elongación  $\eta_i$  de las partículas, solo que en otras coordenadas en donde la apariencia algebraica de las ecuaciones es más simple. De hecho la Ec. (15.52) nos muestra que la matriz  $\mathbf{A}$  de transformación (matriz de vectores propios de la Ec. 15.17) nos lleva de un sistema coordinado al otro. Lo que es imaginario, es la interpretación de las coordenadas normales  $\zeta_k$  como elongaciones de partículas desacopladas, ya que estas partículas desacopladas sí son imaginarias.

## 15.7. Vibraciones forzadas y amortiguadas

### 15.7.1. Vibraciones forzadas

Ocurre con frecuencia que el sistema no se deja oscilar libremente sino que es sometido a alguna fuerza externa en forma continua a partir de un cierto tiempo  $t = 0$ . Cuando dicha fuerza es también de tipo oscilatorio, el sistema termina vibrando con la frecuencia de la fuerza externa en lugar de su frecuencia natural. No obstante, las coordenadas normales siguen siendo de gran importancia para la determinación de las amplitudes de la vibración forzada y continúan siendo las coordenadas que desacoplan a las ecuaciones diferenciales como veremos a continuación.

Por simplicidad, asumiremos que las fuerzas externas si las hay, también son longitudinales al igual que las posiciones y velocidades iniciales. Esto garantiza que los modos vibracionales de la molécula continúan siendo longitudinales. Si denotamos por  $F_j$  a la fuerza generalizada correspondiente a la coordenada  $\eta_j$ , tenemos que en virtud de la Ec. (2.9) la fuerza generalizada  $Q_i$  asociada a la coordenada normal  $\zeta_i$  se escribe como<sup>19</sup>

$$\begin{aligned} Q_i &= F_j \frac{\partial x_j}{\partial \zeta_i} = F_j \frac{\partial (x_j - x_{0j})}{\partial \zeta_i} = F_j \frac{\partial \eta_j}{\partial \zeta_i} = F_j \frac{\partial (a_{jk} \zeta_k)}{\partial \zeta_i} = F_j a_{jk} \frac{\partial \zeta_k}{\partial \zeta_i} = F_j a_{jk} \delta_{ki} \\ Q_i &= a_{ji} F_j = \tilde{a}_{ij} F_j \end{aligned} \quad (15.85)$$

donde hemos utilizado la Ec. (15.48) en componentes. En notación matricial (15.85) se escribe como

$$\mathbf{Q} = \tilde{\mathbf{A}} \mathbf{F} \quad (15.86)$$

las ecuaciones para las coordenadas  $\eta_i$  son idénticas a las Ecs. (15.11) pero agregando las fuerzas generalizadas asociadas  $F_j$  en el miembro derecho como se vé en la Ec. (2.35). Matricialmente escribimos estas ecuaciones en la forma

$$\mathbf{T} \ddot{\boldsymbol{\eta}} + \mathbf{V} \boldsymbol{\eta} = \mathbf{F} \Rightarrow \mathbf{T} \mathbf{A} \ddot{\boldsymbol{\zeta}} + \mathbf{V} \mathbf{A} \boldsymbol{\zeta} = \mathbf{F}$$

donde hemos usado (15.48). Aplicando  $\tilde{\mathbf{A}}$  a la izquierda de la ecuación y utilizando las Ecs. (15.18, 15.19, 15.86) se tiene

$$\left( \tilde{\mathbf{A}} \mathbf{T} \mathbf{A} \right) \ddot{\boldsymbol{\zeta}} + \left( \tilde{\mathbf{A}} \mathbf{V} \mathbf{A} \right) \boldsymbol{\zeta} = \tilde{\mathbf{A}} \mathbf{F} \Rightarrow \ddot{\boldsymbol{\zeta}} + \lambda \boldsymbol{\zeta} = \mathbf{Q}$$

que son las ecuaciones para las coordenadas normales. En componentes, estas ecuaciones quedan

$$\ddot{\zeta}_i + \omega_i^2 \zeta_i = Q_i \quad (15.87)$$

en este caso las  $n$  ecuaciones (15.87) son inhomogéneas y para solucionarlas se debe conocer  $Q_i$  en función de tiempo. Vale recalcar que estas ecuaciones para las coordenadas normales aunque inhomogéneas, siguen estando desacopladas como ya se había anticipado.

A menudo la fuerza externa es de tipo senoidal como ocurre por ejemplo si el forzamiento lo da una onda incidente sobre el sistema (mecánica o electromagnética), o si el forzamiento lo produce un sistema mecánico externo que a su vez también está vibrando. En otros casos aunque el forzamiento no sea senoidal, se puede escribir como una superposición de componentes senoidales, por ejemplo si el forzamiento es periódico se puede escribir como una serie de Fourier y si no es periódico una integral de Fourier podría representar el forzamiento. Teniendo en cuenta que las ecuaciones (15.87) son lineales las soluciones para frecuencias particulares se pueden superponer para generar la solución general para  $Q_i$  dado.

Escribiremos entonces un forzamiento de la forma

$$Q_i = Q_{0i} \cos(\omega t + \delta_i) \quad (15.88)$$

siendo  $\omega$  la frecuencia externa de forzamiento<sup>20</sup>, las ecuaciones de movimiento quedan

$$\ddot{\zeta}_i + \omega_i^2 \zeta_i = Q_{0i} \cos(\omega t + \delta_i) \quad (15.89)$$

de acuerdo con la teoría de ecuaciones diferenciales, la solución de (15.89) consta de una solución general de la ecuación homogénea correspondiente mas una solución particular de la inhomogénea. Por medio de una elección apropiada de condiciones iniciales, es posible hacer que las vibraciones libres superpuestas se cancelen<sup>21</sup>. Por lo tanto vamos a centrar nuestra atención en las soluciones particulares de (15.89), teniendo presente que es de

<sup>19</sup>Puesto que  $\eta_j$  es una coordenada esencialmente cartesiana, las fuerzas generalizadas son las fuerzas newtonianas.

<sup>20</sup>Nótese que  $\omega$  no depende de  $i$ . La Ec. (15.85), nos indica que  $\omega$  tampoco depende de  $j$ . Es decir, estamos asumiendo que el forzamiento tiene la misma frecuencia para cada partícula, aunque la amplitud del forzamiento sí puede depender de la partícula.

<sup>21</sup>Las vibraciones libres son básicamente los transientes generados por la aplicación de la fuerza externa. Si consideramos que el sistema está inicialmente en reposo en la configuración de equilibrio y que la fuerza crece desde cero en forma continua, estos transientes no aparecen. Alternativamente, las fuerzas disipativas cuando están presentes amortiguan las oscilaciones libres.

esperarse que el sistema termine oscilando con la frecuencia del forzamiento externo. Por tanto, emplearemos un ansatz de la forma

$$\zeta_i = B_i \cos(\omega t + \delta_i) \quad (15.90)$$

las amplitudes  $B_i$  se obtienen sustituyendo este ansatz en (15.89)

$$B_i = \frac{Q_{0i}}{\omega_i^2 - \omega^2} \quad (15.91)$$

utilizando de nuevo (15.48) obtenemos la solución para las coordenadas  $\eta_j$

$$\eta_j = a_{ji}\zeta_i = \frac{a_{ji}Q_{0i} \cos(\omega t + \delta_i)}{\omega_i^2 - \omega^2} \quad (15.92)$$

de nuevo, la vibración de cada partícula está compuesta por una combinación lineal de los modos normales, pero en este caso cada oscilación normal posee la frecuencia de la fuerza externa.

A partir de (15.92) se puede observar que la excitación de cada modo está determinada por la amplitud del forzamiento  $Q_{0i}$  y por la diferencia  $\omega_i^2 - \omega^2$ . Si la fuerza sobre cada partícula no tiene componente en la dirección de vibración de algún modo normal particular es obvio que la fuerza generalizada asociada a tal modo será nula i.e.  $Q_{0i} = 0$ , como se puede ver de (15.85) si  $F_j = 0$  para todo  $j$ . En consecuencia, una fuerza externa solo puede excitar un modo normal si tiende a mover a las partículas en la misma dirección que el modo dado. Por otro lado, se puede ver que a medida que disminuye la diferencia  $\omega_i^2 - \omega^2$  aumenta el valor de la amplitud del modo asociado, entre más cerca esté la frecuencia natural de un modo dado a la frecuencia externa, mayor será la intensidad de este modo relativo a los otros. Se puede ver incluso que la ecuación (15.92) predice una divergencia cuando  $\omega_i = \omega$  (resonancia). Sin embargo, hay que tener en cuenta que el formalismo que nos llevó a (15.92) presupone pequeñas oscilaciones alrededor de la configuración de equilibrio, de modo que cuando la fórmula predice amplitudes grandes ya no es válida.

De la Ec. (15.92), podemos examinar la contribución de un modo normal dado ( $i$  fijo) para una partícula específica ( $j$  fijo). Si  $a_{ji}Q_{0i} > 0$  para un par de valores fijos de  $i, j$  tendremos que cuando  $\omega < \omega_i$  ( $\omega > \omega_i$ ) las oscilaciones están en fase (antifase) con la fuerza externa. Similarmente, si  $a_{ji}Q_{0i} < 0$  las oscilaciones están en antifase (fase) con la fuerza externa. cuando  $\omega < \omega_i$  ( $\omega > \omega_i$ ). Es decir, si hacemos una variación continua de la frecuencia externa, encontraremos un cambio brusco en la fase relativa al forzamiento de un modo dado sobre una partícula, cuando pasamos por la resonancia asociada a dicho modo.

### 15.7.2. Vibraciones amortiguadas

En problemas reales de la física es muy usual que los sistemas que vibran estén sujetos a fuerzas disipativas debidas a la interacción con sistemas externos. En muchas situaciones, las fuerzas disipativas son proporcionales a las velocidades de las partículas de tal manera que se pueden generar a través de una función de disipación de Rayleigh como la discutida en la sección 2.5. De su definición Ec. (2.37),  $F$  debe ser una función homogénea de segundo grado en la velocidades

$$F = \frac{1}{2}F_{ij}\dot{\eta}_i\dot{\eta}_j \quad (15.93)$$

los coeficientes  $F_{ij}$  se pueden definir para que sean simétricos  $F_{ij} = F_{ji}$ , y serán en general funciones de las coordenadas. Dado que estamos interesados solo en pequeñas oscilaciones alrededor del equilibrio, expandiremos los coeficientes alrededor del equilibrio reteniendo solo el primer término de la expansión de forma análoga al caso de la energía cinética  $T$ . De aquí en adelante, la notación  $F_{ij}$  se refiere a esos factores constantes

$$F_{ij} \equiv F_{ij}(q_{01}, \dots, q_{0n}) \quad (15.94)$$

esto implica que los coeficientes de viscosidad serán trabajados en aproximación de medio homogéneo. Sin embargo, si los coeficientes  $F_{ij}$  son diferentes para diferentes índices, tendremos que la viscosidad del medio depende de la dirección del desplazamiento, es decir que el medio es anisotrópico.

Recordemos que de acuerdo con la Ec. (5.19),  $2F$  representa la rata de disipación de energía debida a las fuerzas de fricción. Por lo tanto  $F$  no puede ser nunca negativa. Por otra parte, incluso para un medio anisotrópico, es de esperarse que la viscosidad sea diferente de cero en todas las direcciones aunque sus valores específicos sean distintos en cada una. En otras palabras, se espera que haya una disipación diferente de cero para cualquier configuración no nula de velocidades  $\dot{\eta}_i$ . En consecuencia, se espera que la forma bilineal (15.93) sea positiva y solo sea cero para  $\dot{\eta} = 0$ . Por tanto, en la mayor parte de los casos de interés la matriz  $F_{ij}$  será definida positiva. Si se desprecia la viscosidad en ciertas direcciones, es posible modelar a  $F_{ij}$  como una matriz positiva singular, pero esto no ocurre con frecuencia<sup>22</sup>.

Las ecuaciones de Lagrange en este caso se pueden extraer de (2.38) y (15.10)

$$T_{ij}\ddot{\eta}_j + F_{ij}\dot{\eta}_j + V_{ij}\eta_j = 0 \quad (15.95)$$

### Soluciones en coordenadas normales

Si queremos encontrar coordenadas normales en donde las ecuaciones (15.95) desacoplen, es necesario encontrar ejes principales de transformación en los cuales las tres formas cuadráticas  $\mathbf{T}$ ,  $\mathbf{V}$  y  $F$  se diagonalicen simultáneamente. Sin embargo, esto no es posible en el caso más general y solo se pueden encontrar coordenadas normales para ciertos tipos de funciones de disipación. En la sección 12.15.1, Pág. 12.15.1 vimos que si  $\mathbf{T}$  y  $\mathbf{F}$  son definidos positivos y  $\mathbf{V}$  es positivo, los tres se pueden diagonalizar simultáneamente si  $\mathbf{F} = C\mathbf{T}$ , siendo  $C$  una constante. En nuestro contexto, esto se da cuando la fuerza de fricción es proporcional tanto a la velocidad de las partículas como a su masa, de tal manera que

$$F_{ij} = \gamma T_{ij} \quad (15.96)$$

Esto ocurre en el caso en el cual el medio disipador es homogéneo e isotrópico. Nótese sin embargo, que  $F_{ij}$  es anisotrópico ya que es un tensor de segundo rango. Debe observarse sin embargo, que tal anisotropía se debe al sistema de osciladores que está inmerso y no al fluido en el que se encuentra inmerso. Por ejemplo, para un oscilador dado, las componentes de las velocidades en cada dirección son distintas, lo cual es fuente de anisotropía para  $F$ . La isotropía del fluido se manifiesta en el hecho de que  $\gamma$  es un escalar y la homogeneidad en el hecho de que  $\gamma$  es independiente de la posición.

Con la condición (15.96) y empleando las coordenadas normales (15.48), la Ec. (15.95) queda

$$\begin{aligned} \mathbf{T}\ddot{\eta} + \gamma\mathbf{T}\dot{\eta} + \mathbf{V}\eta &= 0 \Rightarrow \mathbf{TA}\ddot{\zeta} + \gamma\mathbf{TA}\dot{\zeta} + \mathbf{VA}\zeta = 0 \Rightarrow \tilde{\mathbf{A}}\mathbf{TA}\ddot{\zeta} + \gamma\tilde{\mathbf{A}}\mathbf{TA}\dot{\zeta} + \tilde{\mathbf{A}}\mathbf{VA}\zeta = 0 \\ &\Rightarrow \ddot{\zeta} + \gamma\dot{\zeta} + \lambda\zeta = 0 \end{aligned}$$

Donde hemos usado las Ecs. (15.18, 15.19). En componentes, esta ecuación se escribe como

$$\ddot{\zeta}_i + \gamma\dot{\zeta}_i + \omega_i^2\zeta_i = 0 \quad (\text{no suma}) \quad (15.97)$$

Vemos entonces que si se cumple la condición (15.96), las coordenadas normales aún desacoplan el problema. Siendo un conjunto de ecuaciones diferenciales lineales con coeficientes constantes, podemos establecer un ansatz de solución para las Ecs. (15.97) en la forma

$$\zeta_i = C_i e^{-i\omega'_i t} \quad (15.98)$$

introduciendo (15.98) en (15.97) resulta una ecuación cuadrática para  $\omega'_i$

$$\omega_i'^2 + i\omega'_i\gamma - \omega_i^2 = 0 \quad (\text{no suma}) \quad (15.99)$$

cuyas soluciones son de la forma

$$\omega'_i = \pm \sqrt{\omega_i^2 - \frac{\gamma^2}{4}} - i\frac{\gamma}{2} \equiv \pm\Omega_i - i\frac{\gamma}{2} \quad (15.100)$$

<sup>22</sup>Si los osciladores están inmersos en un fluido, la viscosidad será aproximadamente isotrópica de modo que la matriz será definida positiva.



la solución para las coordenadas normales es entonces de la forma

$$\zeta_i = e^{-\frac{\gamma}{2}t} C_i e^{\mp i\Omega_i t} \quad (15.101)$$

vemos entonces que la dinámica de las coordenadas normales nos da un factor de amortiguamiento para la amplitud debido a la componente imaginaria de  $\omega'_i$ , que como era de esperarse es causada por la función de disipación i.e. por las fuerzas disipativas. Nótese que este término es siempre decreciente ya que  $\gamma$  no puede ser negativo (en virtud del carácter definido positivo de  $\mathbf{T}$  y  $F$ ). Desde el punto de vista de la energía se ve que cuando las partículas vibran, las fuerzas disipativas hacen un trabajo negativo sobre ellas con lo cual el sistema que vibra pierde energía traduciéndose en una disminución de la amplitud con el tiempo. Por otro lado, la Ec. (15.101) muestra que la parte real de  $\omega'_i$  produce un término oscilatorio, se vé además que la fricción también afecta las frecuencias de oscilación y que como es de esperarse, a medida que disminuye el efecto de la fricción, tales frecuencias se acercan a las frecuencias naturales del sistema. Por otro lado, si  $\gamma/2 > \omega_i$  la cantidad  $\Omega_i$  también se vuelve imaginaria y deja de ser oscilante para convertirse en otro término de decaimiento (sobreamortiguamiento). Cuando  $\gamma/2 = \omega_i$  la forma de la solución cambia ligeramente pero sigue siendo decayente sin oscilación (amortiguamiento crítico), el caso en el cual el amortiguamiento es pequeño y se presenta oscilación se conoce como subamortiguamiento.

Debemos recordar en todo caso, que la solución para las elongaciones de las partículas está dada por las coordenadas  $\eta_i$ , que se obtienen a partir de  $\zeta_i$  por medio de (15.48).

### Solución en el caso general

Si la función de disipación no puede ser diagonalizada junto con  $T$  y  $V$ , la solución es mucho más difícil de obtener debido a la ausencia de desacople en las ecuaciones de movimiento. No obstante, la naturaleza general de las soluciones es similar: un término que decae y un término que oscila (para amortiguamiento pequeño). Aunque no trabajaremos exhaustivamente la solución del problema, podemos extraer mucha información cualitativa e incluso semicuantitativa a partir de ciertas consideraciones generales

Comencemos con un ansatz de solución para (15.95) con estructura similar a la de la solución (15.101)

$$\eta_j = C a_j e^{-i\omega t} = C a_j e^{-\kappa t} e^{-2\pi i\nu t} \quad ; \quad \omega \equiv 2\pi\nu - i\kappa \quad (15.102)$$

introduciendo este ansatz en (15.95) resulta un conjunto de ecuaciones lineales simultáneas

$$V_{ij} a_j - i\omega F_{ij} a_j - \omega^2 T_{ij} a_j = 0 \quad (15.103)$$

definimos

$$\gamma \equiv -i\omega = -\kappa - 2\pi i\nu \quad (15.104)$$

El conjunto de ecuaciones (15.103) se puede escribir en términos de los tensores  $\mathbf{T}$ ,  $F$  y  $\mathbf{V}$  así como de la constante  $\gamma$  en la forma

$$\mathbf{V}\mathbf{a} + \gamma F\mathbf{a} + \gamma^2 \mathbf{T}\mathbf{a} = 0 \quad (15.105)$$

el conjunto de ecuaciones homogéneas (15.103) o (15.105) se puede resolver para las  $a_j$  solo para ciertos valores de  $\omega$  o de  $\gamma$ . Solo podemos encontrar soluciones no triviales para  $\mathbf{a}$ , si  $\gamma$  es solución de la ecuación secular

$$\det(\mathbf{V} + \gamma F + \gamma^2 \mathbf{T}) = 0$$

Sin embargo, es posible demostrar que  $\kappa$  tiene que ser no negativo sin necesidad de resolver la ecuación secular completa. En primer lugar, convertimos la Ec. (15.105) en una ecuación escalar multiplicando a ambos lados por  $\mathbf{a}^\dagger$  a la izquierda de tal ecuación

$$\mathbf{a}^\dagger \mathbf{V}\mathbf{a} + \gamma \mathbf{a}^\dagger F\mathbf{a} + \gamma^2 \mathbf{a}^\dagger \mathbf{T}\mathbf{a} = 0 \quad (15.106)$$

que es una ecuación cuadrática para  $\gamma$  cuyos coeficientes son productos matriciales con una estructura semejante a los que aparecen en la Ec. (12.115). Descompondremos al vector  $\mathbf{a}$  de la misma forma que en (12.114)

$$\mathbf{a} = \boldsymbol{\alpha} + i\boldsymbol{\beta} \quad ; \quad \mathbf{a} = \boldsymbol{\alpha} + i\boldsymbol{\beta} \quad ; \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^n \quad (15.107)$$

y retomaremos el resultado (12.116) válido para cualquier matriz simétrica real

$$\mathbf{a}^\dagger \mathbf{M} \mathbf{a} = \tilde{\alpha} \mathbf{M} \alpha + \tilde{\beta} \mathbf{M} \beta \quad ; \quad \mathbf{a} = \alpha + i\beta \quad ; \quad \alpha, \beta \in \mathbb{R}^n \quad (15.108)$$

Por tanto, en virtud de la simetría de  $\mathbf{T}$ ,  $\mathbf{V}$ , y  $F$  todos los productos matriciales en (15.106) son reales

$$\begin{aligned} a\gamma^2 + b\gamma + c &= 0 \quad ; \quad \gamma_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ a &\equiv \mathbf{a}^\dagger \mathbf{T} \mathbf{a} \quad ; \quad b \equiv \mathbf{a}^\dagger F \mathbf{a} \quad ; \quad c \equiv \mathbf{a}^\dagger \mathbf{V} \mathbf{a} \end{aligned}$$

debido al carácter real de  $a, b, c$ ; tenemos que si  $\gamma$  es una solución de la Ec. (15.106), también lo será  $\gamma^*$ . Es claro que la suma de las dos raíces nos da

$$\gamma_1 + \gamma_2 = -\frac{b}{a} = -\frac{\mathbf{a}^\dagger F \mathbf{a}}{\mathbf{a}^\dagger \mathbf{T} \mathbf{a}}$$

como en este caso nuestras raíces son  $\gamma_1$  y  $\gamma_1^*$  tenemos que

$$\gamma_1 + \gamma_1^* = 2\text{Re}\gamma_1 = -\frac{\mathbf{a}^\dagger F \mathbf{a}}{\mathbf{a}^\dagger \mathbf{T} \mathbf{a}} = -2\kappa_1$$

donde también se tuvo en cuenta la definición de  $\gamma$  Ec. (15.104). Usando (15.108) podemos expresar  $\kappa$  de la forma

$$\kappa_1 = \frac{1}{2} \frac{(\alpha_i F_{ij} \alpha_j + \beta_i F_{ij} \beta_j)}{(\alpha_k T_{kl} \alpha_l + \beta_k T_{kl} \beta_l)}$$

ya vimos que la matriz  $F_{ij}$  debe ser positiva, en tanto que la matriz  $T_{kl}$  es definida positiva, por lo tanto  $\kappa_1$  no puede ser negativo. Es decir, la oscilación del sistema puede decrecer exponencialmente pero no puede crecer en el tiempo<sup>23</sup>. Es de anotar que si  $F_{ij}$  es definida-positiva, como ocurre en la mayoría de los casos, entonces  $\kappa$  sería estrictamente positivo y todos los modos normales estarían amortiguados. Como antes, las frecuencias de oscilación estarían afectadas por el factor de amortiguamiento, aunque naturalmente estas frecuencias deben tender a las frecuencias libres cuando el factor de amortiguamiento es muy pequeño.

### 15.7.3. Vibraciones amortiguadas forzadas

Considerando el caso en el cual el sistema que vibra posee una fuerza armónica externa y fuerzas disipativas externas, podemos plantear más fácilmente las ecuaciones diferenciales usando una extensión compleja. Para ello escribamos la fuerza armónica externa en la forma

$$F_j = F_{0j} e^{-i\omega t} \quad (15.109)$$

donde  $F_{0j}$  es complejo de modo que absorbe cualquier fase inicial que pueda existir. Las ecuaciones de movimiento quedan en la forma

$$V_{ij} \eta_j + F_{ij} \dot{\eta}_j + T_{ij} \ddot{\eta}_j = F_{0i} e^{-i\omega t} \quad (15.110)$$

coloquemos como solución particular (fase estacionaria) un ansatz de la forma

$$\eta_j = A_j e^{-i\omega t} \quad (15.111)$$

que al ser insertado en (15.110) nos da el siguiente conjunto de ecuaciones lineales inhomogéneas para las amplitudes  $A_j$

$$(V_{ij} - i\omega F_{ij} - \omega^2 T_{ij}) A_j = F_{0i} \quad (15.112)$$

<sup>23</sup>No es posible tener una función creciente de la amplitud (que implica aumento de la energía del sistema), si el sistema solo está sometido a fuerzas disipativas.

las soluciones de estas ecuaciones se obtienen por regla de Cramer para los determinantes

$$A_j = \frac{D_j(\omega)}{D(\omega)} \quad (15.113)$$

$D(\omega)$  es el determinante correspondiente a la versión homogénea de (15.110) o (15.112), en tanto que  $D_j(\omega)$  es el determinante que resulta de modificar  $D(\omega)$  cuando la  $j$ -ésima columna se reemplaza por los términos inhomogéneos  $F_{01}, \dots, F_{0n}$ . Para la presente discusión, nos centraremos en las propiedades de  $D(\omega)$  solamente, ya que las características de las resonancias en la amplitud dependen fundamentalmente del denominador en (15.113).  $D(\omega)$  es el determinante que aparece en la ecuación secular correspondiente a la versión homogénea de las Ecs. (15.110) i.e. las Ecs. (15.103), sus raíces son las frecuencias complejas de las vibraciones sin el forzamiento armónico. Por otro lado, el requerimiento de que  $\gamma$  y  $\gamma^*$  sean ambas raíces de la Ec. (15.106) significa teniendo en cuenta la Ec. (15.104) que si  $\omega_i$  es una raíz de  $D(\omega)$  entonces  $-\omega_i^*$  también lo es. Por lo anterior es posible representar  $D(\omega)$  para un sistema de  $n$  grados de libertad en la forma

$$D(\omega) = G(\omega - \omega_1)(\omega - \omega_2) \dots (\omega - \omega_n)(\omega + \omega_1^*)(\omega + \omega_2^*) \dots (\omega + \omega_n^*) \quad (15.114)$$

siendo  $G$  una constante compleja. En notación de productoria y usando la frecuencia  $\nu$  en lugar de la frecuencia angular  $\omega$ , podemos escribir<sup>24</sup>

$$D(\omega) = G \prod_{i=1}^n [2\pi(\nu - \nu_i) + i\kappa_i][2\pi(\nu + \nu_i) + i\kappa_i] \quad (15.115)$$

escribiendo la amplitud (15.113) en notación polar tenemos

$$A_j = \frac{|D_j(\omega)| e^{i\alpha_j(\omega)}}{|D(\omega)| e^{i\alpha(\omega)}} = \frac{|D_j(\omega)|}{|D(\omega)|} e^{i[\alpha_j(\omega) - \alpha(\omega)]}$$

buena parte del comportamiento de las resonancias en la amplitud  $A_j$  se obtiene con el módulo  $|D(\omega)|$  del determinante del sistema (15.115)

$$|D(\omega)| = \sqrt{D^*(\omega) D(\omega)} = \sqrt{GG^* \prod_{i=1}^n [4\pi^2(\nu - \nu_i)^2 + \kappa_i^2] [4\pi^2(\nu + \nu_i)^2 + \kappa_i^2]} \quad (15.116)$$

las amplitudes de la oscilación forzada exhiben entonces resonancia en las cercanías de las frecuencias de las oscilaciones libres  $\pm\nu_i$ . Se vé también de la Ec. (15.116) que la presencia de la constante de amortiguamiento  $\kappa_i$  remueve las divergencias que aparecían en las resonancias forzadas sin amortiguamiento. Las frecuencias del forzamiento externo para las cuales la amplitud alcanza máximos locales ya no coincide exactamente con las frecuencias de los modos libres debido a la dependencia con la frecuencia de términos en  $A_j$  diferentes al denominador. No obstante, de nuevo estas frecuencias de resonancia se aproximan a las frecuencias de vibración libre cuando el factor de amortiguamiento se vuelve muy pequeño.

Ahora bien, la solución más general se obtiene adicionando la solución particular (15.111) con la solución de la ecuación homogénea asociada Ec. (15.102)<sup>25</sup>. Puesto que la solución de la ecuación homogénea decae con el tiempo tendiendo a cero para tiempos suficientemente largos, a esta parte de la solución se le conoce como **fase transitoria**. Es importante definir un **tiempo característico** en el cual la solución transitoria ha decaído en forma significativa. Este se define usualmente como el tiempo en el cual la amplitud asociada a dicha fase ha decaído en un factor  $1/e$  con respecto a su amplitud inicial. De la Ec. (15.102) se observa que este factor de decaimiento  $\tau$  viene dado por

$$\tau \equiv \frac{1}{\kappa}$$

<sup>24</sup> Al pasar de (15.114) a (15.115), hemos supuesto que  $\omega$  en la solución (15.111) es real y representa la frecuencia angular del forzamiento. A diferencia de la solución para el caso amortiguado sin forzamiento Ec. (15.102) en el cual  $\omega$  es complejo generando un factor oscilatorio y otro de decaimiento. La idea es por supuesto, demostrar la consistencia de este ansatz.

<sup>25</sup> Debe tenerse cuidado en distinguir los  $\omega$ 's de las dos ecuaciones.

si el tiempo transcurrido es menor que  $\tau$ , la contribución de la fase transitoria es significativa. Para  $t \gg \tau$  podemos quedarnos solo con la solución particular (15.111) conocida como **fase estacionaria**, puesto que esta parte de la solución predice un movimiento con amplitud estable y la frecuencia del forzamiento externo. Ya habíamos visto que la presencia de decaimiento cuando se consideraba fricción con el medio, se debía a la pérdida de energía por fricción. En consecuencia, para que exista un movimiento estacionario se debe dar un efecto de cancelación entre la energía disipada por los osciladores debida a la fricción y la energía entregada al sistema por el forzamiento externo.

Aunque la presente discusión se realizó en el marco de sistemas mecánicos, hay en realidad diversos sistemas físicos que están descritos por ecuaciones de la forma (15.110). Quizás el ejemplo mas conocido es el de los circuitos eléctricos en donde las Ecs. (15.110) representan la dinámica de un conjunto de  $n$  circuitos acoplados donde los coeficientes  $V_{ij}$  representan coeficientes recíprocos de capacitancia, las  $F_{ij}$  son resistencias, y las  $T_{ij}$  son inductancias. Adicionalmente, el término de forzamiento representaría generadores senoidales de determinada frecuencia aplicada a uno o más de los circuitos (ver ejercicio 4 Pág. 450).

Hay una serie de consideraciones generales que solo mencionaremos por efecto de espacio. Por ejemplo, cuando el forzamiento sobre un sistema oscilante lineal tiene una dependencia temporal de forma arbitraria, la técnica de la transformada de Laplace resulta muy útil para estudiar la respuesta de estos sistemas. Cuando tomamos la expansión del potencial más allá del término cuadrático en (15.5) aparecen oscilaciones no lineales (oscilaciones anarmónicas), que discutiremos brevemente en la sección 15.8. Otro aspecto relevante es la generalización a oscilaciones de sistemas continuos, este tema solo lo trataremos brevemente en la sección 18.2.

## 15.8. Ejemplos de oscilaciones anarmónicas (opcional)

Se denominan oscilaciones anarmónicas a todas aquellas oscilaciones en las cuales se incorporan términos de fuerza no lineales. En el caso del péndulo, se obtienen términos anarmónicos cuando no se realiza la aproximación  $\sin \phi \simeq \phi$ , sino que se considera la función  $\sin \phi$  completa, o se realiza una expansión de dicha función más allá del término lineal. En la presente sección tomaremos el término  $\sin \phi$  completo, y analizaremos soluciones aproximadas de las ecuaciones resultantes cuando el oscilador es además amortiguado forzado. El péndulo de amplias oscilaciones utilizando expansiones de  $\sin \phi$  será estudiado en la sección 17.4.2, Pág. 493.

### 15.8.1. Ecuaciones del péndulo amortiguado forzado de amplias oscilaciones

Consideraremos el caso de un péndulo amortiguado y forzado en el cual no haremos restricción sobre los ángulos de oscilación i.e.  $-\pi \leq \phi \leq \pi$ . Asumiremos por simplicidad que la varilla que lo sostiene no tiene masa, y ésta se concentra en la lenteja.

Para el presente análisis haremos consideraciones de torques en lugar de fuerzas. Por ejemplo, la dinámica del péndulo simple de longitud  $R$  y masa  $m$  está gobernada por el torque gravitacional  $N_g = -mgR \sin \phi$  el cual es de tipo restaurador. Si denotamos  $\alpha$  como la aceleración angular del péndulo y si  $I = mR^2$  es el momento de inercia, tenemos que

$$I\alpha = N_g \quad (15.117)$$

$$mR^2 \frac{d^2\phi}{dt^2} + mgR \sin \phi = 0 \quad (15.118)$$

Para desplazamientos pequeños se hace la aproximación  $\sin \phi \cong \phi$  lo cual linealiza las ecuaciones haciendo que el torque sea proporcional al desplazamiento, y el movimiento será armónico simple con frecuencia característica  $\omega_0$

$$\omega_0 = \sqrt{\frac{g}{R}}$$

Si aplicamos un torque externo adicional  $N$  a un péndulo estacionario, éste oscilará hasta un cierto ángulo  $\phi$ . El valor de este ángulo puede ser acotado gracias a la acción del torque gravitacional restaurador y el péndulo

asume una posición de equilibrio cuando el torque externo se cancela con el torque gravitacional, esto se da a un ángulo  $\phi$  dado por

$$N + N_g = 0 \Rightarrow N = mgR \sin \phi \quad \left( \frac{d\phi}{dt} = 0 \right) \quad (15.119)$$

de modo que  $\phi$  es un ángulo de retorno ( $\dot{\phi} = 0$ ). Por supuesto que entre mayor sea  $N$ , mayor es el ángulo  $\phi$ . Por ejemplo, para  $N = 0$  se tiene que  $\phi = 0$ ; para  $N = \frac{1}{2}mgR$  se tiene  $\phi = \pi/6$  y para  $N = mgR$  se tiene que  $\phi = \pi/2$ , este último caso define un valor crítico  $N_c$  para el torque

$$N_c = mgR \quad ; \quad \phi = \frac{\pi}{2} \quad (15.120)$$

si  $N$  excede este valor crítico, el torque aplicado será siempre mayor al torque restaurador  $N > mgR \sin \phi$ . En consecuencia, el péndulo girará en forma indefinida hasta que dejemos de aplicar dicho torque (o deje de ser válida la condición  $N > N_c$ ) aunque naturalmente el movimiento puede continuar después de que ha cesado el torque.

Si ahora el péndulo está amortiguado debido por ejemplo a algún medio viscoso en el cual está inmerso, es usual considerar que la fuerza viscosa es proporcional a la velocidad lineal, en nuestro caso será mas conveniente (y equivalente<sup>26</sup>) asumir que es proporcional a la velocidad angular  $\omega$

$$F_\eta = -\eta' \omega = -\eta' \frac{d\phi}{dt} \quad ; \quad N_\eta = -\eta' R \omega \equiv -\eta \omega = -\eta \frac{d\phi}{dt}$$

añadimos entonces a las ecuaciones de movimiento (15.117), el torque externo  $N$  y el asociado a la fuerza viscosa con lo cual queda

$$\begin{aligned} N_g + N + N_\eta &= I\alpha \\ mR^2 \frac{d^2\phi}{dt^2} + mgR \sin \phi + \eta \frac{d\phi}{dt} &= N \end{aligned} \quad (15.121)$$

ahora definimos una frecuencia crítica  $\omega_c$  a la cual el torque  $\eta \omega$  se iguala al torque crítico  $mgR$

$$\omega_c = \frac{mgR}{\eta} = \frac{N_c}{\eta} \quad (15.122)$$

con lo cual la Ec. (15.121) se puede escribir en una forma normalizada

$$\begin{aligned} \frac{N}{mgR} &= \frac{R}{g} \frac{d^2\phi}{dt^2} + \sin \phi + \frac{\eta}{mgR} \frac{d\phi}{dt} \\ \frac{N}{N_c} &= \frac{1}{\omega_0^2} \frac{d^2\phi}{dt^2} + \frac{1}{\omega_c} \frac{d\phi}{dt} + \sin \phi \end{aligned} \quad (15.123)$$

### 15.8.2. Soluciones del péndulo amortiguado forzado de amplias oscilaciones

Las soluciones a la ecuación (15.123) poseen comportamiento complejos de  $\phi$  en función del tiempo. Estudiaremos el caso en el cual se aplica un torque constante a un péndulo inicialmente en reposo. En tal sistema hay un comportamiento transitorio inicial que eventualmente cede para dar paso a un comportamiento estacionario. Examinaremos varios casos concentrándonos únicamente en el estado estacionario

1) Para torques *pequeños*  $N \leq N_c$  hay un estado estacionario estático

$$N = N_c \sin \phi \quad (15.124)$$

en el cual se anulan todas las derivadas temporales después de que se han extinguido todas las oscilaciones transitorias iniciales. Esto se ilustra en la Fig. 15.4, que nos muestra un péndulo estacionario en el ángulo  $\phi$ .

<sup>26</sup>Tal equivalencia es válida solo si se desprecia la fuerza viscosa ejercida a lo largo de la varilla, de modo que todo el torque se concentra a una distancia fija  $R$ , i.e. sobre la lenteja.

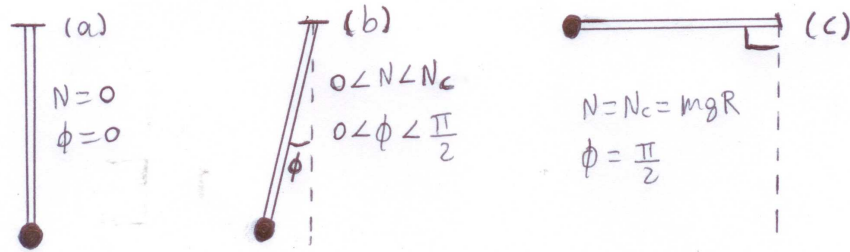


Figura 15.4: Solución para el péndulo amortiguado forzado con torque pequeño i.e.  $N \leq N_c = mgR$ . (a) Es obvio que si el péndulo está inicialmente en reposo entonces  $\phi = 0$  para todo tiempo cuando  $N = 0$ . (b) Para  $0 < N < N_c$  tendremos un ángulo constante en la fase estacionaria tal que  $0 < \phi < \pi/2$ , donde  $\phi$  está determinado por la Ec. (15.119). (c) Cuando  $N = N_c = mgR$  la Ec. (15.119) nos dice que  $\phi = \pi/2$  en todo tiempo cuando se ha extinguido la fase transitoria.

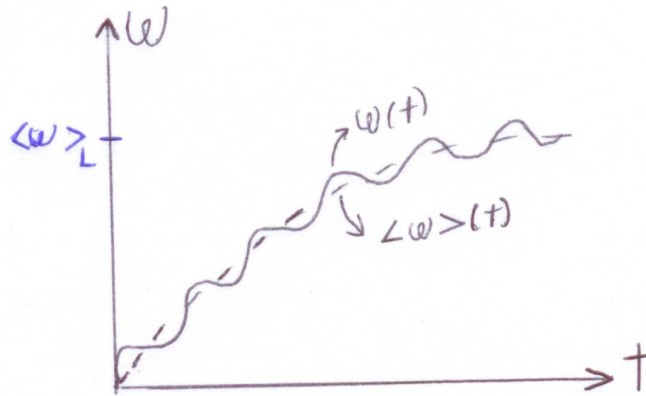


Figura 15.5: Comportamiento de  $\omega$  y  $\langle \omega \rangle$  en función del tiempo para  $N > N_c$  y amortiguamiento pequeño. El valor promedio  $\langle \omega \rangle$  se representa por la línea punteada en tanto que la línea continua representa el valor instantáneo  $\omega(t)$ . La línea punteada presenta un valor asintótico  $\langle \omega \rangle_L$  en presencia de amortiguamiento. Cuando el amortiguamiento se anula, la línea punteada se convierte en línea recta y no hay valor asintótico para  $\langle \omega \rangle$ .

2) Para movimiento no amortiguado ( $\eta = 0$ ) con torque aplicado constante  $N$ , la Ec. (15.121) toma la forma

$$N_T = N - mgR \sin \phi = mR^2 \frac{d^2 \phi}{dt^2} \quad (15.125)$$

que nos muestra que el torque neto  $N_T$  depende del ángulo. Es fácil ver el valor de este torque para ciertos ángulos particulares

$$\begin{aligned} N_T &= N & ; & \phi = 0 \\ N_T &= N - N_c & ; & \phi = \pi/2 \\ N_T &= N & ; & \phi = \pi \\ N_T &= N + N_c & ; & \phi = 3\pi/2 \end{aligned} \quad (15.126)$$

si el torque aplicado excede al valor del torque crítico  $N_c$  el movimiento será de rotación continuamente acelerada, de modo que el péndulo incrementa su energía cinética con el tiempo. La velocidad angular también se incrementará con el tiempo, pero con fluctuaciones que varían cíclicamente. Sin embargo, el promedio sobre

estas oscilaciones nos da una velocidad angular  $\langle \omega \rangle = \langle d\phi/dt \rangle$  que se comporta linealmente con el tiempo. Cuando se introduce un pequeño amortiguamiento, el valor promedio  $\langle \omega \rangle$  será una curva suave monótonamente creciente en el tiempo como se vé en la Fig. 15.5. El valor instantáneo de  $\omega(t)$  fluctuará cíclicamente alrededor de  $\langle \omega \rangle$  como lo muestra la figura 15.5.

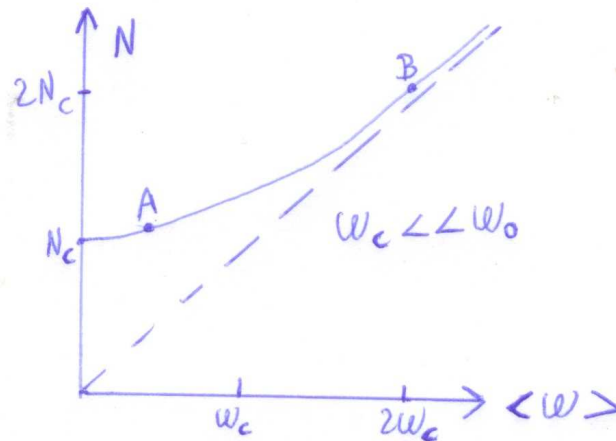


Figura 15.6: Gráfica de  $N$  vs  $\langle \omega \rangle$  con base en las Ecs. (15.127, 15.129) con  $\omega_c \ll \omega_0$ . La línea punteada representa el valor asintótico para  $N \gg N_c$  dado por la Ec. (15.130).

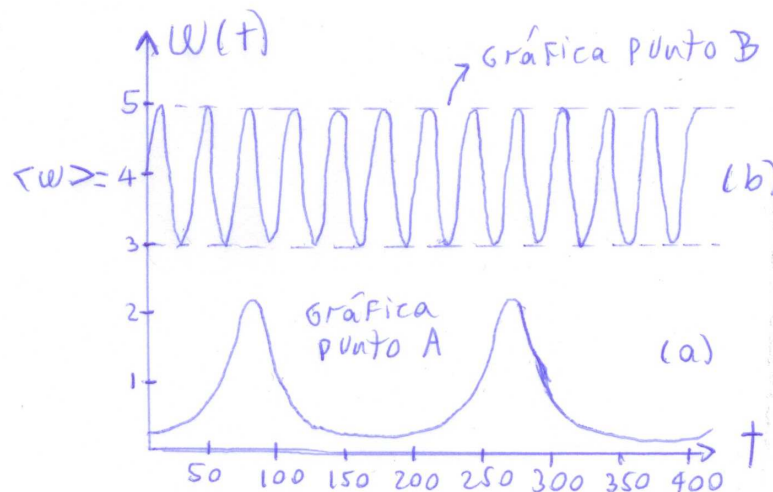


Figura 15.7: Gráfica de  $\langle \omega \rangle$  vs  $t$ , para los puntos A y B de la Figura 15.6. (a) El punto A de la Fig. 15.6 corresponde a  $N = 1,2N_c$  y es la curva en la parte inferior de la gráfica. (b) El punto B de la Fig. 15.6 corresponde a  $N = 2N_c$  y es la curva en la parte superior de la gráfica.

3) Cuando hay atenuación con  $\omega_c \ll \omega_0$  y  $N > N_c$ , la velocidad angular  $\omega$  continúa aumentando hasta que el término de atenuación  $\eta d\phi/dt$  alcanza el valor del torque aplicado. En tal caso, la velocidad angular promedio  $\langle \omega \rangle$  alcanza un valor asintótico  $\langle \omega \rangle_L$  como muestra la Fig. 15.5, y la aceleración fluctúa de tal forma que su valor promedio es nulo  $\langle d^2\phi/dt^2 \rangle = 0$ . En este régimen asintótico, el péndulo experimenta un movimiento cuasi-estático rotando con una velocidad angular  $\omega$  que fluctúa alrededor de  $\langle \omega \rangle_L$  pero que permanece cerca al valor promedio  $\langle \omega \rangle_L$ .

Es fácil obtener información cuantitativa de este régimen asintótico cuasi-estático, ya que solo hay que despreciar el término de aceleración en (15.123)

$$\frac{N}{N_c} = \frac{1}{\omega_c} \frac{d\phi}{dt} + \sin \phi \tag{15.127}$$

que tiene solución analítica de acuerdo con los siguientes casos (a) Si  $N \leq N_c$ , una solución estacionaria es la dada por la Ec. (15.124) de modo que

$$\left\langle \frac{d\phi}{dt} \right\rangle = \langle \omega \rangle = 0 \quad \text{para } N \leq N_c \tag{15.128}$$

(b) Si  $N > N_c$  la Ec. (15.127) tiene la solución estacionaria

$$\langle \omega \rangle = \omega_c \left[ (N/N_c)^2 - 1 \right]^{1/2} \quad \text{para } N > N_c \tag{15.129}$$

(c) Para mirar el comportamiento asintótico de  $N$  versus  $\langle \omega \rangle$ , examinamos el límite  $N \gg N_c$ , para el cual  $(N/N_c)^2 - 1 \simeq (N/N_c)^2$  de modo que la ecuación (15.129) queda

$$\langle \omega \rangle \simeq \omega_c (N/N_c) \quad \text{para } N \gg N_c \tag{15.130}$$

como antes, el valor instantáneo  $\omega(t)$  fluctúa cíclicamente alrededor de su valor promedio  $\langle \omega \rangle(t)$ . La figura 15.6 muestra el comportamiento de  $N$  versus  $\langle \omega \rangle$ , en tanto que la figura 15.7 muestra el comportamiento temporal de las fluctuaciones en los puntos A y B de la figura 15.6. En A se ha tomado  $N = 1,2N_c$  de modo que a partir de las Ecs. (15.126) se vé que el torque neto varía entre  $0,2N_c$  y  $2,2N_c$  alrededor del ciclo, y la velocidad angular es grande en la parte baja y lenta en la parte superior, con las variaciones que se muestran en la parte de abajo de la fig. 15.7. Para el punto B se tomó  $N = 2N_c$  de tal modo que según (15.126), el torque varía entre  $N_c$  y  $3N_c$ , produciendo variaciones más regulares en la rapidez angular que se muestran en la parte superior de la fig. 15.7. En el límite  $N \gg N_c$  que implica  $\langle \omega \rangle \gg \omega_c$ , la rapidez angular se aproxima a una variación sinusoidal en el tiempo

$$\omega(t) \cong \langle \omega \rangle + \alpha \sin \Omega t$$

que es en forma aproximada el perfil que se vé para las fluctuaciones del punto B en la fig. 15.7.

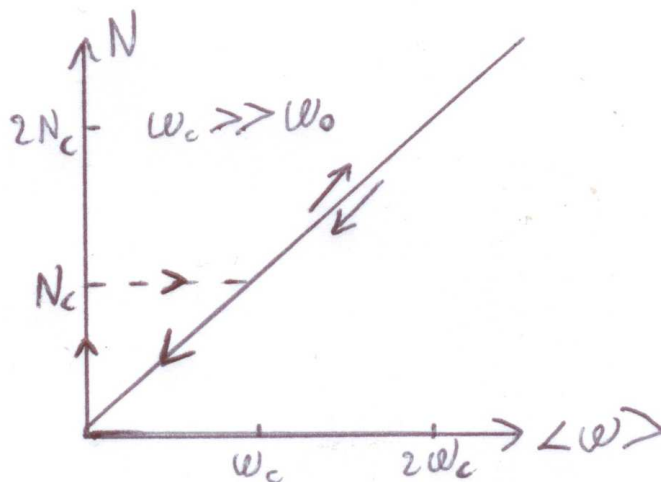


Figura 15.8: Gráfica de  $N$  vs  $\langle \omega \rangle$ , con amortiguamiento despreciable  $\eta \rightarrow 0$  y  $\omega_c \gg \omega_0$ . El sistema presenta histéresis para  $\langle \omega \rangle < \omega_c$ .



4) Para amortiguamiento despreciable en el cual  $\eta \rightarrow 0$  y  $\omega_c \gg \omega_0$ , la solución de estado estacionario dada por (15.128), aún es posible para  $N < N_c$  con el péndulo fijo en el ángulo  $\phi$  dado por (15.124), con lo cual  $\omega = \langle \omega \rangle = 0$ . Por otro lado, la solución (15.130) en la cual el torque se cancela con la fuerza viscosa promedio, es aplicable para todo  $N$  mayor o menor que  $N_c$ . Tenemos entonces dos soluciones de la forma

$$\omega = 0 \quad \text{para} \quad N \leq N_c \tag{15.131}$$

$$\langle \omega \rangle = \frac{N}{N_c} \omega_c \quad \text{para} \quad N \geq 0 \tag{15.132}$$

esto significa que hay dos soluciones diferentes para el mismo intervalo  $0 \leq N \leq N_c$ . Un análisis detallado muestra que esto implica que el sistema exhibe histéresis. Es decir que la respuesta del sistema es diferente cuando el torque se incrementa que cuando se disminuye. Cuando el torque se incrementa desde  $N = 0$  hasta  $N = N_c$  el péndulo se estabiliza en el ángulo  $\phi$  que satisface la relación  $N = N_c \sin \phi$  de la Ec. (15.124), por tanto  $\omega = 0$  durante este incremento, lo cual se muestra con la flecha vertical sobre el eje  $Y$  en la Fig. 15.8. Cuando  $N$  alcanza el valor crítico  $N_c$ , la velocidad angular promedio cambia abruptamente desde cero hasta  $\omega_c$  (transición de fase), como lo indica la flecha horizontal sobre la línea punteada de la Fig. 15.8. Finalmente, incrementando el torque desde  $N_c$  se obtiene un crecimiento lineal de  $\langle \omega \rangle$  como lo muestra la figura 15.8.

Ahora bien, si empezamos con un torque  $N > N_c$  y comenzamos a disminuirlo, se aplica la Ec. (15.132) y por tanto  $\langle \omega \rangle$  decrece linealmente hasta el origen, como lo muestra la Fig. 15.8, cuando la flecha indica decrecimiento de  $N$ . Tenemos entonces que hay histéresis cuando  $\langle \omega \rangle < \omega_c$ .

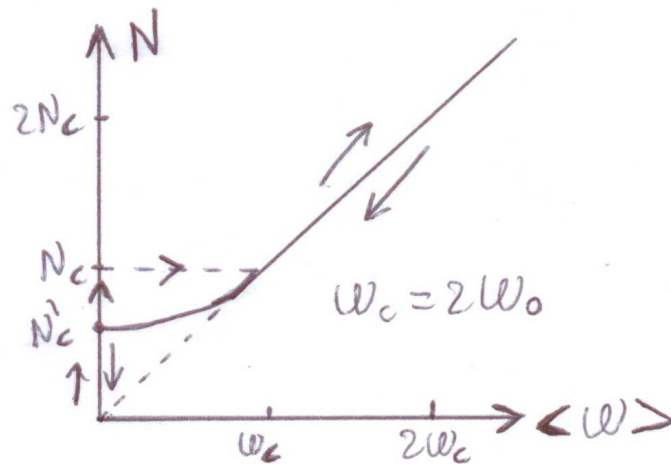


Figura 15.9: Gráfica de  $N$  vs  $\langle \omega \rangle$ , con  $\omega_c \approx \omega_0$ . Para torques pequeños, el péndulo oscila y  $\langle \omega \rangle = 0$ , en tanto que para torques grandes  $N > N_c$  el movimiento es continuo y la gráfica es lineal. El sistema presenta histéresis para  $\langle \omega \rangle < \omega_c$ .

5) Hemos analizado los casos de  $\omega_c$  mucho menor y mucho mayor que  $\omega_0$ . El caso,  $\omega_c \approx \omega_0$  es el más complejo, puesto que no se puede hacer ninguna aproximación y la ecuación (15.123) debe resolverse completa. La Fig. 15.9 muestra el comportamiento de la solución  $N$  versus  $\langle \omega \rangle$  para  $\omega_c = 2\omega_0$ . Una vez más el fenómeno de histéresis se hace presente. Al incrementar el torque desde cero hasta  $N_c$ , el valor de  $\langle \omega \rangle$  permanece nulo. Al alcanzar el valor de  $\omega_c$  se obtiene de nuevo un cambio abrupto de  $\langle \omega \rangle$  desde cero hasta el valor  $\omega_c$  (transición de fase), para continuar con un crecimiento lineal.

La Fig. 15.9, también muestra que cuando empezamos con un valor de  $N > N_c$  nos regresamos por el mismo camino de comportamiento lineal hasta llegar a  $N = N_c$ , aquí el comportamiento de “regreso” se desvía del camino de “ida”, si bien la curva que describe el camino de regreso no es discontinua. Finalmente, se alcanza un valor de  $\langle \omega \rangle = 0$  para un nuevo torque crítico  $N'_c < N_c$ .

### 15.9. Ejercicios

1. Demuestre que las matrices (15.30, 15.32) son ortogonales a primer orden y que  $\tilde{\mathbf{A}}\mathbf{V}\mathbf{A}$  coincide con la matriz de valores propios de  $\mathbf{V}$  en cada caso, a primer orden.
2. Encuentre los modos reales y normales de vibración de una molécula triatómica, emulando el procedimiento de la sección 15.5.3 pero con las condiciones iniciales más generales  $\eta_i(0) = A_i$  y  $\dot{\eta}_i(0) = B_i$ .
3. Partiendo de los modos normales (15.74), encuentre las elongaciones de las partículas de la molécula triatómica lineal Ecs. (15.76) empleando la ecuación de transformación (15.52) entre modos normales y elongaciones, y la matriz de transformación (15.71).
4. Para los conjuntos de circuitos acoplados de la Fig. 15.10, plantee las Ecs. de movimiento y describa las soluciones. Discuta los análogos mecánicos.

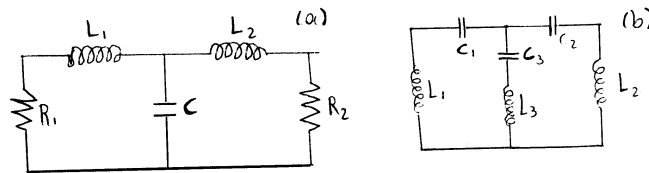


Figura 15.10: (a) Circuitos acoplados por condensador. (b) Circuitos acoplados por inductancia y condensador.

5. Una barra uniforme de longitud  $L$  y masa  $M$  está sostenida por dos resortes iguales de longitud natural  $b$  y constante de restitución  $k$ , como se ilustra en la Fig. 15.11, encuentre los modos normales de pequeñas oscilaciones en el plano.

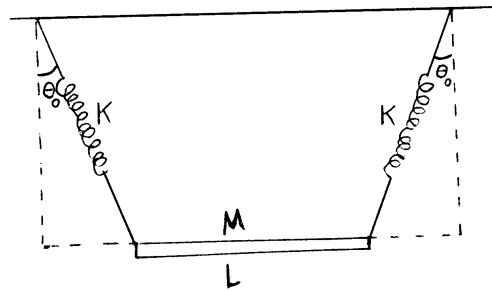


Figura 15.11: Barra suspendida por resortes idénticos que inicialmente hacen un ángulo  $\theta_0$  con la vertical. Se asumen solo modos planos de oscilación.

6. Dos masas puntuales idénticas están unidas entre sí y a puntos fijos mediante tres resortes de longitud natural  $b$  y constantes de restitución  $k$ , como se indica en la Fig. 15.12. Las masas poseen cargas eléctricas  $+q$  de modo que se repelen. Escriba la ecuación secular para las frecuencias propias de oscilación.

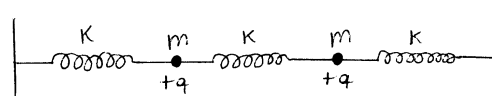


Figura 15.12: (a) Dos masas puntuales cada una de carga  $+q$ , unidas entre sí y a puntos fijos por tres resortes idénticos.

# Capítulo 16

## Relatividad especial

La mecánica newtoniana tiene como premisa que la descripción de las leyes de la naturaleza se puede realizar indistintamente en cualquier sistema de referencia inercial. Si a este postulado le agregamos el postulado de que la velocidad de la luz es la misma en todos los sistemas de referencia inerciales, obtenemos la llamada teoría de la relatividad especial o restringida. En el presente capítulo no discutiremos las motivaciones teóricas o experimentales que pueden inducir a la construcción de estos postulados, ni tampoco desarrollaremos los detalles que nos conducen a las transformaciones de Lorentz con base en tales postulados. Por tanto, es recomendable que el lector haya tenido un contacto previo con los aspectos más básicos de la teoría especial de la relatividad.

En este capítulo tenemos dos objetivos básicos (a) Estudiar las propiedades de las transformaciones de Lorentz con base en el formalismo de las matrices ortogonales y (b) desarrollar una formulación Lagrangiana para la teoría. Con el fin de alcanzar el primer objetivo, utilizaremos el formalismo de eje temporal imaginario. A pesar de que dicho formalismo ha caído en desuso, resulta muy directa su relación con el formalismo de matrices ortogonales que se ha desarrollado en el texto<sup>1</sup>. Adicionalmente, el formalismo de ejes reales y métrica de Minkowski, solo muestra claras ventajas en el marco de la relatividad general que no trataremos aquí. Finalmente, la traducción entre uno u otro formalismo es fácil de realizar, y se estudiará en la sección 16.3.

### 16.1. Propiedades de las transformaciones de Lorentz puras

En vista de lo anterior, las transformaciones de Lorentz serán nuestro punto de partida. Consideremos dos sistemas de referencia  $S$  y  $S'$  cuyo origen es coincidente en  $t = 0$  vistos por ambos sistemas.  $S'$  se mueve a velocidad constante  $\mathbf{v}$  con respecto a  $S$  en una dirección paralela al eje  $X_3$ . En primer lugar, el principio de relatividad nos indica que tales transformaciones deben ser lineales. Para verlo, observemos que si una partícula es libre en  $S$  también lo es en  $S'$  y viceversa, de modo que ambos sistemas de referencia deben predecir movimiento uniforme de la partícula. Si el movimiento uniforme en  $S$  debe transformarse en movimiento uniforme en  $S'$  las transformaciones entre tales sistemas deben ser lineales. Las transformaciones de Lorentz vienen dadas por

$$\begin{aligned}x'_1 &= x_1 \quad ; \quad x'_2 = x_2 \\x'_3 &= \frac{x_3 - vt}{\sqrt{1 - \beta^2}} \quad ; \quad t' = \frac{\left(t - \frac{vx_3}{c^2}\right)}{\sqrt{1 - \beta^2}} \quad ; \quad \beta \equiv \frac{v}{c}\end{aligned}\tag{16.1}$$

siendo  $c$  la velocidad de la luz en el vacío. Estas leyes de transformación cumplen con los postulados de la relatividad especial. Por ejemplo, la velocidad de la luz es la misma en ambos sistemas. Supongamos que con

<sup>1</sup>Adicionalmente, este formalismo nos clarificará la diferencia entre las matrices unitarias y las matrices ortogonales complejas.

respecto al sistema  $S$  se emite una onda esférica desde el origen en  $t = 0$ , la ecuación del frente de onda vista por  $S$  es

$$x_i x_i = c^2 t^2 \quad (16.2)$$

Usando las transformaciones de Lorentz (16.1) y la Ec. (16.2) vemos que

$$\begin{aligned} x'_i x'_i &= x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + \frac{(x_3 - vt)^2}{1 - \beta^2} = x_1^2 + x_2^2 + x_3^2 + \frac{(x_3 - vt)^2}{1 - \beta^2} - x_3^2 \\ &= c^2 t^2 + \frac{(x_3 - vt)^2 - x_3^2 (1 - \beta^2)}{1 - \beta^2} = \frac{c^2 t^2 (1 - \beta^2) + v^2 t^2 - 2x_3 vt + x_3^2 \beta^2}{1 - \beta^2} \\ &= \frac{c^2 t^2 \left(1 - \frac{v^2}{c^2}\right) + v^2 t^2 - 2x_3 vt + \frac{x_3^2 v^2}{c^2}}{1 - \beta^2} = \frac{\left[t^2 (c^2 - v^2) + v^2 t^2 - 2vx_3 t + \frac{v^2 x_3^2}{c^2}\right]}{1 - \beta^2} \\ &= \frac{\left[t^2 c^2 - 2vx_3 t + \frac{v^2 x_3^2}{c^2}\right]}{1 - \beta^2} = \frac{c^2 \left[t^2 - 2\left(\frac{vx_3}{c^2}\right)t + \left(\frac{vx_3}{c^2}\right)^2\right]}{1 - \beta^2} = c^2 \left[\frac{\left(t - \frac{vx_3}{c^2}\right)}{\sqrt{1 - \beta^2}}\right]^2 = c^2 t'^2 \end{aligned}$$

por tanto, la ecuación para el frente de onda transformado (es decir visto por  $S'$ ) es también esférico y se propaga también con velocidad  $c$

$$x'_i x'_i = c^2 t'^2 \quad (16.3)$$

mostrando la consistencia de las transformaciones de Lorentz con el postulado de invarianza de la velocidad de la luz. Si hacemos una expansión de las Ecs. (16.1)

$$x'_1 = x_1 \quad ; \quad x'_2 = x_2 \quad ; \quad x'_3 \approx (x_3 - vt) \left(1 + \frac{1}{2}\beta^2 + \dots\right) \quad ; \quad t' \approx \left(t - \frac{\beta}{c}x_3\right) \left(1 + \frac{1}{2}\beta^2 + \dots\right)$$

a orden cero en  $\beta = v/c$

$$x'_1 = x_1 \quad ; \quad x'_2 = x_2 \quad ; \quad x'_3 \approx x_3 - vt \quad ; \quad t' \approx t$$

con lo cual se obtienen las transformaciones de Galileo. Al ser el movimiento a lo largo de  $x_3$  las coordenadas  $x_1$  y  $x_2$  no se ven afectadas como es de esperarse en virtud de la isotropía del espacio. Como la transformación debe ser igualmente válida para pasar desde  $S'$  hacia  $S$ , se puede ver que la inversa de la transformación debe ser tal que  $T^{-1}(\mathbf{v}) = T(-\mathbf{v})$  lo cual se puede verificar invirtiendo las transformaciones de Lorentz (16.1).

La parte espacial de las transformaciones de Lorentz se puede escribir en forma vectorial teniendo en cuenta que  $\mathbf{v}$  va en la dirección  $x_3$

$$\begin{aligned} (x'_1, x'_2, x'_3) &= \left(x_1, x_2, \frac{x_3 - vt}{\sqrt{1 - \beta^2}}\right) = \left(x_1, x_2, x_3 - x_3 + \frac{x_3 - vt}{\sqrt{1 - \beta^2}}\right) \\ (x'_1, x'_2, x'_3) &= (x_1, x_2, x_3) + \left(0, 0, \frac{x_3 - vt}{\sqrt{1 - \beta^2}} - x_3\right) \\ \mathbf{r}' &= \mathbf{r} + \left(\frac{x_3 - vt}{\sqrt{1 - \beta^2}} - x_3\right) \mathbf{u}_3 \end{aligned}$$

y teniendo en cuenta que  $\mathbf{u}_3 = \mathbf{v}/v$  se obtiene

$$\mathbf{r}' = \mathbf{r} + \left(\frac{x_3 - vt}{\sqrt{1 - \beta^2}} - x_3\right) \frac{\mathbf{v}}{v}$$

usando la relación  $\mathbf{v} \cdot \mathbf{r} = vx_3$ , la parte espacial se puede escribir en forma enteramente vectorial

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + \left(\frac{vx_3 - v^2 t}{\sqrt{1 - \beta^2}} - vx_3\right) \frac{\mathbf{v}}{v^2} = \mathbf{r} + \left(\frac{vx_3}{\sqrt{1 - \beta^2}} - vx_3\right) \frac{\mathbf{v}}{v^2} - \frac{v^2 t}{\sqrt{1 - \beta^2}} \frac{\mathbf{v}}{v^2} \\ \mathbf{r}' &= \mathbf{r} + vx_3 \left(\frac{1}{\sqrt{1 - \beta^2}} - 1\right) \frac{\mathbf{v}}{v^2} - \frac{vt}{\sqrt{1 - \beta^2}} = \mathbf{r} + (\mathbf{v} \cdot \mathbf{r}) \frac{\mathbf{v}}{v^2} \left(\frac{1}{\sqrt{1 - \beta^2}} - 1\right) - \frac{vt}{\sqrt{1 - \beta^2}} \end{aligned}$$

y definiendo

$$\boldsymbol{\beta} \equiv \frac{\mathbf{v}}{c} \quad ; \quad \gamma \equiv \frac{1}{\sqrt{1 - \beta^2}}$$

se obtiene

$$\mathbf{r}' = \mathbf{r} + \frac{(\boldsymbol{\beta} \cdot \mathbf{r})\boldsymbol{\beta}}{\beta^2} (\gamma - 1) - \boldsymbol{\beta}\gamma ct \quad (16.4)$$

y para la transformación de la coordenada temporal

$$t' = \frac{\left(t - \frac{\mathbf{v} \cdot \mathbf{r}}{c^2}\right)}{\sqrt{1 - \beta^2}} = \gamma \left(t - \frac{\mathbf{v} \cdot \mathbf{r}}{c^2}\right)$$

$$t' = \gamma t - (\boldsymbol{\beta} \cdot \mathbf{r}) \frac{\gamma}{c} \quad (16.5)$$

dado que no hay nada especial en la dirección  $x_3$  elegida (por la isotropía del espacio) las ecuaciones vectoriales (16.4, 16.5) son válidas para direcciones arbitrarias de  $\mathbf{v}$  siempre que los ejes de  $S$  y  $S'$  sean paralelos.

### 16.1.1. Transformaciones de Lorentz puras y matrices ortogonales complejas

Las Ecs. (16.4, 16.5) definen transformaciones lineales homogéneas en 4 componentes, por tanto podemos utilizar el formalismo matricial para describir estas transformaciones. Un artificio ideado por Minkowski nos permite construir un sistema coordenado cartesiano con cuatro ejes en el cual el cuarto eje coordenado se elije como  $x_4 \equiv ict$ . El cuadrado del módulo de un vector en este espacio se escribe como

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = x_1^2 + x_2^2 + x_3^2 - c^2 t^2 \quad (16.6)$$

recordemos que esta cantidad es invariante debido a la exigencia de que la velocidad de la luz sea invariante<sup>2</sup>. En consecuencia debemos usar matrices ortogonales en cuatro dimensiones<sup>3</sup>, por tanto las transformaciones de Lorentz se describirán con matrices ortogonales en el espacio de Minkowski.

