



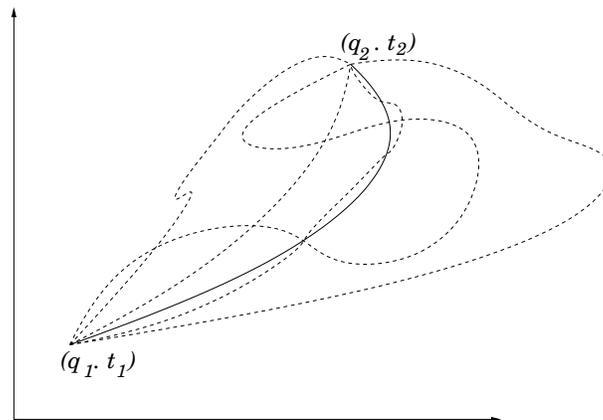
Mecánica Analítica

Universidad de Granada

5º curso de matemáticas

Dr. Bert Janssen

*Departamento de Física Teórica y del Cosmos,
Universidad de Granada, Campus de Fuente Nueva
18071 Granada, España
bjanssen@ugr.es*



junio 2006

En la medida en que las proposiciones matemáticas se refieren a la realidad no son ciertas, y en la medida en que son ciertas no se refieren a la realidad.

A. Einstein

Q: What is the difference between theoretical physics and mathematical physics?

A: Theoretical physics is done by physicists who lack the necessary skills to do real experiments; mathematical physics is done by mathematicians who lack the necessary skills to do real mathematics.

N.D. Mermin (theoretical physicist)

Índice general

Prefacio	5
1. Representación matemática del espacio físico	7
1.1. Espacios vectoriales y duales	7
1.2. La métrica y las transformaciones ortogonales	9
1.3. Algebra de tensores	12
1.4. Operaciones con tensores	13
1.5. Coordenadas curvilíneas	15
2. Repaso de la mecánica newtoniana	19
2.1. Mecánica de una sola partícula	19
2.2. Simetrías y la forma de las leyes de la física	24
2.3. Mecánica de N partículas	27
2.4. Cuerpos continuos	30
2.5. Problemas	31
3. Formalismo lagrangiano	33
3.1. Ligaduras y coordenadas generalizadas	33
3.2. El principio de trabajo virtual y las ecuaciones de Lagrange	35
3.3. El principio de mínima acción	37
3.4. Ejemplos	39
3.5. Interpretación y propiedades del Lagrangiano	41
3.6. Problemas	44
4. Formalismo hamiltoniano y transformaciones canónicas	47
4.1. El Hamiltoniano como transformada de Legendre	47
4.2. Interpretación y cantidades conservadas	50
4.3. Transformaciones canónicas	51
4.4. Corchetes de Poisson	54

4.5. Unas palabras cuánticas	57
4.6. Problemas	59
5. Potenciales centrales	62
5.1. Repaso de coordenadas esféricas	62
5.2. Reducción del problema de dos cuerpos	63
5.3. El lagrangiano y las ecuaciones de movimiento	65
5.4. Estudio cualitativo de las trayectorias	69
5.5. El problema de Kepler	71
5.6. Problemas	74
6. Teoría de Maxwell	77
6.1. El límite a sistemas continuos	77
6.2. Las leyes de Maxwell	80
6.3. Cantidades conservadas	84
6.4. Potenciales electromagnéticos e invariancia gauge	86
6.5. Soluciones de las ecuaciones de Maxwell	88
6.5.1. Ondas electromagnéticas	88
6.5.2. Potenciales retardados	90
6.5.3. El efecto Aharonov-Bohm	91
6.5.4. El monopolo de Dirac	94
6.6. Problemas	98
7. Pequeñas oscilaciones	100
8. Introducción a la mecánica cuántica	101
9. Electromagnetismo avanzado y teorías gauge	102
9.1. Teoría de Maxwell en forma covariante	102
9.2. Soluciones topológicas	102
9.3. Teorías gauge	102

Prefacio

La mecánica analítica es el estudio de la mecánica clásica (newtoniana) en su formalismo matemático. Fue desarrollado durante el siglo XVIII y XIX por físicos y matemáticos, como Euler (1707-1783), D'Alembert (1717-1783), Lagrange (1736-1813) y Hamilton (1805-1865). En cuanto a física fundamental, esta época fue una época calma, hasta relativamente aburrida, puesto que ya un siglo anterior Galilei (1564-1642) y Newton (1642-1727) habían descubierto y desarrollado la mecánica, mientras todavía no había indicios de la física nueva (la mecánica cuántica y la teoría de la relatividad) de los principios del siglo XX.

Sin embargo, la mecánica analítica sí marca un capítulo importante en la historia de la física, puesto que lo que hace es embeber la mecánica de Galilei y Newton en un formalismo matemático, dando así una base más rigurosa a la mecánica newtoniana. Además, muchos de los formalismos desarrollados en la mecánica analítica resultaron ser imprescindibles para el avance de la física moderna: el formalismo lagrangiano resulta mucho más útil para teoría de campos, el formalismo hamiltoniano para la primera cuantización, los corchetes de Poisson son directamente generalizables a los conmutadores de operadores en la mecánica cuántica y fue el principio de mínima acción lo que inspiró a Feynman el formalismo de integrales de caminos. En otras palabras, el desarrollo de la física moderna, tal como la conocemos ahora, hubiera sido imposible sin los avances de la mecánica analítica.

A pesar de ser una asignatura de la carrera de matemáticas, este curso es claramente un curso sobre física. Más que intentar enseñar nueva matemática a los estudiantes, lo que este curso pretende es aplicar las técnicas aprendidas en cursos anteriores a algo fuera del campo de la matemática y intentar crear un poco de intuición más allá del formalismo matemático. En otras palabras, matemática como herramienta para hacer física. En este aspecto, la mecánica analítica es probablemente la parte de la física más adecuada para este objetivo, puesto que forma como un puente entre los dos campos.

El curso está compilado de varios libros de textos sobre mecánica clásica y mecánica analítica, cada uno con su propios temas y su propio estilo. Yo, por mi parte, he hecho una selección personal de temas, tomando en cuenta el historial matemático de los estudiantes. El estilo es inevitablemente el de un físico, puesto que lo soy. Espero que a lo largo del curso, logremos romper la barrera del idioma que separa la física y la matemática y llegar a un diálogo

He empezado a escribir estos apuntes por dos razones diferentes. La primera es personal, como manera (elaborada) de preparar las clases y elegir el temario. La segunda es pedagógico, como gesto al estudiante. Hay un principio didáctico que dice que las clases donde el estudiante copia lo que el profesor apunta en la pizarra son la forma menos eficaz de pasar los apuntes del profesor a los apuntes del alumno sin que pasa por la cabeza de ninguno de los dos. Mi intención ha sido evitar esto y mi esperanza (ingenua?) es que el estudiante tenga más oportunidad de interacción durante las clases y que beneficie de ello.

Como ya he mencionado, existen numerosos libros sobre los temas tratado en este curso. Recomendaría por lo tanto al estudiante interesado que, aparte de estos apuntes, mirara estos libros. Algunos podría coincidir más con el gusto o el estilo a lo que está acostumbrado. Una lista de los libros más conicidos es encuentra al final de estos apuntes.

Capítulo 1

Representación matemática del espacio físico

En este primer capítulo repasaremos las propiedades básicas de álgebra lineal y transformaciones de coordenadas globales y locales. El estudiante de matemáticas no encontrará materia que no haya estudiado en otras asignaturas, de modo que este capítulo sirve básicamente para introducir la notación utilizado en el resto del curso.

Utilizaremos el *convenio de sumación de Einstein*: en cuanto un índice aparezca repetido arriba y abajo, se supone una suma sobre todos los posibles valores de este índice. La descomposición de un vector $|a\rangle$ en una base $\{|e_i\rangle\}$ sería por lo tanto $|a\rangle = a^i|e_i\rangle$ y un producto escalar entre dos vectores $\langle a|b\rangle = a_i b^i$, donde la suma en ambos casos contiene N términos en un espacio vectorial N -dimensional. Dado que un índice repetido sólo es un índice de sumación, da igual el nombre que le demos, de modo que tenemos que $a_i b^i = a_k b^k$. Los índices que no están sumados se llaman *libres*.

1.1. Espacios vectoriales y duales

Consideramos el espacio vectorial N -dimensional \mathbb{R}^N , con un origen O y una base $\{|e_i\rangle\}$, formado por N vectores linealmente independientes. A cada punto x de \mathbb{R}^N se le asigna un vector de posición $|x\rangle$, que se descompone en la base $\{|e_i\rangle\}$ como

$$|x\rangle = x^1|e_1\rangle + x^2|e_2\rangle + \dots + x^N|e_N\rangle = x^i|e_i\rangle. \quad (1.1)$$

(Nótese que estamos utilizando el convenio de sumatoria.) La descomposición $|x\rangle$ en una base es única y los números x^i son las *componentes de $|x\rangle$ en la base $\{|e_i\rangle\}$* .

En vez de trabajar en la base $\{|e_i\rangle\}$, podríamos haber escrito $|x\rangle$ en una base distinta $\{|e_i'\rangle\}$, donde la descomposición sería

$$|x\rangle = x'^i|e_i'\rangle. \quad (1.2)$$

Para encontrar la relación entre las dos bases, descomponemos los vectores de base $|e_i'\rangle$ en la base $\{|e_i\rangle\}$,

$$|e_j'\rangle = (M^{-1})^i_j |e_i\rangle, \quad (1.3)$$

donde $(M^{-1})^i_j$ es la componente i del vector $|e_j'\rangle$ en la dirección $|e_i\rangle$.¹ Desde el punto de vista matemático M^{-1} es la matriz $N \times N$ que parametriza la transformación entre las dos bases.

¹Aquí estamos suponiendo por simplicidad que las dos bases tienen el mismo origen. Si no es así, la transformación es de la forma $|e_j'\rangle = (M^{-1})^i_j |e_i\rangle + |a\rangle$.

Dado que el vector $|x\rangle$ en (1.1) y (1.2) es el mismo y que la relación entre las dos bases viene dada por (1.3), tiene que haber una relación entre las componentes x^i y x'^i . No es difícil ver que las componentes transforman como

$$x'^i = M^i_j x^j, \quad (1.4)$$

donde la matriz M es la inversa de M^{-1} ,

$$M^i_j (M^{-1})^j_k = \delta^i_k = (M^{-1})^i_j M^j_k, \quad (1.5)$$

donde δ^i_j son las componentes de la matriz identidad $\mathbb{1}$

$$(\mathbb{1})^i_j = \delta^i_j = \begin{cases} 0 & \text{cuando } i \neq j, \\ 1 & \text{cuando } i = j. \end{cases} \quad (1.6)$$

Más que elementos de un espacio vectorial, para un físico un vector es un objeto con ciertas reglas de transformación. Por eso, en la física se suele tomar la regla de transformación (1.4) como la definición de un vector: Cualquier objeto con N componentes x^i que bajo un cambio de base (1.3) transforma como (1.4) se le llama un *vector de columna* o un *vector contravariante*.