Dado que la cuarta coordenada es imaginaria, los elementos de la matriz de transformación pueden ser complejos. La representación matricial se puede obtener de las ecuaciones vectoriales (16.4, 16.5). Representando por  $\mathbf{L}$  a la matriz de transformación de Minkowski se tiene que

$$\mathbf{x}' = \mathbf{L}\mathbf{x} \quad (16.7)$$

siendo  $\mathbf{L}_{\mu\nu}$  un elemento genérico. Las letras griegas representarán a las cuatro coordenadas en tanto que las letras latinas representarán solo coordenadas espaciales. Las ecuaciones vectoriales (16.4, 16.5) en componentes se escriben como

$$x'_j = x_j + \frac{\beta_j \beta_k x_k}{\beta^2} (\gamma - 1) + i\beta_j \gamma x_4 \quad (16.8)$$

$$x'_4 = -i\beta_k x_k \gamma + \gamma x_4 \quad (16.9)$$

Con lo cual se pueden determinar los elementos de  $\mathbf{L}$  para una dirección arbitraria de  $\boldsymbol{\beta}$

$$x'_j = \left[ \delta_{jk} + \frac{\beta_j \beta_k}{\beta^2} (\gamma - 1) \right] x_k + i\beta_j \gamma x_4 \quad ; \quad x'_4 = -i\beta_k \gamma x_k + \gamma x_4 \quad (16.10)$$

que podemos escribir en la forma

$$x'_j = L_{jk} x_k + L_{j4} x_4 \quad ; \quad x'_4 = L_{4k} x_k + L_{44} x_4 \quad (16.11)$$

<sup>2</sup>La invarianza se prueba con un procedimiento análogo al que nos llevó a la Ec. (16.3).

<sup>3</sup>Las matrices unitarias mantienen invariante la cantidad  $x_i x_i^*$  en tanto que las matrices ortogonales mantienen invariante la cantidad  $x_i x_i$ . La invarianza de la cantidad expresada en (16.6) es de la forma  $x_i x_i$  con algunas componentes complejas, en consecuencia las matrices asociadas serán ortogonales complejas y no unitarias.

y comparando (16.10) con (16.11) se obtienen los elementos de  $\mathbf{L}$

$$L_{jk} = \delta_{jk} + \frac{\beta_j \beta_k}{\beta^2} (\gamma - 1) \quad ; \quad L_{j4} = i\beta_j \gamma \quad ; \quad L_{4k} = -i\beta_k \gamma \quad ; \quad L_{44} = \gamma \quad (16.12)$$

En el caso particular en el cual  $\mathbf{v}$  va dirigida a lo largo de  $x_3$  tenemos que  $\beta_j = \beta \delta_{j3}$  y  $\mathbf{L}$  adopta la forma

$$\begin{aligned} L_{jk} &= \delta_{jk} + \frac{\beta^2 \delta_{j3} \delta_{k3}}{\beta^2} (\gamma - 1) = \delta_{jk} + \delta_{j3} \delta_{k3} (\gamma - 1) \quad ; \quad L_{j4} = i\beta \delta_{j3} \gamma \\ L_{4k} &= -i\beta \delta_{k3} \gamma \quad ; \quad L_{44} = \gamma \end{aligned} \quad (16.13)$$

que explícitamente se escribe

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix} \quad (16.14)$$

ya hemos visto que en una matriz ortogonal la inversa es igual a la traspuesta. Veamos que ocurre al hacer la traspuesta de la matriz  $\mathbf{L}(\boldsymbol{\beta})$  en (16.12)

$$\begin{aligned} \tilde{L}_{jk}(\boldsymbol{\beta}) &= L_{kj}(\boldsymbol{\beta}) = \delta_{kj} + \frac{\beta_k \beta_j}{\beta^2} [\gamma(\boldsymbol{\beta}) - 1] = \delta_{jk} + \frac{(-\beta_j)(-\beta_k)}{(-\beta)^2} [\gamma(-\boldsymbol{\beta}) - 1] = L_{jk}(-\boldsymbol{\beta}) \\ \tilde{L}_{j4}(\boldsymbol{\beta}) &= L_{4j}(\boldsymbol{\beta}) = -i\beta_j \gamma = i(-\beta_j) \gamma = L_{j4}(-\boldsymbol{\beta}) \\ \tilde{L}_{4k}(\boldsymbol{\beta}) &= L_{k4}(\boldsymbol{\beta}) = i\beta_k \gamma = -i(-\beta_k) \gamma = L_{4k}(-\boldsymbol{\beta}) \quad ; \quad \tilde{L}_{44}(\boldsymbol{\beta}) = L_{44}(\boldsymbol{\beta}) = \gamma(\boldsymbol{\beta}) = \gamma(-\boldsymbol{\beta}) = L_{44}(-\boldsymbol{\beta}) \end{aligned}$$

donde hemos usado el hecho de que  $\gamma(\boldsymbol{\beta}) = \gamma(-\boldsymbol{\beta})$  lo cual es evidente de su definición. Tenemos por tanto que

$$\tilde{L}_{\mu\nu}(\boldsymbol{\beta}) = L_{\mu\nu}(-\boldsymbol{\beta})$$

y como la traspuesta es la inversa llegamos a la propiedad esperada de que  $\mathbf{L}^{-1}(\mathbf{v}) = \mathbf{L}(-\mathbf{v})$ .

Notemos que la submatriz inferior  $2 \times 2$  en (16.14) se asemeja a una rotación en un plano, la cual se escribiría de la forma

$$\begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}$$

en este caso lo que tenemos es una rotación en los ejes  $x_3 x_4$  del espacio de Minkowski, pero en un ángulo  $\phi$  imaginario

$$\cos \phi = \gamma \quad ; \quad \sin \phi = i\beta\gamma \quad (16.15)$$

podemos definir un ángulo real  $\psi$  en la forma  $\phi \equiv i\psi$  con lo cual

$$\cosh \psi = \gamma \quad ; \quad \sinh \psi = \beta\gamma$$

y la matriz (16.14) queda

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \phi & \sin \phi \\ 0 & 0 & -\sin \phi & \cos \phi \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh \psi & i \sinh \psi \\ 0 & 0 & -i \sinh \psi & \cosh \psi \end{pmatrix} \quad (16.16)$$

esta parametrización facilita muchas operaciones matriciales. Por ejemplo, si hacemos dos transformaciones de Lorentz sucesivas en donde ambas poseen velocidades relativas a lo largo de  $x_3$ , la transformación matricial solo es no trivial en el plano  $x_3 x_4$  y se puede ver que simplemente se suman los ángulos  $\phi$  y  $\phi'$  correspondientes como ocurre en una rotación en el plano, de modo que  $\mathbf{L}(\phi') \mathbf{L}(\phi) = \mathbf{L}(\phi + \phi')$ . De las Ecs. (16.15) se tiene que

$$\tan \phi = i\beta \quad ; \quad i\beta'' = \tan \phi'' = \tan(\phi + \phi') = \frac{\tan \phi + \tan \phi'}{1 - \tan \phi \tan \phi'} \quad \Rightarrow \quad i\beta'' = \frac{i\beta + i\beta'}{1 - (i\beta)(i\beta')}$$

de modo que estas dos transformaciones de Lorentz sucesivas corresponden a una sola transformación de Lorentz equivalente de la forma

$$\beta'' = \frac{\beta + \beta'}{1 + \beta\beta'} \quad (16.17)$$

la Ec. (16.17) corresponde a la ley de adición de velocidades para velocidades paralelas. En esta ecuación se vé que la velocidad equivalente no es simplemente la suma de las velocidades de las dos transformaciones en virtud del factor de corrección  $\beta\beta'$  en el denominador. Podemos ver además que incluso tomando valores de  $\beta$  y  $\beta'$  cercanos a la unidad, se tiene que  $\beta'' < 1$ . Esto indica que no se puede obtener una velocidad mayor que  $c$  con transformaciones de Lorentz sucesivas. En conclusión, mientras asumamos la validez de los postulados de la relatividad especial, no hay manera de que un sistema de referencia vaya más rápido que la luz con respecto a otro.

## 16.2. Transformaciones de Lorentz restringidas

Aunque hemos visto que las transformaciones de Lorentz son ortogonales, no hemos demostrado que éstas cubran todas las transformaciones ortogonales posibles en el espacio de Minkowski, de por sí esto no es cierto como se puede demostrar con una transformación descrita por  $L_{44} = 0$ ,  $L_{4i} = L_{i4} = 0$  y los nueve elementos restantes formando una submatriz  $3 \times 3$  ortogonal en el espacio euclídeo tridimensional. Esta matriz es ortogonal en el espacio de Minkowski, pero no produce movimiento relativo entre los dos sistemas, su efecto es una rotación de las coordenadas espaciales. **Las rotaciones espaciales son un subconjunto de las transformaciones ortogonales en el espacio de Minkowski.** Similarmente, (16.7) no define la transformación de coordenadas más general ante la cual deben permanecer invariantes las ecuaciones de la Física, pues es claro que una redefinición de origen espacio temporal tampoco debe afectar a las leyes de la Física. Debemos ver además si existen otro tipo de transformaciones ortogonales en el espacio de Minkowski. La transformación más general en el espacio de Minkowski que mantiene invariante la velocidad de la luz es

$$\mathbf{x}' = \mathbf{L}\mathbf{x} + \mathbf{a} \quad (16.18)$$

donde  $\mathbf{a}$  representa una traslación del origen en el espacio de Minkowski (i.e. de espacio y tiempo) y  $\mathbf{L}$  es una matriz ortogonal. A las transformaciones del tipo (16.18) se les conoce como transformaciones de Poincaré o transformaciones de Lorentz inhomogéneas. La condición de ortogonalidad

$$\tilde{\mathbf{L}}\mathbf{L} = \tilde{\mathbf{L}}\mathbf{L} = \mathbf{1} \quad ; \quad L_{\mu\nu}L_{\mu\rho} = \delta_{\nu\rho} \quad \text{ó} \quad L_{\nu\mu}L_{\rho\mu} = \delta_{\nu\rho} \quad (16.19)$$

representa diez ligaduras sobre los elementos de  $\mathbf{L}$  (cuatro condiciones diagonales y seis no diagonales) de modo que solo hay seis cantidades independientes en  $\mathbf{L}$ . Por otro lado, vemos que las transformaciones de Lorentz (16.12) involucran tres grados de libertad (las tres componentes de la velocidad) en tanto que las rotaciones euclídeas involucran otros tres grados de libertad (e.g. los ángulos de Euler). Esto parece indicarnos que las transformaciones de Lorentz del tipo (16.12) junto con las rotaciones espaciales (o combinaciones de ambas) forman el conjunto más general de transformaciones ortogonales en el espacio de Minkowski. Por otro lado, para la transformación (16.18) existen cuatro grados de libertad adicionales con lo cual la cantidad de elementos independientes será diez. En el presente estudio nos restringimos a las transformaciones de Lorentz homogéneas de modo que requerimos manejar seis elementos independientes de  $\mathbf{L}$

$$\mathbf{x}' = \mathbf{L}\mathbf{x}$$

recordemos que las matrices ortogonales tienen determinante  $\pm 1$

$$(\det \mathbf{L})^2 = \pm 1$$

y ya hemos visto que si nos restringimos a las transformaciones continuas debemos excluir las matrices de determinante  $-1$ . Las matrices  $\mathbf{L}$  de determinante  $+1$  representan entonces **transformaciones de Lorentz**

**propias.** Sin embargo, no hay garantía de que todas las matrices de determinante +1 correspondan a transformaciones continuas. Efectivamente, en el caso de la inversión simultánea de todas las coordenadas espacio temporales, el determinante sigue siendo +1. Necesitamos entonces un criterio para excluir las transformaciones propias no continuas<sup>4</sup>. Examinemos el comportamiento de  $L_{44}$ , usando las Ecs. (16.19) se puede escribir con  $\nu = \rho = 4$

$$L_{4\mu}L_{4\mu} = \delta_{44} \quad \Rightarrow \quad L_{44}^2 + L_{4j}L_{4j} = 1 \quad (16.20)$$

y como los elementos  $L_{4j}$  conectan una coordenada espacial (real) con una temporal (imaginaria), estos elementos deben ser imaginarios puros. En contraste  $L_{44}$  debe ser real porque conecta al eje imaginario consigo mismo, estas características se pueden apreciar en (16.12). En consecuencia  $L_{4j}L_{4j}$  debe ser negativo y  $L_{44}^2$  debe ser positivo de modo que

$$|L_{44}|^2 > \sum_{j=1}^3 |L_{4j}L_{4j}| \quad \text{y} \quad L_{44}^2 \geq 1 \quad (16.21)$$

La Ec. (16.21) plantea dos posibilidades:  $L_{44} \leq -1$  que implica una inversión del tiempo y  $L_{44} \geq 1$  que implica una transformación continua a partir de la identidad<sup>5</sup>. Las transformaciones de Lorentz con  $L_{44} \geq 1$  se denominan **ortócronas** en tanto que las de  $L_{44} \leq -1$  se denominan **no ortócronas**. Solamente las transformaciones ortogonales **propias ortócronas** pueden evolucionar en forma continua a partir de la identidad. De las cuatro subclases solo las transformaciones **propias ortócronas forman un grupo, las otras tres subclases no**. A las transformaciones de Lorentz propias ortócronas se les conoce como transformaciones de Lorentz restringidas, solo ellas pueden generar rotaciones continuas en el espacio y reducirse a las transformaciones de Galileo en el límite de bajas velocidades. En consecuencia, solo trabajaremos transformaciones de Lorentz restringidas denominándolas simplemente transformaciones de Lorentz.

### 16.2.1. Transformaciones de Lorentz restringidas: Boosts y rotaciones

A las transformaciones de Lorentz restringidas que corresponden a dos sistemas de ejes paralelos que se mueven uniformemente uno respecto al otro se les denomina **transformaciones de Lorentz puras (o boosts)**. La matriz descrita por (16.12) corresponde a una transformación de Lorentz pura. La intuición nos indica que una transformación de Lorentz restringida puede descomponerse en una transformación de Lorentz pura junto con una rotación espacial sin movimiento relativo (en uno u otro orden). Veamos como se realizaría tal descomposición. Descompongamos la transformación de Lorentz en un boost  $\mathbf{P}(\boldsymbol{\beta})$  seguido de la transformación remanente  $\mathbf{R}(\boldsymbol{\beta})$

$$\mathbf{L}(\boldsymbol{\beta}) = \mathbf{R}(\boldsymbol{\beta}) \mathbf{P}(\boldsymbol{\beta}) \quad (16.22)$$

mostraremos que la transformación remanente  $\mathbf{R}(\boldsymbol{\beta})$  es una rotación espacial. Las coordenadas del sistema transformado  $x'_\nu$  están relacionadas con las coordenadas no primadas por

$$\mathbf{x}' = \mathbf{L}\mathbf{x} \quad \Rightarrow \quad \mathbf{x} = \mathbf{L}^{-1}\mathbf{x}' \quad \Rightarrow \quad \mathbf{x} = \tilde{\mathbf{L}}\mathbf{x}' \quad \Rightarrow \quad x_\mu = L_{\nu\mu}x'_\nu \quad (16.23)$$

nos preguntamos ahora cual es la velocidad del origen de  $S'$  vista por un observador en  $S$ . En el origen de  $S'$  tenemos que  $x'_j = 0$  y las coordenadas del origen de  $S'$  vistas por el observador en  $S$  se obtienen aplicando (16.23) con  $x'_j = 0$

$$x_j = L_{4j}x'_4 \quad ; \quad x_4 = L_{44}x'_4 \quad (16.24)$$

de las Ecs. (16.24) vemos que la velocidad relativa (normalizada a unidades de  $c$ ) del origen de  $S'$  tiene entonces las siguientes componentes

$$\beta_j = \frac{x_j}{ct} = \frac{ix_j}{x_4} = \frac{iL_{4j}x'_4}{L_{44}x'_4} = \frac{iL_{4j}}{L_{44}} \quad (16.25)$$

<sup>4</sup>Obsérvese que esto se debe al carácter par de la dimensión del espacio. En  $\mathbb{R}^3$  las inversiones espaciales producen determinante  $(-1)^3 = -1$ , de modo que no hay transformaciones discretas con determinante +1 en el espacio cartesiano tridimensional.

<sup>5</sup>La identidad tiene  $L_{44} = 1$  como se puede ver de (16.12) con  $\mathbf{v} = \mathbf{0}$ . Con un argumento similar al que se usó para transformaciones ortogonales impropias en  $\mathbb{R}^3$ , no es de esperarse que para una transformación continua haya un cambio abrupto desde la identidad (con  $L_{44} = 1$ ) hasta un valor de  $L_{44} \leq -1$  sin pasar por estados intermedios. Por lo tanto, las matrices con  $L_{44} \leq -1$  contienen al menos una transformación discreta.



combinando las Ecs. (16.20, 16.25) se obtiene

$$\begin{aligned} L_{44}^2 \left( 1 + \frac{L_{4j}L_{4j}}{L_{44}^2} \right) &= 1 \quad \Rightarrow \quad L_{44}^2 \left[ 1 - \left( i \frac{L_{4j}}{L_{44}} \right) \left( i \frac{L_{4j}}{L_{44}} \right) \right] = 1 \\ L_{44}^2 (1 - \beta_j \beta_j) &= 1 \end{aligned} \quad (16.26)$$

podemos ver que  $|\beta|^2$  está entre cero y uno, usando la primera desigualdad en (16.21) aplicada a (16.25)<sup>6</sup>

$$|\beta|^2 = |\beta_j \beta_j| = \left| \frac{(iL_{4j})(iL_{4j})}{L_{44}^2} \right| = \frac{|L_{4j}L_{4j}|}{|L_{44}|^2} \leq 1 \quad (16.27)$$

mostrando una vez más que el sistema de referencia  $S'$  debe ir a una velocidad menor que  $c$  con respecto a  $S$ . Por otra parte, despejando el valor de  $L_{44}$  en términos de  $\beta$  en (16.26) se obtiene<sup>7</sup>

$$L_{44} = (1 - \beta_j \beta_j)^{-1/2} = \gamma \quad (16.28)$$

por lo tanto,  $L_{44} = \gamma$  para cualquier transformación de Lorentz restringida. Construyendo entonces una transformación de Lorentz pura  $\mathbf{P}(\beta)$  asociada al vector velocidad relativa del origen de  $S'$  Ec. (16.25), vemos que la transformación inversa debe ser  $\mathbf{P}(-\beta)$ . Ahora teniendo en cuenta (16.22) se encuentra entonces que la matriz  $\mathbf{R}$  se puede despejar

$$\begin{aligned} \mathbf{L}(\beta) &= \mathbf{R}(\beta) \mathbf{P}(\beta) \quad \Rightarrow \quad \mathbf{L}(\beta) \mathbf{P}(-\beta) = \mathbf{R}(\beta) \mathbf{P}(\beta) \mathbf{P}(-\beta) \\ &\Rightarrow \quad \mathbf{R}(\beta) = \mathbf{L}(\beta) \mathbf{P}(-\beta) \end{aligned} \quad (16.29)$$

se puede demostrar formalmente que este producto entre  $\mathbf{P}(-\beta)$  y  $\mathbf{L}$  corresponde a una rotación en el espacio usando los elementos matriciales de  $\mathbf{P}(-\beta)$  y la ortogonalidad de  $\mathbf{L}$ . No obstante se puede ver geoméricamente que en la Ec. (16.22), el sistema intermedio de coordenadas definido por  $\mathbf{P}(\beta)$  está en reposo respecto al sistema final de ejes de modo que  $\mathbf{R}$  solo puede girar las coordenadas. Esta descomposición nos permite deducir que los parámetros independientes siempre serán las tres componentes de la velocidad relativa entre los sistemas y los tres grados de libertad de la rotación espacial (por ejemplo los ángulos de Euler).

### 16.2.2. Composición de boosts

Por otro lado, puede demostrarse que la composición de transformaciones de Lorentz puras no es en general otra transformación de Lorentz pura a menos que sean paralelas las velocidades relativas de las transformaciones sucesivas. El caso general es muy complejo y poco ilustrativo, veamos entonces un cálculo sencillo que posee amplias aplicaciones en Física moderna dando origen al efecto llamado **precesión de Thomas**.

Tomaremos tres sistemas inerciales con orígenes  $O_1, O_2, O_3$ . El sistema  $O_1$  es el laboratorio y  $O_2$  tiene velocidad  $\beta$  relativa a  $O_1$ .  $O_3$  se mueve con velocidad  $\beta'$  relativa a  $O_2$ . Sin pérdida de generalidad se puede tomar a  $\beta$  en la dirección de  $x_3$  de  $O_1$  y a  $\beta'$  lo podemos tomar sobre el plano  $x_2x_3$  de  $O_2$  de modo que  $\beta$  y  $\beta'$  definen el plano  $x_2 - x_3$  de  $O_2$ . Supondremos que las componentes de  $\beta'$  son muy pequeñas de modo que solo las conservamos hasta el menor grado no nulo. Con esto,  $\gamma'$  para la transformación entre  $O_2$  y  $O_3$  se puede sustituir por la unidad<sup>8</sup>. Con base en lo anterior la matriz  $\mathbf{L}$  que nos lleva de  $O_1$  a  $O_2$  tiene la forma dada por (16.14) y la matriz que nos lleva de  $O_2$  a  $O_3$  se escribe usando la aproximación  $\gamma' \cong 1$  en (16.12)

$$L_{jk}(\beta') = \delta_{jk} \quad ; \quad L_{j4}(\beta') = i\beta'_j \quad ; \quad L_{4k}(\beta') = -i\beta'_k \quad ; \quad L_{44}(\beta') = 1$$

y recordando que por construcción  $\beta'_1 = 0$ , explícitamente queda

$$\mathbf{L}(\beta') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & i\beta'_2 \\ 0 & 0 & 1 & i\beta'_3 \\ 0 & -i\beta'_2 & -i\beta'_3 & 1 \end{pmatrix} \quad (16.30)$$

<sup>6</sup>La primera desigualdad en (16.21) implica suma sobre  $j$ , pero es obvio que también es válida para  $j$  fijo, como se aplica en (16.27).

<sup>7</sup>Nótese que en (16.28) hemos asumido que  $L_{44} \geq 1$  al tomar la raíz cuadrada positiva.

<sup>8</sup>Puesto que  $\gamma' = (1 - \beta'^2)^{-1/2}$  su expansión de Taylor es  $\gamma'(\beta') \cong 1 + \beta'^2/2 + \dots$  de modo que no hay término lineal en  $\beta'$ .

siendo  $\beta'_2, \beta'_3$  las componentes de  $\beta'$ . Las Ecs. (16.14, 16.30) nos dan la matriz producto, que con la misma aproximación nos da

$$\begin{aligned} \mathbf{L}(\beta'') &= \mathbf{L}(\beta') \mathbf{L}(\beta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & i\beta'_2 \\ 0 & 0 & 1 & i\beta'_3 \\ 0 & -i\beta'_2 & -i\beta'_3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix} \\ \mathbf{L}(\beta'') &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta\gamma\beta'_2 & i\gamma\beta'_2 \\ 0 & 0 & \gamma + \beta\gamma\beta'_3 & i\beta\gamma + i\gamma\beta'_3 \\ 0 & -i\beta'_2 & -i\beta\gamma - i\gamma\beta'_3 & \gamma + \beta\gamma\beta'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta\gamma\beta'_2 & i\gamma\beta'_2 \\ 0 & 0 & \gamma(1 + \beta\beta'_3) & i\gamma(\beta + \beta'_3) \\ 0 & -i\beta'_2 & -i\gamma(\beta + \beta'_3) & \gamma(1 + \beta\beta'_3) \end{pmatrix} \\ \mathbf{L}(\beta'') &= \mathbf{L}(\beta') \mathbf{L}(\beta) \cong \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta\beta'_2\gamma & i\beta'_2\gamma \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & -i\beta'_2 & -i\beta\gamma & \gamma \end{pmatrix} \end{aligned} \quad (16.31)$$

donde se ha despreciado  $\beta'_3$  frente a  $\beta$  por considerar a  $\beta'$  pequeña y se ha usado  $\beta\beta'_3 \ll 1$ . Se puede ver que (16.31) no representa una transformación de Lorentz pura ya que por ejemplo los elementos  $L''_{ij}$  de las coordenadas espaciales no son simétricos como lo demandan las Ecs. (16.12) para transformaciones de Lorentz puras o boosts. Usando la Ec. (16.25), podemos ver que las componentes de la velocidad relativa entre  $O_1$  y  $O_3$  se escriben en la forma

$$\beta''_2 = \frac{iL_{42}(\beta'')}{L_{44}(\beta'')} = \frac{\beta'_2}{\gamma} ; \quad \beta''_3 = \beta ; \quad \beta''_1 = 0 \quad (16.32)$$

con lo cual podemos calcular  $(\beta'')^2$

$$\begin{aligned} (\beta'')^2 &= (\beta''_1)^2 + (\beta''_2)^2 + (\beta''_3)^2 = 0 + \frac{\beta'^2_2}{\gamma^2} + \beta^2 \\ (\beta'')^2 &\simeq \beta^2 \Rightarrow \gamma'' \simeq \gamma \end{aligned}$$

con esto, podemos aproximar la transformación de Lorentz pura  $L_{\mu\nu}(\beta'') \equiv P_{\mu\nu}(\beta'')$ , asociada a la velocidad relativa  $\beta''$  reemplazando estas aproximaciones en (16.12)

$$\begin{aligned} P_{jk}(\beta'') &= \delta_{jk} + \frac{\beta''_j \beta''_k}{\beta''^2} (\gamma'' - 1) \cong \delta_{jk} + \frac{\beta''_j \beta''_k}{\beta^2} (\gamma - 1) ; \quad P_{j4}(\beta'') = i\beta''_j \gamma'' \cong i\beta''_j \gamma \\ P_{4k}(\beta'') &= -i\beta''_k \gamma'' \cong -i\beta''_k \gamma ; \quad P_{44}(\beta'') = \gamma'' \cong \gamma ; \end{aligned} \quad (16.33)$$

combinando las Ecs. (16.32, 16.33) y despreciando los términos de orden  $\beta'^2_2$ , podemos construir  $\mathbf{P}(\beta'')$

$$\mathbf{P}(\beta'') \simeq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\beta'_2}{\beta\gamma}(\gamma - 1) & i\beta'_2 \\ 0 & \frac{\beta'_2}{\beta\gamma}(\gamma - 1) & \gamma & i\beta\gamma \\ 0 & -i\beta'_2 & -i\beta\gamma & \gamma \end{pmatrix} \quad (16.34)$$

Para obtener  $\mathbf{P}(-\beta'')$  que es el que nos interesa, podemos utilizar el hecho de que  $\gamma''(-\beta'') = \gamma''(\beta'')$  y observar de las Ecs. (16.33) que al invertir  $\beta''$  se obtiene

$$\begin{aligned} P_{jk}(-\beta'') &= \delta_{jk} + \frac{(-\beta''_j)(-\beta''_k)}{(-\beta''^2)} (\gamma'' - 1) = P_{jk}(\beta'') ; \quad P_{j4}(-\beta'') = i(-\beta''_j) \gamma'' = -P_{j4}(\beta'') \\ P_{4k}(-\beta'') &= -i(-\beta''_k) \gamma'' = -P_{4k}(\beta'') ; \quad P_{44}(-\beta'') = \gamma'' = P_{44}(\beta'') \end{aligned} \quad (16.35)$$

combinando las ecuaciones (16.34, 16.35)  $\mathbf{P}(-\boldsymbol{\beta})$  queda en la forma

$$\mathbf{P}(-\boldsymbol{\beta}'') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & -i\beta'_2 \\ 0 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & \gamma & -i\beta\gamma \\ 0 & i\beta'_2 & i\beta\gamma & \gamma \end{pmatrix} \quad (16.36)$$

sustituyendo (16.31) y (16.36) en la Ec. (16.29), podemos encontrar la matriz de rotación correspondiente de los ejes de  $O_3$  con respecto a  $O_1$ .

$$\begin{aligned} \mathbf{R}(\boldsymbol{\beta}'') &= \mathbf{L}(\boldsymbol{\beta}'') \mathbf{P}(-\boldsymbol{\beta}'') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \beta\beta'_2\gamma & i\beta'_2\gamma \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & -i\beta'_2 & -i\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & -i\beta'_2 \\ 0 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & \gamma & -i\beta\gamma \\ 0 & i\beta'_2 & i\beta\gamma & \gamma \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (\beta'_2)^2(\gamma-1) - \gamma(\beta'_2)^2 + 1 & \frac{1}{\beta\gamma}\beta'_2(\gamma-1) & -i\beta'_2\beta^2\gamma^2 + i\beta'_2\gamma^2 - i\beta'_2 \\ 0 & \frac{1}{\beta}\beta'_2(\gamma-1) - \beta\gamma\beta'_2 & \gamma^2 - \beta^2\gamma^2 & 0 \\ 0 & i\gamma\beta'_2 - i\beta'_2 - i\beta'_2(\gamma-1) & -\frac{i}{\beta\gamma}(\beta'_2)^2(\gamma-1) & -\beta^2\gamma^2 + \gamma^2 - (\beta'_2)^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - (\beta'_2)^2 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & i\beta'_2[-\beta^2\gamma^2 + \gamma^2 - 1] \\ 0 & \frac{\beta'_2}{\beta}[(\gamma-1) - \beta^2\gamma] & \gamma^2(1 - \beta^2) & 0 \\ 0 & i\beta'_2[\gamma - 1 - (\gamma-1)] & -\frac{i(\beta'_2)^2}{\beta\gamma}(\gamma-1) & (1 - \beta^2)\gamma^2 - (\beta'_2)^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - (\beta'_2)^2 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & i\beta'_2[\gamma^2(1 - \beta^2) - 1] \\ 0 & \frac{\beta'_2}{\beta}[\gamma(1 - \beta^2) - 1] & 1 & 0 \\ 0 & 0 & -\frac{i(\beta'_2)^2}{\beta\gamma}(\gamma-1) & 1 - (\beta'_2)^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - (\beta'_2)^2 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & 0 \\ 0 & \frac{\beta'_2}{\beta}\left[\frac{1}{\gamma} - 1\right] & 1 & 0 \\ 0 & 0 & -\frac{i(\beta'_2)^2}{\beta\gamma}(\gamma-1) & 1 - (\beta'_2)^2 \end{pmatrix} \end{aligned}$$

Suprimiendo los términos de orden superior en  $\beta'_2$  se obtiene<sup>9</sup>

$$\mathbf{R}(\boldsymbol{\beta}'') = \mathbf{L}(\boldsymbol{\beta}'') \mathbf{P}(-\boldsymbol{\beta}'') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & 0 \\ 0 & -\frac{\beta'_2}{\beta\gamma}(\gamma-1) & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (16.37)$$

esta ecuación se puede reescribir en la forma

$$\mathbf{R}(\boldsymbol{\beta}'') = \mathbf{1} + \boldsymbol{\varepsilon}(\boldsymbol{\beta}'') \quad ; \quad \boldsymbol{\varepsilon}(\boldsymbol{\beta}'') = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\beta'_2}{\beta\gamma}(\gamma-1) & 0 \\ 0 & -\frac{\beta'_2}{\beta\gamma}(\gamma-1) & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (16.38)$$

<sup>9</sup>De hecho, para obtener consistentemente los términos cuadráticos en  $\beta'_2$  para  $\mathbf{R}(\boldsymbol{\beta}'')$ , es necesario tener en cuenta los términos de orden  $\beta'^2_2$  en la Ec. (16.34).

como esta rotación es a primer orden en  $\beta'_2$  se puede comparar con una rotación infinitesimal<sup>10</sup>. Comparando entonces la submatriz  $3 \times 3$  superior izquierda de  $\varepsilon(\beta'')$  en (16.38) con la matriz infinitesimal (13.50) Pág. (13.50), se obtiene que (16.37) está asociada a una rotación alrededor del eje  $x_1$  con un ángulo (pequeño) dado por

$$\Delta\Omega_1 = \frac{\beta'_2}{\beta\gamma}(\gamma - 1) = \beta''_2 \frac{(\gamma - 1)}{\beta} \quad (16.39)$$

donde hemos usado (16.32). Este es entonces un ejemplo concreto de dos transformaciones de Lorentz a ejes paralelos sucesivas (boosts de Lorentz) que dan como resultado la combinación de un boost con una rotación. Esta paradoja tiene importantes aplicaciones especialmente en Física atómica como veremos a continuación.

### 16.2.3. Precesión de Thomas

Consideremos una partícula que se mueve en el laboratorio con una velocidad  $\mathbf{v} = v\mathbf{u}_3$  no constante, el sistema en el cual esta partícula está en reposo no es inercial y por tanto no es aplicable el formalismo anterior. Para obviar esta dificultad, consideraremos un conjunto de sistemas inerciales todos coincidentes con el original en  $t = 0$  y que viajan a diferentes valores de velocidad relativa (todos los valores de velocidad que se requieran). En consecuencia, la partícula estará en reposo instantáneo con respecto a alguno de estos sistemas de referencia en cualquier instante de tiempo.

Pensemos que  $O_1$  es el sistema del laboratorio y que  $O_2$  y  $O_3$  son sistemas en los cuales la partícula está en reposo instantáneo en dos tiempos  $t_2$  y  $t_3$  respectivamente. El observador  $O_1$  verá que en el intervalo de tiempo  $\Delta t = t_3 - t_2$  ocurre una variación  $\Delta\mathbf{v}$  en la velocidad de la partícula dada por

$$\frac{\Delta\mathbf{v}}{c} = \frac{\mathbf{v}(t_3) - \mathbf{v}(t_2)}{c} = \beta'' - \beta$$

De acuerdo con las Ecs. (16.32), dicha variación será

$$\begin{aligned} \frac{\Delta\mathbf{v}}{c} &= \left(0, \frac{\beta'_2}{\gamma}, \beta\right) - (0, 0, \beta) = \frac{\beta'_2}{\gamma}\mathbf{u}_2 = \beta''_2\mathbf{u}_2 \\ \Rightarrow \Delta\mathbf{v} &= \beta''_2 c \mathbf{u}_2 \equiv \Delta v \mathbf{u}_2 \end{aligned} \quad (16.40)$$

sustituyendo (16.40) en (16.39) se tiene que

$$\Delta\Omega_1 = \beta''_2 \frac{(\gamma - 1)}{v/c} = \beta''_2 c v \frac{(\gamma - 1)}{v^2} = (\Delta v) v \frac{(\gamma - 1)}{v^2}$$

Dado que  $\mathbf{v} = v\mathbf{u}_3$ , y que  $\Delta\mathbf{v} = \Delta v \mathbf{u}_2$ , la ecuación anterior se puede escribir en forma vectorial. El vector asociado a la rotación (pequeña) durante este tiempo se puede escribir

$$\Delta\boldsymbol{\Omega} = (\gamma - 1) \frac{\Delta\mathbf{v} \times \mathbf{v}}{v^2}$$

de modo que si la partícula tiene alguna dirección específica asociada a ella (como un vector de espín), el sistema del laboratorio observará que esta dirección experimenta una precesión de velocidad angular

$$\boldsymbol{\omega} \simeq \frac{\Delta\boldsymbol{\Omega}}{\Delta t} = (\gamma - 1) \frac{\mathbf{a} \times \mathbf{v}}{v^2} \quad (16.41)$$

siendo  $\mathbf{a}$  la aceleración de la partícula vista desde  $O_1$ . La Ec. (16.41) aparece con frecuencia en la literatura en la forma que posee cuando se toma el límite de velocidad pequeña que permite aproximar a  $\gamma \simeq 1 + v^2/2c^2$

$$\boldsymbol{\omega} = \frac{1}{2c^2} (\mathbf{a} \times \mathbf{v})$$

$\boldsymbol{\omega}$  es la frecuencia de precesión de Thomas.

<sup>10</sup>Las ecuaciones desarrolladas en la sección 13.8, son válidas para matrices infinitesimales, pero también son aproximadamente válidas para rotaciones finitas suficientemente pequeñas como para permitir mantener solo términos de primer orden.

### 16.3. Transformaciones de Lorentz usando espacios de Riemann de cuatro dimensiones

Una forma alternativa de trabajar el espacio en donde ocurren las transformaciones de Lorentz es asumir que el espacio tetradimensional es real de modo que la cuarta coordenada (denominada la coordenada cero en la notación más usual) se escribe como  $x_0 \equiv ct$ . Debemos por supuesto mantener el postulado fundamental de que la luz se propague a la misma velocidad en todos los sistemas, lo cual se manifiesta como la invarianza de la cantidad dada en la ecuación (16.6). Para que esta cantidad siga representando el módulo al cuadrado de un vector en un espacio real, es necesario que el espacio deje de ser euclídeo y se convierta en un espacio de Riemann con tensor métrico diagonal definido por

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (16.42)$$

donde los índices 1230 representan las tres coordenadas espaciales y la coordenada temporal. El módulo al cuadrado de un vector en tal espacio viene dado por

$$\tilde{\mathbf{x}}\mathbf{G}\mathbf{x} = x_i x_i - x_0^2 = x_i x_i - c^2 t^2 \quad (16.43)$$

que nos representa al invariante que queremos. Una transformación de Lorentz homogénea es una transformación lineal en este espacio real que mantiene invariante este módulo de los vectores. Es evidente que la matriz asociada a estas transformaciones debe ser real en este espacio, de modo que la denotaremos por  $\mathbf{\Lambda}$ . La condición de invarianza del módulo de los vectores ante una transformación de Lorentz se escribe matricialmente en la forma

$$\tilde{\mathbf{x}}'\mathbf{G}\mathbf{x}' = \tilde{\mathbf{x}}\mathbf{G}\mathbf{x} \Rightarrow (\widetilde{\mathbf{\Lambda}\mathbf{x}})\mathbf{G}(\mathbf{\Lambda}\mathbf{x}) = \tilde{\mathbf{x}}\mathbf{G}\mathbf{x} \Rightarrow \tilde{\mathbf{x}}\tilde{\mathbf{\Lambda}}\mathbf{G}\mathbf{\Lambda}\mathbf{x} = \tilde{\mathbf{x}}\mathbf{G}\mathbf{x}$$

y como esto es válido para un vector arbitrario en este espacio, la condición para las transformaciones de Lorentz resulta

$$\tilde{\mathbf{\Lambda}}\mathbf{G}\mathbf{\Lambda} = \mathbf{G} \quad (16.44)$$

La Ec. (16.44) es una transformación de congruencia que deja invariante al tensor métrico. Haciendo la analogía con las matrices ortogonales del espacio euclídeo (donde el tensor métrico cartesiano es  $\mathbf{1}$ ), podemos decir que (16.44) es la condición de ortogonalidad de  $\mathbf{\Lambda}$  en el espacio real de Riemann con tensor métrico  $\mathbf{G}$ <sup>11</sup>.

La relación entre las fórmulas expresadas en el espacio de Minkowski y las expresadas en el espacio real de Riemann se logra con las siguientes asociaciones simples

$$x_4 = ix_0 \quad ; \quad \Lambda_{j0} = iL_{j4} \quad ; \quad \Lambda_{0k} = -iL_{4k}$$

en tanto que los demás elementos no varían, lo cual es de esperarse ya que ambos contienen al subespacio  $\mathbb{R}^3$  dotado de la misma estructura. A manera de ejemplo, la transformación de Lorentz pura con velocidad relativa a lo largo de  $x_3$  correspondiente a la Ec. (16.14), tiene la siguiente representación matricial real en este espacio de Riemann

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -\beta\gamma \\ 0 & 0 & -\beta\gamma & \gamma \end{pmatrix}$$

el producto escalar se escribe usando el tensor métrico

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &\equiv \tilde{\mathbf{x}}\mathbf{G}\mathbf{y} = \tilde{\mathbf{y}}\mathbf{G}\mathbf{x} = (\mathbf{y}, \mathbf{x}) \\ \tilde{\mathbf{x}}\mathbf{G}\mathbf{y} &= x_\mu g_{\mu\nu} y_\nu \end{aligned}$$

<sup>11</sup>Es claro que esta condición se reduce a la ortogonalidad usual cuando  $\mathbf{G} = \mathbf{1}$ . Adicionalmente, la Ec. (16.44) es una transformación de similitud puesto que  $\mathbf{\Lambda}$  es real.

donde la igualdad entre  $(\mathbf{x}, \mathbf{y})$  y  $(\mathbf{y}, \mathbf{x})$  viene dada por el carácter real de este producto interno. La condición de ortogonalidad de la Ec. (16.44) garantiza la invarianza del producto escalar ante una transformación de Lorentz  $\mathbf{A}$ .

Es usual escribir estas fórmulas de manera más compacta mediante un conveniente cambio de notación. Supongamos que formamos un vector en el espacio de Riemann con los elementos de coordenadas  $dx_\mu$  y estudiemos su comportamiento ante una transformación general de coordenadas del tipo

$$y_\nu = f_\nu(x_1, x_2, \dots)$$

se encuentra que las propiedades de transformación de  $dx_\mu$  son

$$dy_\nu = \frac{\partial f_\nu}{\partial x_\mu} dx_\mu = \frac{\partial y_\nu}{\partial x_\mu} dx_\mu \quad (16.45)$$

las derivadas son los elementos de la matriz jacobiana de la transformación entre  $(x)$  e  $(y)$ . Cuando la transformación  $\mathbf{A}$  es lineal, serían simplemente los elementos matriciales  $A_{\nu\mu}$ . Por otro lado, las componentes de un vector gradiente se transforman de acuerdo con la ecuación

$$\frac{\partial}{\partial y_\nu} = \frac{\partial x_\mu}{\partial y_\nu} \frac{\partial}{\partial x_\mu} \quad (16.46)$$

nótese que en (16.46) los coeficientes corresponden a los elementos de la matriz jacobiana de la transformación inversa de  $(y)$  hacia  $(x)$ . Los vectores que se transforman de acuerdo con la regla dada por la Ec. (16.45) se denominan vectores contravariantes y se denotan con supraíndices

$$D'^\nu = \frac{\partial y_\nu}{\partial x_\mu} D^\mu$$

en contraste, los vectores que transforman de la manera prescrita por la Ec. (16.46) se denominan covariantes y se denotan con subíndices

$$F'_\nu = \frac{\partial x_\mu}{\partial y_\nu} F_\mu$$

El producto de las matrices jacobianas correspondientes a una transformación y a su inversa debe ser la matriz unidad ya que corresponde a pasar de  $(x)$  a  $(y)$  y volver de nuevo a  $(x)$ . De aquí se desprende que la contracción entre un vector contravariante y un vector covariante queda invariante ante la transformación,

$$D'^\nu F'_\nu = \frac{\partial y_\nu}{\partial x_\mu} \frac{\partial x_\rho}{\partial y_\nu} D^\mu F_\rho = \delta_{\mu\rho} D^\mu F_\rho = D^\mu F_\mu$$

veremos más adelante que la contracción de un vector covariante con otro contravariante equivale al producto interno entre dos vectores contravariantes. En el caso de espacios cartesianos, no hay diferencia entre vectores covariantes y contravariantes ante transformaciones lineales ortogonales. Para verlo, observemos que si la matriz ortogonal  $\mathbf{A}$  describe la transformación, un vector contravariante en el espacio cartesiano transforma como

$$D'^\nu = A_{\nu\mu} D^\mu$$

en tanto que un vector covariante transforma en la siguiente forma

$$F'_\nu = (\mathbf{A}^{-1})_{\mu\nu} F_\mu = (\tilde{\mathbf{A}})_{\mu\nu} F_\mu = A_{\nu\mu} F_\mu$$

de modo que no es necesario distinguir hasta ahora entre los dos tipos de comportamiento ante la transformación. Un argumento similar vale para el espacio de Minkowski con eje temporal imaginario y con la métrica usual, de modo que en dicho espacio no hay diferencia entre vectores covariantes y contravariantes.

De la misma manera en que definimos tensores cartesianos según la prescripción (12.172) Pág. 329, heredada de la transformación de los vectores cartesianos, podemos definir las propiedades de transformación de tensores

de cualquier rango en espacios no euclídeos. Por tanto, un tensor de Lorentz covariante  $\mathbf{R}$  de segundo rango se transforma con la prescripción

$$R'_{\mu\nu} = R_{\rho\lambda} \frac{\partial x_\rho}{\partial y_\mu} \frac{\partial x_\lambda}{\partial y_\nu}$$

y se puede demostrar que la contracción de un tensor de segundo rango covariante con un tensor de segundo rango contravariante (o con dos vectores contravariantes) es invariante ante la transformación. Similarmente, la contracción de un tensor de segundo rango covariante con un vector contravariante transforma como un vector covariante

$$R_{\mu\nu} H^{\mu\nu} = s_1 \quad R_{\mu\nu} R^\mu M^\nu = s_2 \quad ; \quad R_{\mu\nu} D^\nu = F_\mu$$

donde  $s_1$  y  $s_2$  son invariantes ante la transformación (escalares) y  $F_\mu$  es un vector covariante. Veamos la demostración de que  $F_\mu$  de la tercera ecuación es un vector de Lorentz covariante

$$\begin{aligned} F'_\mu &= R'_{\mu\nu} D'^\nu = R_{\rho\lambda} \frac{\partial x_\rho}{\partial y_\mu} \frac{\partial x_\lambda}{\partial y_\nu} \frac{\partial y_\nu}{\partial x_\tau} D^\tau = R_{\rho\lambda} \frac{\partial x_\rho}{\partial y_\mu} \delta_{\lambda\tau} D^\tau = R_{\rho\lambda} D^\lambda \frac{\partial x_\rho}{\partial y_\mu} \\ F'_\mu &= F_\rho \frac{\partial x_\rho}{\partial y_\mu} \end{aligned}$$

En un espacio de Riemann el tensor métrico se construye a través de un elemento diferencial de longitud de arco

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu$$

que se construye de tal manera que sea invariante ante las transformaciones de interés. De esto se desprende que el tensor métrico es covariante de segundo rango. Nótese que en el caso particular de las transformaciones de Lorentz, esto se puede ver directamente de la condición de ortogonalidad (16.44) si la escribimos en la forma  $\mathbf{G} = \tilde{\mathbf{\Lambda}}^{-1} \mathbf{G} \mathbf{\Lambda}^{-1}$  considerada como transformación de congruencia en  $\mathbf{G}$ .

Vemos entonces que en el espacio de Riemann real de cuatro dimensiones, el producto escalar de dos vectores contravariantes  $A^\mu$ ,  $B^\nu$  se puede escribir en la forma

$$g_{\mu\nu} A^\mu B^\nu = (g_{\mu\nu} A^\mu) B^\nu = A_\nu B^\nu \quad (16.47)$$

donde hemos tenido en cuenta el carácter covariante del tensor métrico para obtener el vector covariante  $A_\nu$ . En este punto se puede ver con claridad que si la métrica fuera la usual i.e.  $\delta_{\mu\nu}$ , los vectores covariantes y contravariantes serían idénticos. En particular, el cuadrado del módulo del vector cuatridimensión en el cuatridimensión real se puede escribir en la forma

$$g_{\mu\nu} x^\mu x^\nu = x_\mu x^\mu$$

de esta forma los productos internos se pueden construir sin alusión directa al tensor métrico, teniendo en cuenta que un factor del producto escalar se sustituye por el vector covariante que se obtiene al contraer con el tensor métrico como se vé en (16.47). Si nos interesa el producto escalar de dos vectores covariantes, debemos “subir” el índice por contracción con el inverso del tensor métrico, el cual se puede demostrar que es contravariante. En el caso del cuatridimensión real donde el tensor métrico es diagonal con elementos  $\pm 1$ , el tensor métrico coincide con su inverso y no hay diferencias entre tensores métricos covariantes y contravariantes.

Es claro que esta no es la única forma de construir el tensor métrico, el cual fué diseñado para generar el invariante (16.43) por medio del módulo al cuadrado del vector posición en tal espacio, podemos en cambio construir el invariante en la forma

$$\tilde{\mathbf{x}} \mathbf{G}' \mathbf{x} \equiv x_\mu G'_{\mu\nu} x_\nu = -x_i x_i + c^2 t^2$$

de modo que el tensor métrico queda

$$\mathbf{G}' = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (16.48)$$

es claro que bajo la métrica (16.48) se mantiene invariante la velocidad de la luz y las matrices  $\mathbf{A}$  que describen a las transformaciones de Lorentz no se modifican. Todo el formalismo permanece inalterado excepto que el producto interno cambia de signo<sup>12</sup>. El tensor  $\mathbf{G}$  tiene la signatura  $(+++)$  en tanto que el tensor  $\mathbf{G}'$  tiene la signatura  $(---)$ . También podemos identificarlos por sus trazas  $Tr\mathbf{G} = 2$ ,  $Tr\mathbf{G}' = -2$ .

El uso del formalismo de Minkowski o de Riemann presenta cada uno sus ventajas y desventajas. En teoría general de la relatividad será necesario usar la métrica de un espacio curvo para lo cual es muy adecuado el uso de espacios de Riemann, por otro lado en mecánica cuántica donde las funciones de onda o vectores de estado son complejos, el uso de una coordenada compleja complica la operación de conjugación compleja. Por otro lado, cuando nos restringimos al marco de la relatividad especial, las operaciones en el espacio de Minkowski suelen tener analogías muy cercanas al espacio euclídeo y no es necesaria la distinción entre vectores covariantes y contravariantes, debido a la trivialidad del tensor métrico. En todo caso la mayoría de fórmulas presentan el mismo aspecto en ambos casos o su transición de uno a otro esquema es muy sencilla. Un aspecto común en ambos formalismos es la idea de que el elemento de longitud de arco tiene un carácter indefinido, pues  $(ds)^2$  puede ser positivo, negativo o cero.

## 16.4. El concepto de formulación covariante en Física

El primer postulado de la relatividad especial nos dice que las leyes de la Física deben poseer la misma forma en todos los sistemas inerciales. Por lo tanto, es de gran importancia poder verificar que las leyes de la naturaleza sean invariantes en forma bajo las transformaciones de Lorentz. Esta verificación se facilita enormemente con la introducción del concepto de tensor de Minkowski (o de Lorentz).

Cuando hablamos de la invarianza ante transformaciones de Lorentz, nos referimos tanto a los boosts como a las rotaciones en el espacio ordinario. Como la invarianza ante rotaciones tridimensionales nos es más familiar podemos usar esta invarianza como modelo para establecer un método que se generalice a todas las transformaciones de Lorentz.

Ya hemos definido los tensores euclidianos (o cartesianos) y su comportamiento bajo rotaciones. Para satisfacer el requerimiento de que una ley de la Física sea invariante ante rotaciones tridimensionales es usual escribir las ecuaciones que expresan esa ley de modo que todos sus términos sean escalares o todos vectoriales (en el sentido euclídico). Más en general, todos los términos deben ser tensores del mismo rango y este requisito asegura de manera automática la invarianza de la forma de la ecuación ante una rotación de los ejes espaciales. Por ejemplo una relación escalar tiene la forma general

$$a = b$$

y dado que los dos miembros de la igualdad por ser escalares euclidianos son invariantes ante rotaciones espaciales de los ejes, es evidente que la relación será válida para todos los sistemas de coordenadas con origen común. Una relación vectorial será de la forma

$$\mathbf{F} = \mathbf{G}$$

que se puede escribir en términos de tres relaciones numéricas entre las componentes<sup>13</sup>

$$F_i = G_i \tag{16.49}$$

Claramente, estas componentes no son invariantes ante rotaciones espaciales. En general, se transforman a nuevas componentes  $F'_i, G'_i$  que son las componentes de los vectores transformados (pasivamente)  $\mathbf{F}'$ ,  $\mathbf{G}'$ . Pero como los dos miembros de las ecuaciones se transforman de igual manera, entre las componentes transformadas se debe cumplir la misma relación

$$F'_i = G'_i$$

<sup>12</sup>En ambos casos el tensor  $\mathbf{G}$  describe una pseudométrica ya que la norma de un vector en este espacio no está necesariamente definida positiva.

<sup>13</sup>Nótese que las Ecs. (16.49) son relaciones numéricas pero no son relaciones escalares.



y por tanto la relación vectorial también se preserva con la rotación espacial; en el nuevo sistema coordenado escribimos

$$\mathbf{F}' = \mathbf{G}'$$

Es importante enfatizar que la invarianza en la forma se debe a que ambos miembros de la ecuación son vectores cartesianos. Decimos que los términos de la ecuación son covariantes. Es necesario aclarar que el concepto de covarianza empleado aquí tiene un significado muy distinto al de la covarianza de vectores en el espacio de Riemann. La covarianza en espacios de Riemann se refiere a la propiedad según la cual algunos vectores transforman bajo un cambio de coordenadas según la matriz jacobiana de la transformación, en este escenario el término se usa por contraposición a los vectores (o tensores) contravariantes que transforman con el inverso de la matriz jacobiana bajo el cambio de coordenadas. En el caso que nos ocupa ahora, la covarianza se define para los términos de una ecuación que expresa alguna ley de la Física, para indicar que todos los términos involucrados en la ecuación (escalares, vectores, tensores) transforman en la misma manera de modo que se mantiene la forma de la ecuación.

La covarianza por supuesto se puede generalizar para ecuaciones que involucran tensores de rango arbitrario, si tenemos una ecuación tensorial de la forma  $\mathbf{C} = \mathbf{D}$  los tensores transformados implicarán la misma igualdad  $\mathbf{C}' = \mathbf{D}'$  siempre que los tensores de ambos miembros sean del mismo rango. Por ejemplo, si una ecuación posee términos que son escalares, otros que son vectores etc, no se podrá mantener invariante ante una transformación ortogonal tridimensional. *Podemos concluir que la invarianza de una ley Física ante una rotación del sistema de coordenadas espaciales, exige la covarianza de los términos de la ecuación ante transformaciones ortogonales tridimensionales.*

Vamos ahora al espacio extendido de Minkowski o espacio de universo. El manejo allí es idéntico una vez que hemos caracterizado a las transformaciones ortogonales en este espacio y por ende la estructura de sus tensores de cualquier rango. A los tensores en este espacio los llamamos tensores de Minkowski, de Lorentz o de universo, genéricamente escalares de universo, vectores de universo (cuadrivectores), etc. En consecuencia, la invarianza de una ley Física ante transformaciones de Lorentz será inmediata si se expresa en forma cuatridimensional covariante, de modo que todos los términos son tensores de universo del mismo rango. De lo anterior se deriva que una teoría Física en el marco de la relatividad especial solo tiene validez si es covariante ante transformaciones de Lorentz (boosts y rotaciones espaciales).

Notemos por ejemplo que el producto de un número por un cuadrivector solo será otro cuadrivector si el número es un escalar de universo. Supongamos que  $\alpha$  es un número que no es escalar de universo, en un sistema  $S$  el producto de este número por un cuadrivector es

$$\alpha \mathbf{F} = \mathbf{W}$$

ante una transformación de Lorentz,  $\mathbf{F}$  y  $\mathbf{W}$  transforman como cuadrivectores con una cierta matriz  $\mathbf{M}$  de transformación, por otro lado  $\alpha'$  transforma en la forma  $\alpha' = \mathbf{N}\alpha$  siendo  $\mathbf{N}$  un operador diferente a la identidad (ya que no es escalar de universo). Tenemos entonces

$$\alpha' \mathbf{F}' = (\mathbf{N}\alpha) (\mathbf{M}\mathbf{F}) = \mathbf{N}\mathbf{M}(\alpha \mathbf{F}) \quad ; \quad \mathbf{W}' = \mathbf{M}\mathbf{W} \Rightarrow \alpha' \mathbf{F}' \neq \mathbf{W}'$$

por tanto si  $\mathbf{W}$  es cuadrivector  $\alpha \mathbf{F}$  no lo es y la ecuación no es covariante de Lorentz. La ecuación se vuelve covariante si  $\mathbf{N} = \mathbf{1}$ .

Finalmente, notemos que una ecuación puede ser covariante sin ser **manifiestamente covariante**. Por ejemplo, supongamos que tenemos una ecuación de la forma

$$\mathbf{F}_{\mu\nu} + \mathbf{T}_{\mu\nu} + \mathbf{H}_{\mu\nu} = \mathbf{R}_{\mu\nu}$$

y supongamos que  $\mathbf{F}_{\mu\nu}$ ,  $\mathbf{T}_{\mu\nu}$ ,  $\mathbf{H}_{\mu\nu}$  no son tensores de universo pero que  $\mathbf{R}_{\mu\nu}$  sí lo es. En general esta ecuación no será covariante, pero puede ocurrir que la suma de los tres términos no tensoriales sí transforme como un tensor gracias a ciertos efectos de cancelación, ciertamente si estos términos no son tensores será mucho más complejo demostrar la covarianza de la ecuación (si es que es covariante). Esta anotación es útil, porque a menudo ocurre que se construye una teoría en forma manifiestamente covariante, pero luego para efectos prácticos de cálculo se transforma a una estructura en donde la covarianza no es evidente.

## 16.5. Formulaciones covariantes en el espacio de Minkowski

El ejemplo más simple de cuadrivector de Lorentz es el vector de posición de un “punto” en el espacio de Minkowski o de universo, donde sus componentes se denotan por  $(x_1, x_2, x_3, x_4)$ . Las cuatro coordenadas de un punto de universo nos dicen cuando (tiempo) y donde (espacio) ha ocurrido un suceso, a todo punto de este espacio se le llama entonces un suceso o un evento.

Cuando una partícula en el espacio ordinario sigue una determinada trayectoria, su punto representativo en el espacio de Minkowski describe una trayectoria conocida como **línea de universo**. El cuadrivector  $dx_\mu$  representa la variación del cuadrivector posición para un movimiento diferencial a lo largo de la línea de universo. Este término multiplicado por sí mismo es un invariante de Lorentz de modo que representa un escalar de universo denominado  $d\tau$

$$(d\tau)^2 = -\frac{1}{c^2} dx_\mu dx_\mu \quad (16.50)$$

para elucidar el significado Físico de  $d\tau$  evaluaremos (16.50) en un sistema inercial en el cual la partícula esté en reposo instantáneo. En este sistema el cuadrivector transformado  $dx'_\mu$  asociado a esta partícula está descrito por  $(0, 0, 0, icdt')$  y el invariante  $d\tau$  se escribe

$$(d\tau)^2 = -\frac{1}{c^2} dx'_\mu dx'_\mu = (dt')^2$$

con lo que se vé que  $d\tau$  es el intervalo de tiempo medido por un reloj que se mueve con la partícula, que se denomina el **tiempo propio** o **tiempo de universo**.

Ahora veamos la relación entre  $d\tau$  y el intervalo de tiempo correspondiente a un cierto sistema inercial  $dt$ , usando la Ec. (16.50)

$$(d\tau)^2 = -\frac{1}{c^2} [(dx_1)^2 + (dx_2)^2 + (dx_3)^2 + (dx_4)^2] = -\frac{1}{c^2} [(dx_1)^2 + (dx_2)^2 + (dx_3)^2 - c^2 (dt)^2]$$

$$d\tau = (dt) \left[ \sqrt{1 - \frac{1}{c^2} \left[ \left( \frac{dx_1}{dt} \right)^2 + \left( \frac{dx_2}{dt} \right)^2 + \left( \frac{dx_3}{dt} \right)^2 \right]} \right]$$

que se puede escribir en la forma

$$dt = \frac{d\tau}{\sqrt{1 - \beta^2}} \quad (16.51)$$

debemos tener en cuenta que en este caso  $\beta$  está representando la velocidad de **una partícula** con respecto a un sistema de referencia inercial  $S$ . Este es un uso diferente al que se le ha dado hasta ahora como velocidad relativa (normalizada a  $c$ ) de un cierto sistema de referencia inercial  $S'$  con respecto a otro sistema inercial  $S$ . Por supuesto, se puede pensar en  $\beta$  como la velocidad relativa del sistema  $S'$  (con respecto a  $S$ ) de tal modo que la partícula está en reposo instantáneo con respecto a  $S'$ . La Ec. (16.51) nos dice que el intervalo de tiempo medido por el sistema en el cual la partícula no está en reposo es siempre mayor que el intervalo de tiempo medido en el sistema en donde la partícula está en reposo instantáneo. Este fenómeno se conoce como **dilatación del tiempo** y ha sido comprobado experimentalmente en diversas situaciones, particularmente en la observación de las vidas medias de partículas elementales inestables. La vida media de estas partículas se puede medir cuando éstas están en reposo y se compara con su vida media cuando están en vuelo a velocidades cercanas a la de la luz.

Hemos visto que el cuadrado del módulo de un cuadrivector no es necesariamente definido positivo. Los cuadrivectores cuyo módulo cuadrado sean positivos se denominan del *género espacial* o también se denominan *como de espacio o espacialoides*. Cuando el módulo es cero (lo cual no significa necesariamente que el cuadrivector sea cero) se denominan *como de luz*. Finalmente, cuando su módulo cuadrado es negativo se dice que es *del género temporal, como de tiempo ó temporaloide*. Puesto que este módulo al cuadrado es un escalar de universo, esta denominación no dependerá del sistema inercial utilizado. Los nombres se deben a que el módulo de un vector espacial tridimensional es definido positivo al igual que el cuadrivector del género espacial, adicionalmente un cuadrivector de este género siempre se puede transformar de tal forma que se anule su cuarta

componente (temporal). Por otro lado, un cuadrivector del género temporal tiene su cuarta componente no nula, pero se puede transformar de tal forma que se anulen todas sus tres componentes espaciales. A manera de ilustración veamos el comportamiento del vector diferencia o relativo entre dos puntos de universo. Este vector relativo puede ser del género espacial, temporal o de luz, definiremos a este vector relativo como

$$X_\mu \equiv x_{1\mu} - x_{2\mu} \quad (16.52)$$

donde los subíndices 1 y 2 denotan los dos sucesos. El módulo de este cuadrivector relativo será<sup>14</sup>

$$X_\mu X_\mu = |\mathbf{r}_1 - \mathbf{r}_2|^2 - c^2 (t_1 - t_2)^2$$

de modo que  $X_\mu$  será del género espacial si los dos puntos de universo están separados de modo que

$$|\mathbf{r}_1 - \mathbf{r}_2|^2 > c^2 (t_1 - t_2)^2$$

será como de luz si se cumple la igualdad

$$|\mathbf{r}_1 - \mathbf{r}_2|^2 = c^2 (t_1 - t_2)^2$$

y finalmente será del género temporal si

$$|\mathbf{r}_1 - \mathbf{r}_2|^2 < c^2 (t_1 - t_2)^2$$

la condición para que el vector diferencia sea temporal equivale a decir que se puede cubrir la distancia entre los dos eventos o sucesos mediante una señal luminosa (e incluso algunas señales más lentas que la luminosa), en cuyo caso se habla de sucesos o eventos causalmente conectados. La condición de cuadrivector del género espacial equivale a que estos eventos no podrán conectarse con ninguna onda luminosa o señal que viaje a velocidad menor o igual que  $c$ , decimos que los eventos están causalmente desconectados. Finalmente, si el cuadrivector diferencia es como de Luz, solo una señal que viaje a velocidad  $c$  podrá conectar a estos sucesos (y no se pueden conectar con señales que viajen a velocidades menores), claramente estos son eventos causalmente conectados.

Podemos elegir el eje  $x_3$  de modo que quede alineado con los ejes espaciales  $\mathbf{r}_1 - \mathbf{r}_2$  del cuadrivector relativo. Adicionalmente, podemos asumir sin pérdida de generalidad que  $x_{3(1)} - x_{3(2)} > 0$ . En tal caso<sup>15</sup> se tiene que  $|\mathbf{r}_1 - \mathbf{r}_2| = x_{3(1)} - x_{3(2)}$ . Si realizamos una transformación de Lorentz pura con velocidad  $v$  a lo largo de  $x_3$  podemos aplicar las transformaciones dadas en (16.1) para la cuarta componente de  $X_\mu$

$$\begin{aligned} t'_1 &= \frac{\left(t_1 - \frac{vx_{3(1)}}{c^2}\right)}{\sqrt{1 - \beta^2}} ; & t'_2 &= \frac{\left(t_2 - \frac{vx_{3(2)}}{c^2}\right)}{\sqrt{1 - \beta^2}} \\ t'_1 - t'_2 &= \frac{\left(t_1 - \frac{vx_{3(1)}}{c^2}\right)}{\sqrt{1 - \beta^2}} - \frac{\left(t_2 - \frac{vx_{3(2)}}{c^2}\right)}{\sqrt{1 - \beta^2}} = \frac{\left(t_1 - t_2 - \frac{vx_{3(1)} - vx_{3(2)}}{c^2}\right)}{\sqrt{1 - \beta^2}} \\ c(t'_1 - t'_2) &= \frac{c(t_1 - t_2) - \frac{v}{c}[x_{3(1)} - x_{3(2)}]}{\sqrt{1 - \beta^2}} \end{aligned} \quad (16.53)$$

si  $X_\mu$  es del género espacial y los sucesos son tales que  $t_1 > t_2$  nos queda que

$$c(t_1 - t_2) < x_{3(1)} - x_{3(2)}$$

y será posible encontrar una velocidad  $v < c$  de modo que se anule la cuarta componente  $ic(t'_1 - t'_2) \equiv X'_4$ . Físicamente la anulación de la componente temporal significa que es posible encontrar un sistema inercial

<sup>14</sup>Recordemos que en el espacio de Minkowski con métrica trivial, los vectores covariantes coinciden con los contravariantes de modo que usaremos solo subíndices en las contracciones.

<sup>15</sup>En el espacio cartesiano, el vector relativo suele definirse de manera causal, i.e.  $\Delta \mathbf{r} \equiv \mathbf{r}(t_1) - \mathbf{r}(t_2)$  con  $t_1 > t_2$ . En contraste, la relación (16.52) no se define causalmente, sino de modo que se cumpla la condición  $x_{3(1)} - x_{3(2)} > 0$ .

que viaje a velocidad  $v < c$  en el cual los dos sucesos sean simultáneos. Adicionalmente, también es posible encontrar valores de  $v < c$  que hagan que el miembro de la derecha en (16.53) se vuelva negativo lo cual indicaría que  $t'_2 > t'_1$ , de modo que encontramos un sistema de referencia inercial en el cual se invierte la secuencia de los sucesos. El que pueda invertirse la secuencia de sucesos entre eventos del género espacial no constituye una violación de la causalidad ya que estos eventos están causalmente desconectados y no hay manera de que un suceso pueda influir en el otro. Por ejemplo, nada de lo que ocurra ahora en la tierra puede afectar a la estrella alfa centauri dentro de los siguientes cuatro años en virtud de su distancia a la tierra de unos cuatro años luz.

En contraste, para separaciones del género temporal entre sucesos, no es posible encontrar una transformación de Lorentz que los haga simultáneos y menos aún que pueda invertir el orden temporal de los sucesos. Así debe ser, puesto que estos eventos sí están causalmente conectados y pueden influir el uno sobre el otro. Esto implica que el antes y el después, o la causa y el efecto, son conceptos invariantes de Lorentz y se preserva la causalidad.

### 16.5.1. Cuadrivectores notables

Es importante establecer la generalización relativista de las cantidades Newtonianas más importantes. Por ejemplo la velocidad

$$v_i = \frac{dx_i}{dt}$$

no puede extrapolarse de manera inmediata para construir un cuadrivector de Lorentz ya que la cantidad  $v_\mu = dx_\mu/dt$  es el producto de un cuadrivector ( $dx_\mu$ ) con una cantidad que no es escalar ( $dt$  no es invariante de Lorentz) de modo que el resultado no es un cuadrivector. El invariante más obvio asociado a  $dt$  es el diferencial de tiempo propio  $d\tau$ , de modo que resulta natural definir la cuadrivelocidad  $u_\nu$  como la variación por unidad de tiempo del vector de posición de una partícula (medida en un sistema  $S$ ) con respecto al tiempo propio de dicha partícula (invariante)

$$u_\nu = \frac{dx_\nu}{d\tau} = \frac{dx_\nu}{dt\sqrt{1-\beta^2}} \quad (16.54)$$

cuyas componentes espacial y temporal son

$$u_i = \frac{dx_i}{dt\sqrt{1-\beta^2}} = \frac{v_i}{\sqrt{1-\beta^2}} \quad ; \quad u_4 = \frac{dx_4}{dt\sqrt{1-\beta^2}} = \frac{ic}{\sqrt{1-\beta^2}} \quad (16.55)$$

el módulo al cuadrado de la cuadrivelocidad (o velocidad de universo), es invariante de Lorentz

$$u_\nu u_\nu = \frac{v^2}{1-\beta^2} - \frac{c^2}{1-\beta^2} = \frac{v^2 - c^2}{1-(v^2/c^2)} = -c^2 \quad (16.56)$$

y es además del género temporal. Por supuesto, la cuadrivelocidad no tiene un significado Físico directo ya que para medir  $dx_\nu$  y  $d\tau$  se están usando en general sistemas de referencia diferentes. Sin embargo, la Ec. (16.55) nos muestra que la cuadrivelocidad contiene toda la información sobre la velocidad Física y tiene la ventaja de que si escribimos las expresiones en términos de la cuadrivelocidad, será más fácil chequear la covarianza de las ecuaciones gracias a la naturaleza cuadrivectorial de  $u_\nu$ .

Otro cuadrivector de enorme importancia es el cuadrivector  $j_\mu$  formado con la corriente eléctrica  $\mathbf{j}$  unida con la cantidad  $ic\rho$  siendo  $\rho$  la densidad de corriente eléctrica. Para obtener esta forma cuadrivectorial comenzamos con la ecuación de continuidad

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \quad (16.57)$$

que me expresa la conservación de la carga<sup>16</sup>. Si asumimos que la conservación de la carga es válida en todos los sistemas de referencia inerciales, entonces esta ecuación debe conservar su forma ante una transformación de

<sup>16</sup>La ecuación de continuidad (16.57) puede extenderse para cualquier "carga generalizada" que se conserve. Por ejemplo, la energía o la masa pueden ser también cargas generalizadas, de modo que el desarrollo presente se puede extender a estas cantidades. Para más detalles ver sección 18.1, Pág. 523.

Lorentz. Dado que  $\mathbf{j}$  está asociado en la ecuación de continuidad a derivadas en el espacio es razonable pensar que haga parte de las componentes espaciales de un cuadrivector, similarmente dado que  $\rho$  está asociado a una derivada temporal resulta razonable pensar que hace parte de la componente temporal del cuadrivector. Para escribir esta ecuación en forma manifestamente covariante escribámosla en componentes

$$\begin{aligned} \frac{\partial j_k}{\partial x_k} + \frac{\partial \rho}{\partial t} = 0 &\Rightarrow \frac{\partial j_k}{\partial x_k} + \frac{\partial (ic\rho)}{\partial (ict)} = 0 \Rightarrow \frac{\partial j_\mu}{\partial x_\mu} = 0 \\ &\Rightarrow \partial_\mu j_\mu = 0 \quad ; \quad j_\mu \equiv (j_1, j_2, j_3, ic\rho) \end{aligned} \quad (16.58)$$

en primer lugar, es fácil ver que el operador cuadrigradiante se transforma en el espacio de Minkowski como un cuadrivector<sup>17</sup>

$$\frac{\partial}{\partial x'_\nu} = \frac{\partial x_\mu}{\partial x'_\nu} \frac{\partial}{\partial x_\mu} = L'^{\mu\nu} \frac{\partial}{\partial x_\mu} = L'^{-1}_{\mu\nu} \frac{\partial}{\partial x_\mu} = L_{\nu\mu} \frac{\partial}{\partial x_\mu}$$

donde hemos usado la ortogonalidad de  $\mathbf{L}$ . Dado que  $\partial_\mu$  es un cuadrivector, se tiene que  $j_\mu$  también debe serlo si la ecuación de continuidad ha de ser covariante, es decir si la carga se ha de conservar en todos los sistemas inerciales. Para ver el carácter cuadrivectorial de  $j_\mu$  basta con demostrar que  $j_\mu = \rho_0 u_\mu$  siendo  $\rho_0$  la densidad de carga en el sistema en el cual las cargas están en reposo<sup>18</sup>, es decir es la densidad de carga propia. Se deja esta demostración como ejercicio al lector.

Vemos pues que la cantidad  $\partial_\mu j_\mu$  es invariante ante una transformación de Lorentz (escalar de universo) ya que es la contracción de dos cuadrivectores. Este ejemplo nos muestra una forma de escribir una ley Física en una forma manifestamente covariante.

Veamos otro ejemplo de cuadrivector muy importante en la Física. Es bien conocido de la teoría clásica electromagnética que los potenciales escalar y vectorial obedecen ecuaciones de onda desacopladas

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{c} \mathbf{j} \quad ; \quad \nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -4\pi\rho \quad (16.59)$$

siempre y cuando se imponga la condición de Lorentz.

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} = 0 \quad (16.60)$$

Nótese que la condición de Lorentz es semejante en estructura a la ecuación de continuidad, por ello usando un argumento similar al usado para la ecuación de continuidad es natural pensar que  $\mathbf{A}$  está asociado a las componentes espaciales de un cuadrivector y  $\phi$  a la componente temporal. Esta asociación parece estar reforzada por las Ecs. (16.59) donde  $\mathbf{A}$  tiene como fuente a  $\mathbf{j}$  (que a su vez forma parte de la componente espacial del cuadrivector  $j_\mu$ ) en tanto que  $\phi$  tiene como fuente a  $\rho$  (donde este último es parte de la componente temporal de  $j_\mu$ ). Comencemos por la condición gauge Ec. (16.60) que se puede reescribir como

$$\partial_i A_i + \frac{\partial (i\phi)}{\partial (ict)} = 0 \Rightarrow \partial_\mu \equiv \left( \nabla, \frac{\partial}{\partial x_4} \right) = \left( \nabla, \frac{\partial}{\partial ict} \right) \quad ; \quad A_\mu \equiv (\mathbf{A}, i\phi) \quad (16.61)$$

$$\Rightarrow \partial_\mu A_\mu = 0 \quad (16.62)$$

las Ecs. (16.59) se pueden reescribir en la forma

$$\nabla^2 \mathbf{A} + \frac{\partial^2 \mathbf{A}}{\partial (ict)^2} = -\frac{4\pi}{c} \mathbf{j} \quad ; \quad \nabla^2 i\phi - \frac{1}{c^2} \frac{\partial^2 i\phi}{\partial t^2} = -\frac{4\pi}{c} ic\rho \quad (16.63)$$

<sup>17</sup>Recordemos que en la formulación de espacios de Riemann, este operador se transforma covariantemente y la ecuación (16.58) es el producto escalar de un vector covariante con un contravariante, esto se denota como  $\partial_\mu j^\mu = 0$ . En general los invariantes en el espacio de Riemann son combinaciones de tensores covariantes con tensores contravariantes, de modo que deben escribirse con índice arriba contraído con índice abajo e.g.  $j^\mu k_\mu$ ,  $k^{\mu\nu} p_{\mu\nu}$ .

<sup>18</sup>Debe tenerse en cuenta que la densidad de corriente y de carga son cantidades locales, es decir están definidos en cada punto del espacio-tiempo. Por lo tanto, está bien definido el sistema de referencia en el cual la carga en un punto  $(\mathbf{r}, \mathbf{t})$  está en reposo instantáneo. Es decir que en este sistema de referencia no hay densidad de corriente en dicho punto.

definimos el operador de D'Alembert en la forma

$$\begin{aligned}\square &\equiv \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} = \nabla^2 + \frac{\partial}{\partial (ict)^2} = \partial_i \partial_i + \partial_4 \partial_4 \\ \square &\equiv \partial_\mu \partial_\mu = \frac{\partial^2}{\partial x_\mu \partial x_\mu}\end{aligned}\quad (16.64)$$

las Ecs. (16.63) en el espacio de Minkowski quedan

$$\square \mathbf{A} = -\frac{4\pi}{c} \mathbf{j} \quad ; \quad \square (i\phi) = -\frac{4\pi}{c} ic\rho$$

que se puede condensar en una sola ecuación cuadvivectorial en la forma

$$\square A_\mu = -\frac{4\pi}{c} j_\mu \quad ; \quad j_\mu \equiv (\mathbf{j}, ic\rho) \quad , \quad A_\mu \equiv (\mathbf{A}, i\phi) \quad (16.65)$$

Una vez que se demuestre que la cuadrupla  $A_\mu$  definida en (16.61) es un cuadvivector de Lorentz, se sigue que las Ecs. (16.62) y (16.65) están escritas de manera manifiestamente covariante. Se deja esta demostración como ejercicio al lector. En (16.62) ambos miembros son escalares de universo, en tanto que en (16.65) ambos miembros son vectores de universo, pues el operador de D'Alembert es un escalar de universo. Estas ecuaciones demuestran que la teoría electromagnética de Maxwell es covariante con respecto a las transformaciones de Lorentz de modo que está descrita por la relatividad especial y no por la relatividad de Galileo. El lector puede verificar que el uso del gauge de Coulomb  $\nabla \cdot \mathbf{A} = 0$  hace mucho más difícil el proceso de colocar las ecuaciones de manera manifiestamente covariante.

## 16.6. Fuerza y momento en relatividad especial

Las leyes de Newton son invariantes de Galileo y por tanto no son invariantes de Lorentz. En consecuencia, es necesario encontrar una generalización adecuada de fuerza cuya ley fundamental satisfaga los requisitos de covarianza ante transformaciones de Lorentz. Naturalmente, debemos también exigir que las ecuaciones relativistas se reduzcan a la ecuación dinámica fundamental de Newton en el límite  $\beta \rightarrow 0$

$$\frac{d}{dt} (mv_i) = F_i \quad (16.66)$$

es fácil ver que las componentes espaciales de un cuadvivector forman un vector espacial, ya que las transformaciones de Lorentz contienen a las rotaciones espaciales ( $L_{4i} = L_{i4} = 0$  y  $L_{44} = 1$ ). No obstante, el recíproco no es cierto, las componentes de un vector espacial no se transforman necesariamente como lo harían las componentes espaciales de un cuadvivector. Por ejemplo, se puede multiplicar a las componentes del trivector por una función cualquiera de  $\beta$  y sus propiedades de rotación no se alteran. En cambio, esta multiplicación sí alteraría las propiedades de transformación de las tres componentes espaciales de un cuadvivector ante una transformación de Lorentz. En concordancia con esto, las componentes espaciales de la cuadvivocidad  $u_\nu$  forman un vector espacial  $\mathbf{v}/\sqrt{1-\beta^2}$ . Sin embargo, el  $\mathbf{v}$  no hace parte de ningún cuadvivector, para que lo sea debe dividirse por  $\sqrt{1-\beta^2}$ .

Primero buscaremos una generalización cuadvivectorial del miembro izquierdo en (16.66), es claro que la cuadvivocidad definida en (16.55) posee una parte espacial que se reduce a  $\mathbf{v}$  cuando  $\beta \rightarrow 0$ . Tomaremos a  $m$  como un invariante que lo llamaremos la masa en reposo o masa propia de la partícula. En cuanto al tiempo  $t$ , este no es un invariante relativista pero sabemos que el tiempo propio  $\tau$  sí es un invariante que además se reduce a  $t$  cuando  $\beta \rightarrow 0$ . Los argumentos anteriores sugieren que la generalización de la ley de Newton (16.66) para una partícula tenga la forma

$$\frac{d}{d\tau} (mu_\nu) = K_\nu \quad (16.67)$$

donde  $K_\nu$  debe ser un cuadvivector llamado **fuerza de Minkowski**.

Nótese que en general las componentes espaciales de  $K_\nu$  no tienen que coincidir con las componentes de la fuerza, salvo por supuesto en el límite no relativista con  $\beta \rightarrow 0$ . Podemos pensar por ejemplo que  $K_i$  se puede construir como el producto de  $F_i$  con cierta función  $h(\beta)$  que se reduzca a la unidad en el límite no relativista. Para conocer la forma de  $h(\beta)$  debemos conocer el comportamiento de la fuerza ante una transformación de Lorentz. Utilizaremos dos procedimientos.

En el primer procedimiento, tendremos en cuenta que las fuerzas fundamentales son solo cuatro: las interacciones gravitacional, electromagnética, nuclear débil y nuclear fuerte. La idea sería expresar las leyes que gobiernan a estas interacciones de manera covariante. No obstante, no se conoce teorías covariantes para las fuerzas nucleares, entre otras cosas porque tales interacciones no se pueden modelar clásicamente en forma satisfactoria (en la teoría cuántica la fuerza pierde su significado y es reemplazada por la energía potencial). Sin embargo, en el caso electromagnético clásico es de esperarse que podamos construir una expresión de la fuerza que nos proporcione una ecuación covariante, después de todo la teoría especial de la relatividad fué construída justamente para que las ecuaciones de Maxwell fueran covariantes de Lorentz. Afortunadamente, esta construcción será suficiente ya que las propiedades de transformación de las fuerzas deben ser las mismas independientemente de su origen. Por ejemplo, el hecho de que una partícula esté en equilibrio (suma de fuerzas cero) debe ser independiente del sistema de referencia inercial utilizado y esto solo es posible si las fuerzas transforman todas igual, incluso si cada una es de diferente naturaleza.

Vimos que a partir de la expresión para la fuerza de Lorentz escrita en términos de potenciales en lugar de campos, la fuerza electromagnética que se ejerce sobre una partícula cargada viene dada por

$$F_i = -q \left[ \frac{\partial}{\partial x_i} \left( \phi - \frac{1}{c} \mathbf{v} \cdot \mathbf{A} \right) + \frac{1}{c} \frac{dA_i}{dt} \right]$$

recordando la definición del cuadripotencial (16.61), y de la cuadrivelocidad (16.55) podemos escribir la expresión  $\phi - (1/c) \mathbf{v} \cdot \mathbf{A}$  en forma covariante

$$\begin{aligned} \phi - \frac{1}{c} \mathbf{v} \cdot \mathbf{A} &= i\phi \left( -\frac{1}{c} \sqrt{1-\beta^2} \frac{ic}{\sqrt{1-\beta^2}} \right) - \frac{1}{c} v_i A_i = A_4 \left( -\frac{1}{c} \sqrt{1-\beta^2} u_4 \right) - \frac{\sqrt{1-\beta^2}}{c} \frac{v_i}{\sqrt{1-\beta^2}} A_i \\ \phi - \frac{1}{c} \mathbf{v} \cdot \mathbf{A} &= -\frac{1}{c} \sqrt{1-\beta^2} A_4 u_4 - \frac{1}{c} \sqrt{1-\beta^2} u_i A_i = -\frac{1}{c} \sqrt{1-\beta^2} [A_4 u_4 + u_i A_i] \\ \phi - \frac{1}{c} \mathbf{v} \cdot \mathbf{A} &= -\frac{1}{c} \sqrt{1-\beta^2} u_\nu A_\nu \end{aligned}$$

y las componentes  $F_i$  de las fuerzas son

$$\begin{aligned} F_i &= -q \left[ \frac{\partial}{\partial x_i} \left( -\frac{1}{c} \sqrt{1-\beta^2} u_\nu A_\nu \right) + \frac{1}{c} \sqrt{1-\beta^2} \frac{dA_i}{\sqrt{1-\beta^2} dt} \right] \\ F_i &= \sqrt{1-\beta^2} \left[ \frac{q}{c} \left( \frac{\partial}{\partial x_i} (u_\nu A_\nu) - \frac{dA_i}{d\tau} \right) \right] \end{aligned} \quad (16.68)$$

puesto que  $q/c$  es un escalar de Lorentz, una extensión cuadrivectorial del término entre paréntesis cuadrados es de la forma

$$\frac{q}{c} \left[ \frac{\partial}{\partial x_\mu} (u_\nu A_\nu) - \frac{dA_\mu}{d\tau} \right]$$

este término es claramente un cuadrivector, pues el primer término dentro del paréntesis es la derivada  $\partial_\mu$  (operador cuadrivectorial) de un escalar de universo ( $u_\nu A_\nu$ ), el segundo término es el producto de un cuadrivector  $dA_\mu$  por un escalar de universo  $(d\tau)^{-1}$ . En consecuencia, la expresión en paréntesis cuadrados en (16.68) está asociada a las componentes espaciales de un cuadrivector. Por tanto,  $F_i$  es el producto de  $\sqrt{1-\beta^2}$  por la componente espacial de un cuadrivector, el cual identificamos como la fuerza de Minkowski  $K_\nu$ . Por tanto la relación entre la fuerza ordinaria y la de Minkowski está dada por

$$F_i = K_i \sqrt{1-\beta^2} \quad (16.69)$$

esta relación debe ser general e independiente del origen de las fuerzas. Para el caso de partículas cargadas sometidas a un campo electromagnético, la fuerza de Minkowski se obtiene de la extrapolación de la expresión (16.68)

$$K_\mu = \frac{q}{c} \left[ \frac{\partial}{\partial x_\mu} (u_\nu A_\nu) - \frac{dA_\mu}{d\tau} \right] \quad (16.70)$$

En un segundo procedimiento, se **define** la fuerza como la variación del momento lineal por unidad de tiempo, en todos los sistemas de Lorentz se tiene entonces que

$$F_i = \frac{dp_i}{dt} \quad (16.71)$$

pero para ello será necesario redefinir el momento lineal  $p_i$  de modo que en el límite no relativista se reduzca a  $mv_i$ . Podemos hallar la forma que toma el momento y el significado de  $K_\mu$  haciendo que la Ec. (16.67) se parezca en lo posible a (16.71). A partir de la relación entre  $\tau$  y  $t$  y de la definición de cuadrivelocidad, podemos escribir las componentes espaciales de (16.67) en la forma

$$\frac{d}{dt} \left( \frac{mv_i}{\sqrt{1-\beta^2}} \right) = K_i \sqrt{1-\beta^2} \quad (16.72)$$

y comparando (16.72) con (16.71) vemos que el teorema de conservación del momento lineal (reemplazante más general que la tercera ley de Newton) será invariante de Lorentz si definimos la cantidad de movimiento en la forma

$$p_i = \frac{mv_i}{\sqrt{1-\beta^2}} \quad (16.73)$$

y que  $F_i$  y  $K_i$  estén relacionadas como lo indica la ecuación (16.69). Nótese que la ecuación (16.73) se reduce a  $mv_i$  cuando  $\beta \rightarrow 0$  como se esperaba. Los dos procedimientos conducen entonces a los mismos resultados. Comparando (16.73) con la definición (16.55) de la cuadrivelocidad vemos que  $p_i$  es la parte espacial del llamado cuadrivector momento energía

$$p_\nu \equiv mu_\nu \quad (16.74)$$

la ecuación de movimiento generalizada para una partícula se escribe entonces

$$\frac{dp_\nu}{d\tau} = K_\nu \quad (16.75)$$

## 16.7. Energía y relación momento-energía en relatividad especial

Hasta ahora solo hemos estudiado la parte espacial de las ecuaciones cuadrivectoriales (16.67, 16.75). Para obtener información de la parte temporal hagamos el producto interno de (16.67) por la cuadrivelocidad

$$u_\nu \frac{d}{d\tau} (mu_\nu) = \frac{d}{d\tau} \left( \frac{m}{2} u_\nu u_\nu \right) = K_\nu u_\nu$$

de (16.56) vemos que  $u_\nu u_\nu = -c^2$  y como  $m$  es también constante, vemos que la expresión de la mitad se anula. Luego

$$K_\nu u_\nu \equiv K_i u_i + K_4 u_4 = 0$$

y usando las Ecs. (16.55, 16.69) tenemos

$$K_\nu u_\nu \equiv \frac{F_i}{\sqrt{1-\beta^2}} \frac{v_i}{\sqrt{1-\beta^2}} + K_4 \left( \frac{ic}{\sqrt{1-\beta^2}} \right) = \frac{\mathbf{F} \cdot \mathbf{v}}{1-\beta^2} + \frac{icK_4}{\sqrt{1-\beta^2}} = 0$$

de modo que la cuarta componente de la fuerza de Minkowski será

$$K_4 = \frac{i}{c} \frac{\mathbf{F} \cdot \mathbf{v}}{\sqrt{1-\beta^2}} \quad (16.76)$$



la componente temporal de la Ec. (16.67) se obtiene empleando (16.55) y (16.76)

$$\frac{d}{d\tau}(mu_4) = K_4 \Rightarrow \frac{1}{\sqrt{1-\beta^2}} \frac{d}{dt} \left( m \frac{ic}{\sqrt{1-\beta^2}} \right) = \frac{i}{c} \frac{\mathbf{F} \cdot \mathbf{v}}{\sqrt{1-\beta^2}}$$

de lo cual se obtiene

$$\frac{d}{dt} \left( \frac{mc^2}{\sqrt{1-\beta^2}} \right) = \mathbf{F} \cdot \mathbf{v} \quad (16.77)$$

recordemos ahora el escenario no relativista. En este escenario  $\mathbf{F} \cdot \mathbf{v}$  corresponde al trabajo por unidad de tiempo que se hace sobre la partícula  $dW/dt$ . Teniendo en cuenta además el teorema fundamental del trabajo y la energía resulta  $dW = dT$  siendo  $T$  la energía cinética. De esto se concluye que

$$\mathbf{F} \cdot \mathbf{v} = \frac{dW}{dt} = \frac{dT}{dt} \quad (\text{límite no relativista})$$

Extrapolando esta definición al caso relativista tenemos que

$$\frac{dT}{dt} = \mathbf{F} \cdot \mathbf{v} \quad (\text{escenario relativista}) \quad (16.78)$$

Comparando (16.77) con (16.78) se obtiene la generalización relativista de la energía cinética

$$T = \frac{mc^2}{\sqrt{1-\beta^2}} \quad (16.79)$$

en el límite  $\beta^2 \ll 1$  esta ecuación se expande como

$$T \cong mc^2 \left[ 1 + \frac{\beta^2}{2} \right] = mc^2 + \frac{1}{2}mv^2$$

Este valor no coincide con el límite no relativista esperado. Sin embargo, el valor adicional pareciera a priori ser irrelevante ya que se puede adicionar una constante a la derecha de (16.79) que no afectaría la validez de la Ec. (16.77). Sin embargo, es preferible mantener este valor y conservar la cantidad  $T$  en la forma dada por (16.79), hay dos buenas razones para mantener esta cantidad: (a) En algunos casos como veremos más adelante  $mc^2$  puede cambiar gracias al cambio en la masa en reposo de las partículas, por ejemplo en colisiones inelásticas. Esto nos indica que este tipo de energía se puede intercambiar o transferir y por tanto es físicamente relevante y (b) al comparar (16.74, 16.55) con (16.79) vemos que  $iT/c$  es la cuarta componente del cuadrivector momento energía

$$p_\nu = (p_1, p_2, p_3, p_4) \quad ; \quad p_i = mu_i = \frac{mv_i}{\sqrt{1-\beta^2}} \quad ; \quad p_4 = mu_4 = \frac{imc}{\sqrt{1-\beta^2}} = \frac{iT}{c} \quad (16.80)$$

Sin embargo, para usar una terminología consistente con la newtoniana es preferible definir la energía cinética como la parte de esta energía que se reduce correctamente al valor no relativista

$$K \equiv T - mc^2 = mc^2(\gamma - 1)$$

no existe una única designación para  $T$ . En ocasiones se le llama energía total (si bien esto solo sería apropiado para partícula libre) y en otras simplemente energía. En todo caso  $T$  posee propiedades interesantes. Por ejemplo se puede demostrar que la  $T$  dada por (16.79) se conserva siempre que se conserve el momento lineal espacial definido en (16.73). Para verificar este teorema, podemos tener en cuenta que la conservación del momento espacial debe ser invariante ante una transformación de Lorentz, en realidad esta invarianza está implícita en la definición de sistema inercial dada por Einstein. Las componentes transformadas  $p'_j$  serán funciones lineales de las  $p_i$  pero también de  $p_4$  i.e. de la energía  $T$ . En consecuencia, la conservación de  $p'_j$  para todos los sistemas inerciales exige la conservación conjunta de todas las componentes de  $p_\nu$ . Es fácil calcular el valor del invariante  $p_\nu p_\nu$

$$p_\nu p_\nu = (mu_\nu)(mu_\nu) = m^2 u_\nu u_\nu = -m^2 c^2 \quad (16.81)$$

por otro lado

$$p_\nu p_\nu = \left( \mathbf{p}, \frac{iT}{c} \right) \cdot \left( \mathbf{p}, \frac{iT}{c} \right) = \mathbf{p}^2 - \frac{T^2}{c^2} \quad (16.82)$$

siendo  $\mathbf{p}$  el trimomento espacial definido en (16.73). Combinando (16.81) y (16.82) se obtiene la relación cinemática fundamental para la relatividad especial

$$T^2 = \mathbf{p}^2 c^2 + m^2 c^4 \quad (16.83)$$

la Ec. (16.83) es la análoga a la relación no relativista  $T = mv^2/2$  con la diferencia de que en relatividad  $T$  incluye la energía en reposo  $mc^2$ . La definición de  $T$  Ec. (16.79) muestra que la energía de una partícula con energía en reposo finita tiende a infinito cuando  $v \rightarrow c$ , es decir que se requiere una energía infinita para llevar una partícula material desde el reposo hasta una velocidad  $c$ . Por tanto la teoría predice que no es posible alcanzar o superar la velocidad de la luz en el vacío partiendo de una velocidad menor que  $c$ .

El enunciado anterior no prohíbe la existencia de partículas que **nazcan** con velocidades mayores a las de la luz (taquiones). De acuerdo con la Ec. (16.79) la masa asociada a esta partícula tendría que ser imaginaria para tener una energía real. Esto implica que un taquión está descrito por un parámetro real  $m'$  de modo que  $T = m'c^2/\sqrt{\beta^2 - 1}$ . Las soluciones taquiónicas y en particular sus implicaciones sobre la causalidad han sido motivo de una amplia especulación científica. No obstante, no se han observado partículas taquiónicas hasta el momento.

Ya hemos dicho que en relatividad especial la conservación del trimomento conduce a la conservación de la cuarta componente del cuadrivector momento energía. Esta situación contrasta con la mecánica no relativista, en la cual la conservación del momento lineal y la conservación de la energía cinética son aspectos independientes. Por ejemplo, en un choque inelástico entre dos partículas se conserva el momento lineal pero no la energía cinética, esto se debe a cambios en la energía interna del sistema debido a reconfiguraciones internas. En el caso relativista, la energía  $T$  debe conservarse junto con el momento espacial incluso en choques inelásticos debido a sus propiedades de transformación como cuadrivector. Físicamente, esto se entiende teniendo en cuenta que  $T$  posee dos términos, la energía cinética y la energía en reposo  $mc^2$ . En un choque inelástico la energía cinética relativista cambia también en virtud de reconfiguraciones internas del sistema, pero en este caso estos cambios en la energía interna se traducen en cambios en la energía en reposo y por tanto de la masa en reposo.

Veamos un ejemplo sencillo del hecho de que la conservación de  $T$  implica un cambio en la masa en reposo en colisiones inelásticas. Sean dos partículas idénticas en masa que viajan con respecto al laboratorio a velocidades iguales y opuestas. El momento lineal total es nulo visto por el laboratorio y el cuadrivector momento energía viene dado por

$$P_\mu = p_\mu(1) + p_\mu(2) = \left( \mathbf{p}, \frac{iT}{c} \right) + \left( -\mathbf{p}, \frac{iT}{c} \right) = \left( \mathbf{0}, \frac{2iT}{c} \right) \quad (16.84)$$

siendo  $T$  la energía de cada partícula definida por la Ec. (16.79). Supongamos que el choque es totalmente inelástico de modo que las dos masas quedan unidas (como dos bolas de plastilina). Puesto que en el estado final tenemos una sola partícula y su trimomento es nulo, la partícula resultante debe estar en reposo con respecto al laboratorio. La energía total del estado final  $T_f$  es la energía en reposo del sistema compuesto final, la cual aplicando (16.83) resulta

$$T_f = Mc^2 \quad (16.85)$$

siendo  $M$  la masa del estado final. Es conveniente parametrizar la masa del estado final en la forma

$$M = 2m + \Delta M \quad (16.86)$$

donde  $\Delta M$  nos da información sobre la posible pérdida o ganancia de masa del estado final (de masa  $M$ ) con respecto a la masa del sistema inicial (de masa  $2m$ ). La conservación de  $P_4$  en el choque manifestada en la Ec. (16.84) muestra que

$$T_f = 2T \quad (16.87)$$

la cual combinada con (16.85) nos da

$$2T = Mc^2 \quad (16.88)$$

es claro que se ha perdido energía cinética en el proceso, la energía cinética al principio y al final del proceso están dadas por

$$K_0 = 2T - 2mc^2 \quad ; \quad K_f = 0$$

y usando (16.87), (16.86) y (16.88)

$$\begin{aligned} K_0 - K_f &= 2T - 2mc^2 = Mc^2 - 2mc^2 = (2m + \Delta M)c^2 - 2mc^2 \\ K_0 - K_f &= (\Delta M)c^2 \end{aligned}$$

por lo tanto, el choque inelástico ha convertido toda la energía del movimiento inicial vista por el laboratorio en un incremento en la masa en reposo del sistema. En esta clase de choque inelástico se suele decir que la energía cinética perdida en el choque se convierte en calor (para ver el tratamiento no relativista del choque perfectamente inelástico ver sección 11.1.1, Pág. 263). La relatividad restringida nos dice que la masa en reposo o inercia del sistema aumenta en proporción al calor que se produce. Este incremento de masa se podría detectar poniendo al sistema en movimiento a través de una fuerza conocida, no obstante para sistemas macroscópicos estos cambios de masa son muy difíciles de detectar ya que un joule de energía posee un equivalente de masa de aproximadamente  $1,1 \times 10^{-17} \text{ Kg}$ . No es de extrañarse entonces que las evidencias sobre los cambios de la masa en reposo se hayan visto en sistemas de escala atómica, nuclear o subnuclear. En estos casos no podemos hablar de producción de calor sino de cambios en la energía interna del sistema. A la escala subnuclear, estos cambios en la energía en reposo suelen ser suficientes para permitir la creación de una o más partículas adicionales. Es de anotar además que estos cambios también pueden ocurrir en el sentido opuesto: la energía en reposo se puede convertir en energía en movimiento, fenómeno particularmente visible en las explosiones nucleares, por supuesto en estas explosiones el valor de  $T$  permanece constante durante la explosión. A pesar de la enorme energía liberada en estas explosiones, la pérdida de masa suele ser del orden del 0,1 % de la masa original.

## 16.8. Formulación Lagrangiana de la mecánica relativista

Dado que hemos realizado la generalización adecuada de las leyes de Newton en el marco de la relatividad restringida, podemos intentar establecer una formulación Lagrangiana para la dinámica relativista. Vamos a describir dos procedimientos. En el primero, se pretende reproducir en un sistema inercial particular las ecuaciones de la forma (16.71), resultando en general una formulación que no es manifiestamente covariante. Las fuerzas  $F_i$  podrían estar o no adecuadamente relacionadas con una fuerza de Minkowski  $K_\nu$ . El otro método consiste en construir un principio de Hamilton covariante a partir del cual se llega a las ecuaciones de Lagrange en las cuales el espacio y el tiempo se tratan como coordenadas de un espacio de configuración de cuatro dimensiones. El primer método puede presentar problemas cuando las fuerzas no se formulan bien relativísticamente, pero generalmente las ecuaciones de movimiento así obtenidas son correctas relativísticamente en un sistema inercial dado, aunque no sean manifiestamente covariantes. El segundo método aunque formalmente más correcto, desemboca en grandes dificultades incluso para sistemas de una partícula. Para sistemas de varias partículas falla casi desde el principio y no hay una formulación covariante satisfactoria para la mecánica clásica relativista de muchas partículas. Este sigue siendo un tema activo de investigación.

## 16.9. Formulación no manifiestamente covariante

La idea es encontrar un Lagrangiano que nos lleve a las ecuaciones de movimiento relativistas que constituyeron nuestra generalización de las Leyes de Newton. Estas ecuaciones serán obtenidas en función de las coordenadas de un determinado sistema inercial. Nótese que el principio de D'Alembert resulta poco fructífero en este caso, ya que aunque tal principio sigue siendo válido, las deducciones básicas se basaban en el hecho de que  $p_i = m_i v_i$  relación que ya no es válida en relatividad restringida. Por lo tanto, elegiremos el camino de tomar como punto de partida el principio de Hamilton

$$\delta I = \delta \int_{t_1}^{t_2} L dt = 0 \quad (16.89)$$

y obtener con base en las ecuaciones de Euler Lagrange ecuaciones de movimiento que concuerden con las generalizaciones obtenidas para el formalismo Newtoniano Ec. (16.71). Estudiaremos el caso de una partícula sometida a fuerzas conservativas que no dependen de la velocidad, en cuyo caso escribimos

$$L = -mc^2\sqrt{1-\beta^2} - V \quad (16.90)$$

siendo  $V$  un potencial que solo depende de la posición y  $\beta^2 = v^2/c^2$  donde  $v$  es la velocidad de la partícula en el sistema inercial particular que se toma. Veamos que este Lagrangiano nos conduce a las ecuaciones correctas, partiendo de las ecuaciones de Lagrange

$$\frac{d}{dt} \left( \frac{\partial L}{\partial v_i} \right) - \frac{\partial L}{\partial x_i} = 0$$

y teniendo en cuenta la relación

$$\frac{\partial L}{\partial v_i} = \frac{\partial}{\partial v_i} \left( -mc^2\sqrt{1-\frac{v_k v_k}{c^2}} \right) = -\frac{mc^2}{2\sqrt{1-\frac{v_k v_k}{c^2}}} \left( -2\frac{v_k}{c^2} \delta_{ki} \right) = \frac{mv_i}{\sqrt{1-\beta^2}} = p_i \quad (16.91)$$

se obtiene

$$\frac{dp_i}{dt} + \frac{\partial V}{\partial x_i} = 0 \quad \Rightarrow \quad \frac{dp_i}{dt} = -\frac{\partial V}{\partial x_i} \quad \Rightarrow \quad \frac{dp_i}{dt} = F_i$$

que concuerda con (16.71). Nótese que el lagrangiano NO es de la forma  $L = T - V$ . No obstante, la expresión  $\partial L/\partial v_i$  sigue siendo el momento lineal como ocurre en la formulación no relativista con coordenadas cartesianas y potencial independiente de las velocidades. En realidad es este hecho lo que garantiza la corrección adecuada de las ecuaciones de Lagrange. Por tanto, hubiéramos podido proceder hacia atrás desde (16.91) para obtener al menos la dependencia de la velocidad del Lagrangiano.

La generalización de (16.90) a sistemas de muchas partículas o a sistemas de coordenadas generalizadas  $q_j$  es directa. Las cantidades de movimiento canónicas siguen definiéndose en la forma

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \quad (16.92)$$

de modo que se mantiene la relación entre coordenadas cíclicas y la conservación de los momentos asociados a ellas. Adicionalmente, si el Lagrangiano no depende explícitamente del tiempo se sigue manteniendo a la función  $h$  como constante de movimiento

$$h = \dot{q}_j p_j - L \quad (16.93)$$

hay sin embargo, una diferencia importante con el caso no relativista: debido al factor  $\sqrt{1-\beta^2}$  en el Lagrangiano (16.90), dicho Lagrangiano no es una función homogénea de la velocidad, de modo que la demostración realizada en el caso no relativista para llegar a que  $h$  es la energía del sistema (en el caso de potenciales dependientes de la posición y coordenadas que no dependen explícitamente del tiempo) no es válida en el caso relativista. Veremos sin embargo que para potenciales que solo dependen de la posición,  $h$  continúa siendo la energía total del sistema

$$\begin{aligned} h &= \dot{x}_i p_i - L = \frac{mv_i v_i}{\sqrt{1-\beta^2}} + mc^2\sqrt{1-\beta^2} + V \\ h &= \frac{\sqrt{1-\beta^2}}{m} \left( \frac{mv_i}{\sqrt{1-\beta^2}} \right) \left( \frac{mv_i}{\sqrt{1-\beta^2}} \right) + mc^2\sqrt{1-\beta^2} + V \\ h &= \sqrt{1-\beta^2} \left[ \frac{p_i p_i}{m} + mc^2 \right] + V \end{aligned} \quad (16.94)$$

por otro lado de la Ec. (16.83) vemos que

$$p_i p_i = \mathbf{p}^2 = \frac{T^2}{c^2} - m^2 c^2 \quad (16.95)$$

y reemplazando (16.95) en (16.94) resulta

$$h = \sqrt{1 - \beta^2} \left[ \frac{T^2}{mc^2} - mc^2 + mc^2 \right] + V = \frac{\sqrt{1 - \beta^2}}{mc^2} T^2 + V = \left( \frac{1}{T} \right) T^2 + V$$

la función energía queda entonces

$$h = \frac{mc^2}{\sqrt{1 - \beta^2}} + V = T + V = E \quad (16.96)$$

de modo que para el caso de potenciales dependientes solo de la posición,  $h$  se conserva y es la energía del sistema (naturalmente hemos usado coordenadas cartesianas de modo que la transformación de coordenadas a las coordenadas generalizadas es trivial).

La introducción de potenciales dependientes de la velocidad no conlleva ninguna dificultad particular y se puede efectuar en forma análoga al caso no relativista. De esta forma, para el caso de una partícula inmersa en un campo electromagnético, el lagrangiano se obtiene reemplazando  $V$  por el potencial de Lorentz en (16.90)

$$L = -mc^2 \sqrt{1 - \beta^2} - q\phi + \frac{q}{c} \mathbf{A} \cdot \mathbf{v} \quad (16.97)$$

puede verse que el momento canónico ya no es  $mu_i$ , pues en este caso hay términos adicionales debidos a la parte del potencial que depende de la velocidad

$$p_i = mu_i + \frac{q}{c} A_i \quad (16.98)$$

esta relación es análoga a la Ec. (5.2) obtenida para el caso no relativista. El Lagrangiano (16.97) no es manifiestamente covariante pues no está escrito en términos de tensores de Lorentz. Sin embargo, en este caso se espera que estos resultados sean válidos en cualquier sistema de referencia inercial en virtud de la covarianza relativista de la fuerza de Lorentz, de la cual proviene el potencial dependiente de la velocidad que se usa en (16.97).

De lo anterior se desprende que muchas de las estrategias y propiedades desarrolladas para la mecánica no relativista se pueden aplicar en un escenario relativista como veremos en los siguientes ejemplos

### 16.9.1. Movimiento relativista bajo una fuerza constante (hiperbólico)

Sin pérdida de generalidad se puede tomar el eje  $x_1$  a lo largo de la fuerza constante. El Lagrangiano tiene la forma

$$L = -mc^2 \sqrt{1 - \beta^2} + max \quad ; \quad \beta \equiv \frac{\dot{x}}{c}$$

Las ecuaciones de Lagrange quedan

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) &= \frac{d}{dt} \left[ \frac{\partial}{\partial \dot{x}} \left( -mc^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} \right) \right] = \frac{d}{dt} \left( \frac{m\dot{x}}{\sqrt{1 - \beta^2}} \right) = c \frac{d}{dt} \left( \frac{m\beta}{\sqrt{1 - \beta^2}} \right) \\ \frac{\partial L}{\partial x} &= ma \end{aligned}$$

y se obtiene

$$c \frac{d}{dt} \left( \frac{m\beta}{\sqrt{1 - \beta^2}} \right) - ma = 0 \quad ; \quad \frac{d}{dt} \left( \frac{\beta}{\sqrt{1 - \beta^2}} \right) = \frac{a}{c}$$

la primera integración se escribe como

$$\frac{\beta}{\sqrt{1 - \beta^2}} = \frac{at + \alpha}{c}$$

siendo  $\alpha$  una constante de integración. Podemos despejar  $\beta$  de esta ecuación elevando al cuadrado a ambos lados

$$\begin{aligned} \frac{\beta^2}{1-\beta^2} &= \frac{(at+\alpha)^2}{c^2} \Rightarrow \beta^2 = \frac{(at+\alpha)^2}{c^2} (1-\beta^2) \Rightarrow \beta^2 \left[ 1 + \frac{(at+\alpha)^2}{c^2} \right] = \frac{(at+\alpha)^2}{c^2} \\ \Rightarrow \beta^2 \left[ \frac{c^2 + (at+\alpha)^2}{c^2} \right] &= \frac{(at+\alpha)^2}{c^2} \Rightarrow \beta \left[ \sqrt{c^2 + (at+\alpha)^2} \right] = (at+\alpha) \\ \Rightarrow \beta &= \frac{at+\alpha}{\sqrt{c^2 + (at+\alpha)^2}} \end{aligned} \quad (16.99)$$

la Ec. (16.99) será útil para examinar el límite no relativista, por el momento continuamos manipulando la expresión

$$\Rightarrow \frac{\dot{x}}{c} = \frac{at+\alpha}{\sqrt{c^2 + (at+\alpha)^2}} \Rightarrow dx = c \frac{at+\alpha}{\sqrt{c^2 + (at+\alpha)^2}} dt \quad (16.100)$$

Una segunda integración con  $t'$  entre 0 y  $t$  y con  $x'$  entre  $x_0$  y  $x$  nos da

$$\int_{x_0}^x dx' = c \int_0^t \frac{(at'+\alpha) dt'}{\sqrt{c^2 + (at'+\alpha)^2}}$$

cuya solución es

$$x = x_0 + \frac{c}{a} \left[ \sqrt{c^2 + (at+\alpha)^2} - \sqrt{c^2 + \alpha^2} \right] \quad (16.101)$$

la velocidad se obtiene de (16.100)

$$\dot{x} = v = \frac{c(\alpha + at)}{\sqrt{c^2 + 2at\alpha + \alpha^2 + a^2t^2}} \quad (16.102)$$

y al evaluarla en  $t = 0$  tenemos

$$v_0 = \frac{\alpha c}{\sqrt{c^2 + \alpha^2}} \quad (16.103)$$

la Ec. (16.103) muestra que  $\alpha$  está directamente relacionado con la velocidad inicial. Si la partícula parte del reposo en el origen, las condiciones iniciales quedan  $x_0 = v_0 = \alpha = 0$ , y la Ec. (16.101) se puede escribir en la forma

$$\begin{aligned} x &= \frac{c}{a} \left[ \sqrt{c^2 + (at)^2} - c \right] \Rightarrow x + \frac{c^2}{a} = \frac{c}{a} \left[ \sqrt{c^2 + (at)^2} \right] \Rightarrow \left( x + \frac{c^2}{a} \right)^2 = \frac{c^2}{a^2} [c^2 + a^2t^2] \\ \Rightarrow \left( x + \frac{c^2}{a} \right)^2 - c^2t^2 &= \frac{c^4}{a^2} \end{aligned} \quad (16.104)$$

que describe la ecuación de una hipérbola en el plano  $x-t$ .

Veamos ahora como se recobra el límite no relativista. La Ec. (16.99) se puede reescribir como

$$\beta = \frac{1}{\sqrt{\left(\frac{c}{at+\alpha}\right)^2 + 1}}$$

y como el límite no relativista corresponde a  $\beta \rightarrow 0$ , se ve que esto es equivalente a la condición  $[c/(at+\alpha)]^2 \gg 1$ , o lo que es lo mismo

$$\frac{at+\alpha}{c} \ll 1 \quad \text{si} \quad \beta \ll 1 \quad (16.105)$$

adicionalmente, tomando  $\beta_0 \rightarrow 0$  en (16.103) se obtiene la relación no relativista entre  $\alpha$  y  $v_0$

$$\begin{aligned} v_0^2 &= \frac{c^2 \alpha^2}{c^2 + \alpha^2} \Rightarrow c^2 v_0^2 + \alpha^2 v_0^2 = \alpha^2 c^2 \Rightarrow v_0^2 + \alpha^2 \beta_0^2 = \alpha^2 \Rightarrow v_0^2 = \alpha^2 (1 - \beta_0^2) \\ &\Rightarrow v_0^2 \simeq \alpha^2 \quad \text{si } \beta_0 \ll 1. \end{aligned} \quad (16.106)$$

Reemplazando los límites no relativistas (16.105, 16.106) en (16.101) se obtiene

$$\begin{aligned} x &\simeq x_0 + \frac{c^2}{a} \left[ \sqrt{1 + \left(\frac{at + v_0}{c}\right)^2} - \sqrt{1 + \frac{v_0^2}{c^2}} \right] \simeq x_0 + \frac{c^2}{a} \left\{ \left[ 1 + \frac{1}{2} \left(\frac{at + v_0}{c}\right)^2 \right] - \left[ 1 + \frac{1}{2} \frac{v_0^2}{c^2} \right] \right\} \\ &= x_0 + \frac{c^2}{a} \left[ \frac{1}{2} \frac{(a^2 t^2 + 2atv_0)}{c^2} \right] \\ &\Rightarrow x \simeq x_0 + \frac{1}{2} at^2 + v_0 t \quad \text{si } \beta \ll 1 \end{aligned}$$

y se obtiene la parábola típica del movimiento no relativista.

El movimiento relativista bajo una fuerza constante, puede describir por ejemplo la trayectoria de electrones que se aceleran con un campo eléctrico constante y uniforme. Pues las velocidades típicas de los electrones son al menos del orden de la velocidad de la luz en el vacío.

### 16.9.2. Oscilador armónico unidimensional relativista

En este caso el Lagrangiano (16.90) toma la forma

$$L = -mc^2 \sqrt{1 - \beta^2} - \frac{1}{2} kx^2$$

dado que  $L$  no es función explícita del tiempo entonces la función energía y por tanto la energía total del sistema es una constante de movimiento. Si despejamos la velocidad en la Ec. (16.96) teniendo en cuenta que  $h$  es la energía, se tiene

$$\begin{aligned} E = \frac{mc^2}{\sqrt{1 - \beta^2}} + V &\Rightarrow (E - V)^2 = \frac{m^2 c^4}{1 - \beta^2} \Rightarrow 1 - \beta^2 = \frac{m^2 c^4}{(E - V)^2} \Rightarrow \beta^2 = 1 - \frac{m^2 c^4}{(E - V)^2} \\ \frac{1}{c^2} \left( \frac{dx}{dt} \right)^2 &= 1 - \frac{m^2 c^4}{(E - V)^2} \end{aligned} \quad (16.107)$$

podemos generalizar un poco antes de entrar en el potencial del oscilador armónico. Sea un potencial tal que  $V(x) = V(-x)$  y tal que  $V(0)$  es un mínimo local. Si la energía  $E$  está entre  $V(0)$  y el máximo de  $V$ , el movimiento será oscilatorio entre los límites  $x = \pm b$  donde  $b$  está determinado por

$$V(\pm b) = E$$

el periodo del movimiento oscilatorio se puede obtener a partir de (16.107)

$$\frac{1}{c} \left( \frac{dx}{dt} \right) = \sqrt{1 - \frac{m^2 c^4}{(E - V)^2}} \Rightarrow \frac{1}{c} \left( \frac{dx}{\sqrt{1 - \frac{m^2 c^4}{(E - V)^2}}} \right) = dt$$

un periodo consistirá en ir y volver desde  $-b$  hasta  $b$ . Por simetría esto se puede escribir como cuatro veces la integral entre 0 y  $b$

$$\tau = \frac{4}{c} \int_0^b \frac{dx}{\sqrt{1 - \frac{m^2 c^4}{[E - V(x)]^2}}} \quad (16.108)$$

cuando (16.108) se aplica al potencial de Hooke, se puede expresar en términos de integrales elípticas. No obstante, será más ilustrativo examinar las correcciones relativistas de primer orden cuando  $V(x) \ll mc^2$ . Escribiremos la energía total  $E$  de la forma

$$E = mc^2 (1 + \mathcal{E})$$

con lo cual se tiene que

$$\frac{E - V(x)}{mc^2} = \frac{mc^2 (1 + \mathcal{E}) - \frac{1}{2}kx^2}{mc^2} = 1 + \mathcal{E} - \lambda x^2 \quad ; \quad \lambda \equiv \frac{k}{2mc^2} \quad (16.109)$$

$$\frac{E - V(x)}{mc^2} = 1 + \lambda (b^2 - x^2) \equiv 1 + y \quad ; \quad b^2 \equiv \frac{\mathcal{E}}{\lambda} \quad , \quad y \equiv \lambda (b^2 - x^2) \quad (16.110)$$

y el término en el interior del radical en (16.108) resulta

$$1 - \frac{m^2 c^4}{[E - V(x)]^2} = 1 - \left[ \frac{mc^2}{E - V(x)} \right]^2 = 1 - \frac{1}{(1 + y)^2} = \frac{(1 + y)^2 - 1}{(1 + y)^2} = \frac{y^2 + 2y}{(1 + y)^2} = \frac{2y (1 + \frac{1}{2}y)}{(1 + y)^2}$$

de modo que el integrando en (16.108) resulta

$$\begin{aligned} \frac{1}{\sqrt{1 - \frac{m^2 c^4}{[E - V(x)]^2}}} &= \frac{1}{\sqrt{\frac{2y(1 + \frac{1}{2}y)}{(1 + y)^2}}} = \frac{(1 + y)}{\sqrt{2y}} \frac{1}{\sqrt{1 + \frac{1}{2}y}} \simeq \frac{(1 + y)}{\sqrt{2y}} \left[ 1 - \frac{1}{4}y + \mathcal{O}(y^2) \right] \\ &\simeq \frac{1 + \frac{3}{4}y}{\sqrt{2y}} + \mathcal{O}(y^2) \end{aligned}$$

y retornando a la definición de  $y$

$$\frac{1}{\sqrt{1 - \frac{m^2 c^4}{[E - V(x)]^2}}} \simeq \frac{1 + \frac{3}{4}\lambda (b^2 - x^2)}{\sqrt{2\lambda (b^2 - x^2)}}$$

reemplazando esta expresión en (16.108) resulta

$$\tau \cong \frac{4}{c} \int_0^b \frac{dx}{\sqrt{2\lambda (b^2 - x^2)}} \left[ 1 + \frac{3\lambda}{4} (b^2 - x^2) \right] \quad (16.111)$$

la integral (16.111) se puede evaluar con el cambio de variable  $x = b \sin \phi$  resultando

$$\begin{aligned} \tau &\cong \frac{4}{c} \int_0^{\pi/2} \frac{b \cos \phi \, d\phi}{\sqrt{2\lambda b^2 \cos^2 \phi}} \left[ 1 + \frac{3\lambda b^2}{4} \cos^2 \phi \right] = \frac{4}{bc\sqrt{2\lambda}} \int_0^{\pi/2} \frac{b \cos \phi \, d\phi}{\cos \phi} \left[ 1 + \frac{3\lambda b^2}{4} \cos^2 \phi \right] \\ &= \frac{4}{c\sqrt{2\lambda}} \int_0^{\pi/2} d\phi \left[ 1 + \frac{3\lambda b^2}{4} \cos^2 \phi \right] = \frac{4}{c\sqrt{2\lambda}} \left\{ \frac{\pi}{2} + \frac{3\lambda b^2}{4} \int_0^{\pi/2} \cos^2 \phi \, d\phi \right\} \\ &= \frac{4}{c\sqrt{2\lambda}} \left\{ \frac{\pi}{2} + \frac{3\lambda b^2}{4} \left( \frac{\pi}{4} \right) \right\} = \frac{2\pi}{c} \frac{1}{\sqrt{2\lambda}} \left[ 1 + \frac{3}{8}\lambda b^2 \right] \end{aligned}$$

recordando las definiciones de  $\lambda$  y  $b$ , Ecs. (16.109, 16.110) resulta

$$\tau \cong \frac{2\pi}{c} \frac{1}{\sqrt{2\frac{k}{2mc^2}}} \left[ 1 + \frac{3}{8}\mathcal{E} \right] = 2\pi \sqrt{\frac{m}{k}} \left[ 1 + \frac{3}{8}\mathcal{E} \right] = \tau_0 \left[ 1 + \frac{3}{8}\mathcal{E} \right]$$

donde  $\tau_0$  es el periodo en el caso no relativista. Vemos entonces que las correcciones relativistas introducen una dependencia con la amplitud, dada aproximadamente por

$$\frac{\Delta\nu}{\nu_0} = -\frac{\Delta\tau}{\tau_0} \cong -\frac{3}{8}\mathcal{E} \quad (16.112)$$



### 16.9.3. Movimiento de partícula cargada en un campo magnético

Este problema se puede trabajar aplicando el Lagrangiano (16.97) usando  $\phi = 0$  y el potencial vectorial adecuado para un campo magnético constante  $\mathbf{A} = 1/2(\mathbf{r} \times \mathbf{B})$ . No obstante, es más sencillo emplear directamente el formalismo Newtoniano de fuerzas y escribir

$$\mathbf{F} = \frac{q}{c}(\mathbf{v} \times \mathbf{B}) \quad (16.113)$$

con lo cual la ecuación de movimiento es

$$\frac{d\mathbf{p}}{dt} = \frac{q}{c}(\mathbf{v} \times \mathbf{B}) = \frac{q}{mc\gamma}(\mathbf{p} \times \mathbf{B})$$

La expresión (16.113) nos garantiza que la fuerza de Lorentz magnética no efectúa trabajo sobre la partícula de modo que  $\mathbf{F} \cdot \mathbf{v} = 0$ . Este hecho junto con las Ecs. (16.77, 16.78) nos dice que  $T$  permanece constante, en tanto que la expresión (16.83) nos dice que  $\mathbf{p}$  y  $\gamma$  también lo son. Adicionalmente, (16.113) nos indica que  $\mathbf{F}$  es perpendicular a  $\mathbf{B}$  de modo que la componente del momento a lo largo de  $\mathbf{B}$  se debe conservar. Finalmente, la ortogonalidad entre  $\mathbf{F}$  y  $\mathbf{v}$  nos indica que la partícula no cambia su rapidez.

Por lo anterior será posible sin pérdida de generalidad asumir que  $x_3$  es la dirección de  $\mathbf{B}$  y que el movimiento es en el plano  $x_1x_2$ . Descompondremos a  $\mathbf{p}$  en la forma  $\mathbf{p} = p_3\mathbf{u}_3 + \mathbf{p}_\perp$ . Con base en lo anterior sabemos que  $p_3$  es constante así como el módulo de  $\mathbf{p}$ . Por tanto, el módulo de  $\mathbf{p}_\perp$  es claramente constante de modo que  $\mathbf{p}$  realiza una precesión alrededor de la dirección del campo magnético con una frecuencia dada por

$$\Omega = \frac{qB}{mc\gamma} \quad (16.114)$$

al ser  $\gamma$  constante se deduce que la proyección de la velocidad en el plano  $x_1x_2$  tiene módulo constante y gira con la misma frecuencia. Si no hay componente de la velocidad en  $\mathbf{u}_3$ , la partícula se mueve entonces en un plano y describe una órbita circular uniforme con velocidad angular  $\Omega$ . De esto se obtiene el módulo de  $\mathbf{p}_\perp$

$$p_\perp = m\gamma v_\perp = m\gamma r\Omega$$

siendo  $r$  el radio de la circunferencia. Si combinamos esta ecuación con (16.114) obtenemos el radio de la circunferencia en función del momento lineal

$$r = \frac{p_\perp}{qB/c} \quad (16.115)$$

el radio de curvatura solo depende de las propiedades de la partícula a través del factor  $pc/q$  ( $= Br$ ), que se conoce como la rigidez magnética de la partícula. Se puede ver que aunque  $\Omega$  presenta correcciones relativistas contenidas en el factor  $\gamma$ , la relación entre radio y momento es la misma que en el caso no relativista (justamente porque el momento a su vez se redefine con el mismo factor  $\gamma$ ). Debe tenerse en cuenta que aunque  $r$  solo depende de  $p_\perp$ , en el cálculo de  $\gamma$  debe usarse tanto la componente perpendicular como la paralela a  $\mathbf{B}$  a fin de calcular  $\beta$ .

## 16.10. Formulaciones lagrangianas covariantes

El método abordado en la sección anterior predice correctamente las ecuaciones de movimiento relativistas de los problemas en cuestión pero no es clara la covarianza de las ecuaciones de la mecánica. En particular el parámetro tiempo se ha tratado como parámetro totalmente distinto de las coordenadas espaciales, en tanto que una formulación covariante exige que el espacio y el tiempo se traten como coordenadas análogas del espacio de universo. Esto implica que en lugar del tiempo, se debe utilizar algún parámetro invariante para trazar el punto figurativo que nos muestra la evolución de un sistema en el espacio de configuración. Es necesario partir de un principio de Hamilton manifiestamente covariante. Ahora, si el parámetro de integración es un invariante de Lorentz entonces el Lagrangiano debe ser un escalar de universo en cualquier formulación

covariante. Finalmente, en vez de ser funciones de  $x_i$  y  $\dot{x}_i$  el lagrangiano debe ser función de las coordenadas del espacio de Minkowski y de sus derivadas respecto al parámetro invariante.

Veamos el caso de un partícula. La elección natural del parámetro invariante es aparentemente el tiempo propio de la partícula. Sin embargo, las componentes de la velocidad generalizada, deben entonces cumplir la relación

$$u_\nu u_\nu = -c^2 \quad (16.116)$$

que demuestra que no todas sus componentes son independientes. Por lo tanto, supondremos que el parámetro invariante es más bien una función  $\theta$  que crece monótonamente a medida que el punto de universo evoluciona a lo largo de la línea de universo de la partícula y que es además invariante de Lorentz. La derivada con respecto al parámetro  $\theta$  y con respecto al parámetro  $t$  serán simbolizadas en la forma

$$\bar{x}_\nu \equiv \frac{dx_\nu}{d\theta} \quad ; \quad \dot{x}_\nu \equiv \frac{dx_\nu}{dt}$$

Un principio de Hamilton covariante debe ser entonces del tipo

$$\delta I = \delta \int_{\theta_1}^{\theta_2} \Lambda(x_\mu, \bar{x}_\mu) d\theta = 0 \quad (16.117)$$

siendo  $\Lambda$  un escalar de universo. Nótese que la dependencia explícita del tiempo está contenida en el hecho de que  $\Lambda$  depende de  $x_4$ . Las ecuaciones de Euler Lagrange que surgen de aplicar el principio de Hamilton a (16.117) son

$$\frac{d}{d\theta} \left( \frac{\partial \Lambda}{\partial \bar{x}_\mu} \right) - \frac{\partial \Lambda}{\partial x_\mu} = 0 \quad (16.118)$$

de modo que debemos encontrar un  $\Lambda$  tal que las Ecs. (16.118) sean equivalentes a las Ecs. (16.67). Una manera de hacerlo es partir de la integral de acción usual con parámetro  $t$  y transformando dicha integral de modo que el parámetro tiempo sea reemplazado por el parámetro invariante  $\theta$  y el tiempo en cambio aparezca no como parámetro sino como otra coordenada generalizada. Como  $\theta$  debe ser función monótona de  $t$  medida en cierto sistema inercial, se tiene

$$\frac{dx_i}{dt} = \frac{dx_i}{d\theta} \frac{d\theta}{dt} = \frac{dx_i/d\theta}{dt/d\theta} \equiv ic \frac{\bar{x}_i}{\bar{x}_4} \quad ; \quad dt = \frac{dt}{d\theta} d\theta = \frac{d(ict)}{ic d\theta} d\theta = -\frac{i}{c} \frac{dx_4}{d\theta} d\theta = -\frac{i}{c} \bar{x}_4 d\theta$$

con lo cual la integral de acción usual (con parámetro  $t$ ) se transforma de la manera siguiente

$$I = \int_{t_1}^{t_2} L(x_j, t, \dot{x}_j) dt = -\frac{i}{c} \int_{\theta_1}^{\theta_2} L\left(x_\mu, ic \frac{\bar{x}_j}{\bar{x}_4}\right) \bar{x}_4 d\theta \quad (16.119)$$

comparando la acción definida en (16.117) con (16.119), vemos que una receta adecuada para el lagrangiano transformado  $\Lambda$  sería de la forma

$$\Lambda(x_\mu, \bar{x}_\mu) = -\frac{i\bar{x}_4}{c} L\left(x_\mu, ic \frac{\bar{x}_j}{\bar{x}_4}\right) \quad (16.120)$$

este Lagrangiano tiene una característica no vista hasta ahora en los Lagrangianos, depende del cociente entre derivadas de las coordenadas generalizadas  $\bar{x}_j/\bar{x}_4$  y de un factor global  $\bar{x}_4$ . La primera característica hace que  $\Lambda$  sea una función homogénea de primer grado de las velocidades generalizadas independientemente de la forma funcional de  $L$

$$\begin{aligned} \Lambda(x_\mu, a\bar{x}_\mu) &= -\frac{i(a\bar{x}_4)}{c} L\left(x_\mu, ic \frac{(a\bar{x}_j)}{(a\bar{x}_4)}\right) = -a \frac{i\bar{x}_4}{c} L\left(x_\mu, ic \frac{\bar{x}_j}{\bar{x}_4}\right) \\ \Lambda(x_\mu, a\bar{x}_\mu) &= a\Lambda(x_\mu, \bar{x}_\mu) \end{aligned} \quad (16.121)$$

este no es un fenómeno de la física relativista como tal, sino de la forma en que se amplía el espacio de configuración para incluir a  $t$  como coordenada generalizada, y usar otro parámetro para seguir la evolución

del punto figurativo de la partícula en el espacio de universo. Un Lagrangiano de la forma (16.121) recibe el nombre de Lagrangiano homogéneo y el correspondiente problema homogéneo del cálculo de variaciones exige un tratamiento especial. La dificultad más seria surge en la formulación Hamiltoniana. Para verlo, observemos que al ser  $\Lambda$  homogénea de primer grado en  $\bar{x}_\mu$ , podemos aplicar el teorema de Euler Ec. (5.16) Pág. 77, para obtener

$$\Lambda = \bar{x}_\mu \frac{\partial \Lambda}{\partial \bar{x}_\mu} \quad (16.122)$$

con esto vemos que la función energía  $h$  definida en (5.13) Pág. 76, es idénticamente nula<sup>19</sup>. Por otro lado, de la Ec. (16.122) vemos también que se satisface la relación

$$\left[ \frac{d}{d\theta} \left( \frac{\partial \Lambda}{\partial \bar{x}_\mu} \right) - \frac{\partial \Lambda}{\partial x_\mu} \right] \bar{x}_\mu = 0 \quad (16.123)$$

la cual nos indica que si se satisfacen tres de las Ecs. (16.118) entonces la cuarta se satisface automáticamente. Nótese que esto es consecuencia de la homogeneidad de  $\Lambda$ .

Veamos como se escribe esta transformación para una partícula libre. La Ec. (16.90) nos indica la expresión “no covariante” para el Lagrangiano de una partícula libre

$$L(x_i, \dot{x}_i) = -mc\sqrt{c^2 - \dot{x}_i \dot{x}_i} \quad (16.124)$$

y aplicando la transformación (16.120) a (16.124), resulta un posible Lagrangiano covariante

$$\begin{aligned} \Lambda(x_\mu, \bar{x}_\mu) &= -\frac{i\bar{x}_4}{c} L\left(x_\mu, ic\frac{\bar{x}_j}{\bar{x}_4}\right) = -\frac{i\bar{x}_4}{c} \left[ -mc\sqrt{c^2 - \left(ic\frac{\bar{x}_i}{\bar{x}_4}\right)\left(ic\frac{\bar{x}_i}{\bar{x}_4}\right)} \right] = i\bar{x}_4 \left[ m\sqrt{c^2 + c^2\frac{\bar{x}_i\bar{x}_i}{\bar{x}_4^2}} \right] \\ \Lambda(x_\mu, \bar{x}_\mu) &= imc\bar{x}_4 \left[ \sqrt{1 + \frac{\bar{x}_i\bar{x}_i}{\bar{x}_4^2}} \right] = imc\sqrt{\bar{x}_4^2 + \bar{x}_i\bar{x}_i} \end{aligned}$$

Para llegar a  $\Lambda(x_\mu, \bar{x}_\mu)$  hay una ambigüedad de signo que debe fijarse para que  $L dt$  tenga el mismo valor que  $\Lambda d\theta$ . El paso final se escribe en la forma

$$\Lambda = -mc\sqrt{(-i)^2\sqrt{\bar{x}_4^2 + \bar{x}_j\bar{x}_j}} = -mc\sqrt{-\bar{x}_4^2 - \bar{x}_j\bar{x}_j}$$

quedando finalmente

$$\Lambda(x_\mu, \bar{x}_\mu) = -mc\sqrt{-\bar{x}_\mu\bar{x}_\mu} \quad (16.125)$$

con la cual las ecuaciones de Euler-Lagrange (16.118), equivalen a

$$\begin{aligned} \frac{d}{d\theta} \left[ \frac{\partial}{\partial \bar{x}_\mu} (-mc\sqrt{-\bar{x}_\nu\bar{x}_\nu}) \right] - \frac{\partial}{\partial x_\mu} (-mc\sqrt{-\bar{x}_\nu\bar{x}_\nu}) &= 0 \quad \Rightarrow \quad -mc\frac{d}{d\theta} \left[ \frac{\partial}{\partial \bar{x}_\mu} (\sqrt{-\bar{x}_\nu\bar{x}_\nu}) \right] = 0 \\ \frac{d}{d\theta} \left( \frac{mc\bar{x}_\mu}{\sqrt{-\bar{x}_\nu\bar{x}_\nu}} \right) &= 0 \end{aligned}$$

como el parámetro  $\theta$  debe ser función monótona del tiempo propio  $\tau$ , las derivadas con respecto a  $\theta$  están relacionadas con las derivadas en función de  $\tau$  a través de la expresión

$$\frac{d}{d\theta} = \frac{d\tau}{d\theta} \frac{d}{d\tau} \quad ; \quad \bar{x}_\nu \equiv \frac{dx_\nu}{d\theta} = \frac{d\tau}{d\theta} \frac{dx_\nu}{d\tau} = \frac{d\tau}{d\theta} u_\nu$$

y las ecuaciones de Lagrange quedan de la forma

$$\begin{aligned} \frac{d\tau}{d\theta} \frac{d}{d\tau} \left[ \frac{mc \left( \frac{d\tau}{d\theta} u_\mu \right)}{\sqrt{-\left( \frac{d\tau}{d\theta} \right)^2 u_\nu u_\nu}} \right] &= 0 \quad \Rightarrow \quad \frac{d\tau}{d\theta} \frac{d}{d\tau} \left[ \frac{m c u_\mu}{\sqrt{-u_\nu u_\nu}} \right] = 0 \quad \Rightarrow \quad \frac{d}{d\tau} \left[ \frac{m c u_\mu}{\sqrt{c^2}} \right] = 0 \\ &\Rightarrow \quad \frac{d(m u_\mu)}{d\tau} = 0 \end{aligned}$$

<sup>19</sup>De hecho se observa que la función energía definida en la Ec. (5.17) Pág. 77, no posee parte homogénea de primer grado.

que coincide con las Ecs. (16.67) aplicadas a una partícula libre. La cuarta de estas ecuaciones no es independiente, lo cual también se vió anteriormente ya que esta ecuación nos manifiesta la conservación de la energía cinética, que se puede deducir de las tres ecuaciones anteriores.

Nótese que este método resulta laborioso incluso para una partícula libre. Una forma más directa y práctica propuesta por Dirac, evita la introducción del parámetro  $\theta$  y permite trabajar con el tiempo propio  $\tau$ . Para ello basta con observar que la relación que nos previene de usar el tiempo propio como parámetro invariante Ec. (16.116) no es una verdadera ligadura dinámica de movimiento<sup>20</sup> sino una consecuencia geométrica del modo de definir  $\tau$ , esta ecuación nos dice básicamente que hay una región permitida en el espacio tetradimensional  $u_\nu$  y que no todos los puntos de esta región son accesibles, la región accesible es una hipersuperficie de 3 dimensiones. A las Ecs. del tipo (16.116) Dirac las denomina ecuaciones débiles. Podemos en consecuencia, trabajar las  $u_\nu$  como cantidades sin ligaduras y solo después de efectuar todas las operaciones de derivación necesitaremos imponer la ligadura (16.116). Este procedimiento se puede aplicar al Lagrangiano libre. En cuyo caso se podría haber hecho  $\theta$  igual a  $\tau$  desde el principio y se aplica la Ec. (16.116) solo en el último paso. De esta forma se puede escribir directamente en función de  $\tau$  las ecuaciones de Lagrange covariantes

$$\frac{d}{d\tau} \left( \frac{\partial \Lambda}{\partial u_\nu} \right) - \frac{\partial \Lambda}{\partial x_\nu} = 0 \quad (16.126)$$

por otro lado, no es necesario que la integral de acción del principio de Hamilton deba tener el mismo valor cuando se expresa en función de  $t$  que cuando se expresa en función de  $\theta$  o de  $\tau$ , de modo que el Lagrangiano no tiene que estar descrito por la Ec. (16.120). Solo se le exige a  $\Lambda$  que sea un escalar de universo y que conduzca a las ecuaciones de movimiento correctas, después de todo el principio de Hamilton se escribió con esta idea.  $\Lambda$  no tiene que ser necesariamente homogénea de primer grado en las velocidades generalizadas. Por ejemplo, un Lagrangiano adecuado para partícula libre es el siguiente

$$\Lambda = \frac{1}{2} m u_\nu u_\nu \quad (16.127)$$

y más en general  $\Lambda$  puede tener la forma  $m f(u_\nu u_\nu) \equiv m f(y)$  donde  $f(y)$  es una función arbitraria con la condición de que

$$\left. \frac{\partial f}{\partial y} \right|_{y=-c^2} = \frac{1}{2} \quad (16.128)$$

donde (16.128) es una forma de tener en cuenta la ligadura (16.116) en la función  $f(u_\nu u_\nu)$ . Nótese que las funciones  $f(u_\nu u_\nu) = \frac{1}{2} u_\nu u_\nu$  y  $f(u_\nu u_\nu) = -c \sqrt{-u_\nu u_\nu}$  cumplen esta condición y conducen a los Lagrangianos (16.125) y (16.127) respectivamente. De aquí en adelante, usaremos la Ec. (16.127) para la parte de la “energía cinética” del Lagrangiano.