Considera ahora el espacio de las aplicaciones lineales $\langle y|$ que llevan los vectores de \mathbb{R}^N a los números reales \mathbb{R} : la aplicación $\langle y|$ actúa sobre el vector contravariante $|x\rangle$ como $\langle y|x\rangle \in \mathbb{R}$. Dado que estas aplicaciones son lineales (por construcción) tenemos que

$$\langle y|(\alpha|x_1\rangle + \beta|x_2\rangle) = \alpha\langle y|x_1\rangle + \beta\langle y|x_2\rangle. \quad (1.7)$$

Además una combinación lineal de dos de estas aplicaciones, también es una aplicación

$$(\alpha\langle y_1| + \beta\langle y_2|)|x\rangle = \alpha\langle y_1|x\rangle + \beta\langle y_2|x\rangle, \quad (1.8)$$

de modo que el espacio de las aplicaciones lineales también tiene la estructura de un espacio vectorial, usualmente llamado el *espacio dual*. Por lo tanto podemos considerar también los $\langle y|$ como un tipo de vectores (aunque distinto de los $|x\rangle$, como veremos en seguida) y construir una base dual $\{\langle e^i|\}$, en el cual las aplicaciones lineales se descomponen como

$$\langle y| = y_1\langle e^1| + y_2\langle e^2| + \dots + y_N\langle e^N| = y_i\langle e^i|. \quad (1.9)$$

Dado la estructura de espacio vectorial dual, es habitual llamar a los $\langle y|$ *vectores covariantes*, o *uno-formas*. Obsérvese que anotamos las componentes de los vectores contravariantes (es decir elementos de \mathbb{R}^N) con índice arriba, mientras las componentes de los vectores covariantes (los elementos del espacio dual) con índice abajo.

Los vectores covariantes actúan sobre los contravariantes mediante el *producto escalar*. Lo más sencillo es definir el producto escalar utilizando los vectores de base: por definición el producto escalar entre el vector de base dual $\langle e^i|$ y el vector de base $|e_k\rangle$ es

$$\langle e^i|e_j\rangle = \delta^i_j, \quad (1.10)$$

donde δ^i_j es la *delta de Kronecker*

$$\delta^i_j = \begin{cases} 0 & \text{cuando } i \neq j, \\ 1 & \text{cuando } i = j. \end{cases} \quad (1.11)$$

Utilizando las propiedades de linealidad (1.7) y (1.8), es obvio que en general el producto escalar entre un vector covariante $\langle y|$ y un vector contravariante $|x\rangle$ viene dado por

$$\langle y|x\rangle = y_i x^k \langle e^i|e_k\rangle = y_i x^k \delta^i_k = y_i x^i, \quad (1.12)$$

lo que efectivamente es un elemento de \mathbb{R} .

También en el espacio dual podíamos haber escogido otra base $\langle e'^i |$, relacionado con la base $\langle e^i |$ como

$$\langle e'^i | = \tilde{M}^i_j \langle e^j |. \quad (1.13)$$

Sin embargo, si queremos que bajo un cambio de coordenadas (1.3) en \mathbb{R}^N y un cambio de coordenadas (1.13) en el espacio dual, se sigue manteniendo la aplicación (1.10)

$$\langle e'^i | e'_k \rangle = \delta^i_k, \quad (1.14)$$

las matrices M^{-1} y \tilde{M} tienen que estar relacionadas. Efectivamente, sustituyendo (1.3) y (1.13) en (1.14), implica que \tilde{M} es la matriz inversa de M^{-1} . En otras palabras, la relación entre los vectores de la base dual viene dada por

$$\langle e'^i | = M^i_j \langle e^j |, \quad (1.15)$$

y por lo tanto, con un argumento similar que en el caso de los vectores contravariantes, las componentes y_i del vector covariante $\langle y |$ se transforman como

$$y_i = (M^{-1})^j_i y_j \quad (1.16)$$

También aquí podemos tomar esta propiedad como la definición práctica de un vector covariante: Cualquier objeto con N componentes y_i que bajo un cambio de base (1.3) transforma como (1.16) se le llama un *vector covariante*. Fijaos que, debido al hecho de que anotamos el vector contravariante x^i con índice arriba y el vector covariante x_i con índice abajo, las reglas de transformación son (ligeramente) diferentes para cada caso.

Finalmente, obsérvese que el producto escalar (1.12) entre dos vectores $\langle y |$ y $|x \rangle$ es independiente de la base en que se calcula, justo debido al hecho de que los vectores covariantes transforman con la métrica inversa que los contravariantes. Evaluada en las bases $\{|e'_i \rangle\}$ y $\{\langle e'^i | \}$ el producto escalar (1.12) viene dada por

$$\langle y | x \rangle = y'_i x'^i = (M^{-1})^k_i y_k M^i_l x^l = \delta^k_l y_k x^l = y_k x^k. \quad (1.17)$$

Esto es lo que uno esperaría, puesto que el producto escalar es un número real, cuyo valor no depende de la elección de base.

1.2. La métrica y las transformaciones ortogonales

En la sección anterior hemos considerado \mathbb{R}^N como un espacio vectorial, con los vectores contravariantes como elementos y espacio dual con los vectores covariantes. Hemos visto cómo transforma cada uno bajo cambios de base y hemos definido un producto escalar entre vectores covariantes y contravariantes que es independiente de la elección de base.

Sin embargo, esta estructura todavía es bastante pobre. No podemos dar una interpretación (física) al producto escalar, aparte de una aplicación lineal resultando en un número real. Tampoco podemos calcular la norma de un vector, o el ángulo entre dos vectores en \mathbb{R}^N o en el espacio dual. Para saber por ejemplo si una base $\{|e_i \rangle\}$ es ortogonal y normalizada, tenemos que calcular el producto escalar entre dos vectores de la base $|e_i \rangle$ y $|e_j \rangle$, mientras el producto escalar (1.10) y (1.12) sólo está definido entre un vector covariante y uno contravariante.

Lo que necesitamos claramente es una estructura nueva que nos permita relacionar un vector contravariante $|x \rangle$ unívocamente con su correspondiente vector covariante $\langle x |$. Esto es lo que va hacer la métrica y gracias a ella podremos definir normas, ángulos y distancias en \mathbb{R}^N .

La *métrica* g_{ij} se puede ver como una operación que actuando sobre un vector covariante da un vector contravariante. En los vectores de base actúa como

$$|e_i\rangle = g_{ij}\langle e^j|. \quad (1.18)$$

Por otro lado la *métrica inversa* g^{ij} transforma un vector covariante en uno contravariante

$$\langle e^i| = g^{ij}|e_j\rangle. \quad (1.19)$$

Nótese que la unicidad de la decomposición de vectores en una base implica que g_{ij} y g^{ij} están relacionados a través de la relación $g_{ij}g^{jk} = \delta_i^k$ (eerc.). De allí que g^{ij} se llama la métrica inversa.

En general podemos relacionar las componentes del vector contravariante $|x\rangle$ con las componentes del vector covariante $\langle x|$ a través de la métrica de la siguiente manera: en la base ortonormal $\{|e_i\rangle\}$, la descomposición del vector $|x\rangle$ es

$$x^i|e_i\rangle = x^i g_{ij}\langle e^j|, \quad (1.20)$$

donde hemos utilizado la relación (1.18) entre las dos bases. Ahora, también podemos pensar en el lado derecho de (1.20) como la descomposición de $\langle x|$ en la base $\{\langle e^i|\}$, lo que implica que las componentes x_i de $\langle x|$ se pueden escribir en función de las x^i como

$$x_i = g_{ij}x^j. \quad (1.21)$$

Con un argumento similar podemos también invertir esta relación, escribiendo las componentes x_i en función de las x^i (ejerc.)

$$x^i = g^{ij}x_j. \quad (1.22)$$

Por lo tanto vemos que la métrica y la inversa “suben y bajan índices”, convirtiendo vectores covariantes en contravariantes y vice versa. Una vez introducida la métrica, se puede por lo tanto pensar en los vectores covariantes como un “truco matemático” para poder definir bien el producto escalar entre dos vectores contravariantes, puesto en el fondo (1.21) y (1.22) están diciendo que un vector contravariante y su correspondiente vector covariante tienen la misma información física. Es más, resultará útil relajar un poco la distinción estricta entre los dos tipos de vectores e introducir la notación unificadora \vec{x} , siendo

$$\vec{x} = x^i|e_i\rangle = x_i\langle e^i| \quad (1.23)$$

Para que las relaciones (1.21)-(1.22) sean consistentes con las reglas de transformación (1.4) y (1.16) de los vectores covariantes y contravariantes (o, equivalentemente, para que (1.18) y (1.19) sean consistentes con las reglas de transformación de los vectores de base), la métrica y su inversa tienen que transformarse bajo un cambio de coordenadas M como

$$g'_{ij} = (M^{-1})^k{}_i (M^{-1})^l{}_j g_{kl}, \quad g'^{ij} = M^i{}_k M^j{}_l g^{kl}, \quad (1.24)$$

donde g'_{ij} y g'^{ij} son la métrica y su inversa en las nuevas bases. En general un cambio de coordenadas M no preserva la forma de la métrica, de modo que en general g'_{ij} puede tener una forma muy distinta que g_{ij} .

La introducción de la métrica permite ampliar la definición del producto escalar (1.12), tal que también incluye el producto escalar entre dos vectores contravariantes. Utilizando la definición (1.18), tenemos que

$$\vec{x} \cdot \vec{y} = \langle x|y\rangle = x_i y^i = g_{ij} x^i y^j. \quad (1.25)$$

De este producto escalar podemos entonces definir la norma $\|\vec{x}\|$ de un vector \vec{x} como

$$\|\vec{x}\|^2 \equiv \vec{x} \cdot \vec{x} = x_i x^i = g_{ij} x^i x^j, \quad (1.26)$$

Finalmente el ángulo θ entre dos vectores \vec{x} y \vec{y} se define como

$$\cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}, \quad (1.27)$$

de modo que dos vectores son ortogonales cuando el producto escalar es cero. Dejamos como ejercicio demostrar que las reglas de transformacion son tales que (1.25), (1.26) y (1.27) tienen el mismo valor en todas las bases.