Si la partícula no es libre, es necesario introducir términos escalares de universo que den cuenta de la interacción con el exterior correspondientes a ciertas fuerzas de Minkowski. Usualmente estos términos serán la contracción de un cuadvivector exterior con un cuadvivector del sistema que representan la interacción de la partícula con un campo exterior. Por ejemplo, si el término de interacción es de la forma  $\phi_\mu x_\mu$  siendo  $\phi_\mu$  un cuadvivector exterior y si además en algún sistema de referencia  $\phi_1 = ma$  y las demás componentes son nulas, tendríamos un ejemplo de fuerza constante. La forma concreta dependerá de la formulación covariante que se utilice para la teoría del campo. Ya tenemos a mano un ejemplo cual es la interacción de una carga puntual con un campo electromagnético, para el cual, un Lagrangiano adecuado es

$$\Lambda(x_\mu, u_\mu) = \frac{1}{2} m u_\mu u_\mu + \frac{q}{c} u_\mu A_\mu(x_\lambda) \quad (16.129)$$

y las ecuaciones de Lagrange (16.126) asociadas al Lagrangiano (16.129) son

$$\begin{aligned} \frac{d}{d\tau} \left[ \frac{\partial}{\partial u_\nu} \left( \frac{1}{2} m u_\mu u_\mu + \frac{q}{c} u_\mu A_\mu(x_\lambda) \right) \right] - \frac{\partial}{\partial x_\nu} \left( \frac{1}{2} m u_\mu u_\mu + \frac{q}{c} u_\mu A_\mu(x_\lambda) \right) &= 0 \\ \frac{d}{d\tau} \left[ m u_\nu + \frac{q}{c} A_\nu(x_\lambda) \right] - \frac{q}{c} u_\mu \frac{\partial A_\mu(x_\lambda)}{\partial x_\nu} &= 0 \end{aligned}$$

<sup>20</sup>La Ec. (16.55), Pág. 468, nos muestra que la cuarta componente de  $u_\mu$  es totalmente espúrea, y solo se agregó con el fin de construir un cuadvivector que contuviera la información Física de la velocidad.

$$\frac{d}{d\tau}(mu_\nu) = \frac{q}{c} \left[ \frac{\partial}{\partial x_\nu} (u_\mu A_\mu) - \frac{dA_\nu}{d\tau} \right]$$

que corresponden a las ecuaciones de movimiento generalizadas (16.67) donde la fuerza de Minkowski  $K_\nu$  está dada por la Ec. (16.70). Como antes, el cuadrimento lineal  $p_\mu$  es diferente del momento canónico asociado a  $x_\nu$  que denotaremos por  $\mathcal{P}_\mu$

$$\mathcal{P}_\mu = \frac{\partial \Lambda}{\partial u_\mu} = mu_\mu + \frac{q}{c} A_\mu = p_\mu + \frac{q}{c} A_\mu \quad (16.130)$$

las tres primeras componentes corresponden a las dadas en la Ec. (5.2). La cantidad canónica asociada a  $x_4$  se obtiene sustituyendo (16.61, 16.80) en (16.130) y viene dada por

$$\mathcal{P}_4 = p_4 + \frac{q}{c} A_4 = \frac{iT}{c} + \frac{iq\phi}{c} = \frac{i}{c} E$$

siendo  $E = T + q\phi$  la energía total de la partícula. Una vez más, el momento conjugado a la coordenada tiempo es proporcional a la energía total<sup>21</sup>. La relación entre el módulo del momento lineal y la energía  $T$  sigue estando dada por (16.83). Es notable el hecho de que dado que  $\mathbf{p}$  y  $\mathbf{A}$  son trivectores cartesianos entonces las tres componentes espaciales de  $\vec{\mathcal{P}}$  dadas por

$$\vec{\mathcal{P}} = \mathbf{p} + \frac{q}{c} \mathbf{A}$$

forman un trivector cartesiano (recordemos que en general, los momentos canónicos no forman vectores). Por esta razón es útil escribir la Ec. (16.83) en términos del trivector canónico  $\vec{\mathcal{P}}$

$$T^2 = \left( \vec{\mathcal{P}} - \frac{q}{c} \mathbf{A} \right)^2 c^2 + m^2 c^4 \quad (16.131)$$

relación muy útil entre la energía  $T$  y el momento canónico  $\vec{\mathcal{P}}$ . El ejemplo de la carga en campo electromagnético es un ejemplo de interacción con un campo vectorial. Podríamos tener por ejemplo interacción con un campo escalar  $\phi(x_\mu)$  usualmente acoplado a alguna cantidad escalar del sistema, tales como  $q\phi(x_\mu)$  ó  $m\phi(x_\mu)$  siendo  $m$  masa propia y carga respectivamente, hay por supuesto otras posibilidades como la contracción de dos tensores (uno externo y otro interno) de la forma  $F^{\mu\nu} K_{\mu\nu}$  etc. La única restricción es que los términos de interacción formen un escalar de universo. La forma explícita de estos términos depende entonces de la teoría de campo específica que se está considerando.

## 16.11. Ejercicios

1. Demuestre que  $j_\mu = \rho_0 u_\mu$ , siendo  $j_\mu$  el cuadrivector densidad de corriente,  $\rho_0$  la densidad de carga propia, y  $u_\mu$  la cuadrivelocidad.
2. Demuestre que la cantidad  $A_\mu$  definida en 16.65, es un cuadrivector.
3. Encuentre el equivalente energético (energía en reposo o autoenergía) de una masa de 1Kg. Compare este valor con cantidades energéticas típicas a la escala humana.
4. Para el movimiento relativista de una partícula sometida a una fuerza constante, encuentre la ecuación de la trayectoria para valores arbitrarios de la velocidad y posición inicial.
5. Demuestre que si el Lagrangiano  $\Lambda$  satisface las ecuaciones de Lagrange (16.118), entonces se cumple la relación (16.123) como consecuencia de la homogeneidad de primer grado de  $\Lambda$  con respecto a  $\bar{x}_\mu$  expresada en (16.121).

<sup>21</sup>En el caso no relativista, esta relación de conjugación era mas bien artificial ya que el tiempo era un parámetro y no una coordenada generalizada. En el caso relativista, el tiempo se convierte en coordenada y por tanto podemos hablar en firme de su momento conjugado.

## Teoría canónica de perturbaciones

Es un hecho que la mayoría de problemas físicos reales no poseen soluciones matemáticas exactas. A manera de ilustración el problema de Kepler para dos cuerpos es soluble de manera exacta, pero para el movimiento clásico de tres cuerpos puntuales bajo interacción gravitacional no ha sido posible una solución analítica exacta. Lo anterior nos motiva a estudiar técnicas que permitan encontrar soluciones aproximadas analíticas o numéricas.

Por otro lado, ocurre con frecuencia que el Hamiltoniano del problema que no se puede resolver de forma exacta, difiere solo ligeramente de un Hamiltoniano que se puede resolver rigurosamente. Escribimos entonces

$$H(q, p, t) = H_0(q, p, t) + \Delta H(q, p, t) \quad (17.1)$$

siendo  $H_0$  el **Hamiltoniano no perturbado** del cual conocemos su solución exacta,  $H$  es el **Hamiltoniano perturbado** del cual queremos conocer una solución aproximada y  $\Delta H$  es el **Hamiltoniano de perturbación**. La hipótesis de trabajo básica de la técnica de perturbaciones es la pequeñez de  $\Delta H$  con respecto a  $H_0$ . Genéricamente decimos que el problema más complicado es una perturbación del sistema soluble.

Vale la pena tener en cuenta que incluso cuando el Hamiltoniano de perturbación  $\Delta H$  sea pequeño con respecto al Hamiltoniano no perturbado  $H_0$ , el efecto de la perturbación sobre la solución puede ser considerable. Por ejemplo, para un oscilador bidimensional isotrópico i.e. de frecuencias iguales en  $x$  e  $y$ , la trayectoria es una curva cerrada simple, más específicamente una elipse. Supongamos que aplicamos una perturbación que cambie ligeramente la constante de rigidez en la dirección  $y$  de tal forma que las frecuencias son ahora diferentes e inconmensurables. Como consecuencia de esta pequeña perturbación de  $H$ , la trayectoria ya no será cerrada y en el transcurso del tiempo llenará todo el espacio conformado por el rectángulo que se forma con las amplitudes en  $x$  e  $y$ . Otro ejemplo ilustrativo es el efecto gravitatorio debido al achatamiento de la tierra visto como desviación del problema no perturbado con la tierra perfectamente esférica, las pequeñas perturbaciones producidas por este achatamiento producen un cambio apreciable en la trayectoria de un satélite puntual. En particular el campo central sin perturbar produce un movimiento en un plano en tanto que toda perturbación cuadrupolar gravitatoria por pequeña que sea hace que el movimiento se salga del plano inicial a menos que el eje del cuadrupolo sea perpendicular a dicho plano.

La mayor parte del desarrollo formal de la teoría clásica de perturbaciones se realizó para la mecánica celeste. Por ejemplo, el efecto gravitacional del sol sobre la luna se puede considerar una perturbación con respecto a la interacción gravitacional dominante debida a la tierra. La necesidad de predecir con precisión las órbitas de satélites ha impulsado fuertemente el desarrollo de métodos perturbativos.

En general podemos dividir la teoría de perturbaciones en dos tipos de teoría: teoría de perturbaciones dependiente del tiempo e independiente del tiempo. Buena parte de las técnicas aquí desarrolladas poseen analogía con la teoría de perturbaciones para la mecánica cuántica, si bien la teoría de perturbaciones clásica es mucho más complicada que su contraparte cuántica. En general, la teoría de perturbaciones clásica se puede tratar con cualquiera de los formalismos desarrollados hasta ahora: Newtoniano, Lagrangiano y Hamiltoniano. Sin embargo, la formulación más conveniente es la de Hamilton-Jacobi. Comenzaremos con la teoría de perturbaciones dependiente del tiempo que es la forma más simple

## 17.1. Método de variación de constantes para perturbaciones dependientes del tiempo

Sea  $H_0(q, p, t)$  el Hamiltoniano no perturbado cuya solución exacta se conoce. El Hamiltoniano perturbado  $H(q, p, t)$  será la suma del Hamiltoniano no perturbado  $H_0(q, p, t)$  más el Hamiltoniano de perturbación  $\Delta H(q, p, t)$  como se aprecia en la Ec. (17.1). Supongamos que hemos obtenido la solución para el problema no perturbado mediante la función principal de Hamilton Ec. (9.5)

$$H_0\left(q_1 \dots q_n; \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n}; t\right) + \frac{\partial S}{\partial t} = 0$$

recordemos que la solución  $S(q, \alpha, t)$  de esta ecuación representa la función generatriz de tipo 2 que genera una transformación canónica que nos lleva del Hamiltoniano original  $H_0$  a un Hamiltoniano nulo  $K_0 = 0$ , de modo que todas las nuevas variables canónicas serán constantes (ver sección 9.1). Las nuevas variables canónicas  $(\beta, \alpha)$ , donde las  $\beta$ 's son las nuevas coordenadas y las  $\alpha$ 's los nuevos momentos, son entonces **constantes en el problema no perturbado**.

Ya se ha mencionado que el carácter canónico de una transformación es independiente del problema i.e. de la forma específica del Hamiltoniano, lo cual se puede ver por ejemplo de la condición simpléctica Ec. (7.56) la cual solo depende de la transformación en sí y no del Hamiltoniano al cual se le aplique esta transformación. De lo anterior se deduce que la transformación definida por

$$(q, p) \rightarrow (\beta, \alpha)$$

generada por  $S(q, \alpha, t)$  sigue siendo una transformación canónica para el problema perturbado. No obstante, en general esta transformación canónica no hará que el nuevo Hamiltoniano perturbado sea nulo y las nuevas variables canónicas no serán constantes en el problema perturbado. Teniendo en cuenta las Ecs. (17.1) y (9.2) el nuevo Hamiltoniano perturbado será

$$K(\alpha, \beta, t) = H_0 + \Delta H + \frac{\partial S}{\partial t} \quad (17.2)$$

por supuesto que  $H_0, \Delta H$  y  $S$  se deben escribir como funciones de  $\alpha, \beta, t$ . Las ecuaciones de Hamilton para las nuevas variables canónicas del sistema perturbado vienen dadas por

$$\dot{\alpha}_i = -\frac{\partial K(\alpha, \beta)}{\partial \beta_i} \quad ; \quad \dot{\beta}_i = \frac{\partial K(\alpha, \beta, t)}{\partial \alpha_i} \quad (17.3)$$

las ecuaciones (17.3) son formalmente exactas de modo que no hemos hecho ninguna aproximación aún. Tales ecuaciones nos dan formalmente los valores de  $\beta_i, \alpha_i$  en función del tiempo. Por otro lado, las ecuaciones de transformación desde  $(q, p)$  hacia  $(\beta, \alpha)$  nos permitirán a su vez encontrar a  $q, p$  en función del tiempo a través del inverso de dicha transformación, lo cual resuelve el problema completo. No obstante, la solución para el nuevo Hamiltoniano perturbado  $K(\alpha, \beta, t)$  no es usualmente más sencilla que la del Hamiltoniano perturbado original  $H(q, p, t)$ , la razón es que aunque la transformación se mantiene canónica, ya no trivializa las ecuaciones de Hamilton como lo hacía con el Hamiltoniano no perturbado, por lo tanto no es claro que tal transformación canónica simplifique el problema perturbado. De lo anterior se desprende que las Ecs. (17.3) no serán especialmente fructíferas si buscamos una solución exacta.

Sin embargo, estas ecuaciones serán de gran utilidad para obtener soluciones aproximadas si tenemos en cuenta que  $\Delta H$  debe ser pequeño con respecto a  $H_0$ . Es de esperarse entonces que los valores de  $\alpha$  y  $\beta$  en el nuevo problema no difieran significativamente de su valor constante original de modo que no varíen rápidamente, al menos en comparación con la dependencia explícita de  $K$  con respecto al tiempo. Una primera aproximación a la dependencia temporal de  $(\beta, \alpha)$  se obtendrá entonces asumiendo la validez de las ecuaciones (17.3) pero con los miembros derechos de cada ecuación evaluados en los valores constantes no perturbados  $\alpha_{0i}$  y  $\beta_{0i}$ . Denotando  $\alpha_{1i}$  y  $\beta_{1i}$  las soluciones **de primer orden** en teoría de perturbaciones para  $\alpha_i, \beta_i$  tenemos que

$$\dot{\alpha}_{1i} = -\left.\frac{\partial K(\alpha, \beta)}{\partial \beta_i}\right|_{\alpha_i=\alpha_{0i}, \beta_i=\beta_{0i}} \quad ; \quad \dot{\beta}_{1i} = \left.\frac{\partial K(\alpha, \beta, t)}{\partial \alpha_i}\right|_{\alpha_i=\alpha_{0i}, \beta_i=\beta_{0i}} \quad (17.4)$$

de modo que después de la derivación se deben sustituir las variables  $\alpha, \beta$  por sus valores constantes no perturbados. Como estas son ecuaciones de Hamilton, se pueden escribir matricialmente en forma simpléctica como en (6.33)

$$\dot{\gamma}_1 = \mathbf{J} \left. \frac{\partial K(\gamma, t)}{\partial \gamma} \right|_{\gamma=\gamma_0} ; \quad \tilde{\gamma}_1 \equiv (\beta_{11}, \beta_{12}, \dots, \beta_{1n}, \alpha_{11}, \alpha_{12}, \dots, \alpha_{1n}) \quad (17.5)$$

donde la matriz  $\mathbf{J}$  está definida por la Ec. (6.31). Las Ecs. (17.4) o (17.5) se pueden integrar directamente para obtener la dependencia temporal de las  $\alpha_{1i}, \beta_{1i}$  en aproximación de primer orden. Usando las ecuaciones de transformación se obtiene la dependencia temporal para la correspondiente aproximación de primer orden en las  $q, p$ .

Ahora bien, si queremos proceder a la aproximación de segundo orden, debemos utilizar la aproximación de primer orden de la dependencia temporal de  $\alpha$  y  $\beta$  en los miembros de la derecha de las Ecs. (17.3)

$$\dot{\alpha}_{2i} = - \left. \frac{\partial K(\alpha, \beta)}{\partial \beta_i} \right|_{\alpha_i=\alpha_{1i}, \beta=\beta_{1i}} ; \quad \dot{\beta}_{2i} = \left. \frac{\partial K(\alpha, \beta, t)}{\partial \alpha_i} \right|_{\alpha_i=\alpha_{1i}, \beta=\beta_{1i}} \quad (17.6)$$

en general, si queremos proceder a la aproximación de orden  $m$ , debemos utilizar la aproximación de orden  $m-1$  de la dependencia temporal de  $\alpha$  y  $\beta$  en los miembros de la derecha de las Ecs. (17.3), en notación matricial simpléctica esto se escribe como

$$\dot{\gamma}_m = \mathbf{J} \left. \frac{\partial K(\gamma, t)}{\partial \gamma} \right|_{\gamma=\gamma_{m-1}} ; \quad \tilde{\gamma}_m \equiv (\beta_{m1}, \beta_{m2}, \dots, \beta_{mn}, \alpha_{m1}, \alpha_{m2}, \dots, \alpha_{mn}) \quad (17.7)$$

## 17.2. Teoría de perturbación dependiente del tiempo en términos de los parámetros de movimiento

Las variables transformadas  $(\alpha, \beta)$  contienen información sobre los parámetros del movimiento sin perturbar. En general, los parámetros de movimiento suelen ser más convenientes que los valores de  $\alpha$  y  $\beta$  para trabajar desde el punto de vista teórico y/o experimental. A manera de ejemplo, en el problema de Kepler en tres dimensiones sin perturbar con el formalismo de variables acción ángulo, las variables  $\alpha$  y  $\beta$  corresponden a las variables de acción constantes  $J_i$  y las variables angulares correspondientes  $w_i$  (usualmente conocidas como variables de Delauney). Por otro lado, las variables de Delauney están relacionadas de manera sencilla con los parámetros del sistema que son más fácilmente observables o medibles: semieje mayor, excentricidad, inclinación etcétera (ver sección 10.16). La teoría de perturbación que se construyó anteriormente permite encontrar a cualquier orden las correcciones a los valores de  $\alpha$  y  $\beta$ . Ahora bien, si pretendemos relacionar las cantidades  $\alpha$  y  $\beta$  perturbados con los valores de los parámetros de movimiento perturbados, podemos encontrarnos con la dificultad de que la perturbación puede alterar no solo los valores de las  $\alpha$  y  $\beta$  sino también la relación funcional de éstas con los parámetros de movimiento.

Por lo anterior, sería deseable construir una teoría de perturbaciones en la cual se obtenga la aproximación a cierto orden de los parámetros de movimiento directamente, sin recurrir a la determinación de los  $\alpha$  y  $\beta$  como paso intermedio. Esto es posible a través del formalismo de los corchetes de Poisson. Sea  $c_i$  un conjunto de  $2n$  funciones de las constantes  $(\alpha, \beta)$  del sistema sin perturbar siendo todas las  $c_i$  independientes entre sí<sup>1</sup>

$$c_i = c_i(\alpha, \beta) \quad ; \quad i = 1, \dots, 2n \quad (17.8)$$

las  $c_i$  son arbitrarias desde el punto de vista matemático excepto por la exigencia de que sean independientes e invertibles. Desde el punto de vista físico es usualmente deseable que una o más de dichas funciones representen los parámetros de movimiento usados. Como cualquier función del espacio de fase  $c_i(\alpha, \beta)$  obedece a la ecuación de movimiento (8.25) y asumiendo que las  $c_i$  no son funciones explícitas del tiempo, se tiene que

$$\dot{c}_i = [c_i, K] \quad (17.9)$$

<sup>1</sup>Nótese que las  $c_i$  no se pueden interpretar en general como nuevas coordenadas canónicas puesto que la transformación (17.8) no es necesariamente canónica. Solo pedimos que el jacobiano de la transformación sea no nulo para garantizar la independencia de las funciones  $c_i$  y la invertibilidad del sistema de ecuaciones.



por otro lado, las inversas de las ecuaciones (17.8) permiten a su vez considerar a  $K(\alpha, \beta, t)$  como función de los  $c_i$  y el tiempo i.e.  $K = K(c_i, t)$

$$\begin{aligned} [c_i, K] &\equiv \frac{\partial c_i}{\partial \eta} \mathbf{J} \frac{\partial K}{\partial \eta} = \frac{\partial c_i}{\partial \eta} \mathbf{J} \frac{\partial K}{\partial c_j} \frac{\partial c_j}{\partial \eta} \Rightarrow \\ [c_i, K] &= [c_i, c_j] \frac{\partial K}{\partial c_j} \end{aligned}$$

donde hay suma sobre el índice repetido  $j$ . Reemplazando esta ecuación en (17.9) se tiene

$$\dot{c}_i = [c_i, c_j] \frac{\partial K}{\partial c_j} \quad (17.10)$$

al igual que las ecuaciones de Hamilton (17.3) las ecuaciones (17.10) son exactas y no se ha realizado todavía ninguna aproximación. De manera análoga a la teoría de perturbaciones para los  $\alpha, \beta$ , el cálculo de primer orden consiste en evaluar a la derecha de la ecuación (17.10) incluyendo los corchetes de Poisson con los valores de los parámetros correspondientes al movimiento no perturbado. En la aproximación de orden  $m$ , se hacen todas las derivadas a la derecha de (17.10) luego de lo cual se reemplazan las expresiones resultantes, por los valores de los parámetros obtenidos en la perturbación de orden  $m - 1$ .

$$\dot{c}_{mi} = \left( [c_i, c_j] \frac{\partial K}{\partial c_j} \right)_{c_i=c_{m-1,i}; c_j=c_{m-1,j}} \quad ; \quad i = 1, \dots, 2n \quad (17.11)$$

esta ecuación es el análogo de (17.7). En mecánica celeste aparece a menudo la Ec. (17.11) en términos de corchetes de Lagrange. Multipliquemos la Ec. (17.10) por el corchete de Lagrange  $\{c_k, c_i\}$  y sumamos sobre índices repetidos

$$\{c_k, c_i\} \dot{c}_i = \{c_k, c_i\} [c_i, c_j] \frac{\partial K}{\partial c_j} \quad (17.12)$$

usando la propiedad (8.15) se tiene

$$-\frac{\partial K}{\partial c_j} = \{c_k, c_i\} \dot{c}_i \quad (17.13)$$

de nuevo en la terminología de la mecánica celeste es usual definir  $R \equiv -K$  de modo que la Ec. (17.13) queda en la forma

$$\frac{\partial R}{\partial c_j} = \{c_k, c_i\} \dot{c}_i \quad (17.14)$$

las ecuaciones (17.11) o sus formas equivalentes (17.13, 17.14) se conocen como **ecuaciones de perturbación de Lagrange**.

### 17.3. Variación periódica y variación secular de una perturbación

Un aspecto implícito en la formulación perturbativa es que los parámetros de movimiento del sistema no perturbado nos sirven para describir al sistema perturbado. En general el efecto de una perturbación es hacer variar los parámetros del sistema con el tiempo. Si el movimiento no perturbado es periódico y la perturbación es pequeña, la variación de los parámetros durante un periodo **del movimiento no perturbado** también será pequeña<sup>2</sup>. La teoría de perturbaciones dependiente del tiempo (variación de constantes) implica una imagen en la cual el sistema perturbado se mueve en intervalos cortos de tiempo en una trayectoria de igual forma funcional que el sistema sin perturbar, aunque los parámetros de la trayectoria variarán con el tiempo. La órbita no perturbada a lo largo de la cual viaja instantáneamente el sistema perturbado se denomina usualmente “órbita osculatriz”, la cual en posición y dirección de la tangente coincide instantáneamente con la trayectoria verdadera.

<sup>2</sup>En la mayoría de los casos la periodicidad exacta se perderá con el movimiento perturbado, razón por la cual la cuasi periodicidad del sistema usualmente tendrá como referente el periodo del movimiento no perturbado.

Hay dos maneras en las cuales pueden variar los parámetros en una órbita osculatriz visto en el esquema perturbativo. Es posible que haya una variación periódica en la cual el parámetro vuelve a su valor inicial en un intervalo de tiempo que en primera aproximación suele ser el periodo de la órbita no perturbada. En contraste, es posible que se mantenga un incremento (o decremento) neto del valor del parámetro al final de cada periodo orbital sucesivo. En el primer caso hablamos de una **variación periódica** en tanto que en el segundo caso se habla de una **variación secular**. Cuando la variación de un parámetro dado es periódica, su valor promediado sobre un periodo suele ser muy similar a su valor no perturbado, y si todos los parámetros tienen variación periódica la trayectoria perturbada tiene un aspecto muy similar a la no perturbada. Si en cambio la variación es secular, por pequeña que sea la desviación en cada periodo, los valores de los parámetros pueden diferir fuertemente respecto a los valores no perturbados para un número suficientemente grande de periodos. Correspondientemente, la perturbación llevará a una trayectoria muy diferente a la no perturbada para tiempos mucho mayores al periodo. Por esta razón, para múltiples propósitos suele eliminarse los efectos periódicos<sup>3</sup> y quedarse solo con los efectos seculares en los cálculos perturbativos.

Vale la pena mencionar que el uso de la órbita osculatriz requiere de rectificaciones en periodos de tiempo determinados por el problema específico, ya que para tiempos suficientemente grandes la órbita no perturbada y la perturbada pueden diferir significativamente, en especial si hay efectos seculares. En consecuencia se suele rectificar la órbita osculatriz de referencia pasando a otra órbita osculatriz y tal vez otro tipo de rectificaciones a medida que transcurre el tiempo.

## 17.4. Ejemplos del uso de la teoría de perturbación dependiente del tiempo

### 17.4.1. Oscilador armónico como perturbación de la partícula libre

Consideremos como sistema no perturbado el movimiento libre unidimensional de una partícula puntual  $m$ . Aunque el ejemplo es trivial, nos permitirá ilustrar las principales características del método. El Hamiltoniano no perturbado es

$$H_0 = \frac{p^2}{2m} \quad (17.15)$$

para este sistema el momento lineal  $p$  se conserva, denominaremos  $\alpha$  a su valor constante. La ecuación de Hamilton-Jacobi para este sistema es

$$\frac{1}{2m} \left( \frac{\partial S}{\partial x} \right)^2 + \frac{\partial S}{\partial t} = 0$$

dado que  $x$  es cíclica y el Hamiltoniano no depende explícitamente del tiempo, la función principal de Hamilton queda en la forma

$$S(x, \alpha, t) = W(x, \alpha) - Et = \alpha x - \frac{\alpha^2}{2m} t \quad (17.16)$$

elegiremos al nuevo momento canónico  $P$  como la constante de integración  $\alpha$  (tal como en la Ec. 9.8). La nueva coordenada viene dada por la Ec. (9.10)

$$Q \equiv \beta = \frac{\partial S}{\partial \alpha} = x - \frac{\alpha t}{m}$$

que se puede escribir en la forma

$$x = \frac{\alpha t}{m} + \beta \quad (17.17)$$

que es la solución bien conocida para el movimiento unidimensional libre de fuerzas. En síntesis, la transformación canónica para la partícula libre (sistema no perturbado) es

$$(x, p) \rightarrow (\beta, \alpha) \quad ; \quad p = \alpha \quad ; \quad x = \frac{\alpha t}{m} + \beta \quad (17.18)$$

<sup>3</sup>Dependiendo del grado de aproximación requerido se pueden aproximar los parámetros de variación periódica a su valor no perturbado o a su promedio sobre un periodo no perturbado.

Aunque la solución es inmediata, el uso de la técnica de Hamilton-Jacobi nos garantiza que  $\alpha$  y  $\beta$  son variables canónicas.

Ahora agregaremos una interacción tipo Hooke que consideraremos como perturbación. El Hamiltoniano de perturbación será entonces

$$\Delta H = \frac{m\omega^2 x^2}{2} \quad (17.19)$$

siendo  $\omega$  una constante. El Hamiltoniano perturbado será

$$H = H_0 + \Delta H = \frac{1}{2m} (p^2 + m^2\omega^2 x^2) \quad (17.20)$$

Debemos proceder ahora a calcular el nuevo Hamiltoniano perturbado  $K$ , en términos de  $\alpha, \beta, t$ , para lo cual sustituimos las ecuaciones (17.15), (17.16), (17.19), en (17.2)

$$\begin{aligned} K(\alpha, \beta, t) &= H_0 + \Delta H + \frac{\partial S}{\partial t} = \frac{\alpha^2}{2m} + \frac{m\omega^2 x^2}{2} - \frac{\alpha^2}{2m} \\ K(\alpha, \beta, t) &= \frac{m\omega^2 x^2}{2} \end{aligned}$$

y usamos (17.17) para escribir el miembro derecho en términos de  $\alpha, \beta, t$

$$K(\alpha, \beta, t) = \frac{m\omega^2}{2} \left( \frac{\alpha t}{m} + \beta \right)^2 \quad (17.21)$$

para este caso muy particular es posible obtener la solución exacta sustituyendo el Hamiltoniano (17.21) en las Ecs. (17.3)

$$\dot{\alpha} = -m\omega^2 \left( \frac{\alpha t}{m} + \beta \right) \quad ; \quad \dot{\beta} = \omega^2 t \left( \frac{\alpha t}{m} + \beta \right) \quad (17.22)$$

derivando la primera de las Ecs. (17.22) con respecto al tiempo resulta

$$\ddot{\alpha} = -\omega^2 \alpha - m\omega^2 \left( \dot{\beta} + \frac{\dot{\alpha} t}{m} \right) \quad (17.23)$$

y sustituyendo las dos Ecs. (17.22) en (17.23) se tiene

$$\begin{aligned} \ddot{\alpha} &= -\omega^2 \alpha - m\omega^2 \left[ \omega^2 t \left( \frac{\alpha t}{m} + \beta \right) - m\omega^2 \left( \frac{\alpha t}{m} + \beta \right) \frac{t}{m} \right] \\ \ddot{\alpha} &= -\omega^2 \alpha \end{aligned}$$

lo cual nos da una variación armónica simple con el tiempo para  $\alpha$  en el sistema perturbado. Reemplazando (17.17) en la primera de las ecuaciones (17.22) resulta

$$\dot{\alpha} = -m\omega^2 x \Rightarrow x = -\frac{\dot{\alpha}}{m\omega^2}$$

que fácilmente se puede ver que conduce a la solución armónica exacta. Es importante enfatizar en que este es un caso excepcional en el cual se puede obtener la solución exacta usando la transformación canónica obtenida del sistema no perturbado. El punto es que aunque esta misma transformación canónica no trivializa al Hamiltoniano perturbado, lo deja con una estructura suficientemente simple.

Veamos ahora como se obtendrían las soluciones aproximadas en cada orden en teoría de perturbaciones. Vale decir que para poder tratar a esta interacción como perturbación, es necesario que la energía potencial armónica ( $\Delta H$ ) sea mucho menor que la energía cinética de la partícula ( $H_0$ ). Es claro que esto implicaría que el método podría fallar por ejemplo en los instantes en los cuales la elongación es muy cercana a la amplitud, ya que en estos casos la energía cinética es muy pequeña. Veremos sin embargo que para este caso muy particular,

la solución completa se podrá obtener via teoría de perturbaciones, aunque esto constituye de nuevo un caso excepcional.

Para la aproximación de primer orden, procedemos a escribir las Ecs. (17.4)

$$\dot{\alpha}_1 = - \left. \frac{\partial K(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} ; \quad \dot{\beta}_1 = \left. \frac{\partial K(\alpha, \beta, t)}{\partial \alpha} \right|_{\alpha_0, \beta_0}$$

donde los subíndices  $\alpha_0, \beta_0$  son los valores no perturbados. Por simplicidad tomemos  $x_0 = 0$  lo cual según la Ec. (17.17) implica que  $\beta_0 = 0$ <sup>4</sup>. Para el Kamiltoniano (17.21) resulta

$$\begin{aligned} \dot{\alpha}_1 &= -m\omega^2 \left( \frac{\alpha t}{m} + \beta \right) \Big|_{\alpha_0, \beta_0} ; & \dot{\beta}_1 &= \omega^2 t \left( \frac{\alpha t}{m} + \beta \right) \Big|_{\alpha_0, \beta_0} \\ \dot{\alpha}_1 &= -\omega^2 \alpha_0 t ; & \dot{\beta}_1 &= \omega^2 \frac{\alpha_0 t^2}{m} \end{aligned}$$

estas ecuaciones tienen soluciones inmediatas

$$\alpha_1 = \alpha_{10} - \frac{\omega^2 \alpha_0 t^2}{2} ; \quad \beta_1 = \frac{\alpha_0 \omega^2 t^3}{3m} + \beta_{10} \quad (17.24)$$

donde  $\alpha_{10}$  y  $\beta_{10}$  denotan los valores iniciales de  $\alpha_1$  y  $\beta_1$ . Debemos tener en cuenta que las condiciones iniciales son las mismas para el sistema perturbado que para el sistema sin perturbar. La diferencia es que en el sistema sin perturbar estos valores iniciales  $\alpha_0, \beta_0$  permanecen inalterados en el tiempo, en tanto que en el sistema perturbado ellos evolucionarán en el tiempo. En consecuencia, los valores iniciales de los  $\alpha$  y  $\beta$  en cualquier orden en teoría de perturbaciones corresponden a los valores sin perturbar, por tanto

$$\alpha_{n0} = \alpha_0 ; \quad \beta_{n0} = \beta_0 = 0$$

con lo cual la Ec. (17.24) queda

$$\alpha_1 = \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} ; \quad \beta_1 = \frac{\alpha_0 \omega^2 t^3}{3m} \quad (17.25)$$

las soluciones para  $x$  y  $p$  a primer orden se obtienen de la transformación canónica para el sistema no perturbado (partícula libre) descrito por la Ec. (17.18). Por tanto, tenemos a primer orden

$$\begin{aligned} p_1 &= \alpha_1 = \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} ; & x_1 &= \frac{\alpha_1}{m} t + \beta_1 = \frac{\left( \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} \right)}{m} t + \frac{\alpha_0 \omega^2 t^3}{3m} = \frac{\alpha_0}{m} t - \frac{\omega^2 \alpha_0 t^3}{2m} + \frac{\alpha_0 \omega^2 t^3}{3m} \\ p_1 &= \alpha_0 \left( 1 - \frac{\omega^2 t^2}{2} \right) ; & x_1 &= \frac{\alpha_0}{m\omega} \left[ \omega t - \frac{\omega^3 t^3}{6} \right] \end{aligned}$$

ahora escribamos la solución de segundo orden, para lo cual usamos las ecuaciones (17.6)

$$\begin{aligned} \dot{\alpha}_2 &= - \left. \frac{\partial K(\alpha, \beta)}{\partial \beta} \right|_{\alpha_1, \beta_1} ; & \dot{\beta}_2 &= \left. \frac{\partial K(\alpha, \beta, t)}{\partial \alpha} \right|_{\alpha_1, \beta_1} \\ \dot{\alpha}_2 &= -m\omega^2 \left( \frac{\alpha t}{m} + \beta \right) \Big|_{\alpha_1, \beta_1} ; & \dot{\beta}_2 &= \omega^2 t \left( \frac{\alpha t}{m} + \beta \right) \Big|_{\alpha_1, \beta_1} \\ \dot{\alpha}_2 &= -m\omega^2 \left( \frac{\alpha_1 t}{m} + \beta_1 \right) ; & \dot{\beta}_2 &= \omega^2 t \left( \frac{\alpha_1 t}{m} + \beta_1 \right) \end{aligned} \quad (17.26)$$

<sup>4</sup>Nótese que el subíndice "0" en  $x_0$  y  $\beta_0$  tiene significados muy diferentes, ya que  $x_0$  indica la posición del sistema en  $t = 0$  en tanto que  $\beta_0$  indica el valor (constante) de  $\beta$  en el sistema no perturbado i.e. el valor de  $\beta$  para la perturbación de orden cero.

y reemplazando los valores de (17.25) en (17.26) resulta

$$\begin{aligned}\dot{\alpha}_2 &= -m\omega^2 \left[ \left( \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} \right) \frac{t}{m} + \frac{\alpha_0 \omega^2 t^3}{3m} \right] ; \quad \dot{\beta}_2 = \omega^2 t \left[ \left( \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} \right) \frac{t}{m} + \frac{\alpha_0 \omega^2 t^3}{3m} \right] \\ \dot{\alpha}_2 &= -\alpha_0 \omega^2 \left[ \left( t - \frac{\omega^2 t^3}{2} \right) + \frac{\omega^2 t^3}{3} \right] ; \quad \dot{\beta}_2 = \frac{\alpha_0 \omega^2}{m} \left[ \left( t^2 - \frac{\omega^2 t^4}{2} \right) + \frac{\omega^2 t^4}{3} \right] \\ \dot{\alpha}_2 &= -\alpha_0 \omega^2 \left[ t - \frac{\omega^2 t^3}{6} \right] ; \quad \dot{\beta}_2 = \frac{\alpha_0 \omega^2}{m} \left[ t^2 - \frac{\omega^2 t^4}{6} \right]\end{aligned}$$

cuyas soluciones son

$$\alpha_2 = \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} + \frac{\omega^4 \alpha_0 t^4}{24} ; \quad \beta_2 = \frac{\alpha_0 \omega^2}{m} \left[ \frac{t^3}{3} - \frac{\omega^2 t^5}{30} \right]$$

para las soluciones a segundo orden en  $x$  y  $p$  recurrimos nuevamente a las transformaciones canónicas derivadas del Hamiltoniano no perturbado, Ecs. (17.18)

$$\begin{aligned}p_2 &= \alpha_2 ; \quad x_2 = \frac{\alpha_2 t}{m} + \beta_2 \\ p_2 &= \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} + \frac{\omega^4 \alpha_0 t^4}{24} ; \quad x_2 = \left( \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2} + \frac{\omega^4 \alpha_0 t^4}{24} \right) \frac{t}{m} + \frac{\alpha_0 \omega^2}{m} \left[ \frac{t^3}{3} - \frac{\omega^2 t^5}{30} \right] \\ p_2 &= \alpha_0 \left[ 1 - \frac{\omega^2 t^2}{2} + \frac{\omega^4 t^4}{24} \right] ; \quad x_2 = \frac{\alpha_0}{m\omega} \left( \omega t - \frac{\omega^3 t^3}{2} + \frac{\omega^5 t^5}{24} \right) + \frac{\alpha_0}{m\omega} \left[ \frac{\omega^3 t^3}{3} - \frac{\omega^5 t^5}{30} \right] \\ p_2 &= \alpha_0 \left[ 1 - \frac{\omega^2 t^2}{2!} + \frac{\omega^4 t^4}{4!} \right] ; \quad x_2 = \frac{\alpha_0}{m\omega} \left( \omega t - \frac{\omega^3 t^3}{6} + \frac{\omega^5 t^5}{120} \right)\end{aligned} \quad (17.27)$$

adicionalmente, de las Ecs. (17.18) se observa que  $\alpha_0 = p$ , siendo  $p$  el momento lineal constante asociado al sistema sin perturbar (partícula libre). Es más conveniente reescribir  $\alpha_0 = p_0$  para indicar que está asociado a la condición inicial en el momento lineal tanto del sistema perturbado como sin perturbar. Por otra parte, redefiniremos  $p$  como el momento lineal del sistema perturbado, el cual ya no es constante. Con estas consideraciones, reescribimos  $x$  y  $p$  a segundo orden en la Ec. (17.27) en la forma

$$p_2 = p_0 \left[ 1 - \frac{(\omega t)^2}{2!} + \frac{(\omega t)^4}{4!} \right] ; \quad x_2 = \frac{p_0}{m\omega} \left[ \omega t - \frac{(\omega t)^3}{3!} + \frac{(\omega t)^5}{5!} \right]$$

con estos términos es suficiente para observar la secuencia que nos lleva a la contribución del término  $n$ -ésimo, aunque la demostración formal debe realizarse por inducción (ver ejercicio 1, Pág. 522). En las expresiones a segundo orden vemos que los términos entre corchetes para  $p_2$  y  $x_2$  son los tres primeros términos en la expansión de coseno y seno respectivamente. En el límite de orden infinito en la perturbación resulta claramente

$$p = \lim_{n \rightarrow \infty} p_n = p_0 \cos \omega t \quad ; \quad x = \lim_{n \rightarrow \infty} x_n = \frac{p_0}{m\omega} \sin \omega t \quad (17.28)$$

que son las soluciones exactas compatibles con la condiciones iniciales.

### 17.4.2. Péndulo plano con amplitud finita

Un péndulo con pequeñas oscilaciones se considera como un oscilador armónico simple, este sistema tiene la propiedad de ser isócrono, es decir el periodo es independiente de la amplitud. No obstante, en el régimen de amplias oscilaciones la energía potencial del sistema se aleja del valor del potencial armónico y puesto que la solución del sistema armónico simple es exacta, es lógico considerar que el problema no perturbado sea el péndulo con potencial armónico de modo que la perturbación correspondería a la desviación del potencial del péndulo con respecto al potencial armónico. Deduiremos el cambio en la frecuencia debido a esta perturbación

a través de la variación temporal del ángulo de fase perturbado. Un péndulo plano consistente en una masa puntual  $m$  situado al extremo de una varilla sin peso de longitud  $l$  posee el Hamiltoniano

$$H = \frac{p^2}{2ml^2} + mgl(1 - \cos \theta) \quad (17.29)$$

siendo  $p$  el momento conjugado a  $\theta$ , ( $p$  tiene unidades de momento angular). Este es el Hamiltoniano perturbado ya que incluye el efecto de amplias oscilaciones. Podemos obtener una forma simplificada aunque aproximada del Hamiltoniano de perturbación, realizando una expansión en serie de Taylor del  $\cos \theta$  en el Hamiltoniano

$$\begin{aligned} H &= \frac{p^2}{2ml^2} + mgl \left[ 1 - \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots \right) \right] = \frac{p^2}{2ml^2} + mgl \left[ \frac{\theta^2}{2} - \frac{\theta^4}{24} + \frac{\theta^6}{720} - \dots \right] \\ H &= \frac{p^2}{2ml^2} + \frac{mgl\theta^2}{2} \left[ 1 - \frac{\theta^2}{12} + \frac{\theta^4}{360} - \dots \right] \end{aligned} \quad (17.30)$$

el límite de pequeñas oscilaciones se obtiene manteniendo solo el primer término de la expansión de la energía potencial en (17.30), con lo cual se vé claramente un potencial de tipo armónico. Es claro que el Hamiltoniano es la energía del sistema incluso para el problema perturbado. En el momento de máxima altura i.e. cuando se alcanza el valor de la amplitud angular  $\theta_1$ , la energía cinética se anula, por tanto en el régimen de pequeñas oscilaciones se tiene que:

$$H_0 = E_0 = \frac{mgl\theta_1^2}{2} \quad (17.31)$$

Lo cual sugiere que  $\theta_1$  se pueda usar como parámetro adimensional para analizar el efecto de los términos perturbativos. Usaremos dos parámetros para realizar dicho análisis

$$\theta_1^2 = \frac{2E_0}{mgl} \quad ; \quad \lambda = \frac{\theta_1^2}{6} = \frac{E_0}{3mgl} \quad (17.32)$$

usando (17.32) el Hamiltoniano (17.30) se escribe entonces en la forma

$$\begin{aligned} H &= \frac{p^2}{2ml^2} + \frac{mgl\theta^2}{2} \left[ 1 - 6 \left( \frac{\theta_1^2}{6} \right) \frac{\theta^2}{12\theta_1^2} + 6^2 \left( \frac{\theta_1^2}{6} \right)^2 \frac{\theta^4}{360\theta_1^4} - \dots \right] \\ H &= \frac{p^2}{2ml^2} + \frac{mgl\theta^2}{2} \left[ 1 - \frac{\lambda}{2} \left( \frac{\theta}{\theta_1} \right)^2 + \frac{\lambda^2}{10} \left( \frac{\theta}{\theta_1} \right)^4 - \dots \right] \end{aligned} \quad (17.33)$$

recordemos que  $\theta_1$  es la amplitud angular máxima de oscilación cuando  $E$  y por tanto la amplitud son pequeñas. Por tanto el cociente  $\theta/\theta_1$  es del orden de uno para valores de  $\theta$  cercanos a  $\theta_1$  i.e. de la amplitud máxima. En consecuencia, el orden de magnitud del factor  $\lambda$  será el que regule la convergencia de la serie (17.33)<sup>5</sup>. Si conservamos solo el primer término de corrección y trabajamos perturbaciones a primer orden, solo se introducirán en los cálculos términos del orden de  $\lambda$ . La perturbación a segundo orden con el mismo Hamiltoniano de perturbación introduce términos de orden  $\lambda^2$ , por lo tanto para obtener una corrección coherente hay que calcular una perturbación de segundo orden en el término proporcional a  $\lambda$  del Hamiltoniano y una perturbación de primer orden en el término  $\lambda^2$  de dicho Hamiltoniano.

A continuación realizaremos un cálculo perturbativo de primer orden. El Hamiltoniano no perturbado lo podemos escribir en la forma

$$\begin{aligned} H_0 &= \frac{p^2}{2ml^2} + \frac{mgl\theta^2}{2} = \frac{p^2}{2(ml^2)} + \frac{mgl\theta^2}{2(ml^2)} (ml^2) = \frac{p^2}{2I} + \frac{1}{2} \left( \frac{mgl}{I} \right) \theta^2 I = \frac{p^2}{2I} + \frac{1}{2} \omega^2 \theta^2 I \\ H_0 &= \frac{1}{2I} (p^2 + I^2 \omega^2 \theta^2) \quad ; \quad I \equiv ml^2 \quad ; \quad \omega^2 \equiv \frac{mgl}{I} = \frac{g}{l} \end{aligned} \quad (17.34)$$

<sup>5</sup>De (17.32) podemos observar que  $\theta_1$  y  $\lambda$  son pequeños cuando la energía total  $E$  es mucho menor que la energía potencial máxima  $2mgl$  que podría alcanzar el péndulo.

donde  $I$  y  $\omega$  denotan el momento de inercia del péndulo y su frecuencia angular respectivamente. Como ya vimos en la sección 9.9.1, un sistema adecuado de variables canónicas que nos conduce a un  $K_0$  completamente cíclico es el sistema de variables acción ángulo  $(J, w)$  definido por las Ecs. (9.110, 9.113) Pág. 195

$$w = vt + \beta \quad ; \quad v = \frac{\omega}{2\pi} \quad ; \quad J = \frac{2\pi E}{\omega} \quad (17.35)$$

en el sistema no perturbado la variable canónica  $J$  es constante y la variación temporal de  $w$  es lineal, lo cual se traduce en que las cantidades  $J, v, \beta$  son constantes en el sistema no perturbado. Lo que hace el Hamiltoniano de perturbación es introducir una variación temporal para estas cantidades. En el presente análisis vamos a absorber toda la modificación de la dependencia temporal de  $w$  en el parámetro  $\beta$  con lo cual escribiremos<sup>6</sup>

$$\dot{w}_1 = \dot{v}_1 t + v_1 + \dot{\beta}_1 \equiv v + \dot{\beta}'_1 \quad (17.36)$$

de modo que estudiaremos la variación temporal generada por la perturbación de  $J$  y del término de fase  $\beta'_1$  que absorbe la dependencia temporal de  $w_1$ .

Las Ecs. (9.115) nos daban la transformación canónica entre las variables originales  $q, p$  y las finales  $w$  y  $J$  del oscilador armónico unidimensional. Hay no obstante una diferencia sutil entre el Hamiltoniano (9.109) y el Hamiltoniano (17.29) ya que en el primero las coordenadas  $q, p$  tienen dimensiones de longitud y momento lineal en tanto que en el segundo las coordenadas  $\theta, p$  tienen unidades de ángulo (adimensional) y momento angular. Para escribir adecuadamente las transformaciones canónicas basta con reemplazar  $m \rightarrow I = ml^2$  como se ve comparando los Hamiltonianos (9.109) y (17.34), de esta forma la transformación canónica (9.115) queda

$$\theta = \sqrt{\frac{J}{\pi I \omega}} \sin 2\pi w \quad ; \quad p = \sqrt{\frac{I J \omega}{\pi}} \cos 2\pi w \quad ; \quad w = vt + \beta \quad (17.37)$$

estas son ecuaciones exactas para el sistema no perturbado. Ahora bien, dentro de la filosofía de la teoría de perturbaciones dependiente del tiempo (variación de constantes), la transformación canónica continúa siendo válida pero los valores de  $J$  y  $w$  adquieren una dependencia temporal.

El nuevo Hamiltoniano no perturbado  $K_0$  se puede extraer de las Ecs. (9.111, 9.112) de modo que  $K_0 = Jv$ . Por otro lado, el nuevo Hamiltoniano perturbado coincide numéricamente con el Hamiltoniano perturbado original puesto que la función generatriz  $W$  no depende explícitamente del tiempo<sup>7</sup>, solo hay que escribirlo en términos de las nuevas variables canónicas  $J, w$ . Usando el Hamiltoniano (17.33) hasta orden  $\lambda$  se obtiene

$$\begin{aligned} K &= K_0 + \Delta K \\ \Delta K &= -\frac{mgl\theta^2}{2} \frac{\lambda}{2} \left(\frac{\theta}{\theta_1}\right)^2 = -\frac{mgl\theta^2}{4} \frac{\theta_1^2}{6} \left(\frac{\theta}{\theta_1}\right)^2 = -\frac{mgl\theta^4}{24} = -\frac{mgl}{24} \left(\frac{J}{\pi I \omega}\right)^2 \sin^4 2\pi w \\ \Delta K &= -\frac{mgl}{24} \left(\frac{J}{\pi ml^2 \omega}\right)^2 \sin^4 2\pi w = -\frac{gJ^2}{24\pi^2 ml^3 \omega^2} \sin^4 2\pi w = -\frac{gJ^2}{24\pi^2 ml^3 (g/l)} \sin^4 2\pi w \end{aligned}$$

donde hemos usado (17.32, 17.34) y (17.37), con lo anterior queda finalmente

$$K = K_0 + \Delta K \quad ; \quad K_0 = Jv \quad ; \quad \Delta K = -\frac{J^2}{24\pi^2 ml^2} \sin^4 2\pi w \quad ; \quad w = vt + \beta \quad (17.38)$$

ahora aplicamos las Ecs. (17.4, 17.36) para obtener

$$\dot{J}_1 = -\left. \frac{\partial K(J, w)}{\partial w} \right|_{J_0, w_0} \quad ; \quad \dot{w}_1 = v + \dot{\beta}'_1 = \left. \frac{\partial K(J, w)}{\partial J} \right|_{J_0, w_0} \quad (17.39)$$

<sup>6</sup>Esto equivale a usar la frecuencia del sistema no perturbado para los cálculos en el sistema perturbado.

<sup>7</sup>Recordemos que el formalismo de variables acción-ángulo desarrollado en la sección 9.8, supone que el Hamiltoniano es constante y se utiliza el formalismo de Hamilton-Jacobi para la función característica de Hamilton de la sección 9.3, en la cual la función generatriz  $W$  no depende explícitamente del tiempo y el nuevo Hamiltoniano  $K$  coincide numéricamente con el antiguo Hamiltoniano  $H$ .

con lo que obtenemos

$$\dot{\beta}'_1 = \frac{\partial}{\partial J} \left[ Jv - \frac{J^2}{24\pi^2 ml^2} \sin^4 2\pi w \right] \Big|_{J_0, w_0} - v = \left[ v - \frac{J_0}{12\pi^2 ml^2} \sin^4 2\pi w \right] - v$$

la expresión para  $\dot{\beta}'_1$  queda finalmente

$$\dot{\beta}'_1 = -\frac{J_0}{12\pi^2 ml^2} \sin^4 2\pi (vt + \beta_0) \quad (17.40)$$

donde hemos usado la notación  $J_0$  y  $\beta_0$  para enfatizar que corresponde a los valores no perturbados. La ecuación (17.40) nos dice que en primera aproximación  $\dot{\beta}'_1$  varía a lo largo del ciclo de la oscilación sin perturbar. Nótese sin embargo que  $\dot{\beta}'_1$  posee un valor promedio no nulo a lo largo de un periodo no perturbado, haciendo el cambio de variable  $x = 2\pi vt + 2\pi\beta_0$

$$\begin{aligned} \overline{\dot{\beta}'_1} &= -\frac{J_0}{12\pi^2 ml^2} \frac{1}{(1/v)} \int_0^{1/v} \sin^4 2\pi (vt + \beta_0) dt = -\frac{J_0}{12\pi^2 ml^2} \frac{1}{(1/v)} \frac{1}{2\pi v} \int_{2\pi\beta_0}^{2\pi+2\pi\beta_0} \sin^4 x dx \\ \overline{\dot{\beta}'_1} &= -\frac{J_0}{24\pi^3 ml^2} \int_0^{2\pi} \sin^4 x dx = -\frac{J_0}{24\pi^3 ml^2} \left( \frac{3}{4}\pi \right) \end{aligned}$$

quedando finalmente

$$\overline{\dot{\beta}'_1} = -\frac{J_0}{32\pi^2 ml^2} \quad (17.41)$$

por lo cual la variación de  $\beta'_1$  corresponde a una perturbación secular. A primer orden, el incremento por periodo se da a una razón constante dada por la Ec. (17.41). Para tiempos mucho mayores a un periodo no perturbado, la variación temporal de  $\beta'_1$  se puede aproximar en la forma

$$\beta'_1 \simeq \overline{\dot{\beta}'_1} t + \beta_0 \quad \text{para } t \gg \frac{1}{\nu} \quad (17.42)$$

insertando (17.42) en la Ec. (17.37) obtenemos

$$\theta \simeq \sqrt{\frac{J_1}{\pi I \omega}} \sin 2\pi \left( vt + \overline{\dot{\beta}'_1} t + \beta_0 \right) ; \quad p = \sqrt{\frac{I J_1 \omega}{\pi}} \cos 2\pi \left( vt + \overline{\dot{\beta}'_1} t + \beta_0 \right) \quad \text{para } t \gg \frac{1}{\nu} \quad (17.43)$$

con lo cual vemos que al menos en promedio la solución de primer orden sigue siendo armónica simple con frecuencia

$$v' = v + \overline{\dot{\beta}'_1} \quad (17.44)$$

podemos observar que  $v'$  es una corrección para la frecuencia, pero no para el parámetro  $v$ . Nótese que además  $J_1$  introduce una ligera modificación en las amplitudes en (17.43). Usando la tercera de las Ecs. (17.35)

$$J_0 = \frac{2\pi\omega E_0}{\omega^2} = 2\pi\omega \frac{E_0 l}{g}$$

que al reemplazar en (17.41) resulta

$$\overline{\dot{\beta}'_1} = -\frac{2\pi\omega E_0 l}{32g\pi^2 ml^2} = -\frac{\omega E_0}{16\pi mgl} = -\frac{2\pi v E_0}{16\pi mgl} = -\frac{v}{16} \frac{2E_0}{mgl} = -\frac{v\theta_1^2}{16}$$

donde hemos usado (17.32). La variación relativa de primer orden de la frecuencia para una amplitud finita  $\theta_1$  está dada por

$$\frac{\Delta v}{v} = \frac{v' - v}{v} = \frac{\overline{\dot{\beta}'_1}}{v} = -\frac{\theta_1^2}{16} \quad (17.45)$$



a partir de la Ec. (17.38) y de la primera de las Ecs. (17.39), se puede obtener la variación temporal de  $J$  en primera aproximación

$$\dot{J}_1 = - \left. \frac{\partial K(J, w)}{\partial w} \right|_{J_0, w_0} = - \left. \frac{\partial}{\partial w} \left( Jv - \frac{J^2}{24\pi^2 m l^2} \sin^4 2\pi w \right) \right|_{J_0, w_0} = 2\pi \frac{4J_0^2}{24\pi^2 m l^2} \sin^3 2\pi w_0 \cos 2\pi w_0$$

$$\dot{J}_1 = \frac{J_0^2}{3\pi m l^2} \sin^3 [2\pi (vt + \beta_0)] \cos [2\pi (vt + \beta_0)]$$

se ve de inmediato que la variación de  $J_1$  es periódica y no secular ya que su valor promedio se anula incluso para medio periodo. Físicamente esto era de esperarse puesto que  $J$  es una medida de la amplitud de las oscilaciones y dado que la perturbación no es de naturaleza disipativa (o “donativa”), ésta no sería tal que produzca una disminución o aumento sistemático de la amplitud en cada periodo.

Por otro lado, si estamos interesados en la corrección a primer orden para tiempos menores o del orden de un periodo, ya no será válida la aproximación (17.42), y  $\beta_1'$  deberá obtenerse con base en la Ec. (17.40).

### 17.4.3. Perturbaciones en el problema de Kepler acotado

Cuando al potencial Kepleriano se le agrega un potencial proporcional a  $1/r^2$  puede demostrarse que la órbita resultante es una elipse con respecto a un sistema coordenado rotante. La elipse gira y su periápside presenta precesión. Encontraremos la velocidad de precesión usando teoría de perturbaciones de primer orden para potenciales de la forma  $1/r^n$  con  $n$  entero y  $n \geq 2$ . El potencial y el Hamiltoniano de perturbación se escribirán en la forma

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V \quad ; \quad V = -\frac{k}{r} - \frac{h}{r^n} \quad ; \quad \Delta H = -\frac{h}{r^n} \quad ; \quad n \geq 2$$

donde  $h$  posee un valor pequeño tal que el segundo término sea mucho menor al primero y se pueda considerar una perturbación. Nuevamente, consideremos el problema de Kepler no perturbado en el lenguaje de las variables acción-ángulo. En el problema no perturbado la posición angular del periápside en el plano de la órbita viene dada por la constante  $\omega = 2\pi w_2$  (ver Ec. 10.154). La perturbación genera una dependencia temporal para  $\omega$  que generará la precesión del periápside, esta dependencia temporal se determina a través de las ecuaciones

$$\dot{w}_2 = \frac{\dot{\omega}}{2\pi} = \left. \frac{\partial K}{\partial J_2} \right|_{J_{20}, \beta_0} = \left. \frac{\partial (K_0 + \Delta K)}{\partial J_2} \right|_{J_{20}, \beta_0}$$

teniendo en cuenta que los Hamiltonianos perturbado y no perturbado  $H$  y  $H_0$  coinciden numéricamente con los nuevos Hamiltonianos  $K$  y  $K_0$  respectivamente, y que  $K_0$  solo depende de  $J_3$  (ver Ec. 10.136, Pág. 254) tenemos que

$$\begin{aligned} \dot{w}_2 &= \frac{\dot{\omega}}{2\pi} = \frac{\partial (\Delta K)}{\partial J_2} = \frac{\partial (\Delta H)}{\partial J_2} \\ \dot{w}_2 &= \frac{\dot{\omega}}{2\pi} = \frac{1}{2\pi} \frac{\partial (\Delta H)}{\partial l} \end{aligned} \quad (17.46)$$

donde hemos usado la relación  $J_2 = 2\pi l$  siendo  $l$  el módulo del momento angular (ver Ec. 10.137, Pág. 255). En la práctica el valor de  $\dot{\omega}$  es muy pequeño de modo que la variación de  $\omega$  es casi imperceptible a lo largo de un periodo. Por lo tanto, resulta más interesante medir la variación secular de  $\omega$  tomada sobre muchos periodos y es de poco interés el valor instantáneo de la variación de  $\omega$ . De lo anterior se sigue que la cantidad de interés es el promedio de  $\dot{\omega}$  realizada sobre un periodo de la órbita no perturbada, que según la Ec. (17.46) se escribe como

$$\bar{\dot{\omega}} \equiv \frac{1}{\tau} \int_0^\tau \frac{\partial (\Delta H)}{\partial l} dt$$

la derivada se puede sacar de la integral puesto que  $\tau$  solo es función de  $J_3$  como se vé en las Ecs. (9.174, 9.182), en tanto que la derivación es con respecto a  $l = J_2/2\pi$ . Por tanto

$$\bar{\omega} = \frac{\partial}{\partial l} \left( \frac{1}{\tau} \int_0^\tau K dt \right) = \frac{\partial (\overline{\Delta H})}{\partial l} \quad (17.47)$$

ahora bien, el promedio temporal del Hamiltoniano de perturbación es

$$\overline{\Delta H} = -h \overline{\left( \frac{1}{r^n} \right)} = -\frac{h}{\tau} \int_0^\tau \frac{dt}{r^n}$$

usamos ahora la conservación del momento angular en la forma de la Ec. (10.100)  $l dt = mr^2 d\theta$  escribimos entonces la integral en términos de  $\theta$

$$\overline{\Delta H} = -\frac{mh}{l\tau} \int_0^{2\pi} \frac{d\theta}{r^{n-2}} = -\frac{mh}{l\tau} \left( \frac{mk}{l^2} \right)^{n-2} \int_0^{2\pi} [1 + \varepsilon \cos(\theta - \theta')]^{n-2} d\theta \quad (17.48)$$

donde hemos escrito a  $r$  en función de  $\theta$  usando la ecuación de la órbita (10.85). Solo los términos que contengan potencias pares de la excentricidad darán contribuciones no nulas a la integral. Vale decir además que la derivada con respecto a  $l$  también contiene potencias de  $\varepsilon$ , puesto que  $\varepsilon$  solo es función de  $J_2$  y  $J_3$ .

Hay dos casos particularmente interesantes para los posibles valores de  $n$ , el primero es el caso en el cual  $n = 2$ . Reemplazando  $n = 2$  en (17.48), el promedio del Hamiltoniano de perturbación es

$$\overline{\Delta H} = -\frac{2\pi mh}{l\tau} \quad ; \quad n = 2 \quad (17.49)$$

y sustituyendo (17.49) en (17.47) la velocidad promedio de precesión secular es

$$\bar{\omega} = \frac{2\pi mh}{l^2\tau} \quad ; \quad n = 2 \quad (17.50)$$

Otro caso de interés es  $n = 3$  para el cual las cantidades anteriores están dadas por

$$\begin{aligned} \overline{\Delta H} &= -\frac{mh}{l\tau} \left( \frac{mk}{l^2} \right) \int_0^{2\pi} [1 + \varepsilon \cos(\theta - \theta')] d\theta \\ \overline{\Delta H} &= -\frac{2\pi m^2 hk}{l^3\tau} \quad ; \quad \bar{\omega} = \frac{6\pi m^2 hk}{l^4\tau} \end{aligned} \quad (17.51)$$

este caso es de gran interés ya que la teoría de la relatividad general predice una corrección del movimiento Newtoniano con un potencial asociado de la forma  $h/r^3$ . Más específicamente, corresponde a la solución esféricamente simétrica de Schwarzschild de las ecuaciones de campo de Einstein, la cual corresponde a un Hamiltoniano de perturbación para el problema de Kepler de la forma

$$\Delta H = \frac{h}{r^3} \quad ; \quad h = \frac{kl^2}{m^2c^2} \quad (17.52)$$

siendo  $c$  la velocidad de la luz en el vacío. Reemplazando este resultado en (17.51) se obtiene

$$\bar{\omega} = \frac{6\pi k^2}{l^2\tau c^2} \quad (17.53)$$

veamos una estimación numérica para la precesión secular de un cuerpo que gira alrededor del sol. En tal caso  $k = GmM$  y utilizamos la ecuación (10.96) para la elipse no perturbada

$$l^2 = mka(1 - \varepsilon^2)$$

con lo cual la Ec. (17.53) queda en la forma

$$\begin{aligned} \bar{\omega} &= \frac{6\pi k^2}{mka(1-\varepsilon^2)\tau c^2} = \frac{6\pi}{\tau(1-\varepsilon^2)} \frac{k}{mac^2} = \frac{6\pi}{\tau(1-\varepsilon^2)} \frac{GmM}{mac^2} = \frac{6\pi}{\tau(1-\varepsilon^2)} \frac{(GM/c^2)}{a} \\ \bar{\omega} &= \frac{6\pi}{\tau(1-\varepsilon^2)} \left(\frac{R}{a}\right) \quad ; \quad R \equiv \frac{GM}{c^2} \approx 1,4766 \text{ km} \end{aligned} \tag{17.54}$$

donde  $R$  es el llamado radio gravitatorio del sol. Para el planeta Mercurio el valor numérico de (17.54) se estima en

$$\bar{\omega} \approx 42,98''/\text{siglo}$$

la precesión secular de este planeta es mas de 100 veces mayor que este valor, la razón estriba en que el efecto dominante es la precesión de los equinoccios ( $\sim 90\%$ ), seguido por el efecto perturbativo de los otros planetas ( $\sim 10\%$ ), el efecto aquí calculado debido a la relatividad general equivale a  $\sim 1\%$  del efecto total.

Vale la pena aclarar que al hallar  $\omega$  no hemos puesto a actuar sobre  $h$  la derivada con respecto a  $l$  tal como lo sugeriría la Ec. (17.52). La razón es que  $h$  no depende funcionalmente de  $l$  como momento canónico, lo que nos dice la Ec. (17.52) es como se determina el valor de la constante  $h$  en función del valor del parámetro orbital  $l$ . El potencial de perturbación es función de las variables dinámicas solo a través de  $r$ , no está construido en una forma dependiente de la velocidad.

### 17.5. Teoría de perturbación independiente del tiempo a primer orden con un grado de libertad

En la teoría de perturbación dependiente del tiempo, resolvemos el problema no perturbado por el método de Hamilton Jacobi, generando unas constantes de movimiento. La filosofía general era tener en cuenta que en el problema perturbado estas cantidades serán ahora dependientes del tiempo y se determina su dependencia temporal.

En contraposición, la teoría de perturbación independiente del tiempo tiene como objetivo buscar las cantidades que serán constantes en el sistema perturbado. Este formalismo es solo aplicable a sistemas que sean conservativos, separables y periódicos tanto en el estado perturbado como en el no perturbado. A pesar de estas restricciones, hay un gran número de problemas que se pueden tratar con este formalismo. Por ejemplo, las trayectorias planetarias son corregidas por las interacciones conservativas de los otros planetas y en general no se pierde la naturaleza periódica del movimiento. De hecho, es frecuente que la perturbación no altere la naturaleza periódica del movimiento.

Por simplicidad, comenzaremos con un solo grado de libertad para la teoría a primer orden. Describiremos el sistema periódico no perturbado con variables acción ángulo  $(w_0, J_0)$  y un Hamiltoniano  $H_0(J_0)$ . Sabemos que la frecuencia no perturbada y el valor de la variable angular están dadas por las ecuaciones (9.105, 9.107), Pág. 193

$$v_0 = \frac{\partial H_0(J_0)}{\partial J_0} \quad ; \quad w_0 = v_0 t + \beta_0 \tag{17.55}$$

recordemos que en este formalismo cuando se completa un periodo completo, la variable  $w_0$  aumenta en la unidad. Esta periodicidad permite desarrollar a  $q$  y a  $p$  en una expansión de Fourier (ver Ec. 9.134, Pág. 200)

$$q = \sum_{k=-\infty}^{\infty} A_k(J_0) e^{2\pi i k w_0} \quad ; \quad p = \sum_{k=-\infty}^{\infty} B_k(J_0) e^{2\pi i k w_0} \tag{17.56}$$

Agreguemos ahora una perturbación de la forma

$$H = H_0(J_0) + \varepsilon H_1 \tag{17.57}$$

siendo  $\varepsilon$  un parámetro pequeño que podemos variar continuamente a partir de cero. Dado que el sistema perturbado sigue siendo periódico, la idea es encontrar un nuevo conjunto de variables acción ángulo  $(w, J)$  para tal sistema con  $H = H(J)$ , que cumplan relaciones análogas a (17.55)

$$v = \frac{\partial H(J)}{\partial J} \quad ; \quad w = vt + \beta \quad (17.58)$$

debemos tener en cuenta que de todas maneras  $(w_0, J_0)$  siguen siendo variables canónicas para el sistema perturbado, puesto que la transformación  $(q, p) \rightarrow (w_0, J_0)$  es canónica independientemente de la forma particular del Hamiltoniano como ya se ha mencionado en varias ocasiones. No obstante  $J_0$  ya no es constante y  $w_0$  ya no es función lineal del tiempo. Sin embargo, la relación (17.56) sigue siendo válida en el sistema perturbado ya que esta ecuación expresa la transformación canónica  $(q, p) \rightarrow (w_0, J_0)$ , simplemente los valores de  $w_0$  y  $J_0$  deben ser modificados apropiadamente. En conclusión,  $q$  y  $p$  siguen siendo funciones periódicas de  $w_0$  con lo cual se deduce que cuando  $q$  y  $p$  realizan un periodo completo del movimiento, tanto  $w$  como  $w_0$  deben avanzar una unidad.

El hecho de que los dos conjuntos  $(w_0, J_0)$  y  $(w, J)$  formen variables canónicas para el problema perturbado, significa que deben estar conectados por una transformación canónica con función generatriz  $W(w_0, J)$  que se puede hallar mediante la ecuación de Hamilton Jacobi para la función característica de Hamilton. El hecho de que el Hamiltoniano de perturbación sea pequeño se debe manifestar en que la transformación canónica generada por  $W(w_0, J)$  solo se desvíe en una pequeña cantidad con respecto a la identidad. Por tanto, tomando la función generatriz tipo 2 para la identidad Ec. (7.22) Pág. 125, podemos escribir a primer orden en  $\varepsilon$

$$W(w_0, J) = w_0 J + \varepsilon W_1(w_0, J) \quad (17.59)$$

ahora bien, la primera de las Ecs. (7.19), Pág. 124 en este caso queda

$$\begin{aligned} J_0 &= \frac{\partial W}{\partial w_0} = \frac{\partial [w_0 J + \varepsilon W_1(w_0, J)]}{\partial w_0} \\ J_0 &= J + \varepsilon \frac{\partial W_1}{\partial w_0} \equiv J + \varepsilon J_1 \end{aligned} \quad (17.60)$$

la frecuencia perturbada  $v$  se puede calcular simplemente encontrando  $H$  como función de  $J$  y aplicando (17.58) sin resolver el problema del movimiento completo como es característico del formalismo en variables acción ángulo. La ecuación de Hamilton-Jacobi para la función característica de Hamilton Ec. (9.35), Pág. 176 queda en la forma

$$H(w_0, J_0) = H\left(w_0, \frac{\partial W}{\partial w_0}\right) = \alpha \quad (17.61)$$

$H$  será en general función de  $J$  y de  $\varepsilon$ , de modo que podemos escribir a primer orden en  $\varepsilon$

$$H = \alpha(J, \varepsilon) = \alpha_0(J) + \varepsilon \alpha_1(J) \quad (17.62)$$

usando (17.57) y (17.62) en (17.61), la ecuación de Hamilton Jacobi queda en la forma

$$H_0\left(\frac{\partial W}{\partial w_0}\right) + \varepsilon H_1\left(w_0, \frac{\partial W}{\partial w_0}\right) = \alpha_0(J) + \varepsilon \alpha_1(J) \quad (17.63)$$

nótese sin embargo, que  $H_0$  y  $H_1$  siguen dependiendo de  $\varepsilon$  a través de la función generatriz  $W$ , como se aprecia en (17.59), es necesario tener en cuenta que si el miembro derecho de la Ec. (17.63) es una aproximación a primer orden, debemos conservar términos hasta primer orden también en el miembro izquierdo, lo cual implica expandir  $H_0$  hasta primer orden y  $H_1$  hasta orden cero en  $\varepsilon$ . Teniendo en cuenta (17.59) y usando (17.60),

esto implica que

$$\begin{aligned}
 H_0 \left( \frac{\partial W}{\partial w_0} \right) &= H_0 \left( \frac{\partial [w_0 J + \varepsilon W_1(w_0, J)]}{\partial w_0} \right) = H_0 \left( J + \varepsilon \frac{\partial W_1}{\partial w_0} \right) \approx H_0(J) + \varepsilon \frac{\partial W_1}{\partial w_0} \frac{\partial H_0(J)}{\partial J} \\
 &\approx H_0(J) + \varepsilon \frac{\partial W_1}{\partial w_0} \frac{\partial H_0(J_0 - \varepsilon J_1)}{\partial (J_0 - \varepsilon J_1)} \approx H_0(J) + \varepsilon \frac{\partial W_1}{\partial w_0} \frac{\partial H_0(J_0)}{\partial J_0} \\
 H_0 \left( \frac{\partial W}{\partial w_0} \right) &\approx H_0(J) + \varepsilon v_0 \frac{\partial W_1}{\partial w_0} \\
 H_1 \left( w_0, \frac{\partial W}{\partial w_0} \right) &= H_1 \left( w_0, \frac{\partial [w_0 J + \varepsilon W_1(w_0, J)]}{\partial w_0} \right) \approx H_1 \left( w_0, \frac{\partial [w_0 J]}{\partial w_0} \right) = H_1(w_0, J)
 \end{aligned}$$

usando estas expansiones en (17.63) se obtiene

$$H_0(J) + \varepsilon v_0 \frac{\partial W_1}{\partial w_0} + \varepsilon H_1(w_0, J) = \alpha_0(J) + \varepsilon \alpha_1(J)$$

ahora sí tenemos funciones independientes de  $\varepsilon$  a la izquierda de esta ecuación. Por tanto, podemos igualar potencias en  $\varepsilon$  resultando

$$H_0(J) = \alpha_0(J) \quad (17.64)$$

$$v_0 \frac{\partial W_1}{\partial w_0} + H_1(w_0, J) = \alpha_1(J) \quad (17.65)$$

la ecuación (17.64) se puede obtener también tomando el límite cuando  $\varepsilon \rightarrow 0$ . Nótese que en la Ec. (17.65) el miembro de la izquierda depende de  $J$  y  $w_0$  en tanto que el de la derecha depende solo de  $J$ , por lo cual debe existir una cancelación de la dependencia funcional con  $w_0$  en el miembro izquierdo. La ecuación de transformación que determina a  $w$  (ver Ec. 9.104) es

$$w = \frac{\partial W}{\partial J} = w_0 + \varepsilon \frac{\partial W_1}{\partial J} \quad (17.66)$$

donde hemos usado (17.59). De acuerdo con esta ecuación y el requerimiento de que  $w$  y  $w_0$  avancen en la unidad cuando se completa un periodo, es necesario que  $W_1$  sea función periódica de  $w_0$

$$W_1(w_0, J) = \sum_{k=-\infty}^{\infty} a_k(J) e^{2\pi i k w_0} \quad (17.67)$$

en realidad la forma de  $W_1$  dada por (17.67) es la que nos garantiza que las nuevas variables  $(w, J)$  también sean del tipo acción ángulo<sup>8</sup>. De la Ec. (17.67) se observa que el término con  $k = 0$  para la expansión de  $\partial W_1 / \partial w_0$  (es decir el término constante) debe anularse. Si en la Ec. (17.65) tomamos a ambos lados el promedio sobre un periodo completo de  $w_0$ , tenemos que este promedio ya no dependerá de  $w_0$ .

$$v_0 \overline{\left( \frac{\partial W_1}{\partial w_0} \right)} + \overline{H_1(w_0, J)} = \overline{\alpha_1(J)} = \alpha_1(J) \quad (17.68)$$

Donde hemos usado el hecho de que  $\alpha_1$  solo depende de  $J$ , y por tanto dicho promedio no afectará a esta cantidad. Por otra parte, puede verse que el promedio de  $\partial W_1 / \partial w_0$  es nulo<sup>9</sup>

$$\overline{\left( \frac{\partial W_1}{\partial w_0} \right)} = \int_{w'_0}^{w'_0+1} \frac{\partial W_1}{\partial w_0} dw_0 = W_1(w'_0 + 1, J) - W_1(w'_0, J) = 0 \quad (17.69)$$

<sup>8</sup>Estrictamente, la función que debe ser periódica de acuerdo con la Ec. (17.66), es  $\partial W_1 / \partial J$ . En todo caso, puesto que la derivada es con respecto a  $J$ , ello no afecta la periodicidad con respecto a la variable  $w$  ó  $w_0$ .

<sup>9</sup>Nótese que para que el promedio de  $\partial W_1 / \partial w_0$  sea nulo, es necesario que el término constante en la expansión de esta derivada sea nulo como ya se mencionó.

donde hemos usado la periodicidad de  $W_1$  con respecto a  $w_0$  Ec. (17.67). En consecuencia, la Ec. (17.68) queda

$$\alpha_1(J) = \overline{H_1(w_0, J)} \equiv \overline{H_1}(J) \quad (17.70)$$

por otro lado, usando (17.60) tenemos

$$\begin{aligned} \overline{H_1(w_0, J)} &\simeq \overline{H_1(w_0, J_0)} + (J - J_0) \overline{\left( \frac{\partial H_1(w_0, J)}{\partial J} \right)} \\ \overline{H_1(w_0, J)} &\simeq \overline{H_1(w_0, J_0)} - \varepsilon \overline{\left( \frac{\partial W_1}{\partial w_0} \frac{\partial H_1(w_0, J)}{\partial J} \right)} \end{aligned} \quad (17.71)$$

sumando y restando  $\overline{H_1}$  en (17.65) resulta

$$\alpha_1(J) = \overline{H_1(w_0, J)} - \left[ (\overline{H_1} - H_1) - v_0 \frac{\partial W_1}{\partial w_0} \right] \quad (17.72)$$

combinando las Ecs. (17.70, 17.72) resulta

$$(\overline{H_1} - H_1) - v_0 \frac{\partial W_1}{\partial w_0} = 0 \quad (17.73)$$

y de las Ecs. (17.70, 17.71) se obtiene

$$\alpha_1(J) = \overline{H_1(w_0, J_0)} + \mathcal{O}(\varepsilon) \quad (17.74)$$

y teniendo en cuenta que  $\alpha_1(J)$  es de orden cero en  $\varepsilon$ , las Ecs. (17.73, 17.74) se pueden escribir finalmente en la forma

$$\alpha_1(J) = \overline{H_1(w_0, J)} \simeq \overline{H_1(w_0, J_0)} \quad (17.75)$$

$$\frac{\partial W_1}{\partial w_0} = \frac{\overline{H_1} - H_1}{v_0} \quad (17.76)$$

teniendo en cuenta las Ecs. (17.57, 17.62), se observa que la Ec. (17.75) nos dice como “completar” la dependencia funcional de  $H$  con respecto a  $J$  a primer orden en  $\varepsilon$ . El Hamiltoniano de perturbación se puede expresar en función de las variables sin perturbar  $(w_0, J_0)$  y se promedia para un periodo completo de movimiento. La restante dependencia de  $J_0$  es la misma (hasta ese orden) que la de  $J$  correspondiente a  $\alpha_1$ . La nueva frecuencia se escribe usando (17.58, 17.62)

$$\begin{aligned} v &= \frac{\partial \alpha(J)}{\partial J} = \frac{\partial \alpha_0(J)}{\partial J} + \varepsilon \frac{\partial \alpha_1(J)}{\partial J} \approx \frac{\partial \alpha_0(J_0)}{\partial J_0} + \varepsilon \frac{\partial \alpha_1}{\partial J} \Rightarrow \\ v &\simeq v_0 + \varepsilon \frac{\partial \alpha_1}{\partial J} \end{aligned} \quad (17.77)$$

por otro lado, la Ec. (17.76) nos permite encontrar la función generatriz a primer orden, y por lo tanto la transformación canónica que nos lleva de  $(w_0, J_0)$  a  $(w, J)$  con lo cual el problema queda completamente resuelto a primer orden en  $\varepsilon$ . Específicamente, la dependencia de  $q$  y  $p$  con el tiempo se hallaría por medio de las Ecs. (17.56, 17.58). Ahora bien, como es típico en las variables acción-ángulo, si solo nos interesa la frecuencia perturbada basta con solucionar la Ec. (17.75) y aplicar la Ec. (17.77), no necesitamos resolver el problema completo Ec. (17.76). En la teoría de perturbaciones para mecánica cuántica se encuentra un fenómeno análogo ya que la energía del sistema a primer orden (valor propio de un operador Hamiltoniano) se puede calcular sin tener que encontrar la función de onda perturbada (que equivale a resolver el problema completo).

### 17.5.1. Péndulo plano con oscilación finita usando método de perturbación independiente del tiempo

Ya habíamos escrito el Hamiltoniano de perturbación de este problema en la ecuación (17.38), que lo reescribimos en la forma

$$\varepsilon H_1(w_0, J) = -\frac{J^2}{24\pi^2 ml^2} \sin^4 2\pi w_0$$

donde  $l$  corresponde a la longitud del péndulo. Lo más conveniente para la asignación del parámetro  $\varepsilon$  es la cantidad  $\theta_1^2$  que representa el cuadrado de la amplitud angular de oscilación en el caso no perturbado. Por tanto

$$\varepsilon H_1(w_0, J) \equiv \theta_1^2 H_1(w_0, J) \Rightarrow H_1(w_0, J) = -\frac{J^2}{24\pi^2 \theta_1^2 ml^2} \sin^4 2\pi w_0$$

Para aplicar la Ec. (17.75), debemos calcular el valor medio de este Hamiltoniano promediado sobre un periodo no perturbado

$$\begin{aligned} \alpha_1(J) &= \overline{H_1(w_0, J)} = -\frac{J^2}{24\pi^2 \theta_1^2 ml^2} \frac{1}{\tau_0} \int_0^{\tau_0} \sin^4 [2\pi(v_0 t + \beta_0)] dt \\ &= -\frac{J^2 v_0}{24\pi^2 \theta_1^2 ml^2} \int_0^{1/v_0} \sin^4 [2\pi(v_0 t + \beta_0)] dt \\ \alpha_1(J) &= -\frac{J^2}{24\pi^2 \theta_1^2 ml^2} \left(\frac{3}{8}\right) \end{aligned}$$

esta ecuación junto con (17.77) nos da

$$\alpha_1(J) = \overline{H_1(w_0, J)} = -\frac{J^2}{64\theta_1^2 \pi^2 ml^2} \quad (17.78)$$

$$v - v_0 = \varepsilon \frac{\partial \alpha_1}{\partial J} = -\frac{\varepsilon J}{32\pi^2 ml^2 \theta_1^2} \quad (17.79)$$

en virtud de (17.60), vemos que con el fin de calcular  $v - v_0$  hasta primer orden en  $\varepsilon$ , podemos sustituir  $J$  por  $J_0$ <sup>10</sup>

$$\begin{aligned} v - v_0 &= -\frac{\varepsilon (J_0 - \varepsilon J_1)}{32\pi^2 ml^2 \theta_1^2} \\ v - v_0 &= -\frac{\varepsilon J_0}{32\pi^2 ml^2 \theta_1^2} + \mathcal{O}(\varepsilon^2) \end{aligned} \quad (17.80)$$

Por otro lado, de las Ecs. (9.111, 17.31) y (17.34) se deduce que

$$\begin{aligned} J_0 &= \frac{2\pi\alpha_0}{\omega_0} = \frac{2\pi E_0}{\omega_0} = \frac{\pi mgl\theta_1^2}{\omega_0} = \frac{\pi\theta_1^2}{\omega_0} mgl = \frac{\pi\theta_1^2}{\omega_0} \omega_0^2 I = \pi\theta_1^2 \omega_0 ml^2 \\ J_0 &= 2\pi^2 ml^2 \theta_1^2 v_0 \end{aligned}$$

con lo cual la Ec. (17.80) se escribirá en la forma

$$\begin{aligned} v - v_0 &\simeq -\frac{\varepsilon J_0}{32\pi^2 ml^2 \theta_1^2} = -\frac{2\pi^2 ml^2 \theta_1^2 v_0 \varepsilon}{32\pi^2 ml^2 \theta_1^2} = -\frac{v_0 \varepsilon}{16} \\ \Delta v &\simeq -\varepsilon \frac{v_0}{16} = -\theta_1^2 \frac{v_0}{16} \end{aligned}$$

<sup>10</sup>Esto respalda el hecho de que el promedio  $\overline{H_1(w_0, J)}$  se puede reemplazar a primer orden por el promedio sobre variables no perturbadas  $\overline{H_1(w_0, J_0)}$ , como ya se discutió.

con lo cual la variación relativa de  $v$  será

$$\frac{\Delta v}{v_0} \equiv \frac{v - v_0}{v_0} = -\frac{\theta_1^2}{16}$$

lo cual concuerda con la Ec. (17.45). Nótese que hemos encontrado el corrimiento en la frecuencia, sin resolver la ecuación (17.76) para  $W_1$ .

Por supuesto, es necesario extender el presente formalismo al caso en que hay varios grados de libertad. Adicionalmente, las perturbaciones a orden superior a uno pueden ser de gran importancia, bien porque se requiera una gran precisión en los cálculos o bien porque las contribuciones a primer orden sean nulas (este es el caso del oscilador anarmónico como veremos más adelante). Por estas razones, procederemos a estudiar el caso de órdenes superiores para sistemas de varios grados de libertad.

## 17.6. Teoría de perturbación independiente del tiempo para orden superior al primero y varios grados de libertad

Las condiciones son naturalmente las mismas de la sección anterior: sistemas conservativos periódicos y separables (en realidad la condición de separabilidad solo es no trivial para el caso con varios grados de libertad). Para el sistema no perturbado tenemos el conjunto de variables acción ángulo  $(w_{0i}, J_{0i})$  de modo que

$$H_0 = H_0(J_{0i}) \quad ; \quad w_{0i} = v_{0i}t + \beta_i$$

las variables  $q_k, p_k$  debe ser funciones periódicas múltiples de las  $w_{0i}$ , lo cual en forma compacta se puede escribir en la forma

$$q_k = \sum_{\mathbf{j}} A_{\mathbf{j}}^{(k)}(\mathbf{J}_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} \quad ; \quad p_k = \sum_{\mathbf{j}} B_{\mathbf{j}}^{(k)}(\mathbf{J}_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0}$$

siendo  $\mathbf{j}$ ,  $\mathbf{w}_0$ ,  $\mathbf{J}_0$  vectores  $n$ -dimensionales de los índices enteros, variables acción y coordenadas angulares respectivamente. Nuevamente el conjunto  $(\mathbf{w}_0, \mathbf{J}_0)$  sigue siendo canónico en el sistema perturbado. El Hamiltoniano perturbado en función de estas variables se puede desarrollar en serie con respecto al parámetro  $\varepsilon$

$$H(\mathbf{w}_0, \mathbf{J}_0, \varepsilon) = H_0(\mathbf{J}_0) + \varepsilon H_1(\mathbf{w}_0, \mathbf{J}_0) + \varepsilon^2 H_2(\mathbf{w}_0, \mathbf{J}_0) + \dots \quad (17.81)$$

de nuevo buscamos una transformación canónica que nos lleve desde el conjunto  $(\mathbf{w}_0, \mathbf{J}_0)$  hasta el conjunto  $(\mathbf{w}, \mathbf{J})$  de modo que las  $\mathbf{J}$  sean constantes y las  $\mathbf{w}$  sean lineales en el tiempo. En este nuevo sistema, el Hamiltoniano perturbado solo es función de las  $\mathbf{J}$  y  $\varepsilon$

$$\alpha(\mathbf{J}, \varepsilon) = \alpha_0(\mathbf{J}) + \varepsilon \alpha_1(\mathbf{J}) + \varepsilon^2 \alpha_2(\mathbf{J}) + \dots \quad (17.82)$$

con los valores de  $\alpha_0, \alpha_1, \alpha_2$  etc. se pueden hallar las frecuencias perturbadas hasta un orden dado de  $\varepsilon$ . Con un procedimiento análogo al que nos llevó a la Ec. (17.77) tenemos que

$$\mathbf{v} = \mathbf{v}_0 + \varepsilon \frac{\partial \alpha_1}{\partial \mathbf{J}} + \varepsilon^2 \frac{\partial \alpha_2}{\partial \mathbf{J}} + \dots \quad (17.83)$$

la función generatriz de la transformación canónica  $(\mathbf{w}_0, \mathbf{J}_0) \rightarrow (\mathbf{w}, \mathbf{J})$  se escribe como  $W(\mathbf{w}_0, \mathbf{J}, \varepsilon)$  y su desarrollo en potencias de  $\varepsilon$  se escribe

$$W(\mathbf{w}_0, \mathbf{J}, \varepsilon) = \mathbf{w}_0 \cdot \mathbf{J} + \varepsilon W_1(\mathbf{w}_0, \mathbf{J}) + \varepsilon^2 W_2(\mathbf{w}_0, \mathbf{J}) + \dots$$

esta función generatriz naturalmente es solución de la ecuación de Hamilton Jacobi adecuada

$$H\left(\mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0}, \varepsilon\right) = \alpha(\mathbf{J}, \varepsilon) \quad (17.84)$$



posteriormente se expanden los dos miembros de (17.84) en potencias de  $\varepsilon$  y se igualan coeficientes de las mismas potencias en  $\varepsilon$ . Ilustremos el caso de segundo orden para el cual la expansión de ambos miembros en (17.84) se escribe

$$H_0 \left( \frac{\partial W}{\partial \mathbf{w}_0} \right) + \varepsilon H_1 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) + \varepsilon^2 H_2 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) = \alpha_0(\mathbf{J}) + \varepsilon \alpha_1(\mathbf{J}) + \varepsilon^2 \alpha_2(\mathbf{J}) \quad (17.85)$$

donde hemos usado (17.81, 17.82). Nuevamente cada término del miembro izquierdo depende de  $\varepsilon$  a través de  $\partial W / \partial \mathbf{w}_0$ . El análogo a la Ec. (17.60), se escribe como

$$\mathbf{J}_0 = \frac{\partial W}{\partial \mathbf{w}_0} = \mathbf{J} + \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0}$$

ahora desarrollamos los términos  $H_i$  en serie de Taylor en torno a  $\mathbf{J}_0 = \mathbf{J}$  conservando en  $H_0$  los términos hasta  $\varepsilon^2$ , en  $H_1$  los términos hasta orden  $\varepsilon$  y en  $H_2$  los términos de orden cero, lo último equivale a sustituir  $\mathbf{J}_0 \rightarrow \mathbf{J}$  en  $H_2$ . Escribamos  $H_0$  y  $H_1$  en notación matricial

$$\begin{aligned} H_0 \left( \frac{\partial W}{\partial \mathbf{w}_0} \right) &= H_0 \left( \mathbf{J} + \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \right) = H_0(\mathbf{J}) + \left( \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \right) \cdot \frac{\partial H_0}{\partial \mathbf{J}} + \\ &\quad + \frac{1}{2} \left( \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \right) \cdot \widetilde{\frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}}} \cdot \left( \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \right) + \dots \\ H_0 \left( \frac{\partial W}{\partial \mathbf{w}_0} \right) &= H_0(\mathbf{J}) + \left( \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \right) \cdot \frac{\partial H_0}{\partial \mathbf{J}} + \frac{\varepsilon^2}{2} \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \cdot \widetilde{\frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}}} \cdot \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) + \mathcal{O}(\varepsilon^3) \\ \varepsilon H_1 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) &= \varepsilon H_1(\mathbf{w}_0, \mathbf{J}) + \varepsilon^2 \frac{\partial W_1}{\partial \mathbf{w}_0} \cdot \frac{\partial H_1}{\partial \mathbf{J}} + \mathcal{O}(\varepsilon^3) \\ \varepsilon^2 H_2 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) &= \varepsilon^2 H_2(\mathbf{w}_0, \mathbf{J}) + \mathcal{O}(\varepsilon^3) \end{aligned} \quad (17.86)$$

reemplazando (17.86) en el miembro izquierdo de (17.85) y agrupando las potencias de  $\varepsilon$  se tiene

$$\begin{aligned} H_0 \left( \frac{\partial W}{\partial \mathbf{w}_0} \right) + \varepsilon H_1 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) + \varepsilon^2 H_2 \left( \mathbf{w}_0, \frac{\partial W}{\partial \mathbf{w}_0} \right) &= H_0(\mathbf{J}) + \varepsilon \frac{\partial W_1}{\partial \mathbf{w}_0} \cdot \frac{\partial H_0(\mathbf{J})}{\partial \mathbf{J}} + \varepsilon H_1(\mathbf{w}_0, \mathbf{J}) \\ &\quad + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{w}_0} \cdot \frac{\partial H_0(\mathbf{J})}{\partial \mathbf{J}} + \varepsilon^2 \frac{\partial W_1}{\partial \mathbf{w}_0} \cdot \frac{\partial H_1}{\partial \mathbf{J}} \\ &\quad + \frac{1}{2} \varepsilon^2 \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \cdot \widetilde{\left( \frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}} \right)} \cdot \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \\ &\quad + \varepsilon^2 H_2(\mathbf{w}_0, \mathbf{J}) \end{aligned}$$

Introduciendo estas expansiones en (17.85) e igualando potencias en  $\varepsilon$  se obtiene

$$\alpha_0(\mathbf{J}) = H_0(\mathbf{J}) \quad (17.87)$$

$$\alpha_1(\mathbf{J}) = \mathbf{v}_0 \cdot \frac{\partial W_1}{\partial \mathbf{w}_0} + H_1(\mathbf{w}_0, \mathbf{J}) \quad (17.88)$$

$$\alpha_2(\mathbf{J}) = \mathbf{v}_0 \cdot \frac{\partial W_2}{\partial \mathbf{w}_0} + \Phi_2(\mathbf{w}_0, \mathbf{J}) \quad (17.89)$$

$$\Phi_2(\mathbf{w}_0, \mathbf{J}) \equiv H_2(\mathbf{w}_0, \mathbf{J}) + \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \cdot \left( \frac{\partial H_1}{\partial \mathbf{J}} \right) + \frac{1}{2} \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \cdot \widetilde{\left( \frac{\partial^2 H_0}{\partial \mathbf{J} \partial \mathbf{J}} \right)} \cdot \left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right) \quad (17.90)$$

la ecuación que relaciona a  $\mathbf{w}$  con  $\mathbf{w}_0$ , análoga a la Ec. (17.66) está dada por

$$\mathbf{w} = \frac{\partial W}{\partial \mathbf{J}} = \mathbf{w}_0 + \varepsilon \frac{\partial W_1}{\partial \mathbf{J}} + \varepsilon^2 \frac{\partial W_2}{\partial \mathbf{J}} + \dots \quad (17.91)$$

cuando se completa un periodo de movimiento en el plano  $(q_m, p_m)$ , tanto  $w_m$  como  $w_{0m}$  aumentan en la unidad. Por tanto, cuando se cumple un ciclo en  $(q_m, p_m)$ , las cantidades  $\partial W_k / \partial \mathbf{J}$ , deben retornar a sus valores iniciales para que se cumpla la Ec. (17.91). En consecuencia, todos los términos  $W_k$  deberán ser funciones periódicas de  $\mathbf{w}_0$ , por tanto

$$W_k(\mathbf{w}_0, \mathbf{J}) = \sum_{\mathbf{j}} B_{\mathbf{j}}^{(k)}(\mathbf{J}) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} \quad (17.92)$$

de modo que las expansiones de todas las derivadas  $\partial W_k / \partial w_{0i}$  carecerán de término constante. Tomando promedio sobre un periodo de todas las  $\mathbf{w}_0$  a ambos lados de las Ecs. (17.88, 17.89) y teniendo en cuenta que las  $\alpha_i$  no dependen de  $\mathbf{w}_0$  se tiene que

$$\overline{\alpha_1(\mathbf{J})} = \alpha_1(\mathbf{J}) = \mathbf{v}_0 \cdot \overline{\left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right)} + \overline{H_1(\mathbf{w}_0, \mathbf{J})} \quad (17.93)$$

$$\overline{\alpha_2(\mathbf{J})} = \alpha_2(\mathbf{J}) = \mathbf{v}_0 \cdot \overline{\left( \frac{\partial W_2}{\partial \mathbf{w}_0} \right)} + \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})} \quad (17.94)$$

con el mismo argumento que nos llevó a la Ec. (17.69), tenemos que

$$\overline{\left( \frac{\partial W_1}{\partial \mathbf{w}_0} \right)} = \overline{\left( \frac{\partial W_2}{\partial \mathbf{w}_0} \right)} = 0 \quad (17.95)$$

sustituyendo (17.95) en (17.93, 17.94) resulta

$$\alpha_1(\mathbf{J}) = \overline{H_1(\mathbf{w}_0, \mathbf{J})} \quad (17.96)$$

$$\alpha_2(\mathbf{J}) = \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})} \quad (17.97)$$

Sumando y restando los promedios  $\overline{H_1}, \overline{\Phi_2}$ , las Ecs. (17.87, 17.88, 17.89, 17.90) se pueden escribir en la forma

$$\alpha_0(\mathbf{J}) = H_0(\mathbf{J}) \quad (17.98)$$

$$\alpha_1(\mathbf{J}) = \left[ \mathbf{v}_0 \cdot \frac{\partial W_1}{\partial \mathbf{w}_0} - \overline{H_1} + H_1(\mathbf{w}_0, \mathbf{J}) \right] + \overline{H_1(\mathbf{w}_0, \mathbf{J})} \quad (17.99)$$

$$\alpha_2(\mathbf{J}) = \left[ \mathbf{v}_0 \cdot \frac{\partial W_2}{\partial \mathbf{w}_0} - \overline{\Phi_2} + \Phi_2(\mathbf{w}_0, \mathbf{J}) \right] + \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})} \quad (17.100)$$

y teniendo en cuenta las Ecs. (17.96, 17.97), los términos entre paréntesis cuadrados en las Ecs. (17.99, 17.100) deben anularse y se obtiene

$$\alpha_0(\mathbf{J}) = H_0(\mathbf{J}) \quad ; \quad \alpha_1(\mathbf{J}) = \overline{H_1(\mathbf{w}_0, \mathbf{J})} \quad ; \quad \alpha_2(\mathbf{J}) = \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})} \quad (17.101)$$

$$\mathbf{v}_0 \cdot \frac{\partial W_1}{\partial \mathbf{w}_0} = \overline{H_1} - H_1(\mathbf{w}_0, \mathbf{J}) \quad ; \quad \mathbf{v}_0 \cdot \frac{\partial W_2}{\partial \mathbf{w}_0} = \overline{\Phi_2} - \Phi_2(\mathbf{w}_0, \mathbf{J}) \quad (17.102)$$

en forma unificada, estas ecuaciones se pueden escribir en la forma

$$\alpha_i(\mathbf{J}) = \overline{\Phi_i(\mathbf{w}_0, \mathbf{J})} \quad , \quad \Phi_0 \equiv H_0 \quad , \quad \Phi_1 \equiv H_1 \quad , \quad \Phi_2 \equiv \Phi_2 \quad ; \quad i = 0, 1, 2 \quad (17.103)$$

$$\mathbf{v}_0 \cdot \frac{\partial W_i}{\partial \mathbf{w}_0} = \overline{\Phi_i} - \Phi_i(\mathbf{w}_0, \mathbf{J}) \quad ; \quad i = 1, 2 \quad (17.104)$$

las Ecs. (17.104) son periódicas en  $\mathbf{w}_0$  con media nula. Obsérvese que las ecuaciones (17.83, 17.103, 17.104) se reducen a la forma que se obtuvo para un grado de libertad a primer orden excepto por el hecho de que contienen todas las frecuencias del movimiento no perturbado. Por ejemplo, la Ec. (17.83) es una ecuación para cada componente del arreglo vectorial de frecuencias  $\mathbf{v}_0$ , una componente dada  $\nu_{0k}$  es el periodo correspondiente al plano  $(q_k, p_k)$ , pero como ya se discutió, el movimiento completo no es necesariamente periódico.

La Ec. (17.83) nos muestra que la perturbación de segundo orden en  $\varepsilon$  para las frecuencias depende de  $\alpha_2$ . Por otro lado,  $\alpha_2$  viene dado por las Ecs. (17.103, 17.90), de modo que  $\alpha_2 = \overline{\Phi_2(\mathbf{w}_0, \mathbf{J})}$ . La Ec. (17.90) nos

muestra que  $\Phi_2$  contiene productos de  $\partial W_1/\partial \mathbf{w}_0$  con funciones periódicas en  $\mathbf{w}_0$ , y aunque el promedio de  $\partial W_1/\partial \mathbf{w}_0$  es nulo, no hay garantía de que el promedio de los productos se anule. Esto nos da como consecuencia que  $\alpha_2 = \overline{\Phi_2}$  es en general no nulo. Ahora bien, puesto que  $\Phi_2$  es función de  $W_1$ , vemos que la corrección de segundo orden en las frecuencias requiere conocer la transformación canónica a primer orden. No obstante, la Ec. (17.90) es independiente de  $W_2$  de modo que la corrección de  $W$  a segundo orden no es necesaria para calcular la corrección a segundo orden de la frecuencia<sup>11</sup>. En principio los coeficientes  $B_{\mathbf{j}}^{(1)}(\mathbf{J})$  que definen a  $W_1$  a través de la ecuación (17.92) se pueden hallar directamente de (17.104) con  $i = 1$ .

$$\mathbf{v}_0 \cdot \frac{\partial W_1}{\partial \mathbf{w}_0} = \overline{H}_1 - H_1(\mathbf{w}_0, \mathbf{J}) \quad (17.105)$$

Ahora bien, dado que  $W_1$  es función periódica de las  $\mathbf{w}_0$  según (17.92), tenemos que  $\partial_{\mathbf{w}_0} W_1$  también es función periódica de las  $\mathbf{w}_0$  con coeficiente constante nulo como ya se discutió. Por tanto, el miembro derecho de (17.105) se puede desarrollar en serie de Fourier similar a la Ec. (17.56) o a la Ec. (17.92) pero sin término constante

$$\overline{H}_1 - H_1 = \sum_{\mathbf{j} \neq 0} C_{\mathbf{j}}(\mathbf{J}) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} \quad (17.106)$$

reemplazando la Ec. (17.92) con  $k = 1$  en (17.105), tenemos que el producto punto a la izquierda de (17.105) se puede escribir como

$$\mathbf{v}_0 \cdot \frac{\partial W_1}{\partial \mathbf{w}_0} = 2\pi i \sum_{\mathbf{j} \neq 0} B_{\mathbf{j}}^{(1)}(\mathbf{J}) (\mathbf{j} \cdot \mathbf{v}_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} \quad (17.107)$$

de acuerdo con las Ecs. (17.105, 17.106) y (17.107), tenemos que

$$\sum_{\mathbf{j} \neq 0} C_{\mathbf{j}}(\mathbf{J}) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} = 2\pi i \sum_{\mathbf{j} \neq 0} B_{\mathbf{j}}^{(1)}(\mathbf{J}) (\mathbf{j} \cdot \mathbf{v}_0) e^{2\pi i \mathbf{j} \cdot \mathbf{w}_0} ; \quad \mathbf{j} \neq 0$$

y recurriendo a la independencia lineal de los exponenciales, tenemos que los coeficientes  $B_{\mathbf{j}}^{(1)}$  de la expansión de  $W_1$  se escriben como

$$B_{\mathbf{j}}^{(1)}(\mathbf{J}) = \frac{C_{\mathbf{j}}(\mathbf{J})}{2\pi i (\mathbf{j} \cdot \mathbf{v}_0)}, \quad \mathbf{j} \neq 0 \quad (17.108)$$

vemos que no hemos obtenido el término constante i.e. con  $\mathbf{j} = 0$ , pero en las expresiones para las  $\alpha_i$  Ecs. (17.87, 17.88, 17.89, 17.90) solo intervienen las derivadas  $\partial W_1/\partial \mathbf{w}_0$  y éstas no contendrán términos constantes.

Para órdenes superiores al segundo el procedimiento es análogo, aunque las transformaciones algebraicas serán más complejas. En general, para la transformación de orden  $k$ -ésimo podemos escribir  $\alpha_k$  en la forma

$$\alpha_k(\mathbf{J}) = \mathbf{v}_0 \cdot \frac{\partial W_k}{\partial \mathbf{w}_0} + \Phi_k(\mathbf{w}_0, \mathbf{J}) \quad (17.109)$$

donde el primer término a la derecha de (17.109) viene del término de primera derivada del desarrollo de Taylor de  $H(\mathbf{J}_0)$  alrededor de  $\mathbf{J}_0 = \mathbf{J}$  donde todos los términos de la diferencia  $\mathbf{J}_0 - \mathbf{J}$  se conservan hasta orden  $\varepsilon^k$ . La función  $W_k$  solo aparece en este término de modo que  $\Phi_k$  solo podrá contener las funciones generatrices  $W_m$  para órdenes menores que  $k$ . Por los razonamientos hechos para primer y segundo orden el primer término del miembro derecho en (17.109) tiene media nula cuando se promedia para ciclos completos de  $\mathbf{w}_0$  y por tanto las Ecs. (17.103, 17.104) serán válidas a todos los órdenes. Es claro que para  $k > 2$ ,  $\Phi_k$  se vuelve mucho más complicado que la Ec. (17.90), y contiene a todas las funciones encontradas en los cálculos de los órdenes precedentes.

Veremos más adelante que en presencia de degeneración de las frecuencias del sistema no perturbado, aparecen problemas con esta formulación. De momento veamos un ejemplo sencillo.

<sup>11</sup>De nuevo esto tiene su contrapartida en la teoría cuántica de perturbaciones, ya que la corrección perturbativa de segundo orden para la energía, requiere conocer la función de onda a primer orden.

### 17.6.1. Oscilador anarmónico unidimensional

Consideremos un término de la forma  $q^3$  que se agrega al potencial armónico. Un término de este tipo se encuentra cuando se expande un potencial alrededor de un mínimo local hasta tercer orden en serie de Taylor (e.g. Ec. 15.4 Pág. 418 hasta tercer orden). El término de tercer orden se conoce como primer término anarmónico. El Hamiltoniano se escribe de la forma

$$H = \frac{1}{2m} \left[ p^2 + m^2 \omega_0^2 q^2 \left( 1 + \varepsilon \frac{q}{q_0} \right) \right] \quad ; \quad \omega_0 = 2\pi v_0 = 2\pi \sqrt{\frac{k}{m}}$$

siendo  $\omega_0$  la frecuencia no perturbada y  $q_0$  es una amplitud de referencia que se puede fijar posteriormente y  $\varepsilon$  es un parámetro adimensional pequeño. Separando las potencias en  $\varepsilon$  se obtienen los términos

$$H = H_0 + \varepsilon H_1 \quad ; \quad H_0 = \frac{1}{2m} [p^2 + m^2 \omega_0^2 q^2] \quad ; \quad H_1 = \frac{m\omega_0^2 q^3}{2q_0} \quad ; \quad H_i = 0 \quad \text{para } i \geq 2 \quad (17.110)$$

usaremos como variables canónicas las variables acción-ángulo  $(J_0, w_0)$  del sistema sin perturbar, los términos no nulos de  $H$  de la Ec. (17.110) se pueden escribir en términos de  $w_0, J_0$  usando las Ecs. (9.115), Pág. 196

$$H_0 = \frac{1}{2m} \left[ \frac{mJ_0\omega_0}{\pi} \cos^2 2\pi w_0 + m^2 \omega_0^2 \frac{J_0}{\pi m \omega_0} \sin^2 2\pi w_0 \right] = \frac{J_0\omega_0}{2\pi} [\cos^2 2\pi w_0 + \sin^2 2\pi w_0]$$

$$H_0(w_0, J_0) = J_0 v_0 \quad ; \quad H_1(w_0, J_0) = \frac{m\omega_0^2}{2q_0} \left( \frac{J_0}{\pi m \omega_0} \right)^{3/2} \sin^3 2\pi w_0 \quad (17.111)$$

y las dos primeras Ecs. (17.101) nos dan las contribuciones de orden cero y uno para  $\alpha(J)$

$$\alpha_0(J) = J v_0 \quad ; \quad \alpha_1 = \overline{H_1(w_0, J)} = \frac{m\omega_0^2}{2q_0} \left( \frac{J}{\pi m \omega_0} \right)^{3/2} \overline{\sin^3 2\pi w_0} = 0$$

la anulación de  $\alpha_1$  es inmediata debido al carácter impar del integrando. Vemos que no hay contribución de primer orden, de tal forma que la primera corrección con respecto al caso no perturbado aparece a segundo orden. Para obtener la contribución de segundo orden  $\alpha_2(J)$  se tiene en cuenta que como  $H_0$  es lineal en  $J$  y  $H_2$  se anula, el valor de  $\Phi_2$  en la Ec. (17.90) se reduce a

$$\Phi_2 = \frac{\partial W_1}{\partial w_0} \frac{\partial H_1}{\partial J}$$

la anulación de  $\alpha_1 = \overline{H_1}$  implica que la Ec. (17.105) adquiere la forma

$$\frac{\partial W_1}{\partial w_0} = -\frac{H_1}{v_0}$$

combinando estos dos resultados se llega a

$$\Phi_2 = -\frac{H_1}{v_0} \frac{\partial H_1}{\partial J} = -\frac{1}{2v_0} \frac{\partial (H_1^2)}{\partial J} \quad (17.112)$$

y usando la segunda de las Ecs. (17.111)

$$H_1^2(w_0, J) = \frac{m^2 \omega_0^4}{4q_0^2} \left( \frac{J}{\pi m \omega_0} \right)^3 \sin^6 2\pi w_0 = \frac{\omega_0}{4q_0^2} \frac{J^3}{\pi^3 m} \sin^6 2\pi w_0 = \frac{2\pi v_0}{4q_0^2} \frac{J^3}{\pi^3 m} \sin^6 2\pi w_0$$

$$H_1^2(w_0, J) = \frac{v_0 J^3}{2\pi^2 m q_0^2} \sin^6 2\pi w_0 \quad (17.113)$$

reemplazando (17.113) en (17.112), se obtiene

$$\Phi_2(w_0, J) = -\frac{3J^2}{4\pi^2 m q_0^2} \sin^6 2\pi w_0 \quad (17.114)$$

y reemplazando (17.114) en la tercera de las Ecs. (17.101) nos queda

$$\begin{aligned}\alpha_2(J) &= \overline{\Phi_2(w_0, J)} = -\frac{3J^2}{4\pi^2 m q_0^2} \overline{\sin^6 2\pi w_0} = -\frac{3J^2}{4\pi^2 m q_0^2} \int_0^1 \sin^6 2\pi w_0 dw_0 = -\frac{3J^2}{4\pi^2 m q_0^2} \left(\frac{5}{16}\right) \\ \alpha_2(J) &= -\frac{15J^2}{64\pi^2 m q_0^2}\end{aligned}\quad (17.115)$$

y la frecuencia perturbada hasta segundo orden en  $\varepsilon$  se obtiene a partir de (17.83)

$$v = v_0 + \varepsilon \frac{\partial \alpha_1}{\partial J} + \varepsilon^2 \frac{\partial \alpha_2}{\partial J} = v_0 - \varepsilon^2 \frac{15J}{32\pi^2 m q_0^2} \quad (17.116)$$

recordando que  $q_0$  es una amplitud de referencia, es natural tomarla como la amplitud máxima que tendría el oscilador con una energía dada  $E_0$  en el sistema sin perturbar i.e.

$$\begin{aligned}\frac{m\omega_0^2 q_0^2}{2} &\equiv E_0 = H_0 = Jv_0 \quad \Rightarrow \quad \frac{m\omega_0^2 q_0^2}{2} \equiv \frac{J\omega_0}{2\pi} \\ &\Rightarrow \quad q_0^2 = \frac{J}{\pi m \omega_0}\end{aligned}$$

con esta amplitud de referencia las Ecs. (17.115, 17.116) quedan

$$\begin{aligned}\alpha_2(J) &= -\frac{15\omega_0 J}{64\pi} \\ v &= v_0 - \varepsilon^2 \frac{15\omega_0}{32\pi} = v_0 - \varepsilon^2 \frac{15 \cdot (2\pi v_0)}{32\pi} \\ v &= v_0 \left[1 - \varepsilon^2 \frac{15}{16}\right]\end{aligned}\quad (17.117)$$

con esta amplitud de referencia<sup>12</sup>, podemos decir a partir de la Ec. (17.117) que la variación relativa de segundo orden de la frecuencia está dada por

$$\frac{\Delta v}{v_0} = \frac{v - v_0}{v_0} = \frac{v_0 \left[1 - \varepsilon^2 \frac{15}{16}\right] - v_0}{v_0}$$

quedando finalmente

$$\frac{\Delta v}{v_0} = -\frac{15}{16}\varepsilon^2$$

de modo que el efecto de la perturbación es disminuir ligeramente la frecuencia armónica asociada.

## 17.7. Teoría de perturbaciones independiente del tiempo en presencia de degeneración

En presencia de degeneración, la primera dificultad obvia que aparece es la anulación exacta o aproximada de  $\mathbf{j} \cdot \mathbf{v}_0$  en el denominador de la Ec. (17.108). En realidad los cálculos en presencia de degeneración son muy complejos y continúan siendo tema de investigación, por tanto solo se mencionarán los aspectos cualitativos más relevantes.

Se dice que la degeneración es exacta o propia cuando las frecuencias sin perturbar  $\mathbf{v}_0$  son tales que existan uno o más conjuntos de enteros  $\mathbf{j}$  para los cuales  $\mathbf{j} \cdot \mathbf{v}_0 = 0$  (ver sección 9.10.4). En la sección 9.10.4 veíamos que se puede pasar del conjunto canónico  $(\mathbf{w}_0, \mathbf{J}_0)$  a otro conjunto de variables tipo acción ángulo para el cual las antiguas frecuencias degeneradas corresponden a nuevas frecuencias nulas y las nuevas frecuencias sin

<sup>12</sup>Naturalmente, si cambiamos la amplitud de referencia  $q_0$ , debe cambiar correspondientemente el valor del parámetro perturbativo  $\varepsilon$ .

perturbar no nulas no son degeneradas. Al introducir la perturbación se remueve la degeneración de manera que las frecuencias nulas del sistema no perturbado adquirirán un pequeño valor y por lo tanto corresponden a periodos grandes<sup>13</sup>. Las variables ángulo asociadas se denominan variables **lentas** en contraste con las variables angulares correspondientes a frecuencias no degeneradas que se denominan variables **rápidas**. Los términos de periodo largo pueden aparecer como términos seculares sobre intervalos de tiempo restringidos, a manera de ejemplo sin  $2\pi vt$  se puede considerar función lineal de  $t$  siempre que  $vt \ll 1$ .

En el caso de degeneración exacta, comenzamos usando las variables canónicas  $(\mathbf{w}_0, \mathbf{J}_0)$  y el Hamiltoniano no perturbado es función solamente de las variables  $J_{0i}$  no degeneradas, de modo que aunque la Ec. (17.81) sigue siendo válida, no aparecerán en ella las  $J_{0i}$  degeneradas. Se realiza entonces la transformación canónica del cálculo de la perturbación pero solo para las variables no degeneradas dejando inalteradas las variables degeneradas. El nuevo Hamiltoniano (Ec. 17.82) tiene la forma

$$\alpha(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0, \varepsilon) = \alpha_0(\mathbf{J}) + \varepsilon\alpha_1(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0) + \varepsilon^2\alpha_2(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0) + \dots$$

donde  $\mathbf{w}'_0$  representa las  $m$  variables degeneradas que tienen valor cero en el problema no perturbado y  $\mathbf{J}'_0$  son naturalmente sus momentos conjugados. Las variables no primadas están entonces asociadas a las variables no degeneradas. Se ve entonces que la transformación canónica elimina a las variables rápidas de modo que todo queda en función de las variables lentas. Por otro lado, puesto que  $\alpha$  es cíclica en  $\mathbf{w}$ , los momentos transformados  $\mathbf{J}$  son verdaderas constantes de movimiento y se puede considerar  $\alpha(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0, \varepsilon)$  como un Hamiltoniano de un sistema con  $m$  grados de libertad. Adicionalmente, dado que  $\alpha_0(\mathbf{J})$  es constante, y por tanto independiente de las demás variables, éste término no será necesario para plantear las ecuaciones de movimiento asociadas a las variables  $(\mathbf{w}'_0, \mathbf{J}'_0)$  y se podrán suprimir en  $\alpha$ . En consecuencia, el nuevo Hamiltoniano efectivo es de orden  $\varepsilon$ , y podemos redefinir el Hamiltoniano “no perturbado” como

$$\varepsilon\alpha_1(\mathbf{J}, \mathbf{J}'_0, \mathbf{w}'_0)$$

y en este problema no perturbado  $\mathbf{w}'_0$  ya no consiste en valores nulos. Si solo hay una condición de degeneración, el problema efectivo es de un solo grado de libertad y formalmente integrable. Con la presencia de más de una degeneración se suele buscar una segunda transformación canónica para eliminar los términos de variables lentas de forma similar a como se eliminaron las variables rápidas, este procedimiento es muy complejo en la práctica.

Puede aducirse a partir de la Ec. (17.108) que incluso cuando las frecuencias no son degeneradas, aparecerán inevitablemente valores pequeños del divisor  $\mathbf{j} \cdot \mathbf{v}_0$  a medida que los índices  $\mathbf{j}$  se hacen más grandes. Este fenómeno se conoce como resonancia, e implica que la amplitud de uno de los términos de Fourier se haga muy grande. Sin embargo, esto no supone un gran problema para la mayoría de Hamiltonianos de perturbación que se usan en la práctica. Para ver esto, notemos que en (17.108) lo que cuenta es el factor completo

$$\frac{C_{\mathbf{j}}}{\mathbf{j} \cdot \mathbf{v}_0} \quad (17.118)$$

donde  $C_{\mathbf{j}}$  es el coeficiente de Fourier del Hamiltoniano de perturbación  $H_1$  como se vé en (17.106). En mecánica celeste ocurre que la mayoría de Hamiltonianos de perturbación poseen la **característica de D'Alembert** que en términos eurísticos nos dice que cuando los valores de los enteros en  $\mathbf{j}$  son mayores que el exponente del parámetro  $\varepsilon$  en el Hamiltoniano, las magnitudes de  $C_{\mathbf{j}}$  caen rápidamente, usualmente en forma exponencial, cuando aumentan los valores de los índices  $\mathbf{j}$ . Los cocientes del tipo dado en (17.118) no crecen entonces demasiado y la expansión converge cuando las frecuencias  $\mathbf{v}_0$  satisfacen una condición de inconmensurabilidad.

El comportamiento resonante cuando se tiene la característica de D'Alembert o en general cuando se cumple la condición

$$\frac{C_{\mathbf{j}}}{(\mathbf{j} \cdot \mathbf{v}_0)} < O(\sqrt{\varepsilon})$$

<sup>13</sup>Debemos recordar que las frecuencias nulas no perturbadas no eran frecuencias físicas sino una manifestación de la degeneración. En tal sentido, sus pequeños valores perturbados están relacionados no con frecuencias físicas perturbadas sino con desdoblamientos de frecuencias que eran degeneradas antes de la introducción de la perturbación.

se conoce como resonancia superficial. Estas resonancias en general no son un problema para la expansión perturbativa. En ocasiones sin embargo, el anterior cociente se hace grande, i.e. mayor que  $O(\sqrt{\varepsilon})$  en cuyo caso hablamos de resonancias profundas. Cuando aparecen resonancias profundas se requiere de técnicas especiales tales como el desarrollo de Bohlin en el cual se expande en potencias de  $\sqrt{\varepsilon}$  en lugar de potencias de  $\varepsilon$ . Un estudio profundo de este tema requiere un tratamiento exhaustivo de los fenómenos de resonancia en oscilaciones alineales.

## 17.8. Algunos aspectos cualitativos generales de la teoría clásica de perturbaciones

Hemos mencionado que la teoría de perturbaciones ha evolucionado junto con la mecánica celeste desde los tiempos de Newton. De otra parte, la exploración espacial y la tecnología desarrollada alrededor de ésta, junto con el rápido desarrollo de los recursos computacionales, ha impulsado enormemente el desarrollo de los métodos perturbativos. A manera de ejemplo, la corrección de la relatividad general estudiada en la sección (17.4.3) resulta lo suficientemente grande para la precisión requerida en los sistemas de posicionamiento global satelital (GPS).

A grandes rasgos se distingue entre los métodos perturbativos generales que conducen a fórmulas analíticas y métodos de perturbación especiales, que nos llevan a respuestas numéricas correspondientes al sistema perturbado. Hay también métodos mixtos de modo que esta clasificación es en realidad aproximada. Los métodos numéricos solo alcanzaron un desarrollo considerable con el desarrollo de las calculadoras y más adelante de los computadores.

Vale mencionar sin embargo, que las soluciones numéricas no pueden resolver cuestiones referentes a la estabilidad a largo plazo de los sistemas perturbados. En esa dirección, un debate de gran interés concierne a la estabilidad del sistema solar. Por ejemplo, las perturbaciones podrían llevar a colisiones interplanetarias o a que un planeta abandone el sistema solar. Por esta razón, desde el principio del desarrollo de los métodos perturbativos existió un gran interés en estudiar los términos perturbativos seculares. Por ejemplo, si se identifica un término de perturbación secular para el semieje mayor de la trayectoria de un planeta, se concluiría que el planeta abandonará el sistema solar o colapsará hacia el sol. Si bien al principio se demostró que no existían términos seculares a primer orden, los trabajos de Poincaré en donde demuestra que las series de Fourier multiperiodicas que aparecen a orden superior solo son semiconvergentes llevó a una indeterminación en el problema de la estabilidad.

La cuestión de la estabilidad solo se aclaró a raíz de la aplicación más bien reciente de técnicas matemáticas avanzadas. Por medio del uso de la topología diferencial es posible examinar el comportamiento global de las posibles órbitas en el espacio de fase. Con base en estos estudios se demostró que es posible el movimiento acotado estable de un sistema de  $n$  cuerpos bajo interacción exclusivamente gravitatoria. Específicamente un grupo de medida finita de órbitas están confinadas a regiones concretas del espacio de fase y permanecen así indefinidamente en movimiento cuasi periódico. Debe decirse sin embargo, que aunque algunas órbitas (un conjunto de medida finita) mantienen un movimiento cuasiperiódico permanente, la inmensa mayoría de condiciones iniciales conducen a movimientos ergódicos es decir que llenan completamente una cierta región del espacio de fase, de modo que la trayectoria en el espacio de fase pasa por un punto dado en el interior de esta región en uno o más instantes de tiempo.

De lo anterior se deduce que el problema de la estabilidad no está del todo resuelto puesto que solo un conjunto de condiciones iniciales permite este movimiento estable y adicionalmente las fuerzas disipativas y otras fuerzas no gravitatorias (e.g. interacciones magnéticas entre los planetas y el sol) deben ser tenidas en cuenta en tal estudio. El sistema puede ser además muy sensible a las condiciones iniciales o a efectos de cuerpos fuera del sistema solar.

## 17.9. Invariantes adiabáticos

Consideremos una lenteja que se suspende de modo que la cuerda pasa por un agujero que se practica en el techo, y le damos el impulso necesario para que forme un péndulo plano. Ahora comenzamos a tirar de la cuerda muy lentamente hacia arriba o hacia abajo, de modo que a lo largo de un periodo hay una ligera variación de la longitud de la cuerda. La energía del péndulo no se conserva ya que el agente externo realiza trabajo sobre el sistema (o lo extrae de él) al cambiar la longitud de la cuerda. Un análisis sencillo muestra que si la variación de la longitud es muy lenta (i.e. adiabática) el cociente entre la energía y la frecuencia de oscilación  $E/\nu$  permanece constante. Este cociente será precisamente la variable de acción  $J$  de una formulación en variables acción ángulo. Decimos entonces que  $J$  es invariante adiabático y extenderemos este fenómeno a otros sistemas Físicos y diremos que hay **invarianza adiabática** de las variables de acción cuando dicha cantidad permanece constante (al menos a primer orden) cuando ocurre una variación lenta de los parámetros del sistema. Esta propiedad era de gran interés para los pioneros de la mecánica cuántica, ya que los primeros procesos de cuantización involucraban la cuantización de la variable de acción para describir estados de sistemas atómicos. Por otro lado, ya se sabía que la variación lenta de perturbaciones externas, por ejemplo de los campos electromagnéticos que rodean al átomo, no inducen transiciones entre estados. La invarianza adiabática resultaba entonces una propiedad muy importante para explicar este fenómeno.

El concepto de invarianza adiabática tiene múltiples aplicaciones en temas como Física del plasma, aceleradores de partículas, astronomía galáctica etc. Este constituye un tema de interés e investigación actual y aquí solo daremos los fundamentos de este desarrollo.

Consideraremos por simplicidad sistemas de un solo grado de libertad que contengan un parámetro  $a$  que se variará adiabáticamente. Implícitamente asumimos que el sistema es inicialmente conservativo con  $a$  constante en el tiempo. En un cierto tiempo, se “enciende” la variación temporal de  $a$ , es necesario aclarar que  $a$  se asumirá como función exclusivamente del tiempo y no dependerá por ejemplo de posiciones o velocidades. Una vez encendida la variación temporal de  $a$ , éste variará lentamente por mucho tiempo y llegará a tener de nuevo un valor constante. Cuando  $a$  sea constante, el movimiento será periódico y la variación lenta del parámetro no alterará la naturaleza periódica del movimiento. Incluso si las variaciones son pequeñas en un periodo completo, pueden haber cambios apreciables para tiempos suficientemente largos. La variación temporal de  $a$  se considera entonces una perturbación y buscamos variaciones seculares del movimiento.

Cuando el parámetro  $a$  es constante en el tiempo, el sistema se describe por variables acción ángulo  $(w_0, J_0)$  y el Hamiltoniano es función de  $J_0$  y  $a$ ,  $H = H(J_0, a)$ . Es conveniente considerar que el par  $(w_0, J_0)$  viene de una transformación canónica a partir de  $(q, p)$  con una función generatriz que denotamos por  $W^*(q, w_0, a)$ , esta función generatriz es de tipo 1. Si la transformación canónica se obtiene por Hamilton Jacobi, lo que se obtiene es una función generatriz de tipo 2 que denotamos por  $W(q, J_0, a)$  y que está conectada con  $W^*$  por una transformación de Legendre Ec. (7.21)

$$W^*(q, w_0, a) = W(q, J_0, a) - J_0 w_0 \quad (17.119)$$

la conveniencia de usar una función de tipo  $F_1$  en lugar de la función tipo  $F_2$  que sale de la solución de Hamilton Jacobi, proviene del hecho de que  $W^*$  es función periódica de  $w_0$  como veremos más adelante<sup>14</sup>. Cuando  $a$  varía en el tiempo  $(w_0, J_0)$  continúan siendo variables canónicas pero la función generatriz ahora depende explícitamente del tiempo por medio de la dependencia temporal de  $a$ . Por tanto el Hamiltoniano perturbado para el sistema  $(w_0, J_0)$  vendrá dado por la Ec. (7.17) Pág. 123

$$\begin{aligned} K(w_0, J_0, a) &= H(J_0, a) + \frac{\partial W^*}{\partial t} = H(J_0, a) + \frac{\partial a}{\partial t} \frac{\partial W^*}{\partial a} \\ K(w_0, J_0, a) &= H(J_0, a) + \dot{a} \frac{\partial W^*}{\partial a} \end{aligned} \quad (17.120)$$

nótese que en el último paso hemos usado el hecho de que  $a$  solo depende del tiempo. Ahora  $J_0$  no es constante y  $w_0$  ya no es función lineal del tiempo. El segundo término a la derecha de (17.120) es un Hamiltoniano

<sup>14</sup>Ver también la discusión en la Pág. 202, después de la Ec. (9.142).



de perturbación. Nótese que hemos usado la misma transformación canónica para generar dos Hamiltonianos diferentes a partir del mismo Hamiltoniano original  $H(q, p)$ . Primero usando la transformación canónica con  $a$  constante llegamos de  $H(q, p, a)$  a  $H(J_0, a)$ ; ambos Hamiltonianos coinciden numéricamente ya que la función característica no depende explícitamente del tiempo. Posteriormente, usamos la misma transformación canónica pero ahora con  $a$  variable en el tiempo para ir desde  $H(q, p, a(t))$  hacia  $K(w_0, J_0, a(t))$ , en este caso los Hamiltonianos no coinciden numéricamente ya que  $W^*$  adquiere dependencia explícita con el tiempo<sup>15</sup> por medio de  $a(t)$ . De lo anterior se concluye que la relación entre los Hamiltonianos  $K(w_0, J_0, a)$  y  $H(J_0, a)$  en (17.120) no proviene de una transformación canónica adicional, sino más bien de la diferencia entre la misma TC cuando  $a$  se considera constante y cuando  $a$  se considera variable en el tiempo.

La dependencia temporal de  $J_0$  viene regida por una de las ecuaciones de Hamilton

$$\dot{J}_0 = -\frac{\partial K}{\partial w_0} = -\dot{a} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right) \quad (17.121)$$

donde hemos aplicado (17.120). Al igual que en el método de variación de constantes, tanto la derivada entre paréntesis como el nuevo Hamiltoniano  $K$  se expresan en términos de  $J_0$ , es decir del valor de  $J$  a orden cero en perturbaciones. Dentro de un desarrollo perturbativo a primer orden buscaremos un término secular, que sería el valor medio de  $\dot{J}_0$  a lo largo de un periodo del movimiento no perturbado para un valor apropiado de  $a$ . La variación adiabática de  $a$ , nos permite considerar al parámetro  $\dot{a}$  como una constante a lo largo de un periodo y el valor medio se puede escribir en la forma

$$\overline{\dot{J}_0} = -\frac{1}{\tau} \int_0^\tau \dot{a} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right) dt = -\frac{\dot{a}}{\tau} \int_0^\tau \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial a} \right) dt + \mathcal{O}(\dot{a}^2, \ddot{a}) \quad (17.122)$$

De la Ec. (9.16) recordamos que  $W$  se puede ver como una acción abreviada indefinida Ec. (9.16) Pág. 171

$$W = \int p dq$$

y teniendo en cuenta la definición de  $J_0$  Ec. (9.102), Pág. 193, vemos que en un periodo de movimiento, la función generatriz  $W$  aumenta en una cantidad  $J_0$ . Similarmente,  $J_0 w_0$  también aumenta en  $J_0$  puesto que  $w_0$  aumenta en la unidad. En consecuencia, los términos a la derecha en la Ec. (17.119) cancelan su variación en un periodo, con lo cual se obtiene que  $W^*$  es función periódica de  $w_0$  y tanto ella como su derivada con respecto a  $a$ , pueden expresarse como expansión de Fourier

$$\frac{\partial W^*}{\partial a} = \sum_k A_k(J_0, a) e^{2\pi i k w_0} \quad (17.123)$$

con lo cual el valor medio de  $\dot{J}_0$  se escribe

$$\overline{\dot{J}_0} = -\frac{\dot{a}}{\tau} \int_0^\tau \sum_{k \neq 0} 2\pi i k A_k(J_0, a) e^{2\pi i k w_0} dt + \mathcal{O}(\dot{a}^2, \ddot{a})$$

la integral se anula puesto que el integrando no tiene ningún término constante. Se obtiene entonces finalmente

$$\overline{\dot{J}_0} = \mathcal{O}(\dot{a}^2, \ddot{a}) \quad (17.124)$$

de modo que  $\overline{\dot{J}_0}$  no tendrá variación secular de primer orden en  $\dot{a}$ , por tanto tenemos la propiedad de invarianza adiabática a primer orden. Es importante insistir en el hecho de que  $a$  debe ser un parámetro que solo dependa del tiempo, de modo que sea independiente de las variables canónicas.

<sup>15</sup>Podría decirse que la dependencia temporal de  $W^*$  es en realidad implícita por medio de  $a(t)$ . No obstante, dado que  $a(t)$  es un parámetro pero no es una variable canónica, esta dependencia se considera explícita para efectos de las Ecuaciones de Hamilton. Es decir, cuando el tiempo se mueve manteniendo constantes las variables  $w_0, J_0$  del problema,  $W^*$  cambia numéricamente (en virtud del cambio en el parámetro  $a(t)$ ), por lo cual hablamos de dependencia explícita. Por tanto, este formalismo requiere que el parámetro  $a(t)$  sea independiente de las variables canónicas.