Obsérvese que con la definición (1.25) el producto escalar entre dos vectores de base resulta

$$\begin{aligned}\vec{e}_i \cdot \vec{e}_j &= g_{ik} \langle e^k | e_j \rangle = g_{ik} \delta_j^k = g_{ik}, \\ \vec{e}^i \cdot \vec{e}^j &= g^{jk} \langle e^i | e_k \rangle = g^{jk} \delta_k^i = g^{ij},\end{aligned}\tag{1.28}$$

de modo que podemos definir una base ortonormal en \mathbb{R}^N . Decimos que la base $\{\vec{e}_i\}$ es ortogonal y normalizada si

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij},\tag{1.29}$$

donde δ_{ij} es la *métrica euclidea*² (en coordenadas cartesianas, es decir en una base ortonormal)

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.\tag{1.30}$$

En otras palabras, una base es ortonormal, si todos los vectores de base tienen norma 1 y son mutuamente ortogonales. Esto es la razón porque la métrica toma una forma tan sencilla en una base ortonormal. Consecuentemente las expresiones para el producto escalar y la norma toman la forma simple:

$$\begin{aligned}\vec{x} \cdot \vec{y} &= x_i y^i = \delta_{ij} x^i y^j, \\ \|\vec{x}\|^2 &= x_i x^i = \delta_{ij} x^i x^j,\end{aligned}\tag{1.31}$$

La norma (1.26) nos permite introducir el concepto de distancia en \mathbb{R}^N , ya que la distancia entre dos puntos x y y es la norma del vector $\overline{(x-y)} = \vec{x} - \vec{y}$. Está claro entonces que la expresión (1.31) para la norma no es más que la generalización a N dimensiones del Teorema de Pythagoras. La métrica δ_{ij} resume las propiedades geométricas del espacio. Concretamente en este caso, (1.30) nos está diciendo que la geometría de \mathbb{R}^N es la geometría plana de Euclides. En este punto podemos introducir el *elemento de línea* ds , la distancia entre un punto \vec{x} y un punto cercano $\vec{x} + d\vec{x}$. Para el caso de \mathbb{R}^N tenemos en coordenadas cartesianas que

$$ds^2 = \delta_{ij} dx^i dx^j = (dx^1)^2 + (dx^2)^2 + \dots + (dx^N)^2.\tag{1.32}$$

En general un cambio de coordenadas M no transforma una base ortonormal en otra ortonormal, de modo que en general la métrica no preserva la forma (1.30) bajo un cambio de base M . Sin embargo si existe un clase de transformaciones que tienen la propiedad que dejan la forma de la métrica invariante. Consideremos la clase de matrices M que tienen la propiedad que

$$(M^{-1})^l_j \delta_{kl} = M^l_k \delta_{lj}.\tag{1.33}$$

Si $\{\vec{e}_i\}$ es una base ortonormal (es decir, la métrica es toma la forma δ_{ij}), entonces, utilizando (1.24) es fácil de ver que en la nueva base, la forma de la métrica es

$$g'_{ij} = (M^{-1})^k_i M^l_k \delta_{lj} = \delta_{ij}.\tag{1.34}$$

En otras palabras, en la nueva base la métrica también es de la forma (1.30), puesto que las transformaciones consideradas dejan su forma invariante:

$$\delta_{ij} = (M^{-1})^k_i (M^{-1})^l_j \delta_{kl},\tag{1.35}$$

²Ojo, no se puede confundir la métrica euclidea δ_{ij} con la matriz identidad δ^i_j . La métrica es un objeto (un tensor) que describe las propiedades geométricas del espacio, mientras la matriz identidad es la transformación trivial, la unidad. La diferencia se hará más claro en la siguiente sección.

Las transformaciones que satisfacen la condición (1.33) se llaman *transformaciones ortogonales*. Tienen la buena propiedad de que transforman una base ortonormal en otra base ortonormal (esto es la razón porque deja la métrica invariante). Hay otras propiedades más de las matrices ortogonales que merecen la pena mencionar. La condición (1.33) está diciendo que la inversa de la matriz M es su traspuesta (en forma matricial $M^T = M^{-1}$). En otras palabras la inversa de la matriz M es su traspuesta. Además las matrices ortogonales M son reales y tienen determinante ± 1 . Esta última propiedad se saca directamente de (1.33), ya que si $M^T = M^{-1}$, tomando el determinante implica que $|M| = |M|^{-1}$, o en otras palabras, $|M| = \pm 1$.

Desde en punto de vista geométrico, las transformaciones ortogonales son las rotaciones (si tienen determinante 1) y las reflexiones (si tienen determinante -1). El grupo de las transformaciones ortogonales en N dimensiones es $O(N)$, que contiene el subgrupo $SO(N)$ de las rotaciones.

1.3. Algebra de tensores

En la sección anterior hemos definido vectores contravariantes x^i y covariantes x_i , que bajo cambios de coordenadas transforman como

$$x'^i = M^i_j x^j, \quad x'_i = (M^{-1})^j_i x_j, \quad (1.36)$$

y que están relacionados a través de la métrica δ_{ij} y la métrica inversa δ^{ij} como

$$x_i = g_{ij} x^j, \quad x^i = g^{ij} x_j. \quad (1.37)$$

Con estas definiciones (1.36), se puede construir objetos más generales: los escalares y los tensores.

Un *escalar* ϕ es un objeto que tiene el mismo valor en todos los sistemas de referencia

$$\phi' = \phi. \quad (1.38)$$

En otras palabras, un escalar es un invariante bajo cambios de coordenadas. Un ejemplo de un escalar es el producto escalar de dos vectores, como hemos visto previamente. Del mismo modo, una función escalar (o un campo escalar) es una función $V(x^i)$ que asigna a cada punto \vec{x} un valor numérico $V(x^i)$. Dado que es una función escalar, V es invariante bajo cambios de coordenadas:

$$V'(x'^i) = V(x^i). \quad (1.39)$$

En general V será una función de x^i diferente que V' de x'^i , pero los dos asignarán el mismo valor numérico al mismo punto del espacio. De la sección anterior está claro que, para que $V(x^i)$ sea una función escalar, V sólo puede depender de la combinación $x_i x^i$.

Considérese ahora el objeto T^{ij} construido de los vectores contravariantes a^i y b^j de la siguiente manera: $T^{ij} = a^i b^j$. De la definición de los vectores, está claro que T^{ij} transforma bajo el cambio de base como

$$T'^{ij} = a'^i b'^j = M^i_k a^k M^j_l b^l = M^i_k M^j_l T^{kl}. \quad (1.40)$$

Definimos un *tensor contravariante de rango 2* como cualquier objeto de N^2 componentes que transforma como (1.40). Obsérvese que un producto exterior (tensorial) de dos vectores contravariantes, como T^{ij} , siempre es un tensor contravariante, pero no todos los tensores contravariantes de rango 2 se pueden escribir como el producto de dos vectores contravariantes.

Del mismo modo podemos definir un *tensor covariante de rango 2* como un objeto T_{ij} que transforma como

$$T'_{ij} = (M^{-1})^k_i (M^{-1})^l_j T_{kl}, \quad (1.41)$$

y en general un *tensor mixto de rango m contravariante y rango n covariante* (o simplemente un *tensor de rango (m, n)*) como

$$T'^{i_1 \dots i_m}_{j_1 \dots j_n} = M^{i_1}_{k_1} \dots M^{i_m}_{k_m} (M^{-1})^{l_1}_{j_1} \dots (M^{-1})^{l_n}_{j_n} T^{k_1 \dots k_m}_{l_1 \dots l_n}. \quad (1.42)$$

En otras palabras, cada índice co(ntra)variante transforma como si fuera el índice de un vector co(ntra)variante.

Comparando esta definición con las reglas de transformación (1.24) de la métrica, vemos que la métrica g_{ij} (y en particular δ_{ij}) y su inversa g^{ij} (δ^{ij}) son en realidad tensores de rango 2 covariantes y contravariantes respectivamente. De la misma manera se puede pensar en un vector co(ntra)variante como un tensor de rango 1 y un escalar como un tensor de rango 0.

Es útil introducir los siguientes tensores importantes: la *delta de Kronecker*, definida en (1.11) y el *tensor de Levi-Civita* $\epsilon^{i_1 \dots i_N}$, definida como

$$\epsilon^{i_1 \dots i_N} = \begin{cases} 1 & \text{cuando } (i_1 \dots i_N) \text{ es una permutación par de } (12 \dots N), \\ -1 & \text{cuando } (i_1 \dots i_N) \text{ es una permutación impar de } (12 \dots N), \\ 0 & \text{en todos los demás casos.} \end{cases} \quad (1.43)$$

Por construcción la delta de Kronecker es un tensor simétrico, mientras el tensor de Levi-Civita es completamente antisimétrico en todos sus índices. La delta de Kronecker es en realidad la variante (1,1) de la métrica euclídea, es decir δ^i_j es lo mismo que δ_{ij} con un índice subido:

$$\delta^i_j = \delta^{ik} \delta_{kj}. \quad (1.44)$$

Insistimos otra vez que no se debería confundir la delta de Kronecker con la matriz identidad (1.6): la matriz es una transformación y la delta es un tensor que transforma bajo transformaciones.

También es útil definir el *producto vectorial* de dos vectores como

$$\vec{z} = \vec{x} \times \vec{y} = \epsilon^{ijk} x_j y_k \vec{e}_i, \quad (1.45)$$

o en componentes

$$z_i = (\vec{x} \times \vec{y})_i = \epsilon_{ijk} x_j y_k. \quad (1.46)$$

De la definición sigue que el vector resultante \vec{z} es ortogonal tanto a \vec{x} como a \vec{y} y el producto vectorial de un vector consigo mismo es cero (ejerc.). Además el producto vectorial es antisimétrico: $\vec{x} \times \vec{y} = -(\vec{y} \times \vec{x})$. Es interesante observar que el producto vectorial sólo está definido en tres dimensiones. En espacios de dimensión más alta o más baja no existe una operación equivalente. Fíjese que en la definición del producto escalar es crucial que el tensor de Levi-Civita lleve tres índices, para que el resultado pueda ser interpretado como un vector. En lenguaje más matemático: tres dimensiones es un caso especial porque aquí un tensor anti-simétrico de rango 2, (una dos-forma) es Hodge dual a un vector (una una-forma). Esto es justo la definición del producto vectorial. Para más detalles referimos al curso de teoría de formas.

La gran ventaja del uso de tensores es que si una relación tensorial $A^{i \dots j \dots} = B^{i \dots j \dots}$ es válida en una base, también es válida en cualquier otra base, puesto que los dos lados de la ecuación transforman bajo un cambio de base de la misma manera. Un ejemplo de esta propiedad veremos en el Capítulo 2 al demostrar la invariancia de la Segunda Ley de Newton bajo rotaciones. Por lo tanto, el hecho que por el Principio de la Relatividad las leyes de la física no dependan del sistema de referencia utilizado, implica que estas leyes deberían escribirse en términos de objetos que transforman bien bajo cambios de coordenadas, es decir escalares, vectores y tensores. En otras palabras, la estructura matemática restringe la forma de las leyes a ecuaciones vectoriales y tensoriales. Este principio se llama el *Principio de Covariancia* y forma la base del formalismo matemático de la teoría de la relatividad. En el siguiente capítulo entraremos en más detalle con ejemplos concretos, pero primero discutiremos algunas propiedades más de los tensores.

1.4. Operaciones con tensores

Una primera propiedad de los tensores es una generalización directa de una propiedad conocida de vectores: de las reglas de transformación (1.36) ((1.42)) está claro que si una componente de

un vector (tensor) es cero en una base, en general no lo será en otra base. Sin embargo, también está claro que si todas las componentes de un vector (tensor) son cero en una base, lo serán en todas las bases.

De la regla de transformación (1.42) también está claro que si $A^{i_1 \dots i_m}_{j_1 \dots j_n}$ y $B^{i_1 \dots i_m}_{j_1 \dots j_n}$ son dos tensores de rango (m, n) , la combinación lineal

$$C^{i_1 \dots i_m}_{j_1 \dots j_n} = \alpha A^{i_1 \dots i_m}_{j_1 \dots j_n} + \beta B^{i_1 \dots i_m}_{j_1 \dots j_n} \quad (1.47)$$

también es un tensor de rango (m, n) .