### 17.9.1. Invarianza adiabática del oscilador armónico

Veamos como se traduce el anterior formalismo para el caso del oscilador armónico

$$H = \frac{1}{2m} (p^2 + m^2 \omega^2 q^2)$$

consideraremos que  $\omega$  puede ser función del tiempo, lo cual en el ejemplo que se explicó en la sección 17.9 equivale a la variación de la longitud del péndulo. Por tanto,  $\omega$  será nuestro parámetro  $a$  de variación adiabática. La transformación canónica que lleva desde  $(q, p)$  hasta  $(w_0, J_0)$  viene dada por las Ecs. (9.115)

$$q = \sqrt{\frac{J_0}{\pi m \omega}} \sin 2\pi w_0 \quad ; \quad p = \sqrt{\frac{m \omega J_0}{\pi}} \cos 2\pi w_0 \quad (17.125)$$

las cuales reescribiremos de forma que faciliten el cálculo de  $W^*$ . Para ello tendremos en cuenta que las ecuaciones diferenciales para una función tipo 1 Ecs. (7.16) se escriben en nuestro contexto en la forma

$$J_0 = -\frac{\partial W^*(q, w_0, \omega)}{\partial w_0} \quad ; \quad p = \frac{\partial W^*(q, w_0, \omega)}{\partial q}$$

por tanto manipularemos las Ecs. (17.125) de tal forma que  $J_0$  se escriba en función de  $q, w_0$  y el parámetro  $\omega$ , y similarmente debemos escribir  $p = p(q, w_0, \omega)$  para poder resolver las ecuaciones de  $W^*$

$$q^2 = \frac{J_0}{\pi m \omega} \sin^2 2\pi w_0 \quad \Rightarrow \quad J_0(q, w_0, \omega) = \frac{\pi m \omega q^2}{\sin^2 2\pi w_0} \quad (17.126)$$

$$p(q, w_0, \omega) = \sqrt{\frac{m \omega}{\pi} \left( \frac{\pi m \omega q^2}{\sin^2 2\pi w_0} \right)} \cos 2\pi w_0 \quad (17.127)$$

las ecuaciones diferenciales para  $W^*$  quedan entonces

$$J_0 = \pi m \omega q^2 \csc^2 2\pi w_0 = -\frac{\partial W^*}{\partial w_0} \quad (17.128)$$

$$p = m \omega q \cot 2\pi w_0 = \frac{\partial W^*}{\partial q} \quad (17.129)$$

$W^*$  se puede hallar por integración directa de (17.128, 17.129) excepto por constantes irrelevantes, pues recordemos que lo que importa es hallar una solución y no se necesita que tal solución sea única (y no lo es).

$$\begin{aligned} W^*(q, w_0, \omega) &= \frac{m \omega q^2}{2} \cot 2\pi w_0 \Rightarrow \\ \frac{\partial W^*}{\partial \omega} &= \frac{m q^2}{2} \cot 2\pi w_0 \end{aligned} \quad (17.130)$$

y reemplazando la primera de las Ecs. (17.126) en (17.130), obtendremos a  $W^*$  en función de  $w_0, J_0$  y  $\omega$

$$\begin{aligned} \frac{\partial W^*}{\partial \omega} &= \frac{m}{2} \left( \frac{J_0}{\pi m \omega} \sin^2 2\pi w_0 \right) \cot 2\pi w_0 = \frac{J_0}{2\pi \omega} \sin^2 2\pi w_0 \frac{\cos 2\pi w_0}{\sin 2\pi w_0} \\ \frac{\partial W^*}{\partial \omega} &= \frac{J_0}{2\pi \omega} \sin 2\pi w_0 \cos 2\pi w_0 = \frac{J_0}{2\pi \omega} \frac{\sin 4\pi w_0}{2} \\ \frac{\partial W^*(w_0, J_0, \omega)}{\partial \omega} &= \frac{J_0}{4\pi \omega} \sin 4\pi w_0 \end{aligned} \quad (17.131)$$

podemos obtener los coeficientes de la expansión de Fourier de esta derivada, reescribiendo el seno y usando la Ec. (17.123)

$$\begin{aligned} \frac{\partial W^*}{\partial \omega} &= \frac{J_0}{4\pi \omega} \frac{e^{4\pi i w_0} - e^{-4\pi i w_0}}{2i} = \frac{J_0}{8\pi i \omega} \left( e^{2\pi i (2) w_0} - e^{2\pi i (-2) w_0} \right) = \sum_k A_k(J_0, \omega) e^{2\pi i k w_0} \\ \Rightarrow A_k(J_0, \omega) &= \frac{J_0}{8\pi i \omega} (\delta_{k2} - \delta_{k,-2}) \end{aligned}$$

es decir su expansión de Fourier consta de solo dos términos. La expresión para  $\dot{J}_0$  se obtiene reemplazando (17.131) en la Ec. (17.121)

$$\begin{aligned}\dot{J}_0 &= -\dot{\omega} \frac{\partial}{\partial w_0} \left( \frac{\partial W^*}{\partial \omega} \right) = -\dot{\omega} \frac{\partial}{\partial w_0} \left( \frac{J_0}{4\pi\omega} \sin 4\pi w_0 \right) \\ \dot{J}_0 &= -\frac{\dot{\omega}}{\omega} J_0 \cos 4\pi w_0\end{aligned}\quad (17.132)$$

La Ec. (17.132) es formalmente exacta. Similarmente, la dependencia temporal exacta de  $w_0$  la determina la ecuación de Hamilton para  $w_0$  usando el Hamiltoniano perturbado (17.120) junto con (17.131)

$$\begin{aligned}\dot{w}_0 &= \frac{\partial K(w_0, J_0, \omega)}{\partial J_0} = \frac{\partial H(J_0, \omega)}{\partial J_0} + \dot{\omega} \frac{\partial}{\partial J_0} \left[ \frac{\partial W^*(w_0, J_0, \omega)}{\partial \omega} \right] = v_0 + \dot{\omega} \frac{\partial}{\partial J_0} \left( \frac{J_0}{4\pi\omega} \sin 4\pi w_0 \right) \\ \dot{w}_0 &= \frac{\omega}{2\pi} + \frac{\dot{\omega}}{4\pi\omega} \sin 4\pi w_0\end{aligned}\quad (17.133)$$

de acuerdo a la Ec. (17.124), si queremos calcular un valor medio de  $\dot{J}_0$  que incluya una corrección no nula, debemos hacer algunas aproximaciones. En primer lugar, asumamos que para un periodo del movimiento perturbado el cociente definido por

$$\frac{\dot{\omega}}{\omega} \equiv \varepsilon \quad (17.134)$$

es constante y cumple la relación  $\varepsilon\tau \ll 1$ . La Ec. (17.134) tiene una solución de la forma

$$\dot{\omega} = \varepsilon\omega \quad \Rightarrow \quad \omega = \omega_0 e^{\varepsilon t} \approx \omega_0 (1 + \varepsilon t) \quad \text{para } t \lesssim \tau \quad (17.135)$$

donde  $t$  se mide desde el inicio del intervalo donde se mide el periodo y  $\omega(0) = \omega_0$ . Reemplazando (17.134) y (17.135) en (17.133) resulta

$$\dot{w}_0 = \frac{\omega}{2\pi} + \frac{\varepsilon}{4\pi} \sin 4\pi w_0 \approx \frac{\omega_0 (1 + \varepsilon t)}{2\pi} + \frac{\varepsilon}{4\pi} \sin 4\pi w_0 \quad (17.136)$$

cuya solución a orden cero en  $\varepsilon$ , es

$$\dot{w}_0^{(0)} = \frac{\omega_0}{2\pi} \quad \Rightarrow \quad 2\pi w_0^{(0)} = \omega_0 t$$

donde se ha hecho cero la constante de integración a través de una elección adecuada de la fase inicial. Con esta elección de fase se tiene que

$$4\pi w_0 = 2\omega_0 t$$

Reemplazando esta expresión en la Ec. (17.136), obtenemos  $\dot{w}_0$  a primer orden en  $\varepsilon$

$$\begin{aligned}\dot{w}_0^{(1)} &= \frac{\omega_0 (1 + \varepsilon t)}{2\pi} + \frac{\varepsilon}{4\pi} \sin 2\omega_0 t \\ 2\pi \dot{w}_0^{(1)} &= \omega_0 + \varepsilon\omega_0 t + \frac{\varepsilon}{2} \sin 2\omega_0 t\end{aligned}$$

cuya solución es

$$2\pi w_0^{(1)}(t) = \omega_0 t + \frac{\varepsilon}{2} \left[ \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right] \quad (17.137)$$

donde por simplicidad hemos elegido  $w_0^{(1)}(0) = 0$ .

Veamos ahora la solución para  $\dot{J}_0$ . De la Ec. (17.132) tenemos que

$$\frac{\dot{J}_0}{J_0} = -\frac{\dot{\omega}}{\omega} \cos 4\pi w_0 \quad \Rightarrow \quad \frac{d \ln J_0}{dt} = -\varepsilon \cos 4\pi w_0 \quad (17.138)$$

donde hemos usado (17.134). Ahora reemplazamos (17.137) en la segunda de las Ecs. (17.138) con lo cual se obtiene

$$\frac{d \ln J_0}{dt} = -\varepsilon \cos \left\{ \frac{4\pi}{2\pi} \left[ \omega_0 t + \frac{\varepsilon}{2} \left( \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right) \right] \right\} + \mathcal{O}(\varepsilon^3)$$

lo cual nos da la ecuación a segundo orden para  $d \ln J_0/dt$ , o equivalentemente, para  $\dot{J}_0/J_0$

$$\begin{aligned} \frac{d \ln J_0}{dt} &= -\varepsilon \cos \left[ 2\omega_0 t + \varepsilon \left\{ \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right\} \right] + \mathcal{O}(\varepsilon^3) \\ \frac{d \ln J_0}{dt} &= -\varepsilon \cos [x + \varepsilon y] + \mathcal{O}(\varepsilon^3) \quad ; \quad x \equiv 2\omega_0 t \quad ; \quad y \equiv \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \end{aligned}$$

expandiendo la suma en el coseno y luego los senos y cosenos en serie, usando a  $\varepsilon$  como parámetro de expansión, se obtiene

$$\begin{aligned} \frac{d \ln J_0}{dt} &= -\varepsilon \cos [x + \varepsilon y] + \mathcal{O}(\varepsilon^3) = \varepsilon \sin x \sin(\varepsilon y) - \varepsilon \cos x \cos(\varepsilon y) + \mathcal{O}(\varepsilon^3) \\ &= (\varepsilon \sin x) [\varepsilon y + \mathcal{O}(\varepsilon^3)] - \varepsilon \cos x [1 + \mathcal{O}(\varepsilon^2)] \\ \frac{d \ln J_0}{dt} &= (\varepsilon^2 y \sin x) - \varepsilon \cos x + \mathcal{O}(\varepsilon^3) \end{aligned}$$

y recordando las definiciones de  $x$  e  $y$  se tiene que

$$\frac{d \ln J_0}{dt} = -\varepsilon \cos 2\omega_0 t + \varepsilon^2 \left\{ \omega_0 t^2 + \frac{1 - \cos 2\omega_0 t}{2\omega_0} \right\} \sin 2\omega_0 t + \mathcal{O}(\varepsilon^3)$$

como se indicó en el formalismo general, el comportamiento secular se obtiene promediando esta ecuación sobre un periodo de movimiento es decir desde  $t = 0$  hasta  $2\pi/\omega_0$ . En el miembro derecho solo sobrevive el término proporcional a  $t^2$  una vez que se hace el promedio.

$$\begin{aligned} \overline{\left( \frac{d \ln J_0}{dt} \right)} &= \frac{1}{\tau} \int_0^\tau \varepsilon^2 \omega_0 t^2 \sin 2\omega_0 t \, dt = \frac{\omega_0 \varepsilon^2 \omega_0}{2\pi} \int_0^{2\pi/\omega_0} t^2 \sin 2\omega_0 t \, dt \\ \overline{\left( \frac{d \ln J_0}{dt} \right)} &= -\frac{\pi \varepsilon^2}{\omega_0} = -\frac{\pi \varepsilon^2 \tau^2}{\omega_0 \tau^2} = -\frac{\pi (\varepsilon \tau)^2}{\omega_0 (2\pi/\omega_0)^2} = -\frac{\omega_0 (\varepsilon \tau)^2}{4\pi} \\ \overline{\left( \frac{d \ln J_0}{dt} \right)} &= \overline{\left( \frac{\dot{J}_0}{J_0} \right)} = -\frac{\omega_0 \delta^2}{4\pi} \quad ; \quad \delta \equiv \varepsilon \tau = \frac{\dot{\omega}}{\omega} \tau \ll 1 \end{aligned}$$

de modo que  $\delta$  es la variación relativa de  $\omega$  a lo largo de un periodo completo  $\tau$ . La variación secular relativa de  $J$  se calcula aproximando el valor de  $J$  al valor constante no perturbado  $J_0$  de modo que

$$\frac{\Delta J}{J} \simeq \frac{\overline{\dot{J}_0}}{J_0} \tau \simeq \overline{\left( \frac{\dot{J}_0}{J_0} \right)} \tau = -\frac{\omega_0 \delta^2}{4\pi} \tau = -\frac{\omega_0 \delta^2}{4\pi} \frac{2\pi}{\omega_0}$$

y la variación secular relativa de  $J$  queda en la forma

$$\frac{\Delta J}{J} \simeq -\frac{\delta^2}{2} \quad \text{para } t \lesssim \tau \quad (17.139)$$

donde la condición  $t \lesssim \tau$ , proviene de la Ec. (17.135). La Ec. (17.139) nos dice que la variación secular de la variable de acción es de segundo orden tal como se vió en el formalismo general, Ec. (17.124). Solo considerando términos del orden  $\varepsilon^2 = (\dot{\omega}/\omega)^2$  podemos encontrar variaciones de  $J$  a largo plazo. Nótese que hemos supuesto que  $\varepsilon$  es constante, con lo cual despreciamos las contribuciones debidas a  $\ddot{\omega}$ , que en principio pueden ser comparables a las de  $\dot{\omega}^2$  como se aprecia en la Ec. (17.124).

### 17.9.2. Variación asintótica de $J$ para el oscilador armónico (opcional)

Una de las enormes utilidades del oscilador armónico consiste en que todos los fenómenos de oscilaciones pequeñas en torno a un punto estacionario se pueden aproximar a un oscilador armónico. Por lo anterior es de mucho interés conocer la variación de  $J$  para el oscilador armónico para variaciones lentas y no tan lentas de un parámetro. Esto no solo posee aplicaciones en Física sino incluso en la solución de ecuaciones diferenciales con coeficientes que dependen del tiempo.

Introduzcamos una variable temporal  $t' = \varepsilon t$  siendo  $\varepsilon$  un parámetro pequeño, es claro que

$$\dot{\omega} = \varepsilon \frac{d\omega}{d(\varepsilon t)} = \varepsilon \frac{d\omega}{dt'}$$

el parámetro temporal reescalado nos mide si  $\omega$  varía lenta o rápidamente. Consideraremos el cambio de  $J$  no sobre un periodo de movimiento sino a lo largo de su valor inicial y final asintótico es decir desde  $t' \rightarrow -\infty$  hasta  $t' \rightarrow \infty$ . En cada estado asintótico se asume que  $J$  adquiere valor constante. Es posible llegar al sorprendente resultado de que para  $\omega(t')$  bien comportado, la variación entre los estados asintóticos es mucho menor de lo que se podría esperar en la Ec. (17.139)<sup>16</sup>. Más específicamente se ha demostrado que si  $\omega(t')$  es real, acotada y analítica sobre y en torno al eje real, la variación de  $J$  es exponencialmente decreciente con  $\varepsilon$  en la forma

$$\frac{J(+\infty) - J(-\infty)}{J(+\infty)} \propto O\left(e^{-d/\varepsilon}\right)$$

siendo  $d$  un número real positivo, que está relacionado con la anchura de la franja de analiticidad de  $\omega$  en el plano  $t'$  complejo.

### 17.9.3. Un invariante exacto del oscilador armónico con frecuencia dependiente del tiempo (opcional)

Es interesante ver que el oscilador armónico con frecuencia dependiente del tiempo, posee un invariante exacto que en el límite de variación lenta se reduce a la variable de acción  $J$ . Supongamos un oscilador armónico lineal con Lagrangiano y Hamiltoniano dados por

$$L = \frac{m}{2} [\dot{x}^2 - \omega^2(t) x^2] \quad ; \quad H = \frac{1}{2m} [p^2 + m^2 \omega^2(t) x^2] \quad (17.140)$$

el invariante (adimensional) que mencionamos tiene la forma

$$I = \frac{1}{2} \left[ \frac{x^2}{r^2} + \frac{(r\dot{x} - x\dot{r})^2}{h_0^2} \right] \quad (17.141)$$

donde  $h_0$  es una constante arbitraria con unidades de longitud al cuadrado sobre tiempo, y la variable  $r(t)$  es una función que satisface la ecuación diferencial

$$\ddot{r} + \omega^2(t) r - h_0^2 r^{-3} = 0 \quad (17.142)$$

por simplicidad podemos elegir  $h_0$  como la unidad en el sistema de unidades en el que estamos trabajando (aunque esto no es obligatorio). Por ejemplo, si estamos en MKS es conveniente elegir  $h_0 = 1m^2seg^{-1}$ , si estamos en cgs conviene tomar  $h_0 = 1cm^2seg^{-1}$ . La deducción matemática de la invarianza de  $I$  es más bien compleja y resulta más ilustrativo hacer algunas consideraciones que nos den su significado físico y la razón de porqué es constante. Consideremos el movimiento del oscilador armónico unidimensional con frecuencia constante, como la proyección sobre  $x$  de un oscilador bidimensional isótropo. Esta consideración permanece

<sup>16</sup>En todo caso la ecuación (17.139) es válida bajo la condición  $t \lesssim \tau$ . Por tanto, no es de esperarse que se pueda extrapolar mucho más allá de esta condición.

válida incluso cuando  $\omega$  depende del tiempo. Consideremos entonces un potencial “central” de tipo armónico<sup>17</sup> pero dependiente del tiempo

$$V = \frac{m}{2} \omega^2(t) r^2 \quad (17.143)$$

su Lagrangiano es

$$L = \frac{m}{2} [\dot{r}^2 + r^2 \dot{\theta}^2 - \omega^2(t) r^2] = \frac{m}{2} \{ [\dot{x}^2 - \omega^2(t) x^2] + [\dot{y}^2 - \omega^2(t) y^2] \} \quad (17.144)$$

en coordenadas cartesianas esto se desdobra en dos osciladores armónicos lineales independientes con la misma frecuencia. Cada uno de ellos cumple con la estructura dada por la Ec. (17.140). En coordenadas polares planas se vé que  $\theta$  es cíclica, de modo que se sigue conservando el momento angular con  $\omega = \omega(t)$ , de esto se sigue que una de las ecuaciones de movimiento se reduce a<sup>18</sup>

$$r^2 \dot{\theta} = \frac{l}{m} \equiv h \quad (17.145)$$

donde hemos introducido la constante de movimiento  $h$ , la cual tiene las mismas unidades que  $h_0$ . Con esto podemos seguir considerando al oscilador lineal como la proyección en  $x$  de un oscilador bidimensional isótropo, donde  $h$  nos da su momento angular por unidad de masa. La ecuación radial asociada al Lagrangiano (17.144) resulta

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{r}} \right) &= m \ddot{r} \quad ; \quad \frac{\partial L}{\partial r} = mr \dot{\theta}^2 - mr \omega^2(t) \\ m \ddot{r} + m \omega^2 r - mr \dot{\theta}^2 &= 0 \quad \Rightarrow \quad \ddot{r} + \omega^2 r - r \dot{\theta}^2 = 0 \quad \Rightarrow \quad \ddot{r} + \omega^2 r - \frac{r^4 \dot{\theta}^2}{r^3} = 0 \end{aligned}$$

y teniendo en cuenta la Ec. (17.145) resulta

$$\ddot{r} + \omega^2 r - \frac{h^2}{r^3} = 0 \quad (17.146)$$

y usando la conservación de  $h$  se puede construir una constante de movimiento dada por

$$I' = \frac{1}{2} \left[ \frac{h^2 x^2}{h_0^2 r^2} + \frac{(r \dot{x} - x \dot{r})^2}{h_0^2} \right] \quad (17.147)$$

para conocer el valor de esta constante hagamos  $x = r \cos \theta$  de lo cual resulta

$$\begin{aligned} r \dot{x} - x \dot{r} &= r (\dot{r} \cos \theta - r \dot{\theta} \sin \theta) - r \dot{r} \cos \theta = -r^2 \dot{\theta} \sin \theta \\ r \dot{x} - x \dot{r} &= -h \sin \theta \end{aligned}$$

y reemplazando en (17.147) tenemos

$$\begin{aligned} I' &= \frac{1}{2h_0^2} \left[ \frac{h^2 r^2 \cos^2 \theta}{r^2} + h^2 \sin^2 \theta \right] = \frac{1}{2h_0^2} [h^2 \cos^2 \theta + h^2 \sin^2 \theta] \\ I' &= \frac{1}{2} \left( \frac{h}{h_0} \right)^2 \end{aligned}$$

<sup>17</sup>Nótese que en nuestro invariante  $I$  la variable  $r(t)$  no se refiere a posición en el tiempo, es simplemente la solución de la ecuación diferencial (17.142). Por otro lado, en la Ec. (17.143) sí estamos considerando a  $r(t)$  como una posición para el problema de la fuerza central equivalente.

<sup>18</sup>Nótese que  $\dot{\theta}$  en este caso es independiente de  $\omega(t)$ . La primera se refiere a la variación de la coordenada angular de posición del problema bidimensional, la segunda se refiere al coeficiente de restitución que puede ser variable en nuestro problema. Mas exactamente, el coeficiente de restitución es  $k(t) = m\omega^2(t)$  como se aprecia en la Ec. (17.143).

por otro lado, dada la arbitrariedad en la amplitud inicial de  $y$  y de la fase inicial del movimiento bidimensional en  $x$  e  $y$  es posible escoger estos valores de tal manera que  $h$  coincida numéricamente con  $h_0$ . Para ver esto, supongamos que durante un cierto periodo de tiempo  $\omega$  permanece constante con valor  $\omega_0$ , en este intervalo la dependencia temporal de  $x$  e  $y$  está dada por

$$x = x_0 \sin \omega_0 t \quad ; \quad y = y_0 \cos (\omega_0 t + \alpha)$$

y teniendo en cuenta que el movimiento bajo potencial central descrito por la ecuación (17.146), es en un plano (el plano  $XY$ ), el momento angular va en la dirección  $Z$  de modo que el módulo del momento angular está dado por  $l = |L_z|$ , y usando la definición de  $h$  se tiene

$$\begin{aligned} h &= \frac{|L_z|}{m} = \frac{m |(\mathbf{r} \times \mathbf{v})_z|}{m} = |x\dot{y} - y\dot{x}| = |[x_0 \sin \omega_0 t] [-\omega_0 y_0 \sin (\omega_0 t + \alpha)] - [y_0 \cos (\omega_0 t + \alpha)] [\omega_0 x_0 \cos \omega_0 t]| \\ h &= |-\omega_0 y_0 x_0 \sin \omega_0 t [\sin \omega_0 t \cos \alpha + \cos \omega_0 t \sin \alpha] - \omega_0 x_0 y_0 [\cos \omega_0 t \cos \alpha - \sin \omega_0 t \sin \alpha] \cos \omega_0 t| \\ h &= |-\omega_0 y_0 x_0 \sin^2 \omega_0 t \cos \alpha - \omega_0 x_0 y_0 \cos^2 \omega_0 t \cos \alpha| = |-\omega_0 y_0 x_0 \cos \alpha [\sin^2 \omega_0 t + \cos^2 \omega_0 t]| \\ h &= \omega_0 x_0 y_0 \cos \alpha \end{aligned}$$

ahora hacemos variar  $\omega$  con el tiempo, en tal caso  $h$  conservará su valor inicial que se puede hacer igual a  $h_0$  eligiendo adecuadamente  $y_0$  ó  $\alpha$  para un valor dado de  $x_0$ , es decir para un valor específico de la amplitud del movimiento unidimensional real. De esto se concluye que siempre es posible asociar a cualquier oscilador unidimensional dependiente del tiempo, un oscilador isótropo bidimensional con  $h/h_0 = 1$ . Nótese que para este problema asociado, la ecuación de movimiento “radial” (17.146) se reduce a la Ec. (17.142) cuando  $h = h_0$ . De la misma forma  $I'$  de la Ec. (17.147) se convierte en el invariante  $I$  de la Ec. (17.141) cuando se toma  $h = h_0$ .

De lo anterior vemos que Físicamente la invarianza exacta de la cantidad  $I$  asociada a un oscilador unidimensional y expresada en la Ec. (17.141), es una manifestación de la conservación del momento angular en el problema del oscilador isótropo bidimensional asociado, que tiene como proyección en  $x$  al oscilador unidimensional en cuestión. A su vez, la conservación del momento angular en el oscilador bidimensional isótropo está relacionada con el hecho de que el carácter central de la fuerza no se modifica cuando se introduce la variación temporal de  $\omega$ .

Veamos ahora la relación de  $I$  con la variable de acción  $J$ . Si  $\omega$  es constante, una solución posible para la ecuación radial (17.146) con  $h = h_0$  es  $r = \sqrt{h_0/\omega}$  que es obviamente constante. Esto corresponde a una trayectoria circular en el problema bidimensional asociado. Dado que en este caso  $\dot{r} = 0$ , la ecuación (17.141) nos da

$$\begin{aligned} I &= \frac{1}{2} \left[ \frac{x^2}{\left(\sqrt{h_0/\omega}\right)^2} + \frac{1}{h_0^2} \left(\sqrt{\frac{h_0}{\omega}}\right)^2 \dot{x}^2 \right] = \frac{1}{2h_0} \left[ \omega x^2 + \frac{\dot{x}^2}{\omega} \right] = \frac{m}{2m\omega h_0} (\dot{x}^2 + \omega^2 x^2) = \frac{E}{m\omega h_0} = \frac{J_0 v_0}{2\pi m v_0 h_0} \\ I &= \frac{J_0}{2\pi m h_0} \end{aligned}$$

de modo que en la aproximación de orden cero el invariante exacto es proporcional a la variable de acción  $J$ , la cual es invariante solo a primer orden en un proceso adiabático.

#### 17.9.4. Invariantes adiabáticos de partículas cargadas en campos electromagnéticos

Supongamos que tenemos un electrón que se mueve en un campo magnético uniforme y constante. Es bien conocido que la partícula cargada describirá una circunferencia o un helicoide alrededor de una línea de campo magnético. Empleando segunda ley de Newton y fuerza de Lorentz resulta

$$\frac{d\mathbf{v}}{dt} = \mathbf{v} \times \frac{q\mathbf{B}}{mc} \quad (17.148)$$

la Ec. (17.148) nos dice que  $\mathbf{v}$  es perpendicular a su razón de cambio y por tanto la rapidez es constante. En particular si  $\mathbf{v}$  y  $\mathbf{B}$  son perpendiculares en un instante dado, el movimiento se realizará en un plano perpendicular a las líneas de campo. En tal caso veremos que el vector  $\mathbf{v}$  gira alrededor de una línea de campo magnético con velocidad angular

$$\omega_c = -\frac{q\mathbf{B}}{mc} \quad (17.149)$$

Elijamos  $\mathbf{u}_z$  a lo largo de  $\mathbf{B}$  y escribamos en coordenadas cilíndricas

$$\mathbf{B} = B\mathbf{u}_z \quad ; \quad \mathbf{v} = \dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta + \dot{z}\mathbf{u}_z \quad ; \quad \mathbf{a} = \frac{d\mathbf{v}}{dt} = (\ddot{r} - r\dot{\theta}^2)\mathbf{u}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathbf{u}_\theta + \ddot{z}\mathbf{u}_z$$

la Ec. (17.148) queda

$$\begin{aligned} (\ddot{r} - r\dot{\theta}^2)\mathbf{u}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathbf{u}_\theta + \ddot{z}\mathbf{u}_z &= \frac{q}{mc} [\dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta + \dot{z}\mathbf{u}_z] \times B\mathbf{u}_z = -\frac{q}{mc}\dot{r}B\mathbf{u}_\theta + \frac{q}{mc}Br\dot{\theta}\mathbf{u}_r \\ \ddot{r} - r\dot{\theta}^2 &= \frac{q}{mc}Br\dot{\theta} \quad ; \quad 2\dot{r}\dot{\theta} + r\ddot{\theta} = -\frac{q}{mc}\dot{r}B \quad ; \quad \ddot{z} = 0 \\ \ddot{r} - r\dot{\theta}^2 - \frac{q}{mc}Br\dot{\theta} &= 0 \quad ; \quad 2\dot{r}\dot{\theta} + r\ddot{\theta} + \frac{q}{mc}\dot{r}B = 0 \quad ; \quad \ddot{z} = 0 \end{aligned} \quad (17.150)$$

antes de analizar las soluciones veremos que el problema lo podemos plantear en el marco de la mecánica lagrangiana usando el siguiente Lagrangiano no relativista para el sistema

$$L = \frac{mv^2}{2} + \mathbf{M} \cdot \mathbf{B} \quad (17.151)$$

siendo  $\mathbf{M}$  el momento dipolar magnético de la partícula en movimiento. En términos del momento angular esto se escribe

$$\mathbf{M} = \frac{q\mathbf{L}}{2mc} \quad (17.152)$$

usando coordenadas cilíndricas, el vector  $\mathbf{L}$  se escribe como

$$\begin{aligned} \mathbf{L} &= \mathbf{r} \times \mathbf{p} = m \mathbf{r} \times \mathbf{v} = m (r\mathbf{u}_r + z\mathbf{u}_z) \times (\dot{r}\mathbf{u}_r + r\dot{\theta}\mathbf{u}_\theta + \dot{z}\mathbf{u}_z) = m (r^2\dot{\theta}\mathbf{u}_z - r\dot{z}\mathbf{u}_\theta + z\dot{r}\mathbf{u}_\theta - rz\dot{\theta}\mathbf{u}_r) \\ \mathbf{L} &= m [r^2\dot{\theta}\mathbf{u}_z + (z\dot{r} - r\dot{z})\mathbf{u}_\theta - rz\dot{\theta}\mathbf{u}_r] \end{aligned}$$

el vector  $\mathbf{M}$  y su componente  $M_z$  son entonces

$$\mathbf{M} = \frac{q}{2c} [r^2\dot{\theta}\mathbf{u}_z + (z\dot{r} - r\dot{z})\mathbf{u}_\theta - rz\dot{\theta}\mathbf{u}_r] \quad ; \quad M_z = \frac{qr^2\dot{\theta}}{2c} \quad (17.153)$$

y teniendo en cuenta que  $\mathbf{B} = B\mathbf{u}_z$ , el Lagrangiano (17.151) queda

$$\begin{aligned} L &= \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) + M_z B \\ L &= \frac{m}{2} (\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) + \frac{q}{2c} Br^2\dot{\theta} \end{aligned}$$

$\theta$  es cíclica de modo que se conserva su momento conjugado

$$p_\theta = mr^2\dot{\theta} + \frac{qBr^2}{2c} \quad (17.154)$$

la ecuación de movimiento radial se obtiene de la ecuación de Lagrange para  $r$

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{r}} \right) = m\ddot{r} \quad ; \quad \frac{\partial L}{\partial r} = mr\dot{\theta}^2 + \frac{qB}{c}r\dot{\theta} \quad \Rightarrow \quad m\ddot{r} - r\dot{\theta} \left( m\dot{\theta} + \frac{qB}{c} \right) = 0$$



$$\ddot{r} - r\dot{\theta} \left( \dot{\theta} + \frac{qB}{mc} \right) = 0 \quad (17.155)$$

la cual coincide con la primera de las Ecs. (17.150). La segunda de las Ecs. (17.150) se obtiene de la ecuación de Lagrange para  $\theta$  o combinando las Ecs. (17.154, 17.155).

Una solución de movimiento estacionario para las Ecs. (17.150) consiste en tomar  $r$  constante<sup>19</sup>, con lo cual la segunda de las Ecs. (17.150) nos conduce a  $\dot{\theta} \equiv \omega_c$  constante. Reemplazando  $r$  y  $\dot{\theta}$  constantes en la primera de las Ecs. (17.150) se obtiene

$$-r\omega_c^2 - \frac{q}{mc}Br\omega_c = 0 \quad \Rightarrow \quad \left( \omega_c + \frac{q}{mc}B \right) \omega_c = 0$$

de modo que el valor de  $\dot{\theta}$  será

$$\dot{\theta} = \omega_c \equiv -\frac{qB}{mc} \quad (17.156)$$

que coincide con (17.149). De la tercera de las Ecs. (17.150), podemos ver que la ecuación para la coordenada  $z$  está totalmente desacoplada y que la velocidad en  $z$  es constante. Si  $\dot{z} = 0$  el movimiento es circular uniforme con frecuencia  $\omega_c$  dada por (17.156), si  $\dot{z} \neq 0$  el movimiento será un helicoide con la misma frecuencia angular y las mismas ecuaciones para  $r$  y  $\theta$ .

Reemplazando (17.156) en (17.154) vemos que  $p_\theta$  adquiere el valor

$$\begin{aligned} p_\theta &= -mr^2 \frac{qB}{mc} + \frac{qBr^2}{2c} = -\frac{qBr^2}{c} + \frac{1}{2} \frac{qBr^2}{c} \\ p_\theta &= -\frac{qBr^2}{2c} \end{aligned}$$

y puesto que  $r$  es constante, la cantidad  $p_\theta$  es constante de movimiento como ya se anticipó. Ahora calcularemos la variable de acción asociada a  $\theta$ , como esta coordenada es cíclica tiene movimiento periódico de rotación en el espacio de fase y su periodo natural será  $2\pi$  (ver discusión de la Ec. 9.122) con lo cual

$$J_\theta = p_\theta \oint d\theta = -\frac{\pi qBr^2}{c} \quad (17.157)$$

si asumimos que  $\dot{z} = z = 0$ , y puesto que  $\dot{r} = \ddot{\theta} = 0$ , el movimiento es circular uniforme y el momento angular (y por tanto el momento dipolar magnético) va a lo largo de  $z$ , de modo que  $M_z = M$ . Usando además (17.153), podemos escribir

$$M = \frac{qr^2\omega_c}{2c} \quad \Rightarrow \quad \frac{qr^2}{c} = \frac{2M}{\omega_c} \quad (17.158)$$

sustituyendo (17.158) en (17.157), y usando (17.156), la variable canónica  $J_\theta$  puede escribirse en la forma

$$J_\theta = -\frac{2\pi MB}{\omega_c} = \frac{2\pi mc}{q} M \quad (17.159)$$

Ahora bien, el campo magnético puede tomarse como parámetro dependiente del tiempo ( $a(t) \equiv B(t)$ ) para examinar la invarianza adiabática de  $J_\theta$ . El teorema de invarianza adiabática Ec. (17.124), dice que ante una variación temporal lenta del campo magnético,  $J_\theta$  permanece constante al menos a primer orden. La Ec. (17.159) nos dice además que el momento magnético  $M$  también es un invariante adiabático. A partir de la Ec. (17.157), se tiene que

$$J_\theta = -\frac{q}{c} B (\pi r^2) = -\frac{q}{c} BA$$

siendo  $A$  el área del círculo descrito por la órbita de la carga. Por tanto, la invarianza adiabática de  $M$  (o de  $J_\theta$ ) se puede enunciar de forma alternativa diciendo que el producto del campo por el área encerrada por la órbita se mantiene constante a primer orden<sup>20</sup>, bajo una ligera variación temporal del campo magnético.

<sup>19</sup>Esto implica (como condición inicial) que la proyección de la velocidad en el plano  $XY$  es totalmente transversal.

<sup>20</sup>Por supuesto, una ligera variación temporal de  $B$  implica que  $A$  debe variar ligeramente para que  $BA$  sea constante. Por tanto la trayectoria solo es aproximadamente circular.

Habíamos enfatizado que la validez del teorema de invarianza adiabática estaba supeditada a que el parámetro en cuestión dependiera exclusivamente del tiempo. Si por ejemplo  $B = B(\mathbf{r})$  entonces el parámetro (campo magnético) será función de las variables canónicas del sistema físico (carga puntual), y por tanto ya no es válida la Ec. (17.124). En consecuencia, es posible que aparezca una variación adiabática de  $M$  si permanece estática la configuración del campo magnético pero éste es ligeramente no uniforme. Si por ejemplo el campo depende de  $z$  y la partícula tiene una pequeña componente  $z$  de la velocidad, ésta se desplazará lentamente a regiones con valor diferente de  $B$ . Las Ecs. (17.152, 17.156) nos dicen que la energía cinética transversal, es decir la correspondiente al movimiento alrededor de las líneas de  $\mathbf{B}$  viene dada por

$$T_\theta = \frac{1}{2}mv_\theta^2 = \frac{mr^2\dot{\theta}^2}{2} = \frac{1}{2}|mr^2\dot{\theta}||\dot{\theta}| = \frac{1}{2}|\mathbf{L}|\omega_c = \frac{1}{2}\left(\frac{2Mmc}{q}\right)\left(\frac{qB}{mc}\right)$$

$$T_\theta = MB \quad (17.160)$$

supongamos que la partícula cargada es llevada hacia una dirección en donde  $B$  crece, la Ec. (17.160) nos dice que en tal caso la energía cinética transversal aumenta. Ahora bien, dado que la energía cinética total se conserva (puesto que la rapidez es constante como se puede ver de la Ec. 17.148), es necesario que disminuya la componente de la velocidad que va a lo largo de las líneas de fuerza i.e. la componente radial, también puede haber variación de la componente en  $z$  pero si asumimos que esta velocidad es muy pequeña, la mayor parte de la variación será en las velocidades radial y transversal. La componente radial de la velocidad se ve entonces disminuída hasta anularse, luego de lo cual se invierte el sentido radial del movimiento. Adicionalmente, si hacemos que  $B$  crezca en sentido contrario, la partícula cargada quedará confinada, moviéndose en uno y otro sentido entre los extremos, fenómeno que se conoce como **confinamiento especular**. Hay en realidad varios efectos adicionales debidos por ejemplo a la radiación de ciclotrón y a efectos inductivos debido a la no uniformidad y posible variación temporal de los campos. Sin embargo, esta descripción simplificada da una buena idea de las implicaciones Físicas de la no invarianza adiabática de  $M$ .

## 17.10. Ejercicios

1. Utilizando el procedimiento perturbativo de la sección 17.4.1, encuentre las ecuaciones de movimiento exactas del oscilador armónico Ecs. (17.28), por inducción.
2. Un oscilador armónico lineal con constante de restitución  $k$  incrementa súbitamente su masa en una cantidad fraccional  $\varepsilon$ . Asumiendo como sistema no perturbado al oscilador con la masa original, utilice teoría de perturbaciones independiente del tiempo a primer orden, para encontrar el corrimiento resultante en la frecuencia del oscilador a primer orden en  $\varepsilon$ . Repita el procedimiento si el incremento fraccional  $\varepsilon$  es para la constante de restitución  $k$ .
3. Calcule la corrección de segundo orden para la frecuencia de un péndulo plano con amplitud finita.
4. Un oscilador armónico isotrópico es sometido a una perturbación descrita por el Hamiltoniano

$$\varepsilon H_1 = bp_x^2 p_y^2$$

siendo  $b$  una constante. Calcule la contribución de primer orden de este Hamiltoniano al corrimiento de las frecuencias, en el esquema de la teoría de perturbaciones independiente del tiempo.

5. El efecto Stark en Física atómica se puede modelar como una perturbación a una órbita plana elíptica de Kepler, donde el potencial de perturbación es  $\Delta H \equiv V = -kx$ . Encuentre perturbativamente el corrimiento de las frecuencias de movimiento a primer orden. Este modelo también constituye una primera aproximación para calcular el efecto de la presión de la luz solar sobre la órbita de un satélite.

## Formulación Lagrangiana y Hamiltoniana para Sistemas Continuos y Campos

La formulación que se ha trabajado hasta el momento incluía un número finito (o a lo mas infinito numerable y discreto) de grados de libertad. No obstante, algunos problemas mecánicos como la vibración de un sólido elástico, involucran sistemas continuos. Cada punto del sólido toma parte en la oscilación, y el movimiento completo puede ser descrito solo si especificamos los desplazamientos de las partículas con respecto al equilibrio en **todos los puntos**. La extensión a un sistema continuo se puede realizar partiendo de un sistema discreto, resolviendo el problema para luego hacer un paso al límite apropiado. El paso al continuo nos lleva de manera natural al concepto de campo y a la teoría de campos, siendo un campo una función que asigna a cada punto del espacio tiempo una cantidad escalar, vectorial, tensorial, etc. con lo cual hablamos de campos escalares, vectoriales, tensoriales etc. Vale decir además que el concepto de campo es de gran utilidad en muchas ramas de la Física incluso en escenarios más allá de la mecánica, quizás el lector ya esté familiarizado con campos tales como la densidad de carga, la densidad de masa, el potencial eléctrico, campos de temperaturas, de presiones etc. los cuales son campos escalares, entre los campos vectoriales se destacan el campo gravitacional, el campo eléctrico, el campo magnético, el potencial vectorial magnético, la densidad de corriente eléctrica etc.

Antes de hacer un estudio detallado del paso al continuo que nos lleva a la teoría de campos, es conveniente recapitular algunos aspectos de la ecuación de continuidad, la cual jugará un papel vital en esta formulación.

### 18.1. Ecuación de continuidad

Existen muchos principios de conservación que nos son útiles en la comprensión de la naturaleza. No obstante, en general es deseable que estos principios de conservación se puedan expresar en la forma de ecuaciones diferenciales. Asumiremos la existencia de un principio de conservación para una **carga generalizada** que es simplemente una cantidad escalar a la cual le podemos asociar una densidad (cantidad de carga generalizada por unidad de volumen) y una propiedad de transporte que se manifiesta en una **densidad de corriente generalizada** (carga generalizada por unidad de área por unidad de tiempo, en la dirección de propagación de la carga generalizada). Hay muchas cantidades escalares a las cuales se les puede asociar un principio de conservación así como una densidad y una densidad de corriente, los ejemplos más comunes son: carga eléctrica, energía, masa de un fluido etc. De aquí en adelante usaremos simplemente los términos carga, densidad de carga y densidad de corriente, sobreentendiendo que estamos trabajando con estos conceptos en forma generalizada.

Si tenemos una región cerrada la carga que sale (entra) se debe manifestar como una disminución (aumento) de la carga en el interior, de no ser así significa que se está creando o destruyendo carga neta de manera espontánea. Queremos en primer lugar calcular la carga que sale por unidad de tiempo del volumen  $V$ . Para ello necesitamos conocer el flujo sobre la superficie  $S$  que delimita a  $V$ . Veamos primero el flujo definido sobre un elemento diferencial de superficie  $dS$ . Definimos  $d\mathbf{S}$  como un vector normal al elemento de superficie  $dS$  que

apunta hacia afuera y cuya magnitud es el área  $dS$  del elemento diferencial, denotando los vectores unitarios normal y tangencial a la superficie como  $\mathbf{u}_N$ ,  $\mathbf{u}_T$  se tiene que

$$d\mathbf{S} = dS \mathbf{u}_N \quad ; \quad \mathbf{J} = J_N \mathbf{u}_N + J_T \mathbf{u}_T$$

nótese que la componente normal de  $\mathbf{J}$  es la única que cruza la superficie, ya que la componente tangencial va a lo largo de ésta. En particular si la corriente es puramente tangencial, la carga se propagará a lo largo de la superficie pero sin cruzarla y por lo tanto esta carga no abandonará el volumen. De lo anterior es claro que la cantidad de carga que sale del volumen en la unidad de tiempo estará definida por la componente de la densidad de corriente que cruza dicha superficie, es decir la componente normal de  $\mathbf{J}$ . Dado que  $\mathbf{J}$  es la cantidad de carga por unidad de tiempo que cruza una superficie unidad perpendicular a la dirección de propagación, se tiene que la cantidad de carga por unidad de tiempo que cruza la superficie  $dS$  estará dada por

$$J_N dS = \mathbf{J} \cdot d\mathbf{S}$$

Nótese que si esta cantidad es negativa, significa que la componente normal de  $\mathbf{J}$  va en dirección  $-\mathbf{u}_N$  es decir, está entrando al volumen. Por tanto, la carga neta que sale por unidad de tiempo del volumen  $V$  que está delimitado por la superficie cerrada  $S$  es<sup>1</sup>

$$\oint_S \mathbf{J} \cdot d\mathbf{S}$$

y esta cantidad debe ser igual a la disminución de carga en el interior por unidad de tiempo

$$\oint_S \mathbf{J} \cdot d\mathbf{S} = -\frac{dq_{int}}{dt} \quad (18.1)$$

siendo  $\mathbf{J}$  la densidad de corriente y  $q_{int}$  la carga contenida en el volumen  $V$ . El signo menos se puede entender teniendo en cuenta que cuando la carga sale (entra) el signo de la integral de superficie es positivo (negativo), esto implica que la carga en el interior debe disminuir (aumentar) es decir debe ser una función decreciente (creciente) del tiempo y por lo tanto su derivada debe ser negativa (positiva). Por tanto, el signo menos garantiza que ambos miembros tengan el mismo signo en ambas circunstancias. La carga en el interior se puede escribir en la forma

$$-\frac{dq_{int}}{dt} = -\frac{d}{dt} \int \rho(\mathbf{r}, t) dV$$

siendo  $\rho(\mathbf{r}, t)$  la densidad de carga en el punto  $\mathbf{r}$  y en el instante  $t$ . La integral de volumen se realiza en un instante fijo de tiempo y la derivada depende de este valor evaluado en  $t$  y en  $t + dt$ . Sin embargo, el volumen y un cierto punto  $x, y, z$  dentro de éste son fijos en el proceso, de modo que esta es realmente una derivada parcial en el tiempo.

$$-\frac{dq_{int}}{dt} = -\int \frac{\partial \rho}{\partial t} dV \Rightarrow \oint_S \mathbf{J} \cdot d\mathbf{S} = -\int \frac{\partial \rho}{\partial t} dV \Rightarrow \oint_S \mathbf{J} \cdot d\mathbf{S} + \int \frac{\partial \rho}{\partial t} dV = 0$$

y aplicando el teorema de la divergencia se obtiene

$$\int \left[ \nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} \right] dV = 0$$

como el volumen en cuestión es arbitrario en forma tamaño y ubicación, llegamos a la ecuación diferencial

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \quad (18.2)$$

<sup>1</sup>Por supuesto los signos de  $\mathbf{J} \cdot d\mathbf{S}$  pueden ser diferentes en cada elemento diferencial. Cuando la integral cerrada es positiva (negativa) es porque sale (entra) más carga de la que entra (sale) en la unidad de tiempo, de modo que el flujo neto es hacia afuera (adentro).

esta ecuación diferencial se conoce como **ecuación de continuidad** y expresa la conservación de la carga generalizada en procesos donde existen corrientes que pueden incluso depender del tiempo. Cuando fluye una cierta cantidad de carga hacia afuera (adentro) del volumen, la cantidad de carga disminuye (aumenta) a la misma rata en que tal carga sale (entra). Vale la pena reiterar que la ecuación diferencial (18.2) es válida para expresar la conservación de muchas cantidades escalares a las que se les pueda definir una densidad y una densidad de corriente.

Por otro lado, también existen principios de conservación asociados a **cargas vectoriales generalizadas**, la conservación del momento es quizás el mayor ejemplo en esta dirección. En tal caso definimos las componentes de esta carga vectorial en coordenadas cartesianas de la forma

$$\mathbf{q} = (q_1, q_2, q_3)$$

y cada componente se debe conservar por separado, es decir habrá una ecuación de continuidad para cada escalar  $q_i$  por separado. Para ello es necesario definir una densidad para la componente  $i$ -ésima de la carga vectorial generalizada  $\rho_i$  y una densidad de corriente asociada  $\mathbf{J}_i$ . La ecuación de continuidad se escribe para cada  $i$  en la forma

$$\nabla \cdot \mathbf{J}_i + \frac{\partial \rho_i}{\partial t} = 0$$

y denotando la componente  $k$ -ésima de  $\mathbf{J}_i$  en la forma  $J_i^k$  podemos escribir esta ecuación en componentes, con lo cual queda (suma sobre índices repetidos)

$$\partial_k J_i^k + \frac{\partial \rho_i}{\partial t} = 0 \quad (18.3)$$

nótese entonces que hay en total tres corrientes, una por cada componente conservada. Podemos formar un tensor  $J_i^k$  que nos da cuenta de las tres corrientes (denotadas por el índice  $i$ ) y sus componentes (denotadas por  $k$ ). Por ejemplo  $J_1^2$  denota que la cantidad escalar conservada  $q_1$  se puede propagar en la dirección  $x_2$ . En otras palabras, cuando el vector  $\mathbf{q}$  se propaga en la dirección  $x_2$  su primera componente  $q_1$  puede cambiar. Puede formularse la conservación de un tensor de un rango arbitrario  $n$  y las corrientes asociadas se pueden organizar en un tensor de rango  $n + 1$ . Por ejemplo, es más útil organizar al momento angular en un tensor antisimétrico de segundo rango que en un vector.

Debemos observar que el número de componentes del vector es arbitrario en realidad. Por ejemplo, ocurre con frecuencia que tenemos un vector y un escalar conservados y resulta cómodo construir un vector de 4 componentes que de cuenta de todas las cargas conservadas

$$\mathbf{q} = (q_0, q_1, q_2, q_3) \quad (18.4)$$

tendremos cuatro ecuaciones de continuidad

$$\partial_k J_\mu^k + \frac{\partial \rho_\mu}{\partial t} = 0 \quad ; \quad \mu = 0, 1, 2, 3 \quad ; \quad k = 1, 2, 3 \quad (18.5)$$

esta ecuación se puede escribir en una notación más compacta si definimos

$$\partial_\nu \equiv (\partial_t, \partial_1, \partial_2, \partial_3) \quad ; \quad J_\mu^\nu \equiv (\rho_\mu, J_\mu^1, J_\mu^2, J_\mu^3) \quad (18.6)$$

$$\partial_\nu J_\mu^\nu = 0 \quad ; \quad \mu, \nu = 0, 1, 2, 3 \quad (18.7)$$

esta formulación es cómoda por ejemplo para incluir la conservación de la energía y el momento lineal en un solo arreglo. Por supuesto, la Ec. (18.7) también puede estar expresando la conservación de cuatro cantidades escalares independientes que no forman ningún vector. En ese sentido la cantidad  $J_\mu^\nu$  no necesariamente expresa un tensor, aunque para la mayoría de casos prácticos, se buscarán cantidades vectoriales y tensoriales para construir ecuaciones de continuidad.

Si una cantidad escalar determinada se conserva, entonces es claro que si definimos un volumen  $\bar{V}$  delimitado por una superficie  $\bar{S}$  tal que la carga generalizada no cruza la superficie i.e. está completamente contenida en

el volumen, entonces la carga total dentro del volumen  $q_{int}$  se conserva ya que no hay pérdidas por “radiación” dentro de este volumen. Similarmente ocurre para cada componente de un arreglo vectorial que se conserve, esto se escribe en la forma

$$q_\mu = \int_{\bar{V}} \rho_\mu dV = K_\mu(\text{constante}) \quad (18.8)$$

Vale la pena observar que la conservación de esta cantidad también se puede ver invirtiendo los pasos que hemos seguido hasta ahora, es decir partimos de la formulación diferencial y llegamos a la forma integral. Integrando la Ec. (18.5) sobre el volumen  $\bar{V}$  resulta

$$\begin{aligned} \int_{\bar{V}} \left[ \partial_k J_\mu^k + \frac{\partial \rho_\mu}{\partial t} \right] dV &= 0 \Rightarrow \int_{\bar{V}} \nabla \cdot \mathbf{J}_\mu dV + \int_{\bar{V}} \frac{\partial \rho_\mu}{\partial t} dV = 0 \\ &\Rightarrow \int_{\bar{S}} \mathbf{J}_\mu \cdot d\mathbf{S} + \frac{d}{dt} \int_{\bar{V}} \rho_\mu dV = 0 \end{aligned} \quad (18.9)$$

donde hemos usado el teorema de la divergencia. Ahora bien, dado que la carga generalizada no cruza la superficie, el valor del flujo  $\mathbf{J}_\mu \cdot d\mathbf{S}$  sobre cada elemento diferencial de área se anula<sup>2</sup>, con lo cual la integral de superficie se anula de modo que

$$\frac{d}{dt} \int_{\bar{V}} \rho_\mu dV = 0 \Rightarrow \int_{\bar{V}} \rho_\mu dV = cte \quad (18.10)$$

una anotación de terminología es pertinente en este punto, hemos visto que si tenemos una carga generalizada  $q$  que se conserva en el sentido de que no se puede crear espontáneamente, la ecuación de continuidad se puede escribir en forma sintética usando una configuración de cuatro componentes

$$J^\nu \equiv (\rho, \mathbf{J}) \Rightarrow \partial_\nu J^\nu = 0$$

es usual llamar a la configuración  $J^\nu$  una **corriente conservada**. No obstante, este es un término un tanto desafortunado ya que lo que se conserva es una propiedad global cual es la carga total cuando se encierra en una superficie donde la carga generalizada no cruza la frontera. El término corriente conservada parece sugerir en forma errónea que las densidades de carga y corriente son constantes de movimiento. Con esta aclaración utilizaremos el término corriente conservada de aquí en adelante.

Por otro lado, ocurre con frecuencia que la carga y corriente generalizadas están presentes en todo el espacio pero tienden a cero en el infinito. En tal caso será necesario calcular la integral de volumen de la densidad de carga generalizada en todo el espacio para asegurar que la carga generalizada total dentro del volumen se conserva. Sin embargo, aún integrando sobre todo el espacio puede ocurrir que la integral de superficie en 18.9 no se anule, ya que aunque  $\mathbf{J}$  tienda a cero la superficie crece como el cuadrado de la distancia, si  $\mathbf{J}$  decrece más rápido que  $1/r^2$  la integral de superficie tenderá a cero, pero si por ejemplo  $\mathbf{J}$  decrece como  $1/r^2$  la integral puede dar un valor finito incluso si se integra sobre todo el espacio. Esto implica que la carga generalizada no se conserva para ningún volumen y que por tanto se presenta pérdida de carga generalizada por radiación.

Un ejemplo notable del fenómeno de radiación ocurre en electrodinámica cuando tenemos cargas eléctricas aceleradas. Tomemos a la densidad de energía asociada al campo electromagnético como nuestra densidad de carga generalizada, la densidad de corriente generalizada será entonces la cantidad de energía por unidad de área por unidad de tiempo en la dirección de propagación de la onda electromagnética, que será también la dirección de propagación de la energía transportada por el campo, esta densidad de corriente de energía asociada al campo se conoce como **vector de Poynting**. Puede demostrarse que para sistemas de cargas eléctricas aceleradas la integral de superficie del vector de Poynting en el infinito no es nula. Esto significa que parte de la energía asociada al campo se aleja indefinidamente de la carga eléctrica sin regresar, hay también una fracción de la energía asociada al campo que se transporta junto con la carga sin alejarse indefinidamente de ella, es decir solo una fracción de la energía electromagnética se pierde por radiación. Es importante enfatizar que en este caso hemos considerado que la carga generalizada es la energía asociada al campo y NO la carga eléctrica. Efectivamente si la carga eléctrica está en un volumen finito, dicha carga se conserva aunque no se conserve nuestra carga generalizada.

<sup>2</sup>Esto no significa necesariamente que no haya corriente sobre la superficie, pero si tal corriente existe, es totalmente tangencial a la superficie en todos los puntos de  $\bar{S}$ , de modo que la corriente evaluada en la superficie va a lo largo de ésta sin cruzarla y por tanto sin abandonar el volumen, de modo que el flujo es nulo.

## 18.2. Transición de un sistema discreto a un sistema continuo

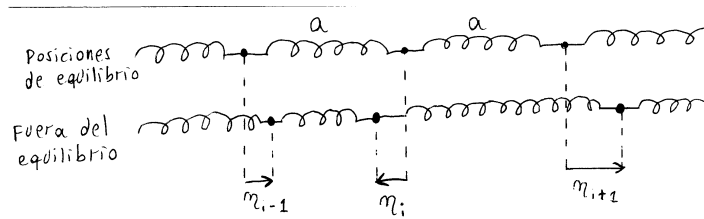


Figura 18.1: Sistema discreto de masas idénticas puntuales conectadas por resortes idénticos. Utilizaremos este sistema para estudiar el paso al continuo en una vara elástica.

Utilizaremos como ejemplo ilustrativo, las vibraciones longitudinales de una varilla infinita. Partiremos entonces del sistema discreto que consiste en una cadena infinita de masas iguales espaciadas una distancia  $a$ , y conectadas a través de resortes sin masa y uniformes (ver Fig. 18.1), todos con la misma constante de restitución  $k$ , la infinitud de la cadena nos evita por el momento trabajar el problema de las condiciones en los extremos. Asumiremos que las masas puntuales solo se pueden mover a lo largo de la cadena, en pequeñas oscilaciones alrededor del equilibrio. Denotando  $\eta_i$  al desplazamiento de la  $i$ -ésima partícula a partir de su posición de equilibrio, la energía cinética es

$$T = \frac{1}{2} \sum_i m \dot{\eta}_i^2$$

donde  $m$  es la masa de cada partícula. La energía potencial correspondiente, resulta ser la suma de las energías potenciales de cada resorte, que surgen como resultado del estiramiento o compresión de los resortes

$$V = \frac{1}{2} \sum_i k (\eta_{i+1} - \eta_i)^2$$

el Lagrangiano queda entonces

$$L = T - V = \frac{1}{2} \sum_i \left[ m \dot{\eta}_i^2 - k (\eta_{i+1} - \eta_i)^2 \right]$$

que también se puede escribir como

$$L = \frac{1}{2} \sum_i a \left[ \frac{m}{a} \dot{\eta}_i^2 - k a \left( \frac{\eta_{i+1} - \eta_i}{a} \right)^2 \right] = \sum_i a L_i \quad (18.11)$$

donde  $a$  es la distancia entre los puntos de equilibrio. Obtengamos las ecuaciones de movimiento de Lagrange

correspondientes a las coordenadas  $\eta_i$

$$\begin{aligned}
\frac{\partial L}{\partial \dot{\eta}_k} &= m\dot{\eta}_k \Rightarrow \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\eta}_k} \right) = m\ddot{\eta}_k \\
\frac{\partial L}{\partial \eta_k} &= \frac{1}{2} \frac{\partial}{\partial \eta_k} \left\{ \sum_i [m\dot{\eta}_i^2 - k(\eta_{i+1}^2 + \eta_i^2 - 2\eta_{i+1}\eta_i)] \right\} = -\frac{k}{2} \frac{\partial}{\partial \eta_k} \left\{ \sum_i [\eta_{i+1}^2 + \eta_i^2 - 2\eta_{i+1}\eta_i] \right\} \\
&= -\frac{k}{2} \sum_i \left[ 2\eta_{i+1}\delta_{i+1,k} + 2\eta_i\delta_{ik} - 2\frac{\partial}{\partial \eta_k} (\eta_{i+1}\eta_i) \right] \\
&= -\frac{k}{2} \sum_i \left[ 2\eta_{i+1}\delta_{i+1,k} + 2\eta_i\delta_{ik} - 2\eta_{i+1}\frac{\partial \eta_i}{\partial \eta_k} - 2\eta_i\frac{\partial \eta_{i+1}}{\partial \eta_k} \right] \\
&= -\frac{k}{2} \sum_i [2\eta_{i+1}\delta_{i+1,k} + 2\eta_i\delta_{ik} - 2\eta_{i+1}\delta_{ik} - 2\eta_i\delta_{i+1,k}] \\
&= -k \sum_i [(\eta_{i+1} - \eta_i)\delta_{i+1,k} + (\eta_i - \eta_{i+1})\delta_{ik}] = -k [(\eta_k - \eta_{k-1}) + (\eta_k - \eta_{k+1})]
\end{aligned}$$

usando las ecuaciones de Lagrange para las coordenadas generalizadas  $\eta_k$  se tiene

$$\begin{aligned}
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\eta}_k} \right) - \frac{\partial L}{\partial \eta_k} &= 0 \Rightarrow m\ddot{\eta}_k + k [(\eta_k - \eta_{k-1}) + (\eta_k - \eta_{k+1})] = 0 \\
m\ddot{\eta}_k - k(\eta_{k+1} - \eta_k) + k(\eta_k - \eta_{k-1}) &= 0
\end{aligned}$$

y dividiendo por  $a$  esta ecuación, las ecuaciones de movimiento de Lagrange correspondientes a las coordenadas  $\eta_i$  son

$$\frac{m}{a}\ddot{\eta}_i - ka \left( \frac{\eta_{i+1} - \eta_i}{a^2} \right) + ka \left( \frac{\eta_i - \eta_{i-1}}{a^2} \right) = 0 \quad (18.12)$$

las ecuaciones (18.11, 18.12) se escogieron de esa forma con el fin de obtener una fácil interpretación de los parámetros en el límite cuando  $a \rightarrow 0$ . Es claro por ejemplo, que  $m/a$  se reduce en el continuo a la densidad lineal de masa que denotaremos por  $\mu$ . No obstante, el límite de  $ka$  no es tan directo. Para una varilla elástica que obedece la ley de Hooke, la elongación de la varilla por unidad de longitud es proporcional a la fuerza o tensión ejercida a lo largo de ésta, lo cual se puede escribir como

$$F = Y\xi$$

donde  $\xi$  es la elongación por unidad de longitud y  $Y$  es el módulo de Young. Por otro lado, la elongación de un segmento  $a$ , por unidad de longitud del sistema discreto en cuestión, se escribe como

$$\xi = \frac{(\eta_{i+1} - \eta_i)}{a}$$

la fuerza necesaria para estirar el resorte en esta cantidad es

$$F = k(\eta_{i+1} - \eta_i) = ka \left( \frac{\eta_{i+1} - \eta_i}{a} \right)$$

con lo cual se vé que  $ka$  debe corresponder al módulo de Young de la varilla continua. Al ir del sistema discreto al continuo, el índice  $i$ , que identifica a una masa puntual específica, se convierte en una coordenada continua de posición  $x$ ; en vez de la variable discreta  $\eta_i$  ahora tendríamos una variable de campo continua  $\eta(x)$ . Adicionalmente, la cantidad

$$\lim_{a \rightarrow 0} \frac{\eta_{i+1} - \eta_i}{a} = \lim_{a \rightarrow 0} \frac{\eta(x+a) - \eta(x)}{a} = \frac{d\eta}{dx}$$



de modo que  $a$  hace el rol de  $dx$ . Finalmente, la suma sobre el índice discreto  $i$  se convierte en una integral sobre  $x$  a lo largo de la longitud de la varilla. El Lagrangiano (18.11) queda

$$L = \frac{1}{2} \int dx \left[ \mu \dot{\eta}^2 - Y \left( \frac{d\eta}{dx} \right)^2 \right] \equiv \int dx \mathcal{L} \quad (18.13)$$

en el límite cuando  $a \rightarrow 0$ , los últimos dos términos de (18.12) quedan

$$\begin{aligned} Y \left( \frac{\eta_i - \eta_{i-1}}{a^2} \right) - Y \left( \frac{\eta_{i+1} - \eta_i}{a^2} \right) &= \frac{Y}{a} \left[ \left( \frac{\eta_i - \eta_{i-1}}{a} \right) - \left( \frac{\eta_{i+1} - \eta_i}{a} \right) \right] \\ &\rightarrow \frac{Y}{a} \left[ \left( \frac{\eta(x) - \eta(x-a)}{a} \right) - \left( \frac{\eta(x+a) - \eta(x)}{a} \right) \right] \end{aligned}$$

escribiendo explícitamente el límite resulta

$$\lim_{a \rightarrow 0} \frac{Y}{a} \left[ \left( \frac{d\eta}{dx} \right)_{x-a} - \left( \frac{d\eta}{dx} \right)_x \right] = -Y \lim_{a \rightarrow 0} \frac{\eta'(x) - \eta'(x-a)}{a} = -Y \frac{d^2\eta}{dx^2}$$

por lo tanto, la extensión al continuo de (18.12), es decir la ecuación de movimiento para la varilla elástica continua es

$$\mu \frac{d^2\eta}{dt^2} - Y \frac{d^2\eta}{dx^2} = 0 \quad (18.14)$$

es decir la ecuación de onda unidimensional, correspondiente a una velocidad de propagación

$$v = \sqrt{\frac{Y}{\mu}} \quad (18.15)$$

la Ec. (18.15) corresponde a la bien conocida expresión para la velocidad de las ondas longitudinales elásticas.

Este ejemplo sencillo ilustra la mayor parte de las características generales que presenta el paso al continuo. Un aspecto muy relevante es el papel de la coordenada de posición  $x$ . Esta variable **no representa una coordenada generalizada**, es simplemente el índice continuo que reemplaza a  $i$ . Así como cada diferente valor de  $i$  corresponde a una coordenada generalizada diferente  $\eta_i$  del sistema discreto, de la misma forma hay una coordenada generalizada  $\eta(x)$  por cada valor de  $x$  en el sistema continuo. Dado que  $\eta$  también es en general función del parámetro continuo tiempo debemos escribir  $\eta(x, t)$ , esto indica que  $x$  al igual que el tiempo se puede considerar como un parámetro que entra en el Lagrangiano. Si el sistema continuo fuera tridimensional en lugar de unidimensional, las coordenadas generalizadas se rotularían con tres índices continuos  $x, y, z$ , y las coordenadas generalizadas se escribirían como  $\eta(x, y, z, t)$ .

Vale la pena enfatizar que en la formulación en el continuo, las cantidades  $x, y, z, t$  son completamente independientes unas de otras, y solo aparecen como variables explícitas en  $\eta$ . Por lo tanto, las derivadas de  $\eta$  con respecto a cualquiera de ellas se pueden escribir como derivadas totales sin ninguna ambigüedad<sup>3</sup>. Por otro lado la Ec. (18.13) nos muestra que el Lagrangiano aparece como una integral sobre el índice continuo  $x$ , hemos denotado al integrando como  $\mathcal{L}$ , en el correspondiente caso tridimensional el Lagrangiano tendría la forma

$$L = \int \int \int \mathcal{L} dx dy dz = \int \mathcal{L} dV \quad (18.16)$$

donde  $\mathcal{L}$  es conocido como **densidad Lagrangiana**. De acuerdo con (18.13), la densidad Lagrangiana asociada a las vibraciones longitudinales de la varilla viene dada por

$$\mathcal{L} = \frac{1}{2} \left[ \mu \left( \frac{d\eta}{dt} \right)^2 - Y \left( \frac{d\eta}{dx} \right)^2 \right] \quad (18.17)$$

que corresponde al paso al límite continuo de la cantidad  $L_i$  que aparece en la Ec. (18.11). Como veremos más adelante, para los sistemas continuos la densidad Lagrangiana será de mucha más utilidad que el Lagrangiano.

<sup>3</sup>La independencia de los parámetros  $x, y, z, t$  se puede ver teniendo en cuenta que ellos representan lugares geométricos (espacio temporales) en el espacio euclidiano, y no posiciones de una determinada partícula. Es decir las variables de posición no tienen ninguna dinámica en esta formulación.

### 18.3. Formulación Lagrangiana en una dimensión para sistemas continuos con una sola variable de campo

De la Ec. (18.17) se observa que  $\mathcal{L}$  para la varilla elástica depende de  $\dot{\eta} \equiv \partial\eta/\partial t$ , pero también de  $\partial\eta/\partial x$ . Esto indica que  $x$  y  $t$  juegan un rol similar como parámetros de la densidad Lagrangiana. Si hubiera fuerzas locales presentes además de las interacciones entre vecinos, podemos tener un  $\mathcal{L}$  que dependa de  $\eta$  como tal, además del gradiente espacial de  $\eta$  y su derivada temporal. En el caso más general la densidad Lagrangiana puede ser también función explícita de  $x, t$  por ejemplo si hay un campo externo que no sea constante ni uniforme. De tal manera que la densidad Lagrangiana para un sistema continuo unidimensional aparece como una función de la variable de campo  $\eta$  en la forma

$$\mathcal{L} = \mathcal{L} \left( \eta, \frac{d\eta}{dx}, \frac{d\eta}{dt}, x, t \right) \quad (18.18)$$

el Lagrangiano total es de acuerdo con (18.16), la integral de  $\mathcal{L}$  sobre el rango de  $x$  que define al sistema. Por otra parte, el principio de Hamilton Ec. (4.2) en el límite de sistemas continuos unidimensionales se escribe como

$$\delta I = \delta \int_{t_1}^{t_2} \int_{x_1}^{x_2} \mathcal{L} dx dt = 0 \quad (18.19)$$

Si el principio de Hamilton para un sistema continuo ha de ser de alguna utilidad, entonces debe ser posible obtener el límite continuo de las ecuaciones de movimiento a partir de dicho principio, como por ejemplo la Ec. (18.14). Las ecuaciones de movimiento deben obtenerse haciendo una variación de la integral (ahora doble) de  $\mathcal{L}$  en (18.19). El procedimiento es muy similar al caso discreto. La variación se hace solo sobre  $\eta$  y sus derivadas; los parámetros  $x, t$  no son afectados por la variación ni directamente ni a través de los límites de integración<sup>4</sup>. Así como la variación de  $\eta$  se toma como nula en los extremos temporales  $t_1$  y  $t_2$ , así mismo la variación de  $\eta$  en los extremos  $x_1, x_2$  de la integración en  $x$  también se tomará como cero, en otras palabras tenemos condición de extremo fijo en los parámetros. Al igual que en la sección 3.2 una familia de caminos variados muy conveniente en el espacio de las  $\eta$  se obtiene escogiendo a  $\eta$  a partir de una familia uniparamétrica de posibles funciones  $\eta$

$$\eta(x, t; \alpha) = \eta(x, t; 0) + \alpha\zeta(x, t) \quad (18.20)$$

donde  $\eta(x, t; 0)$  denota la función correcta que satisface el principio de Hamilton, y  $\zeta$  es cualquier función bien comportada que se anula en los puntos extremos de  $x$  y  $t$ . Si  $I$  se considera una función de  $\alpha$ , de tal modo que tenga un extremo en  $\eta(x, t; 0)$  la derivada de  $I$  con respecto a  $\alpha$  se anula en  $\alpha = 0$ . A través de la diferenciación directa se obtiene<sup>5</sup>

$$\frac{dI}{d\alpha} = \int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left[ \frac{\partial\mathcal{L}}{\partial\eta} \frac{\partial\eta}{\partial\alpha} + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dt}\right)} \frac{\partial}{\partial\alpha} \left( \frac{d\eta}{dt} \right) + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dx}\right)} \frac{\partial}{\partial\alpha} \left( \frac{d\eta}{dx} \right) \right] \quad (18.21)$$

puede verse que a pesar de que  $\mathcal{L}$  puede depender también explícitamente de  $x$  y  $t$ , estos términos no aparecen en (18.21) en virtud de que  $\partial x/\partial\alpha = \partial t/\partial\alpha = 0$  ya que los parámetros no cambian cuando se hace una variación de la coordenada  $\eta(x, t)$ . Ahora bien, debido a que la variación de  $\eta$  es decir  $\alpha\zeta$ , se anula en los extremos de ambos parámetros, una integración por partes en  $x$  y  $t$  nos da las relaciones

$$\int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dt}\right)} \frac{\partial}{\partial\alpha} \left( \frac{d\eta}{dt} \right) dt = - \int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dt}\right)} \right) \frac{\partial\eta}{\partial\alpha} dt$$

<sup>4</sup>Recordemos que los parámetros espaciales son solo lugares geométricos de modo que no se ven afectados por la variación de una coordenada  $\eta$ , además al igual que en el caso discreto el camino variado no implica variación en el parámetro temporal.