Si un tensor tiene la propiedad que $T^{ij \dots} = T^{ji \dots}$, decimos que T es un *tensor simétrico* en i y j . Si tiene la propiedad que $T^{ij \dots} = -T^{ji \dots}$, decimos que es un *tensor anti-simétrico* en i y j . De la propiedad (1.47) se deriva directamente que si un tensor es (anti-)simétrico en una base, lo es en todas (ejerc.).

De la forma en que hemos introducido los tensores también está claro que la composición de un tensor de (m, n) con uno de rango (p, q)

$$C^{i_1 \dots i_m}_{j_1 \dots j_n}{}^{k_1 \dots k_p}_{l_1 \dots l_q} = A^{i_1 \dots i_m}_{j_1 \dots j_n} B^{k_1 \dots k_p}_{l_1 \dots l_q} \quad (1.48)$$

es un tensor de rango $(m + p, n + q)$.

Se puede cambiar el rango de los tensores utilizando la métrica para “bajar índices”. Si $A^{i_1 \dots i_m}_{j_1 \dots j_n}$ es un tensor de rango (m, n) , entonces

$$C^{i_1 \dots i_{m-1}}{}_{k j_1 \dots j_n} = g_{kl} A^{i_1 \dots i_{m-1} l}_{j_1 \dots j_n} \quad (1.49)$$

es un tensor de rango $(m - 1, n + 1)$ (ejerc.). De la misma manera se puede “subir índices” con la métrica inversa g^{ij} y convertir un tensor de rango (m, n) en uno de rango $(m + 1, n - 1)$ (ejerc.). Está claro entonces por (1.37), que se puede cambiar un vector covariante en uno contravariante y vice versa subiendo y bajando los índices.

Se puede bajar el rango de un tensor contrayendo un índice covariante con uno contravariante. Si $A^{i_1 \dots i_m}_{j_1 \dots j_n}$ es un tensor de rango (m, n) ,

$$C^{i_1 \dots i_{m-1}}{}_{j_1 \dots j_{n-1} k} = A^{i_1 \dots i_{m-1} k}_{j_1 \dots j_{n-1} k} \quad (1.50)$$

es un tensor de rango $(m - 1, n - 1)$ (ejerc.). Visto así, el producto escalar es un tensor de rango 0 (un escalar) que surge de la contracción del tensor $x_i x^j$ de rango $(1, 1)$. Lo mismo se puede hacer contrayendo dos índices del mismo tipo con la métrica o la inversa. Por ejemplo

$$D^{i_1 \dots i_{m-2}}{}_{j_1 \dots j_n} = g_{kl} A^{i_1 \dots i_{m-2} kl}_{j_1 \dots j_n} \quad (1.51)$$

es un tensor de rango $(m - 2, n)$ (ejerc.). Otro ejemplo es que la traza $T^i{}_i = g_{ij} T^{ij} = g^{ij} T_{ij}$ de un tensor de rango 2 es un escalar.

El operador ∂ , la derivada con respecto a las coordenadas, se comporta como un vector. Para ver esto, es mejor mirar como actúa sobre un campo escalar. Consideramos la derivada parcial $\partial V / \partial x^i$. Por la regla de la cadena tenemos que

$$\frac{\partial V}{\partial x^i} = \frac{\partial V}{\partial x'^k} \frac{\partial x'^k}{\partial x^i} = \frac{\partial V}{\partial x'^k} M^k{}_i, \quad (1.52)$$

donde en la última igualdad hemos utilizado la relación (1.4) entre las coordenadas x^i y x'^i . En otras palabras la derivada $\partial_i V \equiv \partial V / \partial x^i$ tiene la misma regla de transformación que un vector covariante:

$$\partial'_i V = (M^{-1})^j{}_i \partial_j V. \quad (1.53)$$

Del mismo modo se puede demostrar que la derivada $\partial^i V \equiv \partial V / \partial x_i$ transforma como un vector contravariante (ejerc.). En notación vectorial muchas veces se utiliza el operador *nabla* $\vec{\nabla} = (\partial^i) \vec{e}_i$ y se identifica con el operador matemático de la derivada direccional.

Tomando en cuenta el hecho que ∂_i y ∂^i se comportan como un vector covariante y contravariante respectivamente, se puede utilizar estos operadores para actuar sobre campos tensoriales. Dado que la derivada de un campo escalar es un vector, en general la derivada ∂^i de un campo tensorial de rango (m, n) es un tensor de rango $(m + 1, n)$ y la derivada ∂_i de un campo tensorial de rango (m, n) es un tensor de rango $(m, n + 1)$ (ejerc.).

De este modo se puede interpretar los conocidos operadores diferenciales en \mathbb{R}^3 (la mayoría fácilmente generalizable a N dimensiones) en términos de (contracciones de) vectores y tensores:

1. El *gradiente* $\partial_i \phi$ de un campo escalar ϕ es un vector covariante, al ser la derivada de un escalar. En notación vectorial es $\vec{\nabla} \phi$.
2. La *divergencia* $\partial_i A^i$ de un vector contravariante A^i es un escalar, por ser el producto escalar entre los vectores ∂_i y A^i . En notación vectorial es $\vec{\nabla} \cdot \vec{A}$.
3. El *rotacional* $\epsilon^{ijk} \partial_j A_k$ de un vector covariante es un vector contravariante, por ser la contracción del tensor de Levi-Civita ϵ^{ijk} con el tensor de rango $(0, 2)$ $\partial_i A_j$. Esta operación corresponde al producto vectorial entre los vectores ∂_i y A^i y solamente está definido en \mathbb{R}^3 . En notación vectorial es $\vec{\nabla} \times \vec{A}$.
4. El *laplaciano* $\partial_i \partial^i \phi$ de un campo escalar ϕ es un escalar por ser el producto escalar entre los vectores ∂_i y $\partial^i \phi$. En notación vectorial es $\Delta \phi$.

1.5. Coordenadas curvilíneas

En la física un problema siempre se reduce a la forma más sencilla en el sistema de coordenadas que refleja la simetría del problema y esto no necesariamente son coordenadas cartesianas. Ejemplos conocidos en \mathbb{R}^3 son las coordenadas esféricas

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta. \quad (1.54)$$

o las coordenadas cilíndricas

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi, \quad z = z. \quad (1.55)$$

Los cambios de coordenadas entre cartesianas y esféricas o cilíndricas son muy distintos que los cambios de base que hemos estudiado en las asecciones anteriores. Merece por lo tanto la pena dedicar un poco de tiempo al estudio de las coordenadas curvilíneas y los cambios de coordenadas relacionadas con ellas.

Consideremos N funciones inversibles $y^\mu(x^i)$ de las coordenadas cartesianas x^i . El hecho de que las funciones $y^\mu(x^i)$ sean inversibles implica que a cada punto de \mathbb{R}^N se le puede asignar un único valor de y^μ y que cada valor de y^μ corresponde con un solo punto de \mathbb{R}^N . Las funciones y^μ sirven por lo tanto igual de bien como coordenadas sobre \mathbb{R}^N que las coordenadas cartesianas x^i .³ La transformación

$$y^\mu = y^\mu(x^i) \quad (1.56)$$

se llama una *transformación general de coordenadas*. La gran diferencia entre una transformación (1.56) y la transformación (1.4) es que en (1.56) las reglas de transformación varían de punto a punto, mientras en (1.4) son las mismas para todo \mathbb{R}^N . El cambio de coordenadas (1.4) se llama una *transformación global* y (1.56) una *transformación local*. El carácter local de los cambios generales de coordenadas marca un diferencia muy importante.

³Para evitar confusión utilizamos índices latinos (i, j, k, \dots) para las coordenadas cartesianas y índices griegos (μ, ν, ρ, \dots) para las coordenadas curvilíneas.

Igual que en el sistema cartesiano, podemos definir una base $\{|e_\mu\rangle\}$ y una base dual $\{\langle e^\mu|\}$ para las coordenadas curvilíneas. Un vector \vec{V} , que en una base cartesiana descompone como $\vec{V} = V^i|e_i\rangle = V_i\langle e^i|$, tendrá en coordenadas curvilíneas la forma $\vec{V} = V^\mu|e_\mu\rangle = V_\mu\langle e^\mu|$, donde los $|e_\mu\rangle$ y $\langle e^\mu|$ son los vectores de base (no necesariamente ortogonales o normalizados) definidos como

$$|e_\mu\rangle = \frac{\partial x^i}{\partial y^\mu} |e_i\rangle, \quad \langle e^\mu| = \frac{\partial y^\mu}{\partial x^i} \langle e^i|. \quad (1.57)$$

Por un argumento idéntico al de la sección 1.1, es fácil de averiguar que entonces las componentes de \vec{V} en las distintas coordenadas están relacionadas como

$$V^\mu = \frac{\partial y^\mu}{\partial x^i} V^i, \quad V_\mu = \frac{\partial x^i}{\partial y^\mu} V_i. \quad (1.58)$$

La generalización a tensores de rango (m, n) debería ser obvia.

Obsérvese que la transformación entre la base cartesiana y la curvilínea depende del punto \vec{r} en \mathbb{R}^N en que se haga, debido al carácter local de la transformación (1.56). Esto implica que, en contraste con el sistema cartesiano, una dirección \vec{e}_μ varía de punto en punto. Por ejemplo, la dirección radial \vec{e}_r en coordenadas polares es distinta en puntos distintos. Esto tendrá consecuencias importantes, como veremos en seguida.

Obsérvese también que ya hemos usado que los vectores covariantes y contravariantes están relacionados, lo que implica que en realidad ya hemos usado el hecho de que tenemos definida una métrica $g_{\mu\nu}$ que relaciona los vectores de base como $|e_\mu\rangle = g_{\mu\nu}\langle e^\nu|$ y $\langle e^\mu| = g^{\mu\nu}|e_\nu\rangle$.

La forma de la métrica se puede obtener de varias maneras. Una es a través de la generalización de (1.28),

$$g_{\mu\nu} = \vec{e}_\mu \cdot \vec{e}_\nu. \quad (1.59)$$

Otra posibilidad es dándose cuenta que la métrica es un tensor de rango 2 que transforma por lo tanto bajo un cambio de coordenadas (1.56) como

$$g_{\mu\nu} = \frac{\partial x^i}{\partial y^\mu} \frac{\partial x^j}{\partial y^\nu} \delta_{ij}. \quad (1.60)$$

Por ejemplo, un pequeño ejercicio muestra que la métrica de \mathbb{R}^3 en coordenadas cilíndricas y esféricas vienen dadas por

$$\begin{aligned} ds^2 &= dr^2 + r^2 d\theta^2 + dz^2, \\ ds^2 &= dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \end{aligned} \quad (1.61)$$

La métrica $g_{\mu\nu}$ juega el mismo papel que la métrica euclídea cartesiana δ_{ij} . En particular, el producto escalar de dos vectores en coordenadas curvilíneas está definida como

$$\vec{V} \cdot \vec{W} = V_\mu W^\mu = g_{\mu\nu} V^\mu W^\nu. \quad (1.62)$$

Las operaciones diferenciales son un poco más sutiles en coordenadas curvilíneas que en coordenadas cartesianas. Donde en coordenadas cartesianas la derivada de un vector es simplemente la derivada de cada componente

$$\frac{\partial \vec{V}}{\partial x^i} = \frac{\partial V^j}{\partial x^i} \vec{e}_j, \quad (1.63)$$

en coordenadas curvilíneas el operador diferencial también actúa sobre los vectores de base \vec{e}_μ , puesto que estos cambian de punto en punto. Por lo tanto la derivada de un vector \vec{V} está dada por

$$\partial_\mu \vec{V} = \partial_\mu V^\nu \vec{e}_\nu + V^\nu \partial_\mu \vec{e}_\nu = \left(\partial_\mu V^\rho + V^\nu \Gamma_{\mu\nu}^\rho \right) \vec{e}_\rho. \quad (1.64)$$

En la última igualdad hemos definido la *conexión de Levi-Civita* (o los *símbolos de Christoffel*) $\Gamma_{\mu\nu}^{\rho}$ como

$$\partial_{\mu}\vec{e}_{\nu} = \Gamma_{\mu\nu}^{\rho}\vec{e}_{\rho}, \quad (1.65)$$

o sea, $\Gamma_{\mu\nu}^{\rho}$ es el componente ρ del vector $\partial_{\mu}\vec{e}_{\nu}$ en la base $\{\vec{e}_{\rho}\}$. El objeto

$$\nabla_{\mu}V^{\rho} = \partial_{\mu}V^{\rho} + \Gamma_{\mu\nu}^{\rho}V^{\nu} \quad (1.66)$$

es la *derivada covariante* de V_{μ} y veremos que tiene la propiedad que transforma como un tensor de rango 2 bajo cambio de coordenadas. Aunque no es obvio de la definición (1.65), la conexión de Levi-Civita $\Gamma_{\mu\nu}^{\rho}$ es simétrica en los índices μ y ν (ejerc.) y se puede expresar en función de la métrica como (ejerc.)

$$\Gamma_{\mu\nu}^{\rho} = \frac{1}{2}g^{\rho\lambda}\left(\partial_{\mu}g_{\lambda\nu} + \partial_{\nu}g_{\mu\lambda} - \partial_{\lambda}g_{\mu\nu}\right). \quad (1.67)$$

Es fácil de comprobar que la conexión de Levi-Civita es cero en coordenadas cartesianas, pero no lo es en coordenadas esféricas o cilíndricas.

Una propiedad importante de la derivada covariante (1.66) es que transforma como un tensor bajo cambios generales de coordenadas, en contraste con la derivada parcial. Bajo un cambio general de coordenadas $z^{\alpha} = z^{\alpha}(y^{\mu})$, es fácil de ver que la derivada parcial de un vector transforma como

$$\partial_{\alpha}V^{\beta} = \frac{\partial x^{\mu}}{\partial y^{\alpha}}\partial_{\mu}\left(\frac{\partial y^{\beta}}{\partial x^{\nu}}V^{\nu}\right) = \frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial y^{\beta}}{\partial x^{\nu}}\partial_{\mu}V^{\nu} + \frac{\partial x^{\mu}}{\partial y^{\alpha}}V^{\nu}\frac{\partial^2 y^{\beta}}{\partial x^{\mu}\partial x^{\nu}}. \quad (1.68)$$

El primer término de la derecha sí tiene la forma adecuada para la transformación de un tensor de rango (1, 1), pero tenemos un término extra, proporcional a la segunda derivada del cambio de coordenadas. La presencia de este término es debido al carácter local de la transformación, ya que en el caso de una transformación global las entradas de la matriz M^{β}_{ν} son constantes.

Por otro lado, a través de la expresión de la conexión de Levi-Civita en función de la métrica, podemos calcular que $\Gamma_{\mu\nu}^{\rho}$ tampoco transforma como un tensor bajo el mismo cambio de coordenadas, debido a un término inhomogéneo en las reglas de transformación (ejerc.):

$$\Gamma_{\alpha\beta}^{\gamma} = \frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial x^{\nu}}{\partial y^{\beta}}\frac{\partial y^{\gamma}}{\partial x^{\rho}}\Gamma_{\mu\nu}^{\rho} - \frac{\partial^2 y^{\gamma}}{\partial x^{\mu}\partial x^{\nu}}\frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial x^{\nu}}{\partial y^{\beta}}. \quad (1.69)$$

Sin embargo, es fácil de ver que en la derivada covariante (1.66) los dos términos que rompen el carácter tensorial de la derivada parcial y de la conexión se cancelan, de modo que la derivada covariante transforma como

$$\nabla_{\alpha}V^{\gamma} = \frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial y^{\gamma}}{\partial x^{\rho}}\nabla_{\mu}V^{\rho}. \quad (1.70)$$

Debido a su carácter tensorial, la derivada covariante resulta apropiado para generalizar las definiciones de los operadores diferenciales en coordenadas curvilíneas. La divergencia, el rotacional y el laplaciano para coordenadas curvilíneas se defina a partir de la definición de la derivada covariante (1.66):

$$\vec{\nabla} \cdot \vec{A} = \nabla_{\mu}A^{\mu}, \quad \vec{\nabla} \times \vec{A} = \epsilon^{\mu\nu\rho}\nabla_{\mu}A_{\nu}\vec{e}_{\rho}, \quad \Delta\phi = \nabla_{\mu}\partial^{\mu}\phi. \quad (1.71)$$

Observa que debido al hecho de que la conexión es simétrica en μ y ν , la definición del rotacional no ha cambiado: $\epsilon_{\mu\nu\rho}\nabla_{\mu}A_{\nu}\vec{e}_{\rho} = \epsilon_{\mu\nu\rho}\partial_{\mu}A_{\nu}\vec{e}_{\rho}$.

Dejamos como ejercicio calcular las expresiones de estas operaciones diferenciales en coordenadas esféricas y cilíndricas.

Capítulo 2

Repaso de la mecánica newtoniana

En este capítulo repasaremos la mecánica newtoniana, puesto que antes de meternos en formalismos matemáticos para describir la física, necesitamos primero entender bien esa física. Aparte de la física de una sola partícula, de N partículas y de cuerpos sólidos, revisaremos es estructura matemática de las leyes de la física.

2.1. Mecánica de una sola partícula

Consideramos una partícula moviéndose por el espacio físico. La *posición* de la partícula con respecto al origen o es el vector \vec{r} desde o al punto r donde se encuentra la partícula. El conjunto de posiciones $\vec{r}(t)$ en un intervalo de tiempo $[t_1, t_2]$ es la *trayectoria* de la partícula entre $t = t_1$ y $t = t_2$.

La *velocidad* \vec{v} de la partícula está definida como la derivada del vector de posición con respecto al tiempo¹

$$\vec{v} = \frac{d\vec{r}}{dt}, \quad (2.1)$$

y es por definición un vector tangente a la trayectoria. La *aceleración* \vec{a} es la derivada de la velocidad con respecto a t

$$\vec{a} = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2}, \quad (2.2)$$

y mide el cambio de la velocidad. Nótese que la ecuación (2.2) es una ecuación vectorial, así que un cambio de velocidad puede ser tanto en tamaño como en orientación.

En general sobre una partícula actúan fuerzas, debido a la interacción con otras partículas, con campos gravitatorios o electromagnéticos, o simplemente fuerzas de empuje y de tiro. Una fuerza \vec{F} actuando sobre una partícula induce un cambio en su velocidad y esta aceleración es proporcional (en tamaño y en dirección) a la fuerza que la ha originado:

$$\vec{F} = m\vec{a}. \quad (2.3)$$

La constante de proporcionalidad m es la *masa* de la partícula y es una medida de la inercia de la partícula: cuanto más grande la masa tanto menos cambia la velocidad de la partícula bajo una fuerza dada. Igualmente cuanto más grande la masa, más fuerza es necesaria para generar un cierto cambio de velocidad.

¹Siempre supondremos que todas las funciones encontradas son continuas y diferenciables las veces que haga falta.

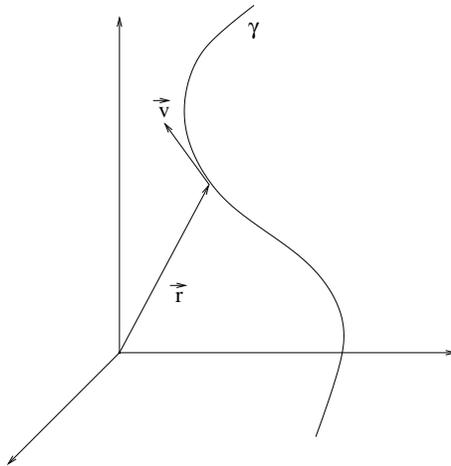


Figura 2.1: La trayectoria γ de una partícula, su vector de posición \vec{r} y de velocidad \vec{v} .

La ecuación (2.3) se llama la (*Segunda*) *Ley de Newton* y se puede considerar como una definición de la masa habiendo definido lo que es una fuerza, o al revés como la definición de fuerza, sabiendo lo que es una masa. Nótese que por (2.2) la Segunda Ley es un sistema de ecuaciones diferenciales de segundo grado en \vec{r} , lo que implica 6 constantes de integración. Dada una fuerza, la evolución del sistema está completamente determinada por las posiciones y las velocidades iniciales.

Una cantidad más importante que la velocidad es el *momento* o cantidad de movimiento \vec{p} , definido como el producto de la velocidad con la masa

$$\vec{p} = m\vec{v}. \quad (2.4)$$

En función de \vec{p} , la Segunda Ley de Newton se escribe como

$$\vec{F} = \frac{d\vec{p}}{dt}, \quad (2.5)$$

lo que implica directamente que el momento sigue constante mientras que no actúe ninguna fuerza sobre la partícula. Si la masa de la partícula no cambia, la ausencia de fuerzas implica que la partícula sigue moviéndose con la misma velocidad. Esta es la *Primera Ley de Newton: una partícula mantiene su estado de movimiento uniforme y rectilíneo a menos que una fuerza actúe sobre ella*.

Las leyes de Newton² no son válidas en cualquier sistema de referencia. Los sistemas en que (2.3) es válida se llaman *sistema inerciales*. La definición exacta de qué sistema es inercial y cuál no, no es obvio (incluso imposible por razones profundas dado por la teoría de la relatividad general), pero aquí nos vale decir que intuitivamente un sistema inercial corresponde a un sistema que no está acelerado o rotando. En la práctica la elección del sistema inercial dependerá del problema concreto que uno quiere resolver: en un experimento de choque de bolas en un laboratorio se puede tomar la Tierra como sistema inercial (aunque se está moviendo alrededor del sol), para el problema de las órbitas planetarias lo más fácil es colocar un sistema inercial en el sol (aunque se mueve en la Vía Láctea), etc.

En los sistemas que no son inerciales las partículas experimentarán aceleraciones extras, debido a la aceleración del sistema de referencia. Un observador no-inercial puede en este caso o bien

²Falta todavía la Tercera Ley de Newton, la de acción y reacción, para completar las leyes de la mecánica. Esta la veremos en la siguiente sección, puesto que aplica a un sistema de dos partículas.

aceptar que la Segunda Ley (2.3) no es válida, o bien atribuir la aceleración extra a una fuerza extra, una llamada *pseudo-fuerza*, puesto que para un observador inercial la aceleración extra y por lo tanto también la fuerza extra no existen. Un ejemplo típico de una pseudo-fuerza es la fuerza de Coriolis, que hace que los anti-ciclones giran en contra del sentido de las agujas del reloj y que hacen girar el péndulo de Foucault.