<sup>5</sup>Nótese que este es un problema variacional multiparamétrico, ya que involucra a los parámetros  $x$  y  $t$ .

y

$$\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dx}\right)} \frac{\partial}{\partial \alpha} \left(\frac{d\eta}{dx}\right) dx = - \int_{x_1}^{x_2} \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dx}\right)} \right) \frac{\partial \eta}{\partial \alpha} dx$$

el principio de Hamilton se puede escribir entonces

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left[ \frac{\partial \mathcal{L}}{\partial \eta} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dt}\right)} \right) - \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dx}\right)} \right) \right] \left( \frac{\partial \eta}{\partial \alpha} \right)_0 = 0$$

y por los mismos argumentos usados en la sección 3.2 la naturaleza arbitraria de los caminos variados implica que el término entre paréntesis cuadrados debe anularse

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dt}\right)} \right) + \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dx}\right)} \right) - \frac{\partial \mathcal{L}}{\partial \eta} = 0 \quad (18.22)$$

las ecuaciones de Euler Lagrange definidas por (18.22) corresponden a las ecuaciones de movimiento que se derivan del principio de Hamilton Ec. (18.19).

Un sistema de  $n$  grados de libertad posee  $n$  ecuaciones de Lagrange, sin embargo, ¡pareciera que para un sistema continuo solo tuviéramos una ecuación de Lagrange!. Debe recordarse sin embargo que en el discreto la ecuación de movimiento para cada  $\eta_i$  es una ecuación diferencial que solo involucra al tiempo, en tal sentido la Ec. (18.22) nos da una ecuación de movimiento para cada valor del “índice” continuo  $x$ . La naturaleza continua de  $x$  se manifiesta en que la Ec. (18.22) es una ecuación diferencial parcial en  $x$  y  $t$ , cuya solución es de la forma  $\eta(x, t)$ . De lo anterior se desprende además que la dimensión del espacio de configuraciones corresponde a la cardinalidad del espacio real (en una dos o tres dimensiones) ya que cada índice continuo  $x$  corresponde a un eje en el espacio de configuraciones.

Para el ejemplo específico de las vibraciones longitudinales de la varilla elástica, la forma de la densidad Lagrangiana (18.17) conduce a las relaciones

$$\left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dt}\right)} \right) = \mu \frac{d\eta}{dt} ; \quad \left( \frac{\partial \mathcal{L}}{\partial \left(\frac{d\eta}{dx}\right)} \right) = -Y \frac{d\eta}{dx} ; \quad \frac{\partial \mathcal{L}}{\partial \eta} = 0 \quad (18.23)$$

de modo que se obtiene la consistencia deseada puesto que al reemplazar (18.23) en las Ecs. (18.22) llegamos efectivamente a las relaciones (18.14).

## 18.4. Formulación Lagrangiana en tres dimensiones para sistemas continuos con un número arbitrario de variables de campo (Teoría clásica de campos)

La formulación Lagrangiana desarrollada aquí debe ser extendida a situaciones en dos y tres dimensiones, como es el caso de un sólido elástico tridimensional. Adicionalmente, en lugar de un solo campo  $\eta$  podríamos tener varias de estas cantidades. A manera de ejemplo, en tres dimensiones un desplazamiento arbitrario con respecto a la posición de equilibrio se describiría por un vector espacial de campo  $\boldsymbol{\eta}$  con tres componentes, en la formulación de Lagrange esto implica tres campos ya que las coordenadas generalizadas deben ser escalares.

La extensión a más dimensiones o más campos no presenta dificultades significativas. Sin embargo, la notación se vuelve altamente engorrosa cuando pasamos por ejemplo a tres dimensiones. Podemos notar sin embargo que el principio de Hamilton (18.19) y las ecuaciones de Euler Lagrange (18.22) presentan una alta simetría con respecto a las variables de espacio y tiempo, las variables espaciales y el tiempo parecen cumplir el mismo rol y los campos son funciones tanto de las coordenadas espaciales como del tiempo y todas ellas

deben ser tratadas como variables independientes. No hay por ejemplo, variación de los campos en los límites de integración en el principio de Hamilton tanto sobre el espacio como sobre el tiempo. Debe tenerse en cuenta sin embargo, que aunque en esta formulación las coordenadas espaciales y temporales entran en forma más simétrica, el tiempo NO corresponde a un índice para la coordenada generalizada  $\eta(\mathbf{r}, t)$ , esto se puede ver del hecho de que en el sistema discreto el campo  $\eta_i(t)$  está rotulado por el índice  $i$  pero no por la variable temporal, la cual simplemente mide la evolución de una coordenada generalizada específica  $\eta_i$ .

En consecuencia de lo anterior, es conveniente introducir una notación en términos de un espacio de cuatro dimensiones con coordenadas  $x^0 = ct$ ,  $x^1 = x$ ,  $x^2 = y$ ,  $x^3 = z$ . Esta notación no tiene ningún significado Físico adicional. El factor  $c$  en  $x^0$  es la velocidad de la luz en el vacío y es usada solo con el fin de que  $x^0$  tenga las mismas unidades que  $x^i$ . El tensor métrico  $g$  es el correspondiente a una geometría Euclidiana, donde las transformaciones del grupo de Galileo son las transformaciones de coordenadas permitidas sobre las componentes del espacio con el tensor métrico restringido por  $g_{ij} = \delta_{ij}$ . Las letras latinas en los supraíndices indicarán índices que involucran solo a las variables espaciales en tanto que las letras griegas como subíndices o supraíndices se referirán a todas las cuatro coordenadas. Se retomará la convención de suma sobre índices repetidos. Las variables de campo se rotularán con el subíndice  $\rho$ , que puede barrer una multitud de formas. En algunas ocasiones simbolizará un índice simple de dos tres, cuatro o más valores, pero también puede simbolizar un conjunto de múltiples índices. Por ejemplo, si la variable de campo es un tensor espacial de segundo rango, entonces  $\rho$  realmente se refiere a dos subíndices. Finalmente, una derivada de las variables de campo con respecto a cualquiera de las cuatro coordenadas  $x^\mu$  se denotará por el subíndice  $\nu$  separado de  $\rho$  por una coma. Si solo hay una variable de campo, el índice  $\rho$  no aparecerá. Algunos ejemplos son los siguientes

$$\eta_{\rho,\nu} \equiv \frac{d\eta_\rho}{dx^\nu} \quad ; \quad \eta_{,j} \equiv \frac{d\eta}{dx^j} \quad ; \quad \eta_{i,\mu\nu} = \frac{d^2\eta_i}{dx^\mu dx^\nu}$$

solo las derivadas de las variables de campo serán simbolizadas de esta manera. En esta notación la densidad Lagrangiana más general que consideraremos se escribe como

$$\mathcal{L} = \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu)$$

el Lagrangiano es una integral sobre el espacio tridimensional de la forma

$$L = \int \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu) dx^i \quad ; \quad dx^i \equiv dx^1 dx^2 dx^3$$

y el principio de Hamilton se escribe como una integral en el espacio tiempo

$$\delta I = \delta \int \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu) dx^\mu = 0 \quad ; \quad dx^\mu \equiv dx^0 dx^1 dx^2 dx^3 \quad (18.24)$$

donde la variación de  $\eta_\rho$  se anula sobre la hipersuperficie  $\sigma^\nu$  que delimita a la región de integración en el espacio y el tiempo. La derivación de las ecuaciones de Euler Lagrange se obtiene haciendo las extensiones adecuadas. Se considera un conjunto uniparamétrico de funciones variadas que se reducen al valor correcto de  $\eta_\rho(x^\nu)$  cuando  $\alpha$  va a cero. Una familia de estas funciones se puede construir como

$$\eta_\rho(x^\nu, \alpha) = \eta_\rho(x^\nu, 0) + \alpha \zeta_\rho(x^\nu)$$

las funciones  $\zeta_\rho(x^\nu)$  son funciones de clase  $C^2$  en todos los parámetros. Dichas funciones se deben anular sobre la hipersuperficie  $\sigma^\nu$  que delimita a la región de integración. La anulación del variacional de  $I$ , es equivalente a establecer que la derivada de  $I$  con respecto a  $\alpha$  sea cero.

$$\frac{dI}{d\alpha} = \int \left[ \frac{\partial \mathcal{L}}{\partial \eta_\rho} \frac{\partial \eta_\rho}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_{\rho,\nu}}{\partial \alpha} \right] (dx^\mu)$$

integrando por partes se obtiene

$$\frac{dI}{d\alpha} = \int \left[ \frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) \right] \frac{\partial \eta_\rho}{\partial \alpha} (dx^\mu) + \int (dx^\mu) \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right) \quad (18.25)$$

la segunda integral se elimina cuando  $\alpha$  se va para cero, lo cual se puede ver de varias formas: si miramos término a término, al realizar la integración para el  $x^\nu$  particular para cada término de derivada, se vé que estos términos se tienen que ir porque la derivada con respecto a  $\alpha$  es cero en los extremos. Por otro lado, la integral se puede transformar por medio de un teorema de la divergencia en 4 dimensiones en una integral sobre la hipersuperficie que delimita la región de integración en el cuadridespacio (**euclidiano**). La integral de superficie también se anula porque la variación de  $\eta_\rho$  en la vecindad de la función de campo correcta es cero sobre la superficie. Para ver esto calculemos la segunda integral a la derecha en (18.25). Recordando que la derivada total en  $x^\nu$  es equivalente a una derivada parcial podemos escribir el segundo término en la forma

$$\int (dx^\mu) \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right) = \int (dx^\mu) \frac{\partial}{\partial x^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right) \equiv \int (dx^\mu) \partial_\nu \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right)$$

lo cual es básicamente una cuadridivergencia del término entre paréntesis. La aplicación del teorema de la divergencia en el espacio de 4 dimensiones convierte la integral de volumen en una integral sobre la hipersuperficie delimitadora donde la variación del término entre paréntesis es cero. En símbolos, la variación relevante se escribe como

$$\delta \int (dx^\mu) \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right) = \delta \int \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \frac{\partial \eta_\rho}{\partial \alpha} \right) d\sigma^\nu = 0$$

donde  $d\sigma^\nu$  representa las componentes de un elemento de hipersuperficie (en el espacio Euclidiano de cuatro dimensiones) orientado a lo largo de la dirección de la normal hacia afuera (para las tres dimensiones espaciales).

Por lo anterior, en el límite  $\alpha \rightarrow 0$ , la Ec. (18.25) se reduce a

$$\left( \frac{dI}{d\alpha} \right)_0 = \int \left[ \frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) \right] \left( \frac{\partial \eta_\rho}{\partial \alpha} \right)_0 (dx^\mu) = 0 \quad (18.26)$$

nuevamente, la naturaleza arbitraria de la variación de cada  $\eta_\rho$  significa que la Ec. (18.26) se satisface solo si se anula el término entre paréntesis cuadrados.

$$\frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right) - \frac{\partial \mathcal{L}}{\partial \eta_\rho} = 0 \quad (18.27)$$

la Ec. (18.27) representa un conjunto de ecuaciones diferenciales parciales para las variables de campo, el número de ecuaciones está determinado por los valores posibles de  $\rho$ . Vale la pena enfatizar de nuevo que dado que cada  $x^i$  corresponde a un índice para las variables de campo, cada una de las Ecs. (18.27) para  $\rho$  fijo, corresponde a un conjunto completo de ecuaciones diferenciales de movimiento de Lagrange en el caso discreto.

Es sencillo ver que las Ecs. (18.27) se reducen al caso particular (18.22) en el caso unidimensional (es decir con un solo parámetro espacial  $x$ ), con una sola variable de campo (i.e.  $\rho = 1$ ). Comparando las dos ecuaciones se puede apreciar la ventaja de la notación introducida (y aún más si se introducen las tres dimensiones). Es necesario insistir que a pesar de la similitud de la notación con aquella usada en relatividad especial, en este caso estamos trabajando en el régimen no relativista con espacio y métrica Euclidianos.

En sistemas discretos, el Lagrangiano tiene un grado de arbitrariedad ya que se puede adicionar una derivada total en el tiempo de una función arbitraria de las coordenadas generalizadas y el tiempo sin alterar las ecuaciones de movimiento. Con sistemas continuos, la densidad Lagrangiana posee una simetría gauge similar bajo la adición de un término de la forma

$$\mathcal{L}' = \mathcal{L} + \frac{dF_\nu(\eta_\rho, x^\mu)}{dx^\nu} \quad (18.28)$$

recordando de nuevo que la derivada total en  $x^\nu$  es equivalente a una derivada parcial, podemos escribir el segundo término en (18.28) en la forma

$$\frac{dF_\nu(\eta_\rho, x^\mu)}{dx^\nu} = \frac{\partial F_\nu(\eta_\rho, x^\mu)}{\partial x^\nu} \equiv \partial_\nu F_\nu$$

con convención de suma sobre índices repetidos. Este término se puede ver como la extensión de una divergencia en 4 dimensiones (en tres dimensiones la divergencia es  $\partial_i F_i$ ).  $F_\nu$  son cuatro funciones arbitrarias de las variables

de campo  $\eta_\rho$  y las coordenadas  $x^\mu$ . La aplicación del teorema de la divergencia en el espacio de 4 dimensiones convierte la integral de hipervolumen en una integral sobre la hipersuperficie  $\sigma^\nu$  donde la variación de  $F_\nu$  es nula

$$\delta \int (dx^\mu) \frac{dF_\nu(\eta_\rho, x^\mu)}{dx^\nu} = \delta \int F_\nu(\eta_\rho, x^\mu) d\sigma^\nu = 0$$

Con lo cual el segundo término a la derecha en (18.28) no contribuye a la variación de la acción y por tanto, tampoco contribuye a las ecuaciones de movimiento. Esto nos indica que la expresión (18.28) es una transformación gauge para la densidad Lagrangiana. Nótese además que esta es una extensión natural del gauge (2.25) para el discreto, con las asignaciones  $q_i(t) \rightarrow \eta_\rho(x^\mu)$ ,  $L \rightarrow \mathcal{L}$ ,  $\Omega \rightarrow F_\nu$ ,  $t \rightarrow x^\mu$  la última asignación tiene que ver con el hecho de que el problema discreto es uniparamétrico (solo el tiempo es un parámetro) y el problema en el continuo es multiparamétrico (con parámetros  $t, x_i$ ).

La formulación Lagrangiana para un conjunto continuo de coordenadas ha sido desarrollado con el fin de trabajar sistemas mecánicos continuos tales como sólidos elásticos en oscilación longitudinal, o un gas vibrando de tal forma que se formen ondas acústicas. Lo interesante consiste en que la formulación se puede emplear para trabajar sistemas no mecánicos que describan ecuaciones que gobiernan a un **campo o campos**. Matemáticamente un campo es simplemente un conjunto de una o más funciones del espacio y el tiempo que son independientes entre sí, obsérvese que esta definición es compatible con la de coordenada generalizada, simplemente extendiendo la posibilidad de que dichas coordenadas puedan depender de más de un parámetro (parámetros espacio temporales en vez de solo el parámetro temporal). No es necesario el requerimiento de que el campo esté relacionado con algún sistema mecánico. Este rompimiento de conexión entre la formulación Lagrangiana en el continuo y la descripción de campos puramente mecánicos es semejante al proceso histórico que ocurrió con la descripción de las ondas electromagnéticas que originalmente fueron descritas como oscilaciones elásticas del misterioso éter. Más adelante, se hace una formulación de campos que permite la ondulación sin la presencia del medio elástico. De la misma manera reconocemos que los procedimientos variacionales desarrollados aquí se pueden independizar de la noción de campos asociados a medios mecánicos continuos. Los campos pueden entonces describir cualquier función del espacio y el tiempo que pueda ser útil en Física (e.g. campos electromagnéticos, potenciales, campos térmicos, campos de presión, densidades generalizadas etc.). En efecto, el principio de Hamilton es una descripción compacta y conveniente para describir la dinámica de los **campos generalizados** que nos lleva a ecuaciones de campo.

Adicionalmente a las ecuaciones para los campos, la densidad Lagrangiana nos puede dar mas información sobre la naturaleza de los campos. Hemos visto que en los sistemas discretos la estructura del Lagrangiano está fuertemente ligada a las cantidades conservadas asociadas al sistema. Análogamente, veremos que en la mecánica del continuo hay algunos teoremas de conservación asociados a la densidad Lagrangiana.

## 18.5. El tensor esfuerzo energía y teoremas de conservación asociados

Un análogo a la conservación de la integral de Jacobi (función energía) para mecánica puntual estudiada en la sección (5.1) se puede derivar para medios continuos. El factor adicional que aparece es que además del tiempo, aparecen las coordenadas  $x^i$  como parámetros en la formulación, ya que todos ellos son parámetros independientes en  $\mathcal{L}$ . Por tanto en vez de comenzar con un derivada temporal en  $\mathcal{L}$ , usaremos una derivada total con respecto a  $x^\mu$

$$\frac{d\mathcal{L}}{dx^\mu} = \frac{\partial\mathcal{L}}{\partial\eta_\rho} \eta_{\rho,\mu} + \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \eta_{\rho,\mu\nu} + \frac{\partial\mathcal{L}}{\partial x^\mu}$$

usando las ecuaciones de movimiento (18.27), esta derivada se escribe

$$\begin{aligned} \frac{d\mathcal{L}}{dx^\mu} &= \left[ \frac{d}{dx^\nu} \left( \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \right) \right] \eta_{\rho,\mu} + \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \frac{d\eta_{\rho,\mu}}{dx^\nu} + \frac{\partial\mathcal{L}}{\partial x^\mu} \\ \frac{d\mathcal{L}}{dx^\mu} &= \frac{d}{dx^\nu} \left( \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \eta_{\rho,\mu} \right) + \frac{\partial\mathcal{L}}{\partial x^\mu} \end{aligned}$$

despejando  $\partial\mathcal{L}/\partial x^\mu$ , esto se puede escribir

$$\frac{d}{dx^\nu} \left( \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_{\mu\nu} \right) = -\frac{\partial\mathcal{L}}{\partial x^\mu} \quad (18.29)$$

supongamos ahora que  $\mathcal{L}$  no depende explícitamente de  $x^\mu$ . Esto usualmente significa que  $\mathcal{L}$  representa un sistema de campos libres, es decir que no contiene fuentes o sumideros de interacción con campos o partículas externas al sistema en un punto explícito del espacio y con una dependencia temporal dada. Por ejemplo si mi sistema es tal que incluye a todos los campos que intervienen en el problema, esto significa no interacción entre los campos y las partículas puntuales que se mueven en el espacio y el tiempo a través de tales campos, también puede por supuesto significar la ausencia de dichas partículas. Bajo estas condiciones, la Ec. (18.29) toma la forma de un conjunto de condiciones de divergencia.

$$\frac{dT_\mu^\nu}{dx^\nu} = \partial_\nu T_\mu^\nu \equiv T_{\mu^\nu, \nu} = 0 \quad (18.30)$$

donde hemos definido

$$T_\mu^\nu \equiv \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_{\mu\nu} \quad (18.31)$$

el cual tiene la forma de un 4-tensor de segundo rango, nótese que dado que  $T_\mu^\nu$  tiene la forma de un tensor de segundo rango, la divergencia sería de la forma de un vector en el 4-espacio. Enfatizamos en que la expresión **tiene la forma** de un tensor de segundo rango pero no necesariamente lo es, dado que el 4-espacio no tiene todavía propiedades de transformación, el espacio y el tiempo son todavía distintos y no tenemos requerimientos para la transformación de  $T_\mu^\nu$ . Sin embargo, las porciones espaciales de estas cantidades se comportan como vectores y tensores en el espacio ordinario (Euclidiano), es decir  $T_{ij}$  son componentes de un tensor tridimensional de segundo rango. Antes de considerar las transformaciones posibles, determinemos el significado Físico de  $T_\mu^\nu$ . Para ello podemos notar que la Ec. (18.30) se puede comparar en forma directa con la ecuación de continuidad para cargas generalizadas vectoriales Ecs. (18.6, 18.7)

$$\partial_\nu J_\mu^\nu = 0 \quad (18.32)$$

y teniendo en cuenta las Ecs. (18.4, 18.6) vemos que  $J_0^0$  es  $\rho_0$ , es decir la densidad del escalar  $q_0$  que se conserva,  $J_0^i$  representa la  $i$ -ésima componente de la corriente asociada con la propagación de  $\rho_0$ .  $J_i^0$  es  $\rho_i$  i.e. la densidad de una de las componentes del vector que se conserva, y  $J_i^k$  es la  $k$ -ésima componente del vector densidad de corriente asociado a  $\rho_i$ .

De lo anterior es claro que la Ec. (18.30) me está expresando la conservación de cuatro cantidades escalares (posiblemente un escalar y un trivector). Veamos la interpretación física de estas cargas generalizadas que se conservan.

### 18.5.1. Interpretación física de $T_0^\mu$ : densidad de energía y vector de Poynting

La similaridad entre la definición de  $T_\mu^\nu$  Ec. (18.31) y la integral de Jacobi (función energía), Ec. (5.13) es obvia, especialmente cuando tomamos la componente  $T_0^0$

$$T_0^0 = \frac{\partial\mathcal{L}}{\partial\dot{\eta}_\rho} \dot{\eta}_\rho - \mathcal{L} \quad (18.33)$$

es lógico pensar entonces que esta es la densidad de función energía en la mecánica del continuo. En muchos sistemas mecánicos, la densidad Lagrangiana tiene la forma  $\mathcal{L} = \mathcal{T} - \mathcal{V}$ , es decir la diferencia entre una densidad de energía cinética y una densidad de energía potencial. Adicionalmente, para muchos de ellos la función energía es la energía del sistema de modo que  $\mathcal{T} + \mathcal{V}$  es la densidad de energía del sistema. En consecuencia, por el mismo argumento usado en la mecánica discreta,  $T_0^0$  se puede identificar como la densidad de la energía total. Veremos que en particular esta interpretación es correcta cuando usamos la densidad Lagrangiana de la varilla elástica Ec. (18.17)

$$\mathcal{L} = \frac{1}{2} [\mu\dot{\eta}^2 - Y \eta_{,x}^2] \quad (18.34)$$

usando (18.33) calculemos  $T_0^0$

$$T_0^0 = \frac{\partial \mathcal{L}}{\partial \dot{\eta}} \dot{\eta} - \mathcal{L} = \mu \dot{\eta} \dot{\eta} - \frac{1}{2} [\mu \dot{\eta}^2 - Y \eta_{,x}^2]$$

$$T_0^0 = \frac{1}{2} \mu \dot{\eta}^2 + \frac{1}{2} Y \eta_{,x}^2 = \mathcal{T} + \mathcal{V}$$

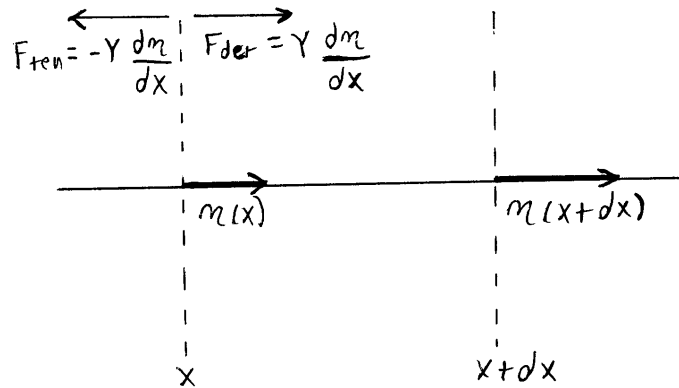


Figura 18.2: Ilustración de las fuerzas sobre un punto  $x$  de la varilla elástica, con el cual se realiza el cálculo de la densidad de corriente de energía en la varilla.

Con base en lo anterior, es de esperarse que el vector  $\mathbf{T}_0$  con componentes  $T_0^i$ , juegue el rol de una densidad de corriente de energía. Nuevamente recurrimos al ejemplo del campo que describe la vibración longitudinal de una varilla elástica. Dividamos la varilla por medio de un corte imaginario en el punto  $x$  (ver figura 18.2). A partir de las consideraciones que nos llevaron al Lagrangiano (18.13) la fuerza ejercida por la parte derecha de la varilla para alargar la parte izquierda del corte es

$$F_{der} = Y \frac{d\eta}{dx}$$

por tanto hay una tensión en  $x$  en la porción izquierda de igual magnitud pero dirección opuesta  $F_{ten} = -F_{der}$ . Además, la porción izquierda está siendo estirada por una cantidad que en  $x$  vale  $\eta(x)$ , y la rata a la cual esta extensión cambia en el tiempo es  $\dot{\eta}$ . Por tanto, la rata de trabajo que la tensión ha hecho en el corte es

$$\frac{dW}{dt} = F_{ten} \dot{\eta} = -\dot{\eta} Y \frac{d\eta}{dx} \quad (18.35)$$

la cual es en consecuencia la rata a la cual la energía está siendo transferida desde la izquierda hacia la derecha por unidad de tiempo. Reemplazando (18.34) en (18.31) podemos calcular  $T_0^1$  y verificar que coincide con el valor dado en (18.35)

$$T_0^1 \equiv \frac{\partial \mathcal{L}}{\partial \eta_{,x}} \eta_{,t} - \mathcal{L} \delta_{tx} = -Y \eta_{,x} \dot{\eta} = -\dot{\eta} Y \frac{d\eta}{dx}$$

Lo cual nos muestra la consistencia de definir  $T_0^i$  como la componente  $i$ -ésima de la densidad de corriente de energía. En la mecánica ondulatoria en tres dimensiones, el vector que representa la propagación de energía por unidad de área por unidad de tiempo se conoce como vector de Poynting. Este vector correspondería entonces a las componentes  $T_0^i$ .

### 18.5.2. Interpretación física de $T_i^j$ : densidad de momento lineal y tensor de esfuerzos

Para poder identificar las otras componentes del tensor, es conveniente escribir (18.30) de la siguiente forma

$$\frac{dT_\mu^0}{dt} + \frac{dT_\mu^j}{dx^j} = 0 \quad (18.36)$$



o

$$T_{\mu,\nu}^{\nu} = \frac{dT_{\mu}^0}{cdt} + \nabla \cdot \mathbf{T}_{\mu} = 0 \quad (18.37)$$

donde  $\mathbf{T}_{\mu}$  cuyas componentes son  $T_{\mu}^{\nu}$  son un conjunto de 4 vectores en el 4-espacio. Ya vimos que  $\mathbf{T}_0$  contiene la densidad de energía y la densidad de corriente de energía. Las Ecs. (18.36, 18.37) son ecuaciones de continuidad, que nos dicen que la rata de cambio en el tiempo de la densidad  $T_{\mu}^0$  mas la divergencia de la densidad de corriente  $\mathbf{T}_{\mu}$  es nula.

Para extraer el significado Físico de las componentes de  $T_i^0$  examinamos de nuevo las vibraciones de la varilla elástica. Si las partículas en la varilla se mueven todas en la misma cantidad a lo largo de ésta, el movimiento será como una traslación de un cuerpo rígido, de modo que no hay oscilaciones. El cambio neto de masa en una longitud  $dx$  de la varilla como resultado del movimiento será claramente cero, dado que pasa la misma cantidad de masa por  $x+dx$  que la que pasa por  $x$ . Habrá sin embargo, una densidad de momento lineal neta  $\mu \mathbf{v}$  para esta traslación rígida, siendo  $\mathbf{v}$  la velocidad de traslación de la barra<sup>6</sup>. Por otro lado, cuando tiene lugar el movimiento ondulatorio, existe un cambio neto de masa en el segmento  $dx$ , que para un cierto tiempo viene dada por

$$dm = \mu [\eta(x) - \eta(x+dx)] = -\mu \frac{d\eta}{dx} dx$$

el momento adicional en el intervalo, que resulta del movimiento ondulatorio es por tanto

$$dP = \dot{\eta} dm = -\mu \dot{\eta} \frac{d\eta}{dx} dx$$

de modo que una densidad adicional de momento (densidad lineal en este caso), mas allá de la que surge del movimiento estacionario, puede ser identificada como la onda o densidad del campo de momento

$$\mathbf{p} \equiv \frac{dP}{dx} = -\mu \dot{\eta} \frac{d\eta}{dx}$$

veamos ahora que esta cantidad es justamente  $-T_1^0$  para la densidad Lagrangiana dada por (18.17).

$$-T_1^0 = -\frac{\partial \mathcal{L}}{\partial \eta_{,t}} \eta_{,x} + \mathcal{L} \delta_{xt} = -\mu \dot{\eta} \eta_{,x} = -\mu \dot{\eta} \frac{d\eta}{dx}$$

Esto nos lleva a identificar  $-T_i^0$  como las componentes del campo de densidad de momento.

Las ecuaciones de continuidad, (18.37) sugieren entonces que  $-\mathbf{T}_i$  debe representar la densidad de flujo vectorial para la  $i$ -ésima componente del campo de densidad de momento. Ascribimos una propiedad de vector a  $\mathbf{T}_i$  porque puede haber por ejemplo, un flujo en la dirección  $y$  de la componente  $x$  de la densidad de momento, que se determinaría por  $-T_x^y \equiv -T_1^2$ . Una interpretación alternativa de  $T_i^j$  proviene de considerar el campo de desplazamiento de un sólido tridimensional elástico. Es bien sabido que en un sólido de este tipo hay también fuerzas de cizalladura tangenciales a la superficie (además de las fuerzas de compresión normales a la superficie) a lo largo del elemento de superficie. El ensamble completo de fuerzas se puede describir diciendo que la fuerza  $d\mathbf{F}$  que actúa sobre un elemento de área  $d\mathbf{A}$  se expresa en términos de un **tensor de esfuerzos**  $\mathbb{T}$  de modo que

$$d\mathbf{F} = \mathbb{T} \cdot d\mathbf{A}$$

que explícitamente se escribe

$$(dF_1, dF_2, dF_3) = \begin{pmatrix} T_1^1 & T_1^2 & T_1^3 \\ T_2^1 & T_2^2 & T_2^3 \\ T_3^1 & T_3^2 & T_3^3 \end{pmatrix} \begin{pmatrix} dA_1 \\ dA_2 \\ dA_3 \end{pmatrix} \quad (18.38)$$

$dA_i$  es un elemento de área en la dirección de  $x_i$  de modo que

$$|dA_1| = dx_2 dx_3 \quad ; \quad |dA_2| = dx_1 dx_3 \quad ; \quad |dA_3| = dx_1 dx_2$$

<sup>6</sup>Nótese que en este caso no hay estiramientos ni encogimientos de los segmentos de modo que nuestros campos  $\eta$  son todos nulos en este caso. Por tanto, aunque el segmento  $dx$  transporta momento, este momento no está asociado al campo.

en consecuencia la fuerza neta digamos en la dirección  $x_1$ , sobre un elemento de volumen rectangular  $dx_1 dx_2 dx_3$  tiene contribuciones que provienen de las fuerzas sobre las superficies paralelas a los planos  $x_2x_3$ ,  $x_1x_3$  y  $x_1x_2$ .

Calculemos primero la contribución en la dirección  $x_1$  que proviene de las fuerzas sobre las dos superficies paralelas al plano  $x_2x_3$ . Estos planos poseen coordenadas  $x_1$  y  $x_1 + dx_1$  respectivamente (ver figura 18.3). Para el plano asociado a la coordenada  $x_1$  el elemento de área (que apunta hacia afuera del volumen) va en la dirección  $-\mathbf{u}_{x_1}$  con lo cual tenemos que para este plano la Ec. (18.38) queda en la forma

$$(dF_1(x_1), dF_2(x_1), dF_3(x_1))^{(23)} = \begin{pmatrix} T_1^1(x_1) & T_1^2(x_1) & T_1^3(x_1) \\ T_2^1(x_1) & T_2^2(x_1) & T_2^3(x_1) \\ T_3^1(x_1) & T_3^2(x_1) & T_3^3(x_1) \end{pmatrix} \begin{pmatrix} -dx_2 dx_3 \\ 0 \\ 0 \end{pmatrix}$$

$$dF_1^{(23)}(x_1) = -T_1^1(x_1) dx_2 dx_3$$

para el plano paralelo a  $x_2x_3$  con coordenada  $x_1 + dx_1$ , el elemento de área apunta hacia  $+\mathbf{u}_{x_1}$  de modo que la suma de las dos contribuciones de los planos paralelos a  $x_2x_3$  nos da

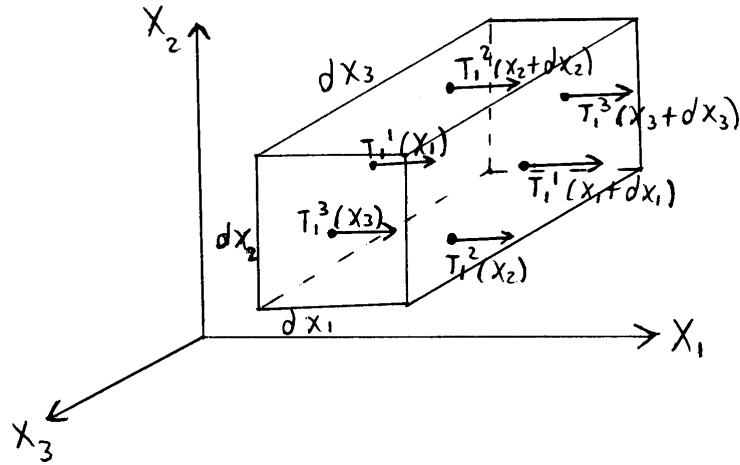


Figura 18.3: Fuerza en la dirección  $x_1$  sobre un elemento de volumen  $dx_1 dx_2 dx_3$  de un sólido elástico.

$$dF_1^{(23)}(x_1) + dF_1^{(23)}(x_1 + dx_1) = [T_1^1(x_1 + dx_1) - T_1^1(x_1)] dx_2 dx_3 = \frac{dT_1^1}{dx_1} dx_1 dx_2 dx_3 \quad (18.39)$$

veamos ahora las contribuciones de los dos planos paralelos a  $x_1x_3$ . Estos planos están asociados a las coordenadas  $x_2$  y  $x_2 + dx_2$  como se aprecia en la Fig. 18.3, para el plano inferior de coordenada  $x_2$  el vector de área es  $-dx_1 dx_3 \mathbf{u}_{x_2}$  de modo que la Ec. (18.38) nos queda en la forma

$$(dF_1(x_2), dF_2(x_2), dF_3(x_2))^{(13)} = \begin{pmatrix} T_1^1(x_2) & T_1^2(x_2) & T_1^3(x_2) \\ T_2^1(x_2) & T_2^2(x_2) & T_2^3(x_2) \\ T_3^1(x_2) & T_3^2(x_2) & T_3^3(x_2) \end{pmatrix} \begin{pmatrix} 0 \\ -dx_1 dx_3 \\ 0 \end{pmatrix}$$

$$dF_1^{(13)}(x_2) = -T_1^2(x_2) dx_1 dx_3$$

para la superficie paralela a  $x_1x_3$  con coordenada  $x_2 + dx_2$  el vector de área es  $dx_1 dx_3 \mathbf{u}_{x_2}$ . Por tanto, la contribución proveniente de las dos superficies paralelas al plano  $x_1x_3$  da

$$dF_1^{(13)}(x_2) + dF_1^{(13)}(x_2 + dx_2) = [T_1^2(x_2 + dx_2) - T_1^2(x_2)] dx_1 dx_3 = \frac{dT_1^2}{dx_2} dx_1 dx_2 dx_3 \quad (18.40)$$

y similarmente para las superficies paralelas a  $x_1x_2$ .

$$dF_1^{(12)}(x_3) + dF_1^{(12)}(x_3 + dx_3) = [T_1^3(x_3 + dx_3) - T_1^3(x_3)] dx_1 dx_2 = \frac{dT_1^3}{dx_3} dx_1 dx_2 dx_3 \quad (18.41)$$

la fuerza infinitesimal resultante es entonces la suma de los términos en las Ecs. (18.39, 18.40, 18.41)

$$dF_{x_1} = \left( \frac{dT_1^1}{dx_1} + \frac{dT_1^2}{dx_2} + \frac{dT_1^3}{dx_3} \right) dV$$

la densidad de fuerza resultante en la dirección  $x_1$  es entonces<sup>7</sup>

$$\frac{dF_{x_1}}{dV} \equiv f_{x_1} = \frac{dT_1^1}{dx_1} + \frac{dT_1^2}{dx_2} + \frac{dT_1^3}{dx_3}$$

La segunda ley de Newton por unidad de volumen podemos escribirla como

$$f_{x_1} = \frac{d\mathbf{p}}{dt}$$

y recordando que la densidad de momento es  $\mathbf{p} = -T_1^0$  tenemos que la segunda ley de Newton por unidad de volumen se puede escribir como

$$\frac{dT_1^1}{dx_1} + \frac{dT_1^2}{dx_2} + \frac{dT_1^3}{dx_3} = -\frac{dT_1^0}{c dt}$$

o en notación sintética

$$\partial_\mu T_1^\mu = 0$$

que es precisamente la componente  $x_1$  de la Ec. (18.37). Para este campo particular  $T_i^j$  puede identificarse como los elementos del tensor de esfuerzos tridimensional, de allí el origen del término **tensor de esfuerzo-energía** para  $T_\mu^\nu$ .

En síntesis, usando consideraciones para medios contínuos mecánicos, hemos podido asociar a cada una de las componentes del tensor de esfuerzo-energía, una interpretación física que se resume en el cuadro 18.1. Adicionalmente, como se vió de la discusión que siguió a la Ec. (18.37),  $T_0$  y  $T_i$  forman vectores 4-espaciales cada uno de los cuales es una corriente conservada similar a el 4- vector densidad de carga-densidad de corriente de la teoría electromagnética, y los denominamos **4-corrientes**. En casi todos los casos el tensor tridimensional

Componentes de $T_\mu^\nu$	Interpretación Física
$T_0^0$	campo de densidad de energía dividido por $c$
trivector $T_0$ con componentes $T_0^j$	campo de densidad de corriente de energía
$-T_i^0$	componente $i$ -ésima del campo de densidad de momento
trivector $-T_i$ con componentes $-T_i^j$	densidad de corriente para la $i$ -ésima componente del campo de densidad de momento
$T_i^j$	tensor de esfuerzos tridimensional

Cuadro 18.1: Interpretación física de las componentes del tensor de esfuerzo energía.

$T$  es simétrico. Esto no solo es físicamente deseable sino que es una característica **casí** necesaria para la parte espacial del tensor de esfuerzo energía.

Debemos recordar que aunque el ejemplo y la nomenclatura provienen de un sistema mecánico, este formalismo es extendible a muchas configuraciones de campos sin importar su naturaleza u origen. Una teoría

<sup>7</sup>Para la contribución a la fuerza a lo largo de  $x_1$ , las contribuciones de los planos paralelos a  $x_2x_3$  son de compresión y las otras contribuciones son de cizalladura.

clásica de campos se puede construir para las vibraciones elásticas de un sólido, campos electromagnéticos, campos térmicos, campos de presión, la función de onda de Schrödinger etc. Debe tenerse presente sin embargo, que en el caso más general, las cuatro ecuaciones de continuidad dadas por (18.30) expresan la conservación de cuatro cargas generalizadas que podrían ser independientes entre sí. Nótese que la interpretación de estas cuatro cargas como energía y momento implican que las cuatro cargas están organizadas como un escalar y un trivector euclidiano (en el caso relativista usualmente representarían un cuadvivector de Lorentz), pero no hay nada en el formalismo general que me garantice este nivel de organización.

Por otro lado, recordemos que en el caso más general la densidad Lagrangiana puede depender de los parámetros  $x^\mu$  en la Ec. (18.29) con lo cual la Ec. (18.32) se convierte en una ecuación de continuidad con fuentes

$$\partial_\nu J_\mu{}^\nu = -\frac{\partial \mathcal{L}}{\partial x^\mu}$$

lo cual corresponde a tener campos adicionales externos al sistema de campos en cuestión, o partículas puntuales interactuando con los campos. Cuando se encuentran ecuaciones de continuidad con fuentes no se conserva la carga generalizada asociada. No obstante, ocurre con frecuencia que es posible extender el sistema Físico para incluir los campos o partículas que actúan de fuentes y sumideros a fin de llegar a una nueva ecuación de continuidad que sea homogénea y que por tanto me conduzca a una carga generalizada conservada. Este es el caso con el campo electromagnético (aunque para las Ecs. de Maxwell se debe usar una formulación relativista), cuando se considera la densidad lagrangiana para el sistema del campo electromagnético se llega a ecuaciones de continuidad para la energía y el momento con fuentes, que corresponden al hecho de que las cargas que interactúan con el campo pueden absorber o transferir energía y momento al campo. Sin embargo, el sistema extendido de cargas y campos sí conserva estas cantidades.

### 18.5.3. Energía y momento total del sistema

Como ya se discutió, la ecuación de continuidad implica la conservación de algunas cantidades integrales (ver Ec. 18.8) siempre que el volumen del campo sea finito, es decir la variable de campo toma valor cero por fuera de cierto volumen acotado<sup>8</sup>, en tal caso se definen cantidades integrales  $R_\mu$  como

$$R_\mu = \int T_\mu{}^0 d\bar{V} \quad (18.42)$$

donde el volumen de integración es tal que el campo está completamente contenido en él, y los flujos de los campos y de todas sus derivadas se anulan en la superficie que delimita al volumen. Es debido a estos teoremas de conservación, derivados a partir de (18.30), que los cuatro arreglos de la forma  $T_\mu{}^\nu$ ,  $\mu = 0, 1, 2, 3$  se conocen como **corrientes conservadas**, en analogía con las ecuaciones de conservación para las corrientes electromagnéticas.

Si  $T_0{}^0$  es una densidad de energía, entonces la cantidad  $R_0$  definida por

$$R_0 = \int_{\bar{V}} T_\mu{}^0 dV$$

se puede identificar como la energía total en el campo siempre que  $T_0{}^i$  se anule en la superficie de delimitación, esto es si el sistema no radía energía al exterior.

Con base en la interpretación de las componentes  $-T_i{}^0$  es claro que la conservación de las otras cargas generalizadas

$$-R_i = \int_{\bar{V}} (-T_i{}^0) dV$$

nos dicen que para un sistema cerrado no interactuante el momento lineal total del campo  $-R_i$  se conserva como se esperaba.

<sup>8</sup>En realidad el volumen podría tomarse en el infinito cuando los campos no están acotados. Pero se requiere que los campos y todas sus derivadas tiendan a cero en el infinito más rápido de lo que crece la superficie de integración, a fin de que la integral (18.9) se anule para que sea válida la Ec. (18.10).

#### 18.5.4. Densidad de momento angular y momento angular total

Por otro lado, debe también existir un teorema correspondiente a la conservación del momento angular total del campo. Es sencillo construir una cantidad que actúe como densidad de momento angular. Dado que el momento angular es un vector axial, se espera que las componentes de una densidad de momento angular sean elementos de un tensor antisimétrico de segundo rango. Teniendo en cuenta la interpretación de  $-T_i^0$  como la densidad de la componente  $i$ -ésima del momento lineal, una forma adecuada para este tensor es

$$\mathcal{M}_i^j = -(x^i T_j^0 - x^j T_i^0) \quad (18.43)$$

donde el momento angular total del campo viene dado por

$$M_i^j = \int \mathcal{M}_i^j dV$$

teniendo en cuenta que  $t$  y  $x^i$  son completamente independientes, la rata de cambio de  $M_i^j$  es

$$\frac{dM_i^j}{dt} = - \int \left( x^i \frac{dT_j^0}{dt} - x^j \frac{dT_i^0}{dt} \right) dV$$

o, a partir de las ecuaciones de continuidad (18.36)

$$\frac{dM_i^j}{dt} = \int \left( x^i \frac{dT_j^k}{dx^k} - x^j \frac{\partial T_i^k}{\partial x^k} \right) dV$$

una integración por partes convierte esta expresión en

$$\frac{dM_i^j}{dt} = \int \frac{d}{dx^k} \left( x^i T_j^k - x^j T_i^k \right) dV - \int (T_i^j - T_j^i) dV \quad (18.44)$$

la primera integral a la derecha tiene la forma de una integral de volumen de una divergencia, con lo cual se puede convertir en una integral sobre la superficie que delimita tal volumen, la cual se anula para un sistema cerrado que no radía. Finalmente, si  $T^{ij} = T^{ji}$ , la segunda integral también es cero. Por lo tanto, el momento angular total del sistema de campos se conserva si  $T$  es simétrico y el sistema no radía por la superficie. Por supuesto es posible aún que la segunda integral de volumen en (18.44) se anule sin que el tensor sea simétrico pero sería menos nítido para demostrar.

En algunos casos resulta que el tensor de esfuerzo (tridimensional) no es simétrico. No obstante, podemos usualmente usar la ambigüedad en la definición de este tensor para restaurar esta simetría. Así como en el caso del Lagrangiano, la forma del tensor esfuerzo-energía, Ec. (18.31), se escogió para que satisficiera ciertas condiciones de divergencia (Ecs. 18.30). Por tanto  $T_\mu^\nu$  está indeterminado puesto que podemos agregar una función cuya 4-divergencia se anule. Nótese que esto conduce por ejemplo al hecho de que la densidad de energía y el vector de Poynting de una onda electromagnética estén definidas salvo cantidades que son divergencias espaciales. La adición de estas cantidades no contribuye a los observables tales como la energía total o la corriente total de energía. Un término de cuádrdivergencia nula se puede parametrizar en forma general en la forma

$$Z_\mu^\nu \equiv \frac{d\psi_\mu^{\nu\lambda}}{dx_\lambda}$$

donde  $\psi_\mu^{\nu\lambda}$  es un conjunto de funciones de las variables de campo que cumplen la condición de antisimetría

$$\psi_\mu^{\nu\lambda} = -\psi_\mu^{\lambda\nu}$$

al calcular la cuádrdivergencia los símbolos  $\nu$  y  $\lambda$  se vuelven índices mudos y la cuádrdivergencia consta de pares de la forma (no hay convenio de suma)

$$\partial_\nu Z_\mu^\nu = \frac{d^2\psi_\mu^{\nu\lambda}}{dx_\nu dx_\lambda} + \frac{d^2\psi_\mu^{\lambda\nu}}{dx_\lambda dx_\nu}$$

que se anulan a causa de la condición de antisimetría. Esta ambigüedad nos permite usualmente simetrizar el tensor esfuerzo energía lo cual se hace a menudo cuando se construye teorías de campos clásicos.

## 18.6. Formulación Hamiltoniana para medios continuos

Es posible obtener una formulación Hamiltoniana para medios continuos de una manera muy similar al caso discreto. Retornaremos al ejemplo de la cadena unidimensional de masas acopladas. Partimos entonces de la definición de momento canónicamente conjugado a una variable de campo  $\eta_i$

$$p_i = \frac{\partial L}{\partial \dot{\eta}_i} = a \frac{\partial L_i}{\partial \dot{\eta}_i} \quad (18.45)$$

donde el  $L_i$  está definido para la cadena de masas acopladas a través de la Ec. (18.11). El Hamiltoniano para el sistema es entonces

$$\begin{aligned} H &\equiv p_i \dot{\eta}_i - L = a \frac{\partial L_i}{\partial \dot{\eta}_i} \dot{\eta}_i - L \\ H &= \sum_i a \left( \frac{\partial L_i}{\partial \dot{\eta}_i} \dot{\eta}_i - L_i \right) \end{aligned}$$

recordando que en el límite  $a \rightarrow 0$ , que nos lleva al continuo, se tiene que  $L_i \rightarrow \mathcal{L}$ ,  $a \rightarrow dx$ , y la suma se convierte en integral

$$H = \int dx \left[ \frac{\partial \mathcal{L}}{\partial \dot{\eta}} \dot{\eta} - \mathcal{L} \right] \quad (18.46)$$

Por otro lado, los momentos canónicos individuales  $p_i$  dados por (18.45), tienden a cero en el límite continuo, pero podemos definir una densidad de momento  $\pi$  que permanece finita

$$\lim_{a \rightarrow 0} \frac{p_i}{a} \equiv \pi = \frac{\partial \mathcal{L}}{\partial \dot{\eta}} \quad (18.47)$$

la Ec. (18.46) nos muestra que el Hamiltoniano también viene dado por una integral de tal forma que se puede definir una **densidad Hamiltoniana**  $\mathcal{H}$

$$\mathcal{H} \equiv \pi \dot{\eta} - \mathcal{L} \quad (18.48)$$

aunque la formulación Hamiltoniana es fácil de introducir para campos clásicos, es de anotar que el procedimiento le otorga a la variable tiempo, un tratamiento diferente a los parámetros de espacio<sup>9</sup>, en contraste con la formulación Lagrangiana en la cual los parámetros espacio temporales aparecen en forma muy simétrica. Por esta razón el formalismo Hamiltoniano al menos como se describe aquí, no es muy adecuado para trabajar una formulación covariante relativista de los campos. El método Hamiltoniano no es entonces tan fructífero como el Lagrangiano para trabajar sistemas continuos relativistas.

La ruta obvia para generalizar a campos en tres dimensiones, consiste en definir las densidades de momento canónico (18.47) para un cierto campo  $\eta_\rho(x^i, t)$  de la forma

$$\pi^\rho(x^\mu) = \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \quad (18.49)$$

las cantidades  $\eta_\rho(x^i, t)$ ,  $\pi^\rho(x^i, t)$  definen el espacio de fase infinito dimensional que describe los campos clásicos y su evolución en el tiempo. Por otro lado, existe una ley de conservación asociada a  $\pi^\rho$  muy similar a la que se obtiene en el caso discreto. Si el campo  $\eta_\rho$  no está contenido explícitamente en  $\mathcal{L}$ , es decir  $\eta_\rho$  es una variable de campo cíclica, entonces la ecuación de Lagrange (18.27) asociada a este campo, se asemeja a una ecuación de continuidad que describe una corriente conservada

$$\begin{aligned} \frac{d}{dx^\mu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, \mu}} \right) &= 0 \Rightarrow \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \right) + \frac{d}{dx^i} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, i}} \right) = 0 \\ \frac{d\pi^\rho}{dt} + \frac{d}{dx^i} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho, i}} \right) &= 0 \end{aligned}$$

<sup>9</sup>Esto proviene del hecho de que la definición de densidad de momento Ec. (18.47), solo depende de  $\dot{\eta}$  y no de las derivadas de  $\eta$  con respecto a las coordenadas espaciales.

se sigue de aquí que si  $\eta_\rho$  es cíclica, hay una integral que define una cantidad conservada

$$\Pi^\rho = \int dV \pi^\rho(x^i, t)$$

es importante enfatizar que no es  $\pi^\rho$  el que se conserva, sino su variable global asociada  $\Pi^\rho$ . La generalización de la Ec. (18.48) sería

$$\mathcal{H}(\eta_\lambda, \eta_{\lambda,i}, \pi^\lambda, x^i) \equiv \pi^\lambda \dot{\eta}_\lambda - \mathcal{L} \quad (18.50)$$

donde se asume que la dependencia funcional sobre  $\dot{\eta}_\lambda$  se puede eliminar por inversión de la ecuación de definición (18.49) tal como en el caso discreto. A partir de la definición (18.50) y usando (18.49) se sigue que

$$\frac{\partial \mathcal{H}}{\partial \pi^\rho} = \dot{\eta}_\rho + \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \pi^\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \pi^\rho} = \dot{\eta}_\rho \quad (18.51)$$

La otra mitad de las ecuaciones canónicas de campo son menos directas de obtener. Cuando se expresa en términos de variables canónicas,  $\mathcal{H}$  es una función de  $\eta_\rho$  a través de la dependencia explícita de  $\mathcal{L}$ , así como a través de  $\dot{\eta}_\rho$ , (nótese que al despejar la Ec. 18.49 tenemos  $\dot{\eta}_\rho$  como función de  $\pi^\rho, \eta_\rho, \eta_{\rho,i}$  y  $x^\mu$ ). Por tanto,

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \eta_\rho} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \eta_\rho} - \frac{\partial \mathcal{L}}{\partial \eta_\rho} = - \frac{\partial \mathcal{L}}{\partial \eta_\rho} \quad (18.52)$$

donde hemos usado (18.49). Ahora usando las Ecuaciones de Lagrange en el continuo Ecs. (18.27) esto queda

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = - \frac{d}{dx^\mu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\mu}} \right) = -\dot{\pi}^\rho - \frac{d}{dx^i} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} \right) \quad (18.53)$$

debido a la presencia de  $\mathcal{L}$  esta forma no es todavía útil. No obstante una derivación paralela de (18.50) muestra que

$$\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} = \pi^\lambda \frac{\partial \dot{\eta}_\lambda}{\partial \eta_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\lambda} \frac{\partial \dot{\eta}_\lambda}{\partial \eta_{\rho,i}} - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} = - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} \quad (18.54)$$

donde de nuevo hemos usado (18.49). Podemos escribir entonces la segunda mitad de ecuaciones de campo reemplazando (18.54) en (18.53) resultando

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} - \frac{d}{dx^i} \left( \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) = -\dot{\pi}^\rho \quad (18.55)$$

las Ecs. (18.51) y (18.55) se pueden colocar en una notación en la cual su apariencia se asemeja mas a la de las ecuaciones de Hamilton en el discreto. Para ello introducimos la noción de **funcional derivativa** definida por

$$\frac{\delta}{\delta \psi_\rho} \equiv \frac{\partial}{\partial \psi_\rho} - \frac{d}{dx^i} \frac{\partial}{\partial \psi_{\rho,i}} \quad (18.56)$$

y teniendo en cuenta que  $\mathcal{H}$  no es función de  $\pi_{,i}^\rho$ , las Ecs. (18.51) y (18.55) se pueden escribir como

$$\dot{\eta}_\rho = \frac{\delta \mathcal{H}}{\delta \pi^\rho} ; \quad \dot{\pi}^\rho = - \frac{\delta \mathcal{H}}{\delta \eta_\rho} \quad (18.57)$$

nótese que en esta notación las Ecuaciones de Lagrange en el continuo (18.27), quedan de la forma

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \right) - \frac{\delta \mathcal{L}}{\delta \eta_\rho} = 0$$

la única ventaja de la funcional derivativa sin embargo es la similaridad de las ecuaciones (tanto de Hamilton como de Lagrange) con respecto a aquellas en el discreto. Sin embargo, en el caso de las Ecs. de Lagrange, esta notación destruye la simetría con la que las coordenadas del espacio y el tiempo aparecían en esta formulación.

### 18.6.1. Propiedades básicas de $\mathcal{H}$

Podemos obtener otras propiedades de  $\mathcal{H}$  haciendo la derivada con respecto al tiempo de (18.50), teniendo en cuenta que hay que considerar a  $\dot{\eta}_\rho$  como funciones de  $\pi^\rho, \eta_\rho, \eta_{\rho,i}$  y  $x_\mu$ , con lo cual se obtiene

$$\frac{d\mathcal{H}}{dt} = \frac{d\pi^\rho}{dt} \dot{\eta}_\rho + \pi^\rho \frac{d\dot{\eta}_\rho}{dt} - \frac{\partial \mathcal{L}}{\partial \eta_\rho} \frac{d\eta_\rho}{dt} - \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \frac{d\dot{\eta}_\rho}{dt} - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} \frac{d\eta_{\rho,i}}{dt} - \frac{\partial \mathcal{L}}{\partial t}$$

el segundo y cuarto término a la derecha se cancelan entre sí en virtud de la definición (18.49) quedando

$$\frac{d\mathcal{H}}{dt} = \dot{\pi}^\rho \dot{\eta}_\rho - \frac{\partial \mathcal{L}}{\partial \eta_\rho} \dot{\eta}_\rho - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} \frac{d\eta_{\rho,i}}{dt} - \frac{\partial \mathcal{L}}{\partial t} \quad (18.58)$$

por otra parte cuando consideramos  $\mathcal{H}$  en términos de sus momentos canónicos, en cuyo caso los argumentos de  $\mathcal{H}$  son  $\pi^\rho, \eta_\rho, \eta_{\rho,i}$  y  $x_\mu$  la derivada temporal total adquiere la forma

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial \pi^\rho} \dot{\pi}^\rho + \frac{\partial \mathcal{H}}{\partial \eta_\rho} \dot{\eta}_\rho + \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \frac{d\eta_{\rho,i}}{dt} + \frac{\partial \mathcal{H}}{\partial t} \quad (18.59)$$

comparando (18.58) con (18.59), se ve que los primeros términos a la derecha de éstas, son iguales en virtud del primer conjunto de ecuaciones de Hamilton Ecs. (18.51), los segundos términos a la derecha también coinciden debido a (18.52), finalmente los terceros términos coinciden en virtud de (18.54). Por lo tanto, los últimos términos deben coincidir

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} \quad (18.60)$$

que es el equivalente de (6.11) cuando el medio es discreto, y nos dice que cuando  $\mathcal{L}$  no contenga explícitamente al tiempo tampoco lo contendrá  $\mathcal{H}$  y viceversa.

Por otro lado, la análoga de la ecuación (6.38) **no** se cumple, es decir, las derivadas temporales parcial y total de la densidad Hamiltoniana no coinciden. Para calcular la relación que hay entre ellas partimos de la Ec. (18.59) y usamos las Ecuaciones de movimiento de Hamilton (18.51, 18.55)

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \frac{\partial \mathcal{H}}{\partial \pi^\rho} \left[ \frac{d}{dx^i} \left( \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) - \frac{\partial \mathcal{H}}{\partial \eta_\rho} \right] + \frac{\partial \mathcal{H}}{\partial \eta_\rho} \left[ \frac{\partial \mathcal{H}}{\partial \pi^\rho} \right] + \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \frac{d}{dx_i} \left( \frac{d\eta_\rho}{dt} \right) + \frac{\partial \mathcal{H}}{\partial t} \\ \frac{d\mathcal{H}}{dt} &= \frac{\partial \mathcal{H}}{\partial \pi^\rho} \left[ \frac{d}{dx^i} \left( \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) \right] + \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \frac{d\dot{\eta}_\rho}{dx_i} + \frac{\partial \mathcal{H}}{\partial t} \end{aligned}$$

hemos intercambiado las derivadas sobre el campo en  $x_i$  y el tiempo dado que las derivadas totales del campo con respecto a los parámetros coinciden con las derivadas parciales. Usando de nuevo (18.51) resulta

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \dot{\eta}_\rho \left[ \frac{d}{dx^i} \left( \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) \right] + \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \frac{d\dot{\eta}_\rho}{dx_i} + \frac{\partial \mathcal{H}}{\partial t} \\ \frac{d\mathcal{H}}{dt} &= \frac{d}{dx^i} \left( \dot{\eta}_\rho \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) + \frac{\partial \mathcal{H}}{\partial t} \end{aligned} \quad (18.61)$$

de modo que las derivadas temporales total y parcial de  $\mathcal{H}$  no coinciden como se predijo<sup>10</sup>. Recordando que el elemento  $T_0^0$  del tensor de esfuerzo energía se interpretaba como densidad de la función energía (y por tanto con la densidad Hamiltoniana) podemos identificar a  $\mathcal{H}$  con  $T_0^0$ . Por otro lado, usando la Ec. (18.54) se deduce que

$$\dot{\eta}_\rho \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} = -\dot{\eta}_\rho \frac{\partial \mathcal{L}}{\partial \eta_{\rho,i}} = -T_0^i$$

<sup>10</sup>Es importante notar que las derivadas parciales y totales con respecto a  $x^\mu$  solo coinciden para funciones que solo dependan de  $x^\mu$  como ocurre con los campos y sus derivadas. En contraste, cantidades tales como las densidades Lagrangiana y Hamiltoniana dependen no solo de  $x^\mu$  sino también de campos y derivadas de los campos, de modo que sus derivadas totales y parciales con respecto a  $x^\mu$  no coinciden en general.



Si asumimos que  $\mathcal{H}$  no depende explícitamente del tiempo, la Ec. (18.61) se reduce a

$$\frac{dT_0^0}{dt} + \frac{dT_0^i}{dt} = 0$$

que es la ecuación asociada a la conservación de la corriente  $\mathbf{T}_0$ . Ya vimos que la existencia de una corriente conservada implica la conservación de una cantidad integral que es la integral de volumen de la componente cero de la 4-corriente (que corresponde a una densidad). En este caso

$$P_0 \equiv H = \int \mathcal{H} dV$$

obsérvese que aún en el caso en el cual  $\mathcal{H}$  no depende explícitamente del tiempo, la cantidad que se conserva es  $H$  y no  $\mathcal{H}$ . Nótese que al contrario del caso Lagrangiano, aquí primero definimos las densidades y luego las cantidades totales. Debe tenerse presente que el volumen de integración debe contener a todos los campos. Finalmente, notemos que en el caso en el cual el Hamiltoniano no sea la energía del sistema, lo que se conserva es el Hamiltoniano y no la energía, pues  $T_0^0$  realmente corresponde a la densidad de la función energía que numéricamente es el Hamiltoniano.

### 18.6.2. Densidades generalizadas y corchetes de Poisson

El Hamiltoniano total  $H$  es solo un caso particular de funciones que se obtienen integrando densidades en el volumen. Veamos entonces el formalismo general para estas cantidades integrales. Consideremos cierta densidad  $\mathcal{U}$  que sea función de las coordenadas  $(\eta_\rho, \pi^\rho)$  del espacio de fase así como de sus gradientes espaciales y posiblemente  $x^\mu$

$$\mathcal{U} = \mathcal{U}(\eta_\rho, \eta_{\rho,i}, \pi^\rho, \pi^{\rho,i}, x_\mu)$$

la cantidad integral correspondiente la denotamos por  $U$

$$U(t) = \int \mathcal{U} dV \quad (18.62)$$

donde la integral se realiza sobre un volumen fuera del cual  $\eta_\rho$  y  $\pi^\rho$  se deben anular (también deben ser nulas estas cantidades en el contorno que define al volumen). Derivando  $U$  respecto al tiempo resulta

$$\frac{dU}{dt} = \int dV \left\{ \frac{\partial \mathcal{U}}{\partial \eta_\rho} \dot{\eta}_\rho + \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \dot{\eta}_{\rho,i} + \frac{\partial \mathcal{U}}{\partial \pi^\rho} \dot{\pi}^\rho + \frac{\partial \mathcal{U}}{\partial \pi^{\rho,i}} \dot{\pi}^{\rho,i} + \frac{\partial \mathcal{U}}{\partial t} \right\} \quad (18.63)$$

consideremos un término tal como

$$\int dV \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \dot{\eta}_{\rho,i} = \int dx^i \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \frac{d\dot{\eta}_\rho}{dx^i}$$

integrando por partes

$$u = \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \quad ; \quad dv = dx^i \frac{d\dot{\eta}_\rho}{dx^i} \Rightarrow$$

$$\int dx^i \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \frac{d\dot{\eta}_\rho}{dx^i} = \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \dot{\eta}_\rho \Big|_S - \int dx^i \dot{\eta}_\rho \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \right)$$

y recordando que  $\eta_\rho$  y las derivadas se anulan en la superficie del contorno, tenemos

$$\int dV \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \dot{\eta}_{\rho,i} = - \int dV \dot{\eta}_\rho \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \right)$$

y para el término con  $\dot{\pi}^{\rho,i}$  se hace una reducción análoga. Con lo cual la Ec. (18.63) queda en la forma

$$\begin{aligned} \frac{dU}{dt} &= \int dV \left\{ \frac{\partial \mathcal{U}}{\partial \eta_\rho} \dot{\eta}_\rho - \dot{\eta}_\rho \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \right) + \frac{\partial \mathcal{U}}{\partial \pi^\rho} \dot{\pi}^\rho - \dot{\pi}^\rho \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \pi^{\rho,i}} \right) + \frac{\partial \mathcal{U}}{\partial t} \right\} \\ \frac{dU}{dt} &= \int dV \left\{ \left[ \frac{\partial \mathcal{U}}{\partial \eta_\rho} - \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \right) \right] \dot{\eta}_\rho + \left[ \frac{\partial \mathcal{U}}{\partial \pi^\rho} - \frac{d}{dx^i} \left( \frac{\partial \mathcal{U}}{\partial \pi^{\rho,i}} \right) \right] \dot{\pi}^\rho + \frac{\partial \mathcal{U}}{\partial t} \right\} \end{aligned} \quad (18.64)$$

ahora introduciendo la notación de derivada funcional descrita por (18.56), la Ec. (18.64) se reduce a

$$\frac{dU}{dt} = \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \dot{\eta}_\rho + \frac{\delta \mathcal{U}}{\delta \pi^\rho} \dot{\pi}^\rho + \frac{\partial \mathcal{U}}{\partial t} \right\} \quad (18.65)$$

por último introducimos las ecuaciones de Hamilton en notación de derivadas funcionales Ecs. (18.57), con lo cual la Ec. (18.65) queda

$$\frac{dU}{dt} = \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \frac{\delta \mathcal{H}}{\delta \pi^\rho} - \frac{\delta \mathcal{H}}{\delta \eta_\rho} \frac{\delta \mathcal{U}}{\delta \pi^\rho} \right\} + \int dV \frac{\partial \mathcal{U}}{\partial t} \quad (18.66)$$

La primera integral tiene la forma de un corchete de Poisson aunque reemplazando al operador derivada ordinaria por la derivada funcional. Si  $\mathcal{U}$  y  $\mathcal{W}$  son dos funciones densidad cualquiera, lo anterior nos sugiere definir como los corchetes de Poisson a las siguientes cantidades integrales

$$[U, W] \equiv \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \frac{\delta \mathcal{W}}{\delta \pi^\rho} - \frac{\delta \mathcal{W}}{\delta \eta_\rho} \frac{\delta \mathcal{U}}{\delta \pi^\rho} \right\} \quad (18.67)$$

definimos el significado de la derivada parcial en el tiempo de  $U$  en la forma evidente

$$\frac{\partial U}{\partial t} = \int dV \frac{\partial \mathcal{U}}{\partial t}$$

con lo cual la Ec. (18.66) se puede escribir en la forma

$$\frac{dU}{dt} = [U, H] + \frac{\partial U}{\partial t} \quad (18.68)$$

que es el análogo a la Ec. (8.25) para sistemas discretos, aunque debe tenerse en cuenta que los corchetes de Poisson han sido redefinidos en forma integral y con la derivada funcional. Haciendo  $U = H$  en (18.68) y teniendo en cuenta que por definición el corchete de Poisson de  $H$  consigo mismo es nulo, se obtiene inmediatamente que

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \quad (18.69)$$

que es la forma integral de (18.61)<sup>11</sup> y la versión en teoría de campos de (6.38). Nótese que aunque las derivadas total y parcial no coinciden para la densidad Hamiltoniana, sí coinciden para el Hamiltoniano total.

Se ve entonces que el formalismo de los corchetes de Poisson aparece como consecuencia de la formulación de Hamilton. Sin embargo, no se puede desarrollar una teoría de campos en términos de corchetes de Poisson de tal manera que guarde una analogía pasa a paso con la de los sistemas discretos. Por ejemplo, aquí definimos los corchetes de Poisson solo en función de un par de densidades. Por tanto no se puede con facilidad plantear corchetes de Poisson fundamentales como lo hacíamos en sistemas discretos. Esto se debe a que aunque  $\pi^\rho$  es una densidad,  $\eta_\rho$  no lo es. Además si  $x_i$  desempeña el papel de índices continuos de las variables mecánicas, los corchetes de Poisson fundamentales contendrían funciones para diferentes valores de  $x_i$ , lo cual no se introduce fácilmente en la presente formulación. Por esta razón las transformaciones canónicas no son muy usadas en teorías de campos clásicos a pesar de ser tan fructíferas en sistemas discretos. También es difícil en este formalismo seguir los pasos de la cuantización, que requieren cambiar los corchetes de Poisson por conmutadores cuánticos.