Un sistema S' que se mueve con una velocidad constante \vec{v} con respecto a un sistema inercial S , también es un sistema inercial y por lo tanto la Segunda Ley (2.3) también será válida en el sistema S' . Dado que la Segunda Ley es el fundamento de la mecánica, todas las leyes físicas que son válidas en el sistema S por lo tanto también serán válidas en el sistema S' . Esto implica que no hay manera de saber cual de los dos sistemas realmente se está moviendo y cual está en reposo: no se puede hacer ningún experimento mecánico en un sistema inercial que nos diga si el sistema está en movimiento o en reposo. No existe por lo tanto una velocidad absoluta, respecto a un sistema de referencia absoluto. Sólo tiene sentido hablar de las posiciones y las velocidades relativas entre sistemas inerciales. Este principio se llama el *Principio de la Relatividad* de Galilei.³

El *momento angular* \vec{L} de una partícula con respecto a un punto O se define como

$$\vec{L} = \vec{r} \times \vec{p} \quad (2.6)$$

y mide la “cantidad de rotación” de la partícula respecto a O . Aquí \vec{r} no necesariamente es el vector de posición, dado que O no necesariamente es el origen del sistema de coordenadas (aunque suele ser muy conveniente elegirlo como tal). El *momento de fuerza* o *torque* \vec{M} con respecto a un punto O se define como

$$\vec{M} = \vec{r} \times \vec{F}. \quad (2.7)$$

Utilizando la Segunda Ley de Newton en la forma (2.5) no es difícil demostrar que

$$\vec{M} = \vec{r} \times \frac{d\vec{p}}{dt} = \frac{d}{dt}(\vec{r} \times \vec{p}) = \frac{d\vec{L}}{dt}, \quad (2.8)$$

donde en la segunda igualdad hemos utilizado que $\frac{d\vec{r}}{dt} \times \vec{p} = 0$ por la definición (2.4) de \vec{p} . El momento angular y el momento de fuerza son el análogo angular del momento y la fuerza y la ecuación (2.8) es el equivalente de (2.5). Deducimos directamente de (2.8) que en ausencia de momento de fuerza, el momento angular se conserva.

Hay una clase de fuerzas, llamadas *fuerzas centrales*, donde el vector \vec{F} siempre apunta al mismo punto del espacio (vease capítulo 5). En estos campos de fuerza el momento de fuerza siempre es cero y el momento angular por lo tanto constante (ejerc.).

El *trabajo* W realizado por una fuerza \vec{F} llevando una partícula de un punto x a un punto y a lo largo de una curva C está definida como

$$W = \int_C \vec{F} \cdot d\vec{s}. \quad (2.9)$$

En general el trabajo realizado depende de la curva entre x y y . Utilizando la segunda ley se puede escribir esta definición como

$$W = m \int_x^y \frac{d\vec{v}}{dt} \cdot \vec{v} dt = \frac{1}{2} m \int_x^y d(v^2) = \frac{1}{2} m v^2(y) - \frac{1}{2} m v^2(x), \quad (2.10)$$

donde $v(x)$ y $v(y)$ son las velocidades de la partícula en los puntos x y y . El trabajo realizado resulta en un cambio de la norma de la velocidad. La cantidad $T = \frac{1}{2} m v^2$ se llama la *energía*

³En realidad se trate de un principio muy profundo, que es válido mucho más allá de la mecánica. No existe ningún experimento físico (no sólo mecánico) que puede dar la velocidad absoluta de un sistema inercial. La aplicación del Principio de la Relatividad a fenómenos electromagnéticos es lo que llevó a Einstein a desarrollar su teoría de la relatividad especial en 1905.

cinética y (2.10) implica que la cantidad de trabajo realizado por la fuerza es igual a la diferencia de energía cinética entre los puntos x y y :

$$W = T(y) - T(x). \quad (2.11)$$

Ya hemos mencionado que en general el trabajo W depende de la curva elegida. Existe una clase de fuerzas especiales, las llamadas *fuerzas conservativas* que tienen la propiedad que W no depende de la curva. En particular tenemos entonces que para cualquier curva cerrada

$$W = \oint \vec{F} \cdot d\vec{s} = 0. \quad (2.12)$$

Ahorra, según el teorema de Stokes cada campo vectorial que satisface (2.12) tiene la propiedad $\vec{\nabla} \times \vec{F} = 0$ y por lo tanto se puede escribir como el gradiente de un campo escalar, que por razones de comodidad denominaremos $-V$:

$$\vec{F} = -\vec{\nabla}V \quad (2.13)$$

El campo escalar V se llama la *energía potencial*. Fijaos que V no está completamente determinada por la ecuación (2.13): cualquier función V' relacionada con V a través de $V' = V + c$, con c una constante arbitraria, también es una solución de (2.13). El punto $V = 0$ por lo tanto no está determinada unívocadamente, sino que es arbitrario y depende del convenio utilizado. En otros palabras: el valor del potencial no tiene un significado bien definido, sólo las diferencias de potencial.

De (2.9) y (2.13) se puede derivar otra expresión para el trabajo realizada por un campo de fuerza conservativo:

$$W = - \int_x^y \frac{\partial V}{\partial \vec{r}} \cdot d\vec{r} = - \int_x^y dV = V(x) - V(y) \quad (2.14)$$

Igualando esta expresión con (2.11), para un campos de fuerzas conservativas obtenemos que $T(y) - T(x) = V(x) - V(y)$, o, poniendo las funciones evaluadas en el mismo punto al mismo lado de la ecuación,

$$T(x) + V(x) = T(y) + V(y). \quad (2.15)$$

En un campo de fuerza conservativo, la suma de la energía cinética y la energía potencial es constante. A esta suma de energía cinética y energía potencial se la llama la *energía mecánica* $E = T + V$, o en problemas de mecánica simplemente la energía. La expresión (2.15) por lo tanto un ejemplo de la *Ley de la conservación de energía mecánica*: la energía mecánica de una partícula sometida a un campo de fuerza conservativo es constante en cada momento.

En un campo de fuerzas no-conservativo esta ley de conservación de energía no se cumple, puesto que una parte de la energía se pierde. Considera un campo de fuerzas \vec{F} que consiste tanto de una parte conservativa \vec{F}_c un una parte no-conservativa \vec{F}_{nc}

$$\vec{F} = \vec{F}_c + \vec{F}_{nc} = -\vec{\nabla}V + \vec{F}_{nc}. \quad (2.16)$$

El trabajo realizado por esta fuerza \vec{F} es

$$W = - \int_x^y \vec{\nabla}V \cdot d\vec{s} + \int_x^y \vec{F}_{nc} \cdot d\vec{s}, \quad (2.17)$$

tal que por (2.11) tenemos que

$$T(y) - T(x) = V(x) - V(y) + \int_x^y \vec{F}_{nc} \cdot d\vec{s}, \quad (2.18)$$

o, en términos de la energía mecánica E

$$E(y) - E(x) = \int_x^y \vec{F}_{nc} \cdot \vec{v} dt. \quad (2.19)$$

Se ve por lo tanto que la diferencia de energía mecánica entre los puntos a y b es debido solamente a la acción de las fuerzas no-conservativas.

Daremos ahora unos ejemplos típicos de fuerzas conservativas y no-conservativas:

- La fuerza gravitatoria de una masa M actuando sobre una masa de prueba m viene dado por la *Ley de la gravitación universal* de Newton:

$$\vec{F} = G \frac{mM}{r^2} \vec{e}_r, \quad (2.20)$$

donde $G = 6,673 \cdot 10^{-11} m^3/kg s^2$ es la constante de Newton, r la distancia entre las dos masas y \vec{e}_r un vector de unidad apuntando desde la posición de m hacia la posición de M . Esta fuerza también es conservativa y el potencial es

$$V = -G \frac{mM}{r}. \quad (2.21)$$

La m que aparece en (2.20) es la masa gravitatoria y estrictamente hablando es distinta a la masa inercial m que aparece en (2.3). Por razones fuera del alcance de la mecánica clásica las dos masas tienen exactamente el mismo valor, lo que hace que todos los objetos caen con la misma velocidad (en el vacío). En la práctica por lo tanto se identifica la masa inercial con la gravitatoria. Esa igualdad de los dos tipos de masas tiene su origen en la *Principio de Equivalencia* de Einstein, que forma la base de la teoría de la relatividad general.

- La fuerza actuando sobre una partícula de masa m en caída libre a altura z cerca de la superficie de la Tierra viene dada por $\vec{F} = -m g \vec{e}_z$, donde g es la aceleración gravitatoria y tiene el valor de $g = 9,8 m/s^2$. En realidad es un caso especial de la fuerza gravitatoria mencionado arriba para $z \ll R$, con R el radio de la Tierra, tal que $g = GM/R^2$.⁴ Esta fuerza es conservativa y el potencial viene dado por $V = mgz$, donde z es la altura de la partícula medido desde el suelo.
- La fuerza ejercida por un muelle de tamaño L es proporcional a la deformación (compresión o alargamiento) del muelle: $\vec{F} = k(L - x)\vec{e}_x$, donde por conveniencia hemos colocado el muelle a lo largo del eje x y k es la constante elástica, característica del muelle. Esta fuerza es claramente conservativa y el potencial viene dado por $V = -\frac{1}{2}k(L - x)^2$. Un sistema con un potencial cuadrático en las coordenadas, como el del muelle, se llama un *oscilador armónico* y estudiaremos los movimientos de estos sistemas en más detalle en el capítulo 7.
- Las fuerzas de rozamiento son fuerzas que aparte de la posición de la partícula, también dependen de su velocidad. Suelen ser en dirección opuesta a la velocidad, puesto que frenan el movimiento. Por lo tanto el último término de (2.19) es distinto de cero y hay una pérdida de energía mecánica. En realidad es energía mecánica está transformada en calor, debido al rozamiento.
- La *fuerza de Lorentz* sobre una partícula con carga eléctrica q y velocidad \vec{v} en un campo eléctrico \vec{E} y magnético \vec{B} es $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$. Aunque la fuerza del campo magnético depende de \vec{v} , es fácil de ver que el trabajo realizado por esta fuerza es cero, debido al hecho que $\vec{v} \times \vec{B}$ es un vector ortogonal a \vec{v} . La fuerza eléctrica es conservadora sólo en el caso que todos los campos sean independientes del tiempo. En este caso se puede escribir $\vec{E} = -\vec{\nabla}\phi$, donde ϕ es el potencial eléctrico y el potencial mecánico V viene dado por $V = q\phi$. En el caso más general de campos no-estáticos, tenemos, según la Leyes de Maxwell, que $\vec{\nabla} \times \vec{E} = -\partial_t \vec{B}$ y $\vec{\nabla} \times \vec{B} = \vec{j} + \partial_t \vec{E}$. Por lo tanto $\vec{\nabla} \times \vec{F}$ es distinto de cero y la energía mecánica de la partícula cargada no se conserva. Sin embargo los campos \vec{E} y \vec{B} también tienen una energía asociada y el total de la energía mecánica de la partícula y de los campos sí se conserva.

⁴El signo surge porque tomamos $\vec{e}_z = -\vec{e}_r$.

La ley de la conservación de la energía mecánica para fuerzas conservativas es sólo un caso especial de la *Ley de la conservación de la energía total*. La energía puede aparecer en muchas formas diferentes (energía cinética, potencial, eléctrica, calor, ...) y transformar de una forma en otra. Pero en estos procesos de transformación nunca se crea o se pierde energía. La cantidad total de la energía del universo es constante. Ejemplos de estas transformaciones ya hemos los dado en los ejemplos de fuerzas no-conservativas.

2.2. Simetrías y la forma de las leyes de la física

Antes de seguir estudiando las leyes de la física relacionadas con sistemas de múltiples partículas y cuerpos sólidos, merece la pena hacer un breve inerezzo sobre simetrías y la estructura que imponen en la forma de las leyes físicas. Estudiaremos principalmente las simetrías del espacio y el tiempo, pero las conclusiones que se pueden sacar son extensibles a cualquier tipode simetrías en una teoría física.

Las leyes de la física son invariantes bajo traslaciones en el tiempo

$$t \rightarrow t + \Delta t, \quad (2.22)$$

por cualquier valor de Δt . En otras palabras, no hay un momento especial $t = 0$ y las leyes son válidas en todos los momentos t . Esta simetría encaja bien con la hipótesis básica de la ciencia de que los resultados son reproducibles, por la sencilla razón de que sino la teoría no es falsificable. Si un científico obtiene cierto resultado en el momento $t = t_1$, otro investigador debería ser capaz de obtener el mismo resultado a $t = t_1 + \Delta t$. La simetría (2.22) se llama la *homogeneidad del tiempo*.

De la misma manera, las leyes de la física son invariantes bajo traslaciones en el espacio

$$\vec{r} \rightarrow \vec{r} + \vec{r}_0, \quad (2.23)$$

para cualquier vector \vec{r}_0 . Esto corresponde al hecho de que no hay un punto especial en el espacio y que podemos elegir el origen del sistema de referencia donde queramos. Esta simetría se llama la *homogeneidad del espacio*.

La *isotropía del espacio* es el hecho de que todas las direcciones son equivalentes y que no hay ninguna dirección preferida. La isotropía implica una invariancia bajo rotaciones ortogonales arbitrarias

$$x'^i = \sum_{j=1}^3 M^i_j x^j, \quad (2.24)$$

donde M^i_j es la matriz ortogonal que representa la rotación.

Estas simetrías imponen restricciones en la forma que pueden tener las leyes físicas. Si una ley es válida por un observador \mathcal{O} que trabaja con ciertas coordenadas \vec{r} en un momento t , también debería ser válida para un observador \mathcal{O}' en un momento t' que usa coordenadas \vec{r}' , relacionadas con \vec{r} a través de (2.23)-(2.24). En otras palabras, las leyes de la física debe poder escribirse en término de objetos que transforman bien bajo las transformaciones (2.23)-(2.24). Por el análisis del capítulo anterior sabemos que en general estos objetos serán escalares, vectores y tensores. Pero las simetrías del espacio no solo determinan la estructura de las leyes físicas, sino por gran parte también su forma.

La isotropía del tiempo causa que un lagrangiano $\mathcal{L} = T - V$,⁵ que describe un sistema de objetos en posiciones x^i con velocidades \dot{x}^i , no puede depender explícitamente de t :

$$\partial_t \mathcal{L}(x^i(t), \dot{x}^i(t)) = 0. \quad (2.25)$$

⁵Para la definición más rigurosa del lagrangiano referimos al Capítulo 3. De momento basta saber que un lagrangiano es una función de las coordenadas que resume la dinámica del sistema.

Como indicado, las variables x^i y \dot{x}^i sí pueden ser dinámicas y depender de t , pero la dependencia del lagrangiano de t sólo puede darse a través de éstas variables. La homogeneidad del espacio implica que las interacciones permitidas en el potencial $V(x^i)$ sólo pueden depender de las distancias entre dos partículas, pero no de las posiciones de las partículas con respecto a cierto origen. Y finalmente la isotropía del espacio implica que el potencial V tiene que ser un potencial central, que depende de la distancia radial r , pero no de los ángulos.

Efectivamente, los potenciales que uno encuentra en teorías de gravedad, electromagnetismo, física molecular o nuclear son típicamente de la forma $V = V(|\vec{r}_1 - \vec{r}_2|)$, que claramente satisface los requisitos mencionados arriba.

Las simetrías no sólo restringen la forma que puede tener una teoría, sino también ayudan a la hora de encontrar soluciones de las ecuaciones de movimiento. Si por ejemplo queremos encontrar la expresión para el campo eléctrico, causado por un electrón en reposo en el origen de un sistema de coordenadas, tenemos que resolver la ecuación de Poisson para el potencial ϕ ,

$$\nabla^2\phi \equiv \frac{1}{r^2}\partial_r(r^2\partial_r\phi) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta\phi) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2\phi = 0. \quad (2.26)$$

Para resolver la ecuación (2.26), se podría aplicar el método de separación de variables y buscar la solución más general en función de las funciones armónicas esféricas $Y_{lm}(\theta, \varphi)$, pero nos ahorraríamos mucho trabajo si utilizamos la simetría del sistema. Dado que el electrón en el origen posee una simetría esférica, el campo eléctrico también tiene que reflejar esta simetría: la solución dependerá sólo de la coordenada radial r y no de los ángulos θ y φ . La ecuación diferencial parcial (2.26) se reduce por lo tanto a una ecuación diferencial ordinaria en r y es fácil averiguar que la solución para el campo eléctrico es la solución conocida de la electrostática:

$$\phi(r) = \frac{Q}{r} \quad \implies \quad \vec{E} = -\vec{\nabla}\phi = -\frac{Q}{r^2}\vec{e}_r. \quad (2.27)$$

Aparte de las simetrías de traslaciones y rotaciones de arriba, el Principio de la Relatividad de Galilei nos proporciona otra simetría, relacionada con observadores en movimiento relativo. Ya que el Principio de la Relatividad impone que un observador no es capaz de medir la velocidad absoluta de su sistema de referencia, sino solamente velocidades relativas entre sistemas de referencia, obviamente este implica que las leyes de la física no pueden ser formuladas en término de velocidades. Sólo cambios de velocidad son admisibles, ya que estos son independientes de los observadores. Efectivamente, las leyes de Newton están formulados en función de la aceleración \vec{a} , y no de la velocidad \vec{v} .

Pero hay más. Si dos observadores observan el mismo suceso, cada uno en su propio sistema de referencia, tiene que haber una manera de relacionar los resultados de un observador con los resultados del otro. La relación entre los resultados de diferentes observadores se llama *cambio de coordenadas*. Las rotaciones y las traslaciones en el tiempo y el espacio son ejemplos de cambios de coordenadas, pero por lo menos igual de importante son los cambios entre dos observadores que están en movimiento uniforme rectilíneo relativo.

Dos observadores \mathcal{O} y \mathcal{O}' tienen sus sistemas de coordenadas orientados de tal manera que los ejes de sus coordenadas cartesianas son paralelos. El observador \mathcal{O}' se mueve con una velocidad constante \vec{v} con respecto al sistema de referencia de \mathcal{O} , así que la posición del origen de \mathcal{O}' con respecto a \mathcal{O} a tiempo t arbitrario viene dado por $\vec{r}_0 + \vec{v}t$, donde \vec{r}_0 es la posición de \mathcal{O}' a $t = 0$ en el sistema de \mathcal{O} (ve Figura 2.2). Una masa m , que se mueve según el observador \mathcal{O}' con velocidad \vec{V}' , se encontrará a tiempo t en la posición $\vec{r}' + \vec{V}'t$ en el sistema de coordenadas de \mathcal{O}' . Por su parte el observador \mathcal{O} verá la misma masa en sus propias coordenadas en la posición $\vec{r} + \vec{V}t$, donde \vec{V} es la velocidad de m según \mathcal{O} . La relación entre las medidas de \mathcal{O} y \mathcal{O}' viene dado por

$$\begin{aligned} \vec{r} + \vec{V}t &= \vec{r}_0 + \vec{v}t + \vec{r}' + \vec{V}'t \\ &= \vec{r}_0 + \vec{r}' + (\vec{v} + \vec{V}')t. \end{aligned} \quad (2.28)$$

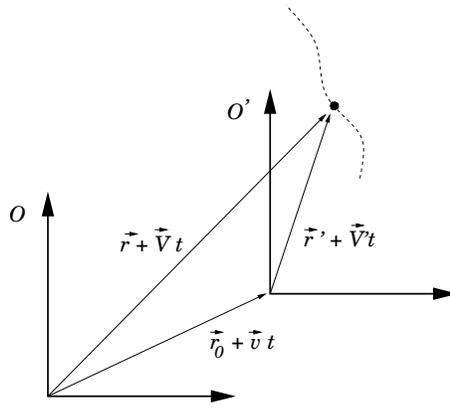


Figura 2.2: *Las transformaciones de Galilei: Un observador \mathcal{O}' se mueve con velocidad \vec{v} con respecto a un observador \mathcal{O} , tal que la posición de \mathcal{O}' en las coordenadas de \mathcal{O} es $\vec{r}_0 + \vec{v}t$. Una partícula, que tiene posición $\vec{r}' + \vec{V}'t$ en las coordenadas de \mathcal{O}' tendrá posición $\vec{r} + \vec{V}t = \vec{r}_0 + \vec{r}' + (\vec{v} + \vec{V}')t$.*

Estas transformaciones se llaman *las transformaciones de Galilei*, y muchas veces se presenta para el caso especial, cuando los dos sistemas de referencia coinciden a $t = 0$ y \mathcal{O}' se mueve a lo largo del eje x positivo ($\vec{r}_0 = 0$ y $\vec{v} = v\vec{e}_x$). En este caso tenemos:

$$x = x' + vt, \quad y = y', \quad z = z'. \quad (2.29)$$

Una conclusión que se saca inmediatamente de las transformaciones de Galilei es que la velocidad \vec{V} que \mathcal{O} mide para la masa m es la suma de la velocidad \vec{V}' que mide \mathcal{O}' y la velocidad \vec{v} de \mathcal{O}' relativa a \mathcal{O} :

$$\vec{V} = \vec{v} + \vec{V}'. \quad (2.30)$$

Esta regla se llama *la regla de la sumatoria de las velocidades (clásica)* y parece confirmar lo que conocemos intuitivamente de la vida cotidiana.

Como hemos visto, el Principio de la Relatividad dice que diferentes observadores inerciales ven la misma física. Así que las leyes de la física que estos observadores formulan tienen que ser invariantes bajo la transformación (2.29) (o en general (2.28)).

Las transformaciones de Galilei, junto con las rotaciones (2.24) y las traslaciones espaciales (2.23) y temporales (2.22) forman un grupo (ejerc.), llamado el *grupo de Galilei*. El grupo de Galilei juega un papel importante en la física clásica, puesto que es el grupo de simetrías de la mecánica newtoniana.