### 18.6.3. Formulación por corchetes de Poisson utilizando descomposición de Fourier

Hay una forma alternativa que nos va a permitir trabajar casi todo el formalismo de Hamilton y de corchetes de Poisson para campos clásicos, que ya conocemos para el caso discreto. La idea de este tratamiento consiste en sustituir la variable espacial continua o el índice continuo, por un índice discreto numerable. Retomemos

<sup>11</sup>La integración del primer término a la derecha de la Ec. (18.61) se anula al tomar un volumen que contenga todos los campos y derivadas, como se puede ver aplicando el teorema de la divergencia.

una vez más el ejemplo de la varilla elástica. Supongamos que dicha varilla posee una longitud  $L = x_2 - x_1$ . Pondremos como condición de contorno que  $\eta$  se anule en los extremos, lo cual se logra colocando a la varilla con sus extremos en contacto con dos paredes rígidas. La amplitud de la oscilación se podrá escribir en una descomposición de Fourier discreta

$$\eta(x) = \sum_{n=0}^{\infty} q_n \sin \frac{2\pi n(x-x_1)}{2L} \quad (18.70)$$

en lugar del índice continuo  $x$  aquí tenemos el índice discreto  $n$ . Esta representación solo es válida si  $\eta(x)$  es una función regular, lo cual usualmente se cumple.

Para el caso general, supondremos por simplicidad, que hay solo una cantidad de campo real  $\eta$  que se puede desarrollar en serie de Fourier tridimensional

$$\eta(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} q_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (18.71)$$

donde  $\mathbf{k}$  es un “vector de onda” que solo puede tomar módulos y direcciones discretos de modo que en una dimensión lineal dada solo encaje un número entero (o tal vez semientero) de “longitudes de onda”. Decimos que  $\mathbf{k}$  tiene un espectro discreto. El índice escalar  $\mathbf{k}$  representa entonces una ordenación del sistema de índices enteros que se utiliza para enumerar los valores discretos de  $\mathbf{k}$ , y  $V$  es el volumen del sistema que actúa como un factor de normalización. Como este factor es real se tiene que  $q_{\mathbf{k}}^* = q_{-\mathbf{k}}$ .

La ortogonalidad de las exponenciales en el volumen se manifiesta como

$$\frac{1}{V} \int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} dV = \delta_{\mathbf{k},\mathbf{k}'} \quad (18.72)$$

en realidad los valores permitidos de  $\mathbf{k}$  son aquellos para los que se cumple la condición de ortogonalidad. Los coeficientes de la expansión se pueden calcular con base en la ortonormalidad de la siguiente forma

$$q_{\mathbf{k}}(t) = \frac{1}{\sqrt{V}} \int e^{-i\mathbf{k}\cdot\mathbf{r}} \eta(\mathbf{r}, t) dV \equiv \int \tilde{\eta}(\mathbf{k}, \mathbf{r}, t) dV \quad ; \quad \tilde{\eta}(\mathbf{k}, \mathbf{r}, t) \equiv \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{r}} \eta(\mathbf{r}, t)$$

donde  $\tilde{\eta}(\mathbf{k}, \mathbf{r}, t)$  es la función densidad correspondiente a la variable global  $q_{\mathbf{k}}$ , como se puede ver por comparación con la Ec. (18.62). De manera análoga, la densidad de cantidad de movimiento se puede escribir como

$$\pi(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} p_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{r}} \quad (18.73)$$

donde también se cumple que  $p_{\mathbf{k}}^* = p_{-\mathbf{k}}$ . Los coeficientes  $p_{\mathbf{k}}(t)$  también se obtienen por ortonormalidad

$$p_{\mathbf{k}}(t) = \frac{1}{\sqrt{V}} \int e^{i\mathbf{k}\cdot\mathbf{r}} \pi(\mathbf{r}, t) dV \equiv \int \tilde{\pi}(\mathbf{k}, \mathbf{r}, t) dV \quad ; \quad \tilde{\pi}(\mathbf{k}, \mathbf{r}, t) \equiv \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \pi(\mathbf{r}, t)$$

siendo  $\tilde{\pi}(\mathbf{k}, \mathbf{r}, t)$  la función densidad asociada a  $p_{\mathbf{k}}(t)$ . Tanto  $q_{\mathbf{k}}$  como  $p_{\mathbf{k}}$  son cantidades integrales de la forma descrita por (18.62). Podemos entonces calcular los corchetes de Poisson de dichas cantidades. Como las exponenciales no contienen variables de campo, tenemos en virtud de la definición (18.67)

$$\begin{aligned} [q_{\mathbf{k}}, p_{\mathbf{k}'}] &= \int dV \left\{ \frac{\delta \tilde{\eta}(\mathbf{k}, \mathbf{r}, t)}{\delta \eta} \frac{\delta \tilde{\pi}(\mathbf{k}', \mathbf{r}, t)}{\delta \pi} - \frac{\delta \tilde{\pi}(\mathbf{k}', \mathbf{r}, t)}{\delta \eta} \frac{\delta \tilde{\eta}(\mathbf{k}, \mathbf{r}, t)}{\delta \pi} \right\} \\ [q_{\mathbf{k}}, p_{\mathbf{k}'}] &= \frac{1}{V} \int dV e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} \left\{ \frac{\delta \eta}{\delta \eta} \frac{\delta \pi}{\delta \pi} - \frac{\delta \pi}{\delta \eta} \frac{\delta \eta}{\delta \pi} \right\} \\ [q_{\mathbf{k}}, p_{\mathbf{k}'}] &= \frac{1}{V} \int dV e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} \end{aligned}$$

y usando la ortonormalidad (18.72)

$$[q_{\mathbf{k}}, p_{\mathbf{k}'}] = \delta_{\mathbf{k}\mathbf{k}'} \quad (18.74)$$

usando la definición de corchetes de Poisson, se puede verificar que

$$[q_{\mathbf{k}}, q_{\mathbf{k}'}] = [p_{\mathbf{k}}, p_{\mathbf{k}'}] = 0 \quad (18.75)$$

por tanto, los corchetes de Poisson de  $q_{\mathbf{k}}, p_{\mathbf{k}}$  forman un conjunto de corchetes fundamentales con las mismas propiedades que aparecen en el discreto Ecs. (8.3). Esto nos sugiere considerar a estas cantidades como coordenadas canónicamente conjugadas. Vale la pena entonces estudiar las ecuaciones de movimiento que surgen para estas coordenadas. Empleando la Ec. (18.68), la evolución temporal de  $q_{\mathbf{k}}(t)$  se escribe como

$$\begin{aligned} \dot{q}_{\mathbf{k}}(t) &= [q_{\mathbf{k}}, H] = \int dV \left\{ \frac{\delta \tilde{\eta}(\mathbf{k}, \mathbf{r}, t)}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \mathcal{H}}{\delta \eta} \frac{\delta \tilde{\eta}(\mathbf{k}, \mathbf{r}, t)}{\delta \pi} \right\} = \frac{1}{\sqrt{V}} \int dV e^{-i\mathbf{k}\cdot\mathbf{r}} \left\{ \frac{\delta \eta}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \mathcal{H}}{\delta \eta} \frac{\delta \eta}{\delta \pi} \right\} \\ \dot{q}_{\mathbf{k}}(t) &= \frac{1}{\sqrt{V}} \int dV e^{-i\mathbf{k}\cdot\mathbf{r}} \frac{\delta \mathcal{H}}{\delta \pi} \end{aligned} \quad (18.76)$$

por otra parte, tenemos

$$\frac{\partial H}{\partial p_{\mathbf{k}}} = \int dV \frac{\partial \mathcal{H}}{\partial \pi} \frac{\partial \pi}{\partial p_{\mathbf{k}}} = \int dV \frac{\delta \mathcal{H}}{\delta \pi} \frac{\partial \pi}{\partial p_{\mathbf{k}}} \quad (18.77)$$

donde hemos tenido en cuenta que  $\mathcal{H}$  no es función del gradiente de  $\pi$  y por tanto su derivada parcial con respecto a  $\pi$  coincide con su derivada funcional. Teniendo en cuenta (18.73) resulta

$$\frac{\partial \pi}{\partial p_{\mathbf{k}}} = \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{r}}$$

con lo cual (18.77) coincide con (18.76) resultando

$$\dot{q}_{\mathbf{k}} = \frac{\partial H}{\partial p_{\mathbf{k}}} \quad (18.78)$$

veamos ahora como se obtiene la ecuación de movimiento para  $p_{\mathbf{k}}$ . Los pasos son similares

$$\begin{aligned} \dot{p}_{\mathbf{k}}(t) &= [p_{\mathbf{k}}, H] = \frac{1}{\sqrt{V}} \int dV e^{i\mathbf{k}\cdot\mathbf{r}} \left\{ \frac{\delta \pi}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \mathcal{H}}{\delta \eta} \frac{\delta \pi}{\delta \pi} \right\} \\ \dot{p}_{\mathbf{k}}(t) &= -\frac{1}{\sqrt{V}} \int dV e^{i\mathbf{k}\cdot\mathbf{r}} \frac{\delta \mathcal{H}}{\delta \eta} \end{aligned} \quad (18.79)$$

y calculamos

$$\frac{\partial H}{\partial q_{\mathbf{k}}} = \int dV \left\{ \frac{\partial \mathcal{H}}{\partial \eta} \frac{\partial \eta}{\partial q_{\mathbf{k}}} + \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \frac{\partial \eta_{,j}}{\partial q_{\mathbf{k}}} \right\} \quad (18.80)$$

la diferencia fundamental con el cálculo anterior es que  $\mathcal{H}$  no depende del gradiente de  $\pi$  pero sí depende del gradiente de  $\eta$ . El término que contiene las componentes del gradiente de  $\eta$  se integra por partes,

$$\begin{aligned} \int dV \left\{ \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \frac{\partial \eta_{,j}}{\partial q_{\mathbf{k}}} \right\} &= \int dx^j \left\{ \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \frac{d}{dx^j} \left( \frac{\partial \eta}{\partial q_{\mathbf{k}}} \right) \right\} \\ u &= \frac{\partial \mathcal{H}}{\partial \eta_{,j}} ; \quad dv = dx^j \frac{d}{dx^j} \left( \frac{\partial \eta}{\partial q_{\mathbf{k}}} \right) \\ \int dV \left\{ \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \frac{\partial \eta_{,j}}{\partial q_{\mathbf{k}}} \right\} &= \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \frac{\partial \eta}{\partial q_{\mathbf{k}}} \Big|_S - \int dx^j \frac{\partial \eta}{\partial q_{\mathbf{k}}} \frac{d}{dx^j} \left( \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \right) \end{aligned}$$

la integral de superficie se anula de modo que la integral (18.80) queda

$$\frac{\partial H}{\partial q_{\mathbf{k}}} = \int dV \left\{ \frac{\partial \mathcal{H}}{\partial \eta} \frac{\partial \eta}{\partial q_{\mathbf{k}}} - \frac{\partial \eta}{\partial q_{\mathbf{k}}} \frac{d}{dx^j} \left( \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \right) \right\} = \int dV \left\{ \frac{\partial \mathcal{H}}{\partial \eta} - \frac{d}{dx^j} \left( \frac{\partial \mathcal{H}}{\partial \eta_{,j}} \right) \right\} \frac{\partial \eta}{\partial q_{\mathbf{k}}}$$

la cual se puede reducir a otra que contenga la derivada funcional

$$\frac{\partial H}{\partial q_{\mathbf{k}}} = \int dV \frac{\delta \mathcal{H}}{\delta \eta} \frac{\partial \eta}{\partial q_{\mathbf{k}}}$$

y teniendo en cuenta (18.71)

$$\frac{\partial \eta}{\partial q_{\mathbf{k}}} = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}$$

quedando finalmente

$$\frac{\partial H}{\partial q_{\mathbf{k}}} = \frac{1}{\sqrt{V}} \int dV \frac{\delta \mathcal{H}}{\delta \eta} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (18.81)$$

igualando (18.79) con (18.81) se encuentra que

$$\dot{p}_{\mathbf{k}}(t) = -\frac{\partial H}{\partial q_{\mathbf{k}}} \quad (18.82)$$

las cantidades  $q_{\mathbf{k}}$  y  $p_{\mathbf{k}}$  obedecen entonces las ecuaciones de movimiento de Hamilton. En general  $\mathbf{k}$  denota un conjunto de índices enteros.

Nótese que en el capítulo se comenzó con un sistema discreto (coordenadas generalizadas numerables) para el cual se estudió el paso al continuo (coordenadas generalizadas no numerables). Posteriormente, hemos descrito al sistema continuo mediante un sistema de coordenadas discreto y numerable que posee una estructura matemática similar a la del sistema discreto del cual se partió. Debido a esta correspondencia entre  $q_{\mathbf{k}}$ ,  $p_{\mathbf{k}}$  con las variables canónicamente conjugadas en el discreto, estas cantidades son buenos candidatos para la cuantización cuando pasemos a una teoría cuántica de campos. Las  $q_{\mathbf{k}}$  corresponden a lo que se denomina “número de ocupación” para el campo.

La razón por la cual se pudo describir el sistema continuo por medio de coordenadas discretas, es que el tamaño finito del sistema y las condiciones de contorno permiten un desarrollo de Fourier discreto<sup>12</sup>. El desarrollo presente se efectuó sobre un espectro discreto de ondas planas. Dado que el vector de onda  $\mathbf{k}$  es en mecánica cuántica directamente proporcional a la cantidad de movimiento de la partícula asociada a la onda plana, los desarrollos que hemos utilizado reciben a veces el nombre de representación por cantidad de movimiento. Sin embargo, este desarrollo se puede realizar en cualquier conjunto ortonormalizable y numerable de funciones que cumplan la propiedad de completitud en el volumen en cuestión. Las ondas planas son solo un caso particular de estas funciones. El empleo de otras bases para estas expansiones puede ser más conveniente, dependiendo del problema a estudiar, pero el procedimiento es análogo al aquí expuesto.

## 18.7. Ejemplos de teorías de campos

### 18.7.1. Un modelo juguete

Consideremos un campo de dos componentes  $\eta_1$  y  $\eta_2$  cuyas ecuaciones de movimiento fuesen

$$\begin{aligned} a\nabla^2 \eta_1 + b \frac{d\eta_1}{dt} - c \frac{d^2 \eta_1}{dt^2} &= f \eta_2 \\ a\nabla^2 \eta_2 - b \frac{d\eta_2}{dt} - c \frac{d^2 \eta_2}{dt^2} &= g \eta_1 \end{aligned} \quad (18.83)$$

el siguiente paso natural es encontrar una densidad Lagrangiana que a través de las ecuaciones de Euler-Lagrange Ecs. (18.27) me genere las ecuaciones de campo (18.83), recordemos que esta densidad Lagrangiana no es única, una de ellas es

$$\mathcal{L} = c \dot{\eta}_1 \dot{\eta}_2 + \frac{b}{2} (\dot{\eta}_1 \eta_2 - \eta_1 \dot{\eta}_2) - a \nabla \eta_1 \cdot \nabla \eta_2 - \frac{1}{2} (f \eta_2^2 + g \eta_1^2)$$

<sup>12</sup>Por ejemplo, si el sistema no es acotado la descomposición de Fourier se vuelve en general continua. Por otro lado, la anulación del campo en los extremos permitió que las sumatorias fueran desde  $n = 0$  hasta  $\infty$ . En general estas sumatorias irán desde  $-\infty$  hasta  $\infty$ .

para esta densidad Lagrangiana, las componentes del tensor esfuerzo energía son

$$\begin{aligned} T_{00} &= c\dot{\eta}_1\dot{\eta}_2 + a\nabla\eta_1 \cdot \nabla\eta_2 + \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \\ T_{0i} &= -a\left(\dot{\eta}_1\frac{d\eta_2}{dx_i} + \dot{\eta}_2\frac{d\eta_1}{dx_i}\right) \\ T_{i0} &= c\left(\dot{\eta}_1\frac{d\eta_2}{dx_i} + \dot{\eta}_2\frac{d\eta_1}{dx_i}\right) + \frac{b}{2}\left(\eta_2\frac{d\eta_1}{dx_i} - \eta_1\frac{d\eta_2}{dx_i}\right) \\ T_{ij} &= -a\left(\frac{d\eta_1}{dx_i}\frac{d\eta_2}{dx_j} + \frac{d\eta_1}{dx_j}\frac{d\eta_2}{dx_i}\right) - \mathcal{L}\delta_{ij} \end{aligned}$$

obsérvese que en este ejemplo, el tensor tridimensional  $T_{ij}$  es simétrico.

Hagamos los cálculos explícitos. Podemos escribir la densidad lagrangiana en una notación más conveniente

$$\begin{aligned} \mathcal{L} &= c\dot{\eta}_1\dot{\eta}_2 + \frac{b}{2}(\dot{\eta}_1\eta_2 - \eta_1\dot{\eta}_2) - a(\partial_i\eta_1)(\partial_i\eta_2) - \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \\ \mathcal{L} &= c\dot{\eta}_1\dot{\eta}_2 + \frac{b}{2}(\dot{\eta}_1\eta_2 - \eta_1\dot{\eta}_2) - a\eta_{1,i}\eta_{2,i} - \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \end{aligned}$$

Usando las ecuaciones de Lagrange para el continuo

$$\frac{d}{dx^\nu}\left(\frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}}\right) - \frac{\partial\mathcal{L}}{\partial\eta_\rho} = 0$$

tenemos que para el campo  $\eta_1$

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\eta}_1}\right) + \frac{d}{dx^k}\left(\frac{\partial\mathcal{L}}{\partial\eta_{1,k}}\right) - \frac{\partial\mathcal{L}}{\partial\eta_1} = 0$$

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\dot{\eta}_1} &= \frac{\partial}{\partial\dot{\eta}_1}\left[c\dot{\eta}_1\dot{\eta}_2 + \frac{b}{2}\dot{\eta}_1\eta_2\right] = c\dot{\eta}_2 + \frac{b}{2}\eta_2 \Rightarrow \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\eta}_1}\right) = c\ddot{\eta}_2 + \frac{b}{2}\dot{\eta}_2 \\ \frac{\partial\mathcal{L}}{\partial\eta_{1,k}} &= \frac{\partial}{\partial\eta_{1,k}}[-a\eta_{1,i}\eta_{2,i}] = -a\delta_{ik}\eta_{2,i} = -a\eta_{2,k} \Rightarrow \frac{d}{dx^k}\left(\frac{\partial\mathcal{L}}{\partial\eta_{1,k}}\right) = -a\eta_{2,kk} \\ -\frac{\partial\mathcal{L}}{\partial\eta_1} &= -\frac{\partial}{\partial\eta_1}\left[-\frac{b}{2}(\eta_1\dot{\eta}_2) - \frac{1}{2}g\eta_1^2\right] = \frac{b}{2}\dot{\eta}_2 + g\eta_1 \end{aligned}$$

la ecuación de Lagrange es

$$\begin{aligned} c\ddot{\eta}_2 + \frac{b}{2}\dot{\eta}_2 - a\eta_{2,kk} + \frac{b}{2}\dot{\eta}_2 + g\eta_1 &= 0 \\ -a\nabla^2\eta_2 + c\frac{d^2\eta_2}{dt^2} + b\frac{d\eta_2}{dt} + g\eta_1 &= 0 \end{aligned}$$

similarmente se obtiene la ecuación de Lagrange para  $\eta_2$

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\eta}_2}\right) + \frac{d}{dx^k}\left(\frac{\partial\mathcal{L}}{\partial\eta_{2,k}}\right) - \frac{\partial\mathcal{L}}{\partial\eta_2} = 0$$

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\dot{\eta}_2} &= \frac{\partial}{\partial\dot{\eta}_2}\left[c\dot{\eta}_1\dot{\eta}_2 - \frac{b}{2}(\eta_1\dot{\eta}_2)\right] = c\dot{\eta}_1 - \frac{b}{2}\eta_1 \Rightarrow \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\eta}_2}\right) = c\ddot{\eta}_1 - \frac{b}{2}\dot{\eta}_1 \\ \frac{\partial\mathcal{L}}{\partial\eta_{2,k}} &= \frac{\partial}{\partial\eta_{2,k}}[-a\eta_{1,i}\eta_{2,i}] = -a\delta_{ik}\eta_{1,i} = -a\eta_{1,k} \Rightarrow \frac{d}{dx^k}\left(\frac{\partial\mathcal{L}}{\partial\eta_{2,k}}\right) = -a\eta_{1,kk} \\ -\frac{\partial\mathcal{L}}{\partial\eta_2} &= -\frac{\partial}{\partial\eta_2}\left[\frac{b}{2}\dot{\eta}_1\eta_2 - \frac{1}{2}f\eta_2^2\right] = -\frac{b}{2}\dot{\eta}_1 + f\eta_2 \end{aligned}$$

la ecuación de Lagrange es

$$\begin{aligned} c\ddot{\eta}_1 - \frac{b}{2}\dot{\eta}_1 - a\eta_{1,kk} - \frac{b}{2}\dot{\eta}_1 + f\eta_2 &= 0 \\ -a\nabla^2\eta_1 + c\frac{d^2\eta_1}{dt^2} - b\frac{d\eta_1}{dt} + f\eta_2 &= 0 \end{aligned}$$

finalmente la densidad de energía se obtiene recordando la definición del tensor esfuerzo energía

$$\begin{aligned} T_\mu^\nu &\equiv \frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}}\eta_{\rho,\mu} - \mathcal{L}\delta_{\mu\nu} \Rightarrow T_0^0 \equiv \frac{\partial\mathcal{L}}{\partial\eta_{\rho,0}}\eta_{\rho,0} - \mathcal{L}\delta_{00} = \frac{\partial\mathcal{L}}{\partial\dot{\eta}_\rho}\dot{\eta}_\rho - \mathcal{L} \\ T_0^0 &= \frac{\partial\mathcal{L}}{\partial\dot{\eta}_1}\dot{\eta}_1 + \frac{\partial\mathcal{L}}{\partial\dot{\eta}_2}\dot{\eta}_2 - \mathcal{L} \\ T_0^0 &= \left(c\dot{\eta}_2 + \frac{b}{2}\eta_2\right)\dot{\eta}_1 + \left(c\dot{\eta}_1 - \frac{b}{2}\eta_1\right)\dot{\eta}_2 - \left[c\dot{\eta}_1\dot{\eta}_2 + \frac{b}{2}(\dot{\eta}_1\eta_2 - \eta_1\dot{\eta}_2) - a\eta_{1,i}\eta_{2,i} - \frac{1}{2}(f\eta_2^2 + g\eta_1^2)\right] \\ T_0^0 &= c\dot{\eta}_2\dot{\eta}_1 + \frac{b}{2}\eta_2\dot{\eta}_1 + c\dot{\eta}_1\dot{\eta}_2 - \frac{b}{2}\eta_1\dot{\eta}_2 - c\dot{\eta}_1\dot{\eta}_2 - \frac{b}{2}(\dot{\eta}_1\eta_2 - \eta_1\dot{\eta}_2) + a\eta_{1,i}\eta_{2,i} + \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \\ T_0^0 &= c\dot{\eta}_2\dot{\eta}_1 - c\dot{\eta}_1\dot{\eta}_2 + c\dot{\eta}_1\dot{\eta}_2 + \frac{b}{2}\eta_2\dot{\eta}_1 - \frac{b}{2}\eta_1\dot{\eta}_2 - \frac{b}{2}(\dot{\eta}_1\eta_2 - \eta_1\dot{\eta}_2) + a\nabla\eta_1 \cdot \nabla\eta_2 + \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \\ T_0^0 &= c\dot{\eta}_2\dot{\eta}_1 + a\nabla\eta_1 \cdot \nabla\eta_2 + \frac{1}{2}(f\eta_2^2 + g\eta_1^2) \end{aligned}$$

De la misma forma se pueden calcular los otros elementos del tensor esfuerzo energía. Nótese que si podemos encerrar todos los campos y sus derivadas en un volumen, podemos garantizar que

$$R_0 = \int T_0^0 dV = cte$$

para saber que volumen es adecuado para incluir la carga conservada generalizada (hay cuatro estrictamente), será necesario conocer el comportamiento del campo y sus condiciones de frontera. Por otro lado, para esta teoría de campo es posible que  $T_0^0$  no esté asociado a una densidad de energía, ni que las otras cargas conservadas estén asociadas a momentos. Todo lo que podemos afirmar es que hay cuatro cargas generalizadas conservadas que incluso pueden ser totalmente independientes unas de otras. Cualquier interpretación adicional depende del modelo de campo específico.

## 18.8. Teoría de campos relativista

Hemos visto en la sección 16.10 que una formulación Lagrangiana o Hamiltoniana manifiestamente covariante de Lorentz resulta particularmente difícil de implementar en la mecánica de partículas puntuales. Esta dificultad está principalmente relacionada con el hecho de que la formulación no relativista de la mecánica del punto es muy asimétrica en su tratamiento de la variable espacial y la temporal. Las variables espaciales son cantidades mecánicas en tanto que el tiempo es un parámetro que se recorre en forma monótona. En contraste, la mecánica no relativista de medios continuos presenta las coordenadas de espacio tiempo de una forma mucho más simétrica, al menos en la formulación Lagrangiana. Ello permitirá realizar una formulación que sea manifiestamente covariante de Lorentz con algunos pequeños cambios de la versión no relativista.

Hay tres aspectos que deben trabajarse para llegar a esta formulación manifiestamente covariante de Lorentz (1) la estructura y métrica del espacio tiempo, (2) las propiedades de transformación de los campos, densidades lagrangianas etc. cuando se realiza una transformación de Lorentz, (3) la descripción covariante de los límites de integración.

En lo que respecta al primer punto, es posible trabajar en el espacio tiempo usando la métrica usual pero con coordenada temporal imaginaria (espacio de Minkowski) o por otro lado, trabajar con el espacio tiempo

usual pero con métrica no trivial (métrica de Minkowski, ver cap. 16). Al igual que en el cap. 16 trabajaremos en el espacio de Minkowski con la métrica usual. Como hasta ahora, los índices latinos indican coordenadas espaciales y los índices griegos nos denotan las coordenadas espacio temporales, pero en este caso la coordenada temporal se escribirá como  $x_4 = ict$ . En primer lugar vemos que las Ecuaciones 18.27 de Lagrange en el continuo no se afectan por este cambio, ya que el término

$$\frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \right)$$

queda invariante ante un cambio de escala de cualquiera de los parámetros  $x_\nu$ , por ejemplo para  $x_4$  tenemos que

$$\frac{d}{dx^4} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,4}} \right) = \frac{d}{idx^0} \left( \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{idx^0} \right)} \right) = \frac{d}{dx^0} \left( \frac{\partial \mathcal{L}}{\partial \eta_{\rho,0}} \right)$$

y por otro lado el otro término en las ecuaciones de Lagrange no contiene coordenadas. En lo que respecta al principio de Hamilton, el cambio de espacio no lo afecta dado que simplemente se introduce una constante multiplicativa.

$$\begin{aligned} \delta I &= \delta \int \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu) dx^\mu = 0 \quad ; \quad dx^\mu \equiv dx^0 dx^1 dx^2 dx^3 = -idx^4 dx^1 dx^2 dx^3 \\ \delta I &= -i\delta \int_{-ix_{4,1}}^{-ix_{4,2}} \int_V \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu) dx'^\mu = 0 \quad ; \quad dx'^\mu \equiv dx^4 dx^1 dx^2 dx^3 \end{aligned}$$

el cambio de variable  $x'_4 = -ix_4$  saca otra constante multiplicativa de modo que el principio variacional permanece inmodificado.

Los campos y las cantidades relacionadas con éstos deben formar tensores de Lorentz de determinado rango para que se puedan escribir ecuaciones de campo manifiestamente covariantes. Si tenemos varios campos cada uno de ellos puede ser un tensor de rango diferente, lo importante es que las ecuaciones de movimiento sean manifiestamente covariantes. Las densidades lagrangiana y Hamiltoniana también deben ser covariantes. En el principio de Hamilton el elemento de cuadrivolumen es invariante de Lorentz. Si hemos de asumir que la acción es escalar de Lorentz (ya que es un escalar euclidiano en el régimen no relativista) esto implica que las densidades Lagrangiana y Hamiltoniana sean escalares de Lorentz. De lo anterior se desprende que el tensor de esfuerzo energía en la Ec. (18.31) debe ser un tensor de Lorentz de segundo rango. No obstante, el cambio de espacio implica que las componentes de tal tensor pueden cambiar su valor. Es claro que para la misma densidad Lagrangiana, el cambio de  $(t, x_1, x_2, x_3)$  a  $(x_1, x_2, x_3, ict)$  no altera a las componentes  $T_i^j$ . Para las otras componentes el cambio en su valor se puede obtener de las Ecs. (18.31, 18.33)

$$T_\mu^\nu \equiv \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_{\mu\nu}$$

$$\begin{aligned} T_0^0 &\equiv \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \dot{\eta}_\rho - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{dt} \right)} \left( \frac{d\eta_\rho}{dt} \right) - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{d(ict)} \right)} \left( \frac{d\eta_\rho}{d(ict)} \right) - \mathcal{L} \\ &= \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{dx^4} \right)} \left( \frac{d\eta_\rho}{dx^4} \right) - \mathcal{L} \equiv T_4^4 \end{aligned}$$

$$\begin{aligned} T_j^0 &= \frac{\partial \mathcal{L}}{\partial \dot{\eta}_\rho} \eta_{\rho,j} = \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{dt} \right)} \frac{d\eta_\rho}{dx^j} = \frac{\partial \mathcal{L}}{ic \partial \left( \frac{d\eta_\rho}{d(ict)} \right)} \frac{d\eta_\rho}{dx^j} = \frac{\partial \mathcal{L}}{ic \partial \left( \frac{d\eta_\rho}{dx^4} \right)} \frac{d\eta_\rho}{dx^j} = \frac{T_j^4}{ic} \\ &\Rightarrow icT_j^0 = T_j^4 \end{aligned}$$



$$T_0^j = \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{dx^j} \right)} \frac{d\eta_\rho}{dt} = ic \frac{\partial \mathcal{L}}{\partial \left( \frac{d\eta_\rho}{dx^j} \right)} \frac{d\eta_\rho}{d(ict)} = ic T_4^j$$

en resumen las equivalencias quedan en la forma<sup>13</sup>

$$T_i^j \rightarrow T_i^j ; T_4^4 \rightarrow T_0^0 ; T_j^4 \rightarrow ic T_j^0 ; T_4^j \rightarrow -\frac{i}{c} T_0^j \quad (18.84)$$

la densidad lagrangiana está definida excepto por una constante multiplicativa. Este factor multiplicativo suele tomarse de tal manera que  $T_{44}$  o su forma simetrizada coincidan con la densidad de energía del campo. En el cuadriespacio de Minkowski las cantidades conservadas de la Ec. (18.8) que aquí denotamos como  $R_\mu$ , quedan en la forma

$$R_\mu = \int T_\mu^4 dV \quad (18.85)$$

definamos ahora los factores  $P_\mu$  asociados a estas cantidades como

$$P_\mu \equiv \frac{i}{c} R_\mu = \frac{i}{c} \int T_\mu^4 dV \quad (18.86)$$

de (18.84) y de la interpretación de  $T_i^0$  como la componente  $i$ -ésima de la densidad de momento del campo, se concluye que  $P_i$  son las componentes del momento total del campo y  $P_4$  es  $iE/c$ , siendo  $E$  la energía total del campo. Esto sugiere que el arreglo  $P_\mu$  puede representar al cuadvectores energía momento definido en la sección 16.7. Para esto es necesario demostrar que  $R_\mu$  y por tanto  $P_\mu$  se transforman como vectores de universo (cuadvectores) bajo una transformación de Lorentz. Como estas expresiones son integrales en tres dimensiones, requeriremos examinar el significado de la integración tridimensional en una formulación covariante y estudiar la forma en que se deben tratar los límites de integración.

La covarianza en los límites de integración será de particular importancia en el principio de Hamilton

$$\delta I = \delta \int \mathcal{L}(\eta_\rho, \eta_{\rho,\nu}, x^\nu) dx^\mu = 0 ; dx^\mu \equiv dx^0 dx^1 dx^2 dx^3 \quad (18.87)$$

si la densidad Lagrangiana es un escalar de universo, la integral (18.87) será claramente covariante de Lorentz pero no así los límites de integración. La integración espacial se realiza sobre un volumen fijo del espacio tridimensional seguido de una integración en el tiempo entre los límites  $t_1$  y  $t_2$ . Pero la integración sobre un volumen extendido para tiempo fijo no es un concepto covariante puesto que la simultaneidad no es un concepto invariante de Lorentz. Por otro lado, si tenemos en cuenta que la conexión causal sí es un invariante de Lorentz (eventos del género espacial del género temporal y como de luz), una descripción covariante adecuada sería decir que la integración se hace sobre una hipersuperficie de tres dimensiones **del género espacial**, es decir una superficie en la cual todos los vectores de universo contenidos en ella sean del género espacial. Los vectores normales a dicha superficie son del género temporal. En particular, los vectores que unan dos puntos de universo en una superficie de tiempo constante serán del género espacial ya que se anula su componente  $x_4$ . Aunque la anulación de  $x_4$  no permanece cuando se hace una transformación de Lorentz (no se conserva la simultaneidad) el vector continuará siendo del género espacial. Por tanto, una superficie de tres dimensiones a tiempo constante es un tipo particular de superficie del género espacial. Análogamente, lo que en un sistema específico es una integración en el tiempo para un punto espacial fijo, se puede describir covariantemente por una integración extendida a una superficie del género temporal.

En un sistema unidimensional en el espacio físico, la integración en el principio de Hamilton descrita en (18.19) se extiende al rectángulo que se ilustra en la Fig. 18.4a. Una transformación de Lorentz es una rotación en el espacio de Minkowski y los lados paralelos a los ejes originales ya no lo serán en los ejes del sistema transformado, pero los lados horizontales seguirán siendo del género espacial y los lados verticales seguirán siendo del género temporal. Por tanto, en todos los sistemas de Lorentz podremos describir la integración como extendida a una región del cuadriespacio contenida entre dos hipersuperficies del género espacial limitadas por superficies del género temporal que cortan a las anteriores.

<sup>13</sup>En algunas ocasiones se usa en el régimen no relativista la coordenada  $x^0 \equiv ct$ . En tal caso las equivalencias son las indicadas en (18.84) pero omitiendo el factor  $c$ .

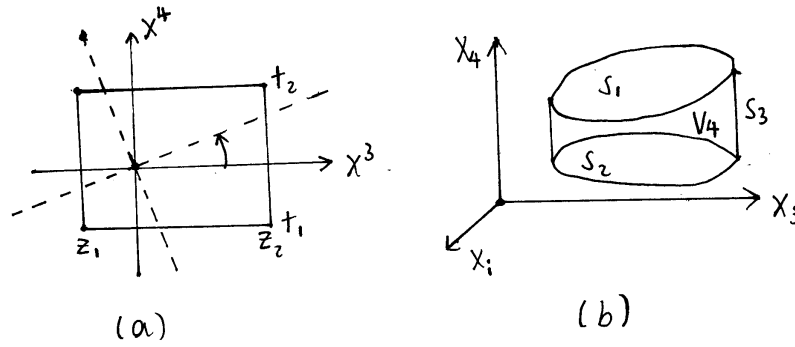


Figura 18.4: (a) Región de integración para el principio de Hamilton en un sistema unidimensional relativista. (b) Volumen de integración esquemático en el espacio tiempo de Minkowski.

Veamos ahora la descripción covariante de cantidades del tipo  $P_\mu$  descritas por las Ecs. (18.85, 18.86). Una forma apropiada para esta descripción covariante es de la forma

$$P_\mu = \frac{i}{c} \int_S T_\mu{}^\nu dS_\nu \quad (18.88)$$

en donde la integración se realiza sobre una hipersuperficie del género espacial para la cual los vectores que describen los elementos de superficie y que son normales a ésta son los  $dS_\nu$ .  $T_{\mu\nu}$  debe ser un tensor de universo de segundo orden para garantizar la covarianza de  $P_\mu$  es decir que  $P_\mu$  defina un vector de universo o cuadvectores. Veremos a continuación que las componentes de  $P_\mu$  en (18.88) se reducen a integrales de volumen en el espacio ordinario tridimensional del tipo descrito en (18.86) siempre que la divergencia del tensor  $T_{\mu\nu}$  sea nula

$$\partial_\nu T_\mu{}^\nu = 0$$

para ilustrarlo tomemos una región volumétrica  $V_4$  del cuadridespacio definida por tres superficies (ver Fig. 18.4b);  $S_1$  y  $S_2$  son del género espacial y  $S_3$  es del género temporal. Usando el teorema de la divergencia en cuatro dimensiones, podemos escribir

$$\int_{V_4} \left( \frac{dT_\mu{}^\nu}{dx^\nu} \right) dx^\nu = \int_{S_1+S_2+S_3} T_\mu{}^\nu dS_\nu \quad (18.89)$$

la integral sobre  $S_3$  es con respecto al tiempo a  $\mathbf{r}$  constante. Tomando el volumen suficientemente grande, la integral extendida a esta superficie mantendrá a  $\mathbf{r}$  fuera del sistema, en donde se anulan todas las cantidades de campo. Por tanto

$$\int_{V_4} \left( \frac{dT_\mu{}^\nu}{dx^\nu} \right) dx^\nu = \int_{S_1+S_2} T_\mu{}^\nu dS_\nu \quad (18.90)$$

Nuestra hipótesis de nulidad de la divergencia de  $T_\mu{}^\nu$  hace que también se anule el miembro izquierdo de (18.90). Por tanto, si tomamos en el mismo sentido las normales a la superficie del género espacial resulta

$$\int_{S_1} T_\mu{}^\nu dS_\nu = \int_{S_2} T_\mu{}^\nu dS_\nu \quad (18.91)$$

en la Fig. 18.4b, hemos tomado a  $S_1$  como una superficie del género espacial arbitraria en tanto que  $S_2$  es una superficie con tiempo constante es decir con  $x_4$  constante. En consecuencia, el vector normal a un elemento de superficie de  $S_2$  es un vector del género temporal que solo tiene componente en  $x_4$  i.e.  $dS_\nu = \delta_{\nu 4} dS_2$  y recordando que  $S_2$  es una hipersuperficie tridimensional, la integración sobre esta superficie es una integral de volumen a tiempo constante. De lo anterior, se vé que la Ec. (18.91) queda

$$\int_{S_1} T_\mu{}^\nu dS_\nu = \int T_\mu{}^4 dV$$

es claro que el miembro izquierdo se transforma como cuadrivector si  $T_{\mu}^{\nu}$  es un tensor de segundo rango, por tanto el segundo miembro también transforma como vector de universo. Teniendo en cuenta (18.85) vemos por tanto que  $R_{\mu}$  transforma como cuadrivector bajo la condición de que el tensor  $T_{\mu}^{\nu}$  tenga divergencia nula. Si adicionalmente las dos superficies  $S_1$  y  $S_2$  son superficies a tiempos constantes  $t_1$  y  $t_2$  respectivamente la Ec. (18.91) conduce a

$$R_{\mu}(t_1) = R_{\mu}(t_2)$$

que corresponde a la forma covariante de demostrar que  $R_{\mu}$  se conserva en el tiempo.

En consecuencia, será posible con un poco de cuidado utilizar las cantidades conservadas ya obtenidas en el régimen no relativista. Sin embargo, la correspondencia detallada no siempre es posible, pero en la mayoría de los casos es suficiente que la integración de volumen se refiera a un sistema de Lorentz específico en el cual la hipersuperficie del género espacial sea una región en el espacio de tres dimensiones a tiempo constante, ya que siempre que tenemos eventos del género espacial es posible encontrar un sistema de referencia en donde los eventos son simultáneos.

Para la densidad de momento angular, notemos que el análogo covariante de  $\mathcal{M}_i^j$  Ec. (18.43) es un cuadritensor de tercer rango

$$\mathcal{M}_{\mu\nu}^{\lambda} = \frac{i}{c} (x_{\mu}T_{\nu}^{\lambda} - x_{\nu}T_{\mu}^{\lambda}) \quad (18.92)$$

el cual es antisimétrico en  $\mu, \nu$ . La integral conservada es

$$M_{\mu\nu} = \int \mathcal{M}_{\mu\nu}^{\lambda} dS_{\lambda} \quad (18.93)$$

y la integración se extiende a lo largo de una hipersuperficie del género espacial. Si en particular se toma un sistema de Lorentz en el cual la hipersuperficie sea a  $t$  constante se obtiene

$$M_{\mu\nu} \rightarrow \int \mathcal{M}_{\mu\nu}^{\lambda} dV$$

que corresponde a la definición en el régimen no relativista. Por tanto, el razonamiento hecho en la sección 18.5.4 sobre la conservación de  $M_i^j$  para tensores esfuerzo-energía simétricos (o simetrizados) puede extrapolarse considerando el sistema de Lorentz particular con hipersuperficie a tiempo constante.

Ya mencionamos en la sección 18.6 que la formulación Hamiltoniana distingue las coordenadas espaciales y temporales de forma explícita. Esto no significa que no se pueda construir una teoría covariante de Lorentz, simplemente implica que la formulación no será manifiestamente covariante. Podemos imaginar la formulación Hamiltoniana en el tiempo como la que vé cada observador particular. Si los campos y sus derivadas transforman adecuadamente, es posible que la formulación sea compatible con la relatividad especial en cada sistema de Lorentz particular.

Otro punto importante se refiere a la descripción de sistemas físicos que involucran un conjunto de campos  $\eta_{\rho}$  que pueden interactuar entre sí. En este caso la densidad Lagrangiana usualmente consistirá en sumas de densidades Lagrangianas que representan a los campos libres más términos que describen la interacción entre los campos. Ya habíamos mencionado que parte de la dificultad para realizar una formulación relativista de la mecánica del punto tenía que ver con el hecho de que la acción a distancia no es compatible con la relatividad especial. La ventaja de las formulaciones de campo es que las interacciones entre campos pueden ser de naturaleza local, es decir pueden tener lugar en un punto lo cual facilita la construcción de una formulación covariante. Más aún, se puede trabajar interacciones de tipo local (en un punto del espacio tiempo) para campos interactuando con partículas. Por tanto, también es factible una formulación relativista de partículas inmersas en uno o más campos y que interactúan con éstos. Un ejemplo notable es la interacción de cargas eléctricas puntuales con campos electromagnéticos.

## 18.9. Algunas teorías de campos relativistas

Estudiaremos algunos ejemplos de teorías relativistas que forman parte de teorías reales de la naturaleza.

### 18.9.1. Campo escalar complejo

Un campo complejo está constituido por dos grados de libertad independientes que se pueden tomar como la parte real y la parte imaginaria del campo o cualquier transformación de estos grados de libertad mediante una matriz no singular, por ejemplo

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \text{Re}\phi \\ \text{Im}\phi \end{pmatrix} = \begin{pmatrix} \text{Re}\phi + i\text{Im}\phi \\ \text{Re}\phi - i\text{Im}\phi \end{pmatrix} = \begin{pmatrix} \phi \\ \phi^* \end{pmatrix}$$

por tanto  $\phi$  y  $\phi^*$  se pueden tomar como variables de campo independientes. Tomando entonces a  $\phi$  y  $\phi^*$  como las variables independientes, la densidad Lagrangiana y todas las funciones asociadas las escribiremos en términos de estas dos variables de campo. Cada uno de estos campos será un escalar de universo. Mas adelante veremos que los campos complejos conducen de forma natural a una densidad asociada de carga y de corriente y esta es la principal razón para considerar este tipo de campos.

Tomaremos en este caso la siguiente densidad Lagrangiana

$$\mathcal{L} = -c^2 \phi_{,\lambda} \phi_{,\lambda}^* - \mu_0^2 c^2 \phi \phi^* \quad (18.94)$$

donde  $\mu_0$  es una constante. Obsérvese que si los campos son escalares de universo, esta densidad Lagrangiana también lo será como se requirió en la sección 18.8. Expresando esta densidad Lagrangiana en términos de espacio y tiempo separados nos queda

$$\begin{aligned} \mathcal{L} &= -c^2 \frac{\partial \phi}{\partial x^\lambda} \frac{\partial \phi^*}{\partial x^\lambda} - \mu_0^2 c^2 \phi \phi^* = -c^2 \frac{\partial \phi}{\partial x^4} \frac{\partial \phi^*}{\partial x^4} - c^2 \frac{\partial \phi}{\partial x^i} \frac{\partial \phi^*}{\partial x^i} - \mu_0^2 c^2 \phi \phi^* \\ \mathcal{L} &= -c^2 \frac{\partial \phi}{\partial (ict)} \frac{\partial \phi^*}{\partial (ict)} - c^2 \nabla \phi \cdot \nabla \phi^* - \mu_0^2 c^2 \phi \phi^* \end{aligned}$$

quedando finalmente

$$\mathcal{L} = \dot{\phi} \dot{\phi}^* - c^2 \nabla \phi \cdot \nabla \phi^* - \mu_0^2 c^2 \phi \phi^* \quad (18.95)$$

encontremos la ecuación de Lagrange asociada al campo  $\phi^*$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \phi_{,\nu}^*} &= -c^2 \phi_{,\nu} \quad ; \quad \frac{d}{dx^\nu} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\nu}^*} \right) = -c^2 \phi_{,\nu\nu} \quad ; \quad \frac{\partial \mathcal{L}}{\partial \phi^*} = -\mu_0^2 c^2 \phi \\ \phi_{,\nu\nu} - \mu_0^2 \phi &= 0 \end{aligned} \quad (18.96)$$

que en notación más explícita queda

$$\begin{aligned} \sum_\nu \frac{d^2 \phi}{dx_\nu^2} - \mu_0^2 \phi = 0 &\Rightarrow \frac{\partial}{\partial (ict)} \left( \frac{\partial \phi}{\partial (ict)} \right) + \sum_k \frac{\partial^2 \phi}{\partial x_k^2} - \mu_0^2 \phi = 0 \\ \nabla^2 \phi - \frac{1}{c^2} \frac{d^2 \phi}{dt^2} - \mu_0^2 \phi &= 0 \end{aligned} \quad (18.97)$$

y recordando la definición del operador D'Alembertiano Ec. (16.64) pag. 470, la ecuación de campo se puede reescribir en forma manifiestamente covariante en la forma

$$(\square^2 - \mu_0^2) \phi = 0 \quad (18.98)$$

la ecuación de Lagrange asociada al campo  $\phi$  se puede encontrar inmediatamente en virtud de la simetría de  $\mathcal{L}$  con  $\phi$  y  $\phi^*$

$$(\square^2 - \mu_0^2) \phi^* = 0 \quad (18.99)$$

A las Ecs. (18.98, 18.99) se les conoce como **ecuaciones de Klein Gordon**. Nótese que éstas ecuaciones se reducen a ecuaciones de onda cuando  $\mu_0 = 0$ . En la literatura, es usual encontrar dichas ecuaciones con un signo positivo en el término proporcional a  $\mu_0^2$ , esto se debe a la utilización de la métrica de Minkowski definida

en (16.42) que produce un cambio de signo en el D'Alembertiano. La ecuación de Klein Gordon es en realidad el análogo relativista de la ecuación de Schrödinger para una partícula cargada de espín cero y masa en reposo  $\mu_0$ .

Veamos la forma del tensor esfuerzo energía definido en la Ec. (18.31), explícitamente se escribe para este caso

$$\begin{aligned} T_{\mu}^{\nu} &\equiv \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\mu} - \mathcal{L} \delta_{\mu}^{\nu} = \frac{\partial \mathcal{L}}{\partial \phi_{,\nu}} \phi_{,\mu} + \frac{\partial \mathcal{L}}{\partial \phi_{,\nu}^*} \phi_{,\mu}^* - \mathcal{L} \delta_{\mu}^{\nu} \\ &= -c^2 \phi_{,\nu}^* \phi_{,\mu} - c^2 \phi_{,\nu} \phi_{,\mu}^* - (-c^2 \phi_{,\lambda} \phi_{,\lambda}^* - \mu_0^2 c^2 \phi \phi^*) \delta_{\mu}^{\nu} \\ T_{\mu}^{\nu} &= -c^2 \phi_{,\mu} \phi_{,\nu}^* - c^2 \phi_{,\mu}^* \phi_{,\nu} + c^2 (\phi_{,\lambda} \phi_{,\lambda}^* + \mu_0^2 \phi \phi^*) \delta_{\mu}^{\nu} \end{aligned} \quad (18.100)$$

el cual es claramente simétrico. Nótese que  $\mathcal{L}$  no depende explícitamente de los  $x^{\mu}$  lo cual nos dice que esta densidad Lagrangiana describe un campo libre sin interacción con el resto del universo. Adicionalmente, la no dependencia de  $\mathcal{L}$  con  $x^{\mu}$  nos dice que se cumplen los teoremas de conservación expresados por la Ec. (18.30).

### Formulación Hamiltoniana para el campo escalar complejo

Para introducir la formulación Hamiltoniana debemos distinguir entre coordenadas espaciales y temporales en un sistema de Lorentz particular. Los momentos canónicamente conjugados a las variables de campo están dados por (18.49) que al aplicarlos a la densidad Lagrangiana 18.95 nos dan

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^* \quad ; \quad \pi^* = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^*} = \dot{\phi} \quad (18.101)$$

y la densidad Hamiltoniana definida en (18.50) y que tiene la misma magnitud de  $T_{44}$  está dada por

$$\begin{aligned} \mathcal{H} &= \pi \dot{\phi} + \pi^* \dot{\phi}^* - \mathcal{L} = \pi \pi^* + \pi^* \pi - (-c^2 \phi_{,\lambda} \phi_{,\lambda}^* - \mu_0^2 c^2 \phi \phi^*) \\ \mathcal{H} &= 2\pi \pi^* + c^2 \frac{\partial \phi}{\partial (ict)} \frac{\partial \phi^*}{\partial (ict)} + c^2 \phi_{,i} \phi_{,i}^* + \mu_0^2 c^2 \phi \phi^* \\ \mathcal{H} &= 2\pi \pi^* - \dot{\phi} \dot{\phi}^* + c^2 \nabla \phi \cdot \nabla \phi^* + \mu_0^2 c^2 \phi \phi^* = 2\pi \pi^* - \pi^* \pi + c^2 \nabla \phi \cdot \nabla \phi^* + \mu_0^2 c^2 \phi \phi^* \\ \mathcal{H} &= \pi \pi^* + c^2 \nabla \phi \cdot \nabla \phi^* + \mu_0^2 c^2 \phi \phi^* \end{aligned} \quad (18.102)$$

las Ecuaciones de Hamilton en el continuo Ecs. (18.57) nos dan

$$\begin{aligned} \dot{\phi} &= \frac{\delta \mathcal{H}}{\delta \pi} = \left( \frac{\partial}{\partial \pi} - \frac{d}{dx^i} \frac{\partial}{\partial \pi_{\rho,i}} \right) \mathcal{H} = \frac{\partial \mathcal{H}}{\partial \pi} = \pi^* \quad ; \quad \dot{\phi}^* = \frac{\delta \mathcal{H}}{\delta \pi^*} = \frac{\partial \mathcal{H}}{\partial \pi^*} = \pi \\ \dot{\pi} &= -\frac{\delta \mathcal{H}}{\delta \phi} = -\left( \frac{\partial}{\partial \phi} - \frac{d}{dx^i} \frac{\partial}{\partial \phi_{,i}} \right) (\pi \pi^* + c^2 \phi_{,k} \phi_{,k}^* + \mu_0^2 c^2 \phi \phi^*) \\ &= -\mu_0^2 c^2 \phi^* + \frac{d}{dx^i} (c^2 \phi_{,i}^*) = c^2 \phi_{,ii}^* - \mu_0^2 c^2 \phi^* = c^2 \nabla^2 \phi^* - \mu_0^2 c^2 \phi^* \\ \dot{\pi}^* &= -\frac{\delta \mathcal{H}}{\delta \phi^*} = c^2 \nabla^2 \phi - \mu_0^2 c^2 \phi \end{aligned}$$

las ecuaciones de Hamilton nos dan entonces

$$\dot{\phi} = \pi^* \quad ; \quad \dot{\phi}^* = \pi \quad ; \quad \dot{\pi} = c^2 \nabla^2 \phi^* - \mu_0^2 c^2 \phi^* \quad ; \quad \dot{\pi}^* = c^2 \nabla^2 \phi - \mu_0^2 c^2 \phi \quad (18.103)$$

las dos primeras ecuaciones de Hamilton reproducen las Ecs. (18.101). Si derivamos las dos primeras ecuaciones (18.103)

$$\ddot{\phi} = \dot{\pi}^* \quad ; \quad \ddot{\phi}^* = \dot{\pi}$$

y reemplazamos en las dos últimas de las Ecs. (18.103)

$$\begin{aligned} \ddot{\phi}^* &= c^2 \nabla^2 \phi^* - \mu_0^2 c^2 \phi^* \quad ; \quad \ddot{\phi} = c^2 \nabla^2 \phi - \mu_0^2 c^2 \phi \\ -\frac{1}{c^2} \ddot{\phi}^* + \nabla^2 \phi^* - \mu_0^2 \phi^* &= 0 \quad ; \quad -\frac{1}{c^2} \ddot{\phi} + \nabla^2 \phi - \mu_0^2 \phi = 0 \end{aligned}$$

y usando de nuevo el D'Alambertiano, vemos que éstas corresponden a las ecuaciones de Klein Gordon.

Veamos ahora la representación de Fourier de momentos canónicos totales. Para ello se introducen las expansiones de Fourier (18.71, 18.73)

$$\phi(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} q_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} \quad ; \quad \pi(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} p_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{r}}$$

en la densidad Hamiltoniana. Debe tenerse en cuenta que como el campo no es real, no se cumple la condición  $q_{\mathbf{k}}^* = q_{-\mathbf{k}}$ . En este caso  $q_{\mathbf{k}}$  y  $q_{\mathbf{k}}^*$  representan dos conjuntos independientes de coordenadas discretas (de la misma forma en que  $\phi$  y  $\phi^*$  se consideran independientes). El Hamiltoniano total es la integral de volumen de la densidad Hamiltoniana dada por (18.102)

$$H = \int (\pi\pi^* + c^2 \nabla\phi \cdot \nabla\phi^* + \mu_0^2 c^2 \phi\phi^*) dV$$

e introduciendo las expansiones de Fourier el Hamiltoniano queda

$$H = \frac{1}{V} \int \left\{ \sum_{\mathbf{k}, \mathbf{j}} p_{\mathbf{k}}(t) p_{\mathbf{j}}^*(t) e^{-i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{j}\cdot\mathbf{r}} + c^2 \left( \sum_{\mathbf{k}} i\mathbf{k} q_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} \right) \cdot \left( \sum_{\mathbf{j}} (-i\mathbf{j}) q_{\mathbf{j}}^*(t) e^{-i\mathbf{j}\cdot\mathbf{r}} \right) + \mu_0^2 c^2 \sum_{\mathbf{k}, \mathbf{j}} q_{\mathbf{k}}(t) q_{\mathbf{j}}^*(t) e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\mathbf{j}\cdot\mathbf{r}} \right\} dV$$

$$H = \sum_{\mathbf{k}, \mathbf{j}} p_{\mathbf{k}}(t) p_{\mathbf{j}}^*(t) \left[ \frac{1}{V} \int e^{i(\mathbf{j}-\mathbf{k})\cdot\mathbf{r}} dV \right] + c^2 \sum_{\mathbf{k}, \mathbf{j}} (\mathbf{k} \cdot \mathbf{j}) q_{\mathbf{k}}(t) q_{\mathbf{j}}^*(t) \left[ \frac{1}{V} \int e^{i(\mathbf{k}-\mathbf{j})\cdot\mathbf{r}} dV \right] + \mu_0^2 c^2 \sum_{\mathbf{k}, \mathbf{j}} q_{\mathbf{k}}(t) q_{\mathbf{j}}^*(t) \left[ \frac{1}{V} \int e^{i(\mathbf{k}-\mathbf{j})\cdot\mathbf{r}} dV \right]$$

y utilizando la ortonormalidad de las ondas planas Ec. (18.72) se obtiene

$$H = \sum_{\mathbf{k}, \mathbf{j}} p_{\mathbf{k}}(t) p_{\mathbf{j}}^*(t) \delta_{\mathbf{j}\mathbf{k}} + c^2 \sum_{\mathbf{k}, \mathbf{j}} (\mathbf{k} \cdot \mathbf{j}) q_{\mathbf{k}}(t) q_{\mathbf{j}}^*(t) \delta_{\mathbf{j}\mathbf{k}} + \mu_0^2 c^2 \sum_{\mathbf{k}, \mathbf{j}} q_{\mathbf{k}}(t) q_{\mathbf{j}}^*(t) \delta_{\mathbf{j}\mathbf{k}}$$

$$H = p_{\mathbf{k}} p_{\mathbf{k}}^* + c^2 \mathbf{k}^2 q_{\mathbf{k}} q_{\mathbf{k}}^* + \mu_0^2 c^2 q_{\mathbf{k}} q_{\mathbf{k}}^*$$

que finalmente se puede escribir en la forma

$$H = p_{\mathbf{k}} p_{\mathbf{k}}^* + \omega_{\mathbf{k}}^2 q_{\mathbf{k}} q_{\mathbf{k}}^* \quad ; \quad \omega_{\mathbf{k}}^2 \equiv c^2 \mathbf{k}^2 + \mu_0^2 c^2 \quad (18.104)$$

donde se usa suma sobre índices repetidos. Nótese que este Hamiltoniano tiene la forma de una suma de osciladores armónicos desacoplados de masa unidad y con frecuencia angular  $\omega_{\mathbf{k}}$ . Esto se puede ver en forma explícita utilizando las ecuaciones de Hamilton globales dadas por (18.78, 18.82)

$$\dot{q}_{\mathbf{j}} = \frac{\partial H}{\partial p_{\mathbf{j}}} = p_{\mathbf{j}}^* \quad ; \quad \dot{q}_{\mathbf{j}}^* = \frac{\partial H}{\partial p_{\mathbf{j}}^*} = p_{\mathbf{j}}$$

$$\dot{p}_{\mathbf{j}} = -\frac{\partial H}{\partial q_{\mathbf{j}}} = -\omega_{\mathbf{j}}^2 q_{\mathbf{j}}^* \quad ; \quad \dot{p}_{\mathbf{j}}^* = -\frac{\partial H}{\partial q_{\mathbf{j}}^*} = -\omega_{\mathbf{j}}^2 q_{\mathbf{j}} \quad (\text{no suma})$$

derivando con respecto al tiempo las dos primeras ecuaciones resulta

$$\ddot{q}_{\mathbf{j}} = \dot{p}_{\mathbf{j}}^* \quad ; \quad \ddot{q}_{\mathbf{j}}^* = \dot{p}_{\mathbf{j}}$$

y reemplazando en las dos últimas

$$\ddot{q}_j^* = -\omega_j^2 q_j^* \quad ; \quad \ddot{q}_j = -\omega_j^2 q_j \quad (\text{no suma})$$

que son ecuaciones de oscilador armónico para cada coordenada discreta  $q_j$ . Por lo tanto, en la representación por expansión de Fourier en ondas planas se sustituyen los campos  $\phi$  y  $\phi^*$  por sistemas discretos de osciladores armónicos. Un proceso similar se realiza cuando el campo sonoro en un sólido finito se construye como una superposición discreta de fonones. El espectro discreto de “vibraciones” de nuestro campo cargado escalar viene dado por la Ec. (18.104) que actúa como una **relación de dispersión**. Es muy sencillo hacer una teoría cuántica (conocida como segunda cuantización) de estos campos cuando se trabaja con el formalismo de expansión de Fourier pues en este caso cada oscilador se cuantiza como se cuantizan los osciladores armónicos en la mecánica cuántica del discreto.

### 18.9.2. Ecuación de seno Gordon

Tomemos el campo anterior como real de modo que  $\phi = \phi^*$  y limitemonos al caso unidimensional. La densidad Lagrangiana (18.95) se reduce a

$$\mathcal{L} = \frac{c^2}{2} \left[ \frac{\dot{\phi}^2}{c^2} - \left( \frac{\partial \phi}{\partial x} \right)^2 - \mu_0^2 \phi^2 \right] \quad (18.105)$$

donde se ha introducido un factor 1/2 por conveniencia, pues claramente este factor no altera las ecuaciones de movimiento. La ecuación de campo asociada (ecuación de Klein Gordon unidimensional para partícula neutra) es

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \mu_0^2 \phi \quad (18.106)$$

esta es una ecuación lineal en el campo  $\phi(x, t)$ . Podemos considerar la densidad Lagrangiana en (18.105) como una aproximación de campo débil de la densidad Lagrangiana dada por

$$\mathcal{L} = \frac{c^2}{2} \left[ \frac{\dot{\phi}^2}{c^2} - \left( \frac{\partial \phi}{\partial x} \right)^2 \right] - \mu_0^2 \phi^2 (1 - \cos \phi)$$

cuya ecuación de campo es

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \mu_0^2 \sin \phi \quad (18.107)$$

a la expresión (18.107) se le conoce como ecuación de seno Gordon. Vale resaltar que la ecuación de seno Gordon (18.107) tiene similitud al péndulo de amplias oscilaciones (por esta razón se le denomina también ecuación del péndulo) en tanto que la Ec. (18.106) se asemeja al oscilador armónico y por tanto, al péndulo de pequeñas oscilaciones.

Para este escenario unidimensional el tensor esfuerzo-energía solo posee cuatro componentes. Como la densidad Lagrangiana no depende explícitamente de los parámetros  $x, t$  los elementos del tensor satisfacen dos ecuaciones de continuidad. La densidad de energía  $T_4^4$  es de especial interés

$$\begin{aligned} T_4^4 &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} - \mathcal{L} = \dot{\phi} \dot{\phi} - \left\{ \frac{c^2}{2} \left[ \frac{\dot{\phi}^2}{c^2} - \left( \frac{\partial \phi}{\partial x} \right)^2 \right] - \mu_0^2 \phi^2 (1 - \cos \phi) \right\} \\ T_4^4 &= \left\{ \frac{1}{2} \left[ \dot{\phi}^2 + c^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right] + \mu_0^2 \phi^2 (1 - \cos \phi) \right\} \end{aligned}$$

que naturalmente coincide con la magnitud de la densidad Hamiltoniana

$$\mathcal{H} = \frac{1}{2} \left[ \pi^2 + c^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right] + \mu_0^2 c^2 (1 - \cos \phi) \quad ; \quad \pi(x, t) = \dot{\phi}$$

por otro lado, la representación por expansión de Fourier en ondas planas del campo de Klein Gordon real unidimensional conduce a la relación

$$q_{\mathbf{k}}(t) = A_0(k) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t)}$$

donde  $\omega_{\mathbf{k}}$  y  $\mathbf{k}$  están enlazadas por la relación de dispersión dada por la segunda de las Ecs. (18.104). Para un campo que sea solución de la ecuación de seno Gordon es mucho más difícil construir una representación por expansión de Fourier debido al término  $\cos \phi$  en la densidad Hamiltoniana. Sin embargo, la relación de seno Gordon se puede resolver con una estructura similar a una onda progresiva. Una solución  $\phi$  de la Ec. (18.107) que tenga la forma de una perturbación que se propaga con velocidad  $v$ , pero que por lo demás conserva su forma, debe ser función exclusiva de  $\tau = t - x/v$  (nótese la similitud con las soluciones generales de la ecuación de onda). Con base en lo anterior la Ec. (18.107) queda en la forma

$$\frac{d^2\phi}{d\tau^2} - A \sin \phi = 0 \quad ; \quad A \equiv \frac{\mu_0^2 c^2 v^2}{c^2 - v^2}$$

y en función de  $\tau$  lo que resulta es la ecuación de un péndulo simple con amplitud finita. Si la amplitud es pequeña sabemos que  $\phi$  describe un movimiento armónico simple en  $\tau$  con  $\omega$  dado por la segunda de las Ecs. (18.104) para un número separado de onda  $k = \omega/v$  independiente de la amplitud. Si la amplitud es finita sabemos que el péndulo deja de ser isócrono, es decir la frecuencia comienza a depender de la amplitud, aún cuando nuestra  $\phi$  continúe siendo periódica. La relación de dispersión también dependerá entonces de la amplitud. Esto es característico de las ecuaciones alineaes, ya que en particular la ecuación de seno Gordon es no lineal. Por otro lado, la ecuación de Klein Gordon sí es lineal pero la relación de dispersión en (18.104) es alineal puesto que  $\omega_{\mathbf{k}}$  no es función lineal de  $\mathbf{k}$  y solo se vuelve lineal cuando  $\mu_0$  se anula, es decir cuando la ecuación de Klein Gordon se convierte en la ecuación de onda lineal.

De lo anterior se vé que la ecuación de seno Gordon es alineal y con relación de dispersión alineal no isócrona. Un estudio más detallado muestra que esta ecuación admite soluciones de características muy peculiares incluso entre las ecuaciones alineaes. Estas soluciones peculiares son del tipo onda progresiva que pueden interactuar entre sí atravesándose unas a otras y que emergen sin variar su forma salvo por un posible cambio de fase. Estas ondas solitarias que conservan su forma incluso a través de interacciones se denominan solitones, los cuales tienen amplias aplicaciones especialmente en Física de Partículas Elementales y en Física del Estado Sólido.



# Bibliografía

- [1] **Classical Mechanics (3rd edition)**. Herbert Goldstein, Charles P. Poole, John L. Safko. Addison-Wesley (2001).
- [2] **Physics Vol. I-Mechanics**. Marcelo Alonso, Edward Finn. Addison-Wesley Publishing Co. Inc. Massachusetts (1967).
- [3] **An introduction to Mechanics**. Daniel Kleppner, Robert Kolenkow. McGraw-Hill International Editions, (1973).
- [4] **On the transformation of torques between the laboratory and center of mass reference frames**. Rodolfo A. Diaz, William J. Herrera. Revista Mexicana de Fisica, vol. E51 #2 p.112 - 115 (2005).
- [5] **Work and energy in inertial and noninertial reference frames**. Rodolfo A. Díaz, William J. Herrera, Diego A. Manjarrés. American Journal of Physics 77(3), 270-273 (2009).
- [6] **Normal Forces Can do Work**. Diego A. Manjarrés, Rodolfo A. Díaz, William J. Herrera, <http://demonstrations.wolfram.com/NormalForcesCanDoWork>.
- [7] **The physical meaning of Lagrange multipliers**. Hasan Karabulut. European Journal of Physics 27, 709 (2006).
- [8] **Moments of inertia for solids of revolution and variational methods**. Rodolfo A. Diaz, William J. Herrera, R. Martinez. European Journal of Physics 27, 183 (2006).
- [9] **Classical Dynamics: A Contemporary Approach**. Jorge V. José, Eugene J. Saletan. Cambridge University Press (1998).
- [10] **Mathematical Methods of Classical Mechanics**, 2nd Ed. Arnol'd V. I. Springer-Verlag, Berlin (1988).
- [11] **Classical Dynamics: A Modern Perspective**. E. C. G. Sudarshan, N. Mukunda. John Wiley & Sons (1974).
- [12] **A theorem on moments of inertia**. J. F. Streib. American Journal of Physics 57, 181 (1989).
- [13] **Using symmetries and generating functions to calculate and minimize moments of inertia**. Rodolfo A. Diaz, William J. Herrera, R. Martinez. arXiv:physics/0404005v2.
- [14] **Another reversing gyroscope**. K. Hara. Journal of Technical Physics vol 49, #1, 27-37 (2008). **On the possible reversal of a satellite spin axis**. K. Hara. J. Tech. Phys. vol 50, #2, 75-85 (2009). **On the possible reversal of an earth-scale top**. K. Hara. J. Tech. Phys. vol 50, #4, 375-385 (2009).

- [15] **New conditional integrable cases of motion of a rigid body with Kovalevskaya's configuration.**  
H.M. Yehia, A.A Elmandouh. *Journal of Physics A: Math. Theor.* 44, 012001 (2011).