Para demostrar esto demostraremos la invariancia de la segunda ley de Newton bajo las distintas transformaciones de simetrías. Concretamente lo haremos para el caso en que las fuerzas son conservativas, es decir que son derivables de un potencial $\vec{F} = -\vec{\nabla}V$. Esto es el caso para todas las interacciones fundamentales ya que los efectos no-conservativos, como el rozamiento, son una descripción efectiva de interacciones conservativas complicadas a nivel molecular.

En componentes la segunda ley de Newton tiene la forma

$$-\partial^i V = m\ddot{x}^i. \quad (2.31)$$

A base de conservación de momento, se puede demostrar que la masa m de una partícula tiene el mismo valor para todos los observadores. En otras palabras, la masa es invariante bajo cambios de coordenadas. También hemos visto que los potenciales típicos de gravedad, electromagnetismo

y física molecular, atómico o nuclear son de la forma $V = V(|\vec{r}_1 - \vec{r}_2|)$, tal que estos potenciales son invariantes bajo rotaciones y traslaciones.⁶

En el lado derecho de (2.31) aparece la segunda derivada de la posición \ddot{x}^i , de forma que la ecuación es invariante bajo las transformaciones (2.22), (2.23) y (2.28). En particular, el hecho de que la segunda ley de Newton venga en función de la aceleración y no de la velocidad es una consecuencia directa del Principio de la Relatividad: las velocidades habrá que decir con respecto a qué sistema de referencia, mientras las aceleraciones son absolutas en la mecánica newtoniana.

La invariancia bajo rotaciones es un poco más sutil: bajo la transformación (2.24), la aceleración transforma como

$$\ddot{x}'^i = \sum_{j=1}^3 M^i_j \ddot{x}^j. \quad (2.32)$$

Aunque, por ser un potencial central, V es invariante bajo rotaciones, el gradiente sí transforma, debido a la transformación de las derivadas. En el Capítulo 1) hemos visto que

$$\frac{\partial V}{\partial x'_i} = \sum_{j=1}^3 M^i_j \frac{\partial V}{\partial x_j}. \quad (2.33)$$

Si para un observador es válida la segunda ley (2.31), otro observador relacionado con el primero a través de una rotación (2.24) verá la segunda ley como

$$-\frac{\partial V}{\partial x'_i} = m \ddot{x}'^i. \quad (2.34)$$

Los dos observadores obtendrán resultados diferentes para los componentes de la fuerza y la aceleración, pero cada uno medirá que la fuerza es proporcional a la aceleración, con la masa como constante de proporcionalidad. Si la ley de Newton es válido para un observador, también lo es para el otro. Si objetos (y leyes) transforman de esta manera decimos que no son invariantes sino *covariantes*.

La razón por qué esto funciona es que tanto la fuerza $\vec{\nabla}V$ como la aceleración $\ddot{\vec{x}}$ transforman de la misma manera bajo rotaciones. Ambas son vectores en \mathbb{R}^3 y transforman como tales bajo el grupo $O(3)$ de rotaciones en \mathbb{R}^3 . La lección que aprendemos de esto es que para que las leyes de la física sean invariantes o covariantes, hay que expresarlas en función de objetos que transformen bien bajo el grupo de simetría. Si no estas leyes sólo tendrían sentido en un sistema de referencia específico.

2.3. Mecánica de N partículas

En un sistema de N partículas, con posiciones \vec{r}_a y masas m_a ($a = 1, \dots, N$), es conveniente definir el *centro de masas* \vec{r}_{cm} como⁷

$$\vec{r}_{cm} = \frac{1}{M} \sum_{a=1}^N m_a \vec{r}_a, \quad (2.35)$$

donde $M = \sum_a m_a$ es la masa total del sistema. Se puede demostrar que el centro de masas está bien definida, es decir, que no depende del sistema de coordenadas elegido. Veremos que en muchos aspectos (aunque no todos), el sistema se comportará como una sola partícula con masa M situada en \vec{r}_{cm} .

⁶Ojo, el hecho de que V sea invariante bajo rotaciones no significa que $\vec{\nabla}V$ lo sea. Es más, en seguida demostraremos que $\vec{\nabla}V$ no es invariante, sino que transforma de manera covariante.

⁷Observa que no asumimos el convenio de sumación en a , aunque aparece repetido en el lado derecha de la ecuación. En principio el convenio de sumación sólo aplica a índices de vectores.

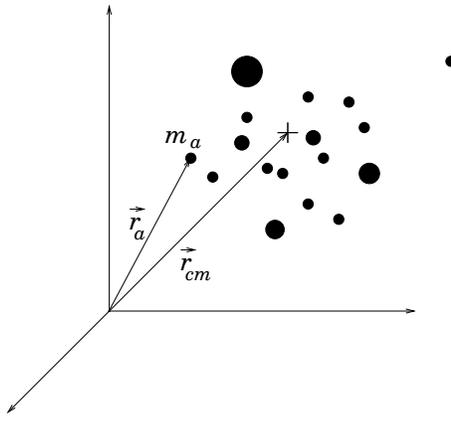


Figura 2.3: Un sistema de N partículas, el vector de posición \vec{r}_a de la a -ésima partícula y el vector de posición \vec{r}_{cm} del centro de masas.

Del mismo modo que para una sólo partícula, se define la velocidad, la aceleración y el momento del centro de masas como

$$\vec{v}_{cm} = \frac{1}{M} \sum_{a=1}^N m_a \frac{d\vec{r}_a}{dt} = \frac{d\vec{r}_{cm}}{dt}, \quad (2.36)$$

$$\vec{a}_{cm} = \frac{1}{M} \sum_{a=1}^N m_a \frac{d\vec{v}_a}{dt} = \frac{d\vec{v}_{cm}}{dt}, \quad (2.37)$$

$$\vec{p}_{cm} = \sum_{a=1}^N m_a \frac{d\vec{r}_a}{dt} = M \frac{d\vec{r}_{cm}}{dt}. \quad (2.38)$$

En cuanto a las fuerzas actuando sobre la a -ésima partícula, hay que distinguir entre *fuerzas internas*, ejercido por otra partícula del sistema, y *fuerzas externas*, proviniendo de una causa exterior al sistema. Por la segunda ley de Newton (2.3) sabemos que el total de fuerzas internas y fuerzas externas causan una aceleración de la a -ésima partícula, dada por

$$\sum_b \vec{F}_{ab}^{(int)} + \vec{F}_a^{(ext)} = m_a \vec{a}_a, \quad (2.39)$$

donde con $F_{ab}^{(int)}$ anotamos la fuerza de la b -ésima partícula del sistema sobre la a -ésima. En cuanto a las fuerzas internas, asumiremos la *tercera ley de Newton*, que la fuerza que la partícula a ejerce sobre b es igual de intensidad y opuesto en dirección a la fuerza ejercida por b sobre a . Sumando (2.39) sobre todas las partículas, obtenemos que

$$\sum_{b,a} \vec{F}_{ab}^{(int)} + \sum_a \vec{F}_a^{(ext)} = \sum_a m_a \vec{a}_a. \quad (2.40)$$

Por la tercera ley está claro que las fuerzas internas aparecen en el primer término de dos en dos, anulándose mutuamente, tal que, llamando la fuerza externa total actuando sobre el sistema $\vec{F}_{tot} = \sum_a \vec{F}_a^{(ext)}$, (2.40) se reduce a

$$\vec{F}_{tot} = M \vec{a}_{cm} = \frac{d\vec{p}_{cm}}{dt}. \quad (2.41)$$

En otras palabras, sean como sean las fuerzas internas y no importa como de complicado resultan los movimientos de las partículas individuales, el centro de masas se comporta como si fuera una

sola partícula de masa M , obedeciendo una “segunda ley de Newton”. Por lo tanto, si el total de las fuerzas externas es cero, el momento total se conserva.

De manera similar se define el momento angular y el momento de fuerza externas del sistema de N partículas como

$$\vec{L}_{tot} = \sum_{a=1}^N \vec{r}_a \times \vec{p}_a, \quad (2.42)$$

$$\vec{M}_{tot}^{(ext)} = \sum_{a=1}^N \vec{r}_a \times \vec{F}_a^{(ext)}. \quad (2.43)$$

Fijaos que en la definición (2.43) sólo entran las fuerzas externas, puesto que el momento de las fuerzas internas es cero. Efectivamente, el momento de las fuerzas internas sería

$$\begin{aligned} \vec{M}_{tot}^{(int)} &= \sum_a \vec{r}_a \times \left(\sum_b \vec{F}_{ab}^{(int)} \right) \\ &= \sum_{a,b} \vec{r}_a \times \vec{F}_{ab}^{(int)} \\ &= \frac{1}{2} \sum_{a,b} \left(\vec{r}_a \times \vec{F}_{ab}^{(int)} + \vec{r}_b \times \vec{F}_{ba}^{(int)} \right) \\ &= \frac{1}{2} \sum_{a,b} (\vec{r}_a - \vec{r}_b) \times \vec{F}_{ab}^{(int)} = 0. \end{aligned} \quad (2.44)$$

En la tercera igualdad hemos agrupados los términos de dos en dos, según las fuerzas entre cada par de partículas a y b . El principio de acción y reacción dice que estas fuerzas son iguales y de sentido contrario, tal que $\vec{F}_{ab}^{(int)} = -\vec{F}_{ba}^{(int)}$. El vector resultante $(\vec{r}_a - \vec{r}_b)$ es un vector que apunta de la partícula a a la partícula b y es por lo tanto paralelo a la fuerza interna $\vec{F}_{ab}^{(int)}$ entre estas dos partículas, lo que implica que el producto vectorial es cero.

También para sistemas de múltiples partículas tenemos que

$$\frac{d\vec{L}_{tot}}{dt} = \sum_{a=1}^N \vec{r}_a \times \frac{d\vec{p}_a}{dt} = \sum_{a=1}^N \vec{r}_a \times \vec{F}_a = \sum_{a=1}^N \vec{r}_a \times \left(\vec{F}_a^{(ext)} + \sum_b \vec{F}_{ab}^{(int)} \right) = \vec{M}_{tot}^{(ext)}, \quad (2.45)$$

tal que, si el momento de las fuerzas externas es cero, el momento angular total se conserva. Sin embargo, en función de las coordenadas del centro de masas, la expresión de \vec{L}_{tot} es un poco más complicado que solamente el momento angular del centro de masas. Si definimos \vec{r}_a y \vec{v}_a como la posición y la velocidad de la a -ésima partícula relativa al centro de masas

$$\vec{r}_a = \vec{r}_a - \vec{r}_{cm}, \quad \vec{v}_a = \vec{v}_a - \vec{v}_{cm}, \quad (2.46)$$

de la definición (2.42) del momento angular total tenemos que

$$\begin{aligned} \vec{L}_{tot} &= \sum_a m_a (\vec{r}_{cm} + \vec{r}_a) \times (\vec{v}_{cm} + \vec{v}_a) \\ &= \sum_a m_a \left((\vec{r}_{cm} \times \vec{v}_{cm}) + (\vec{r}_{cm} \times \vec{v}_a) + (\vec{r}_a \times \vec{v}_{cm}) + (\vec{r}_a \times \vec{v}_a) \right) \\ &= (\vec{r}_{cm} \times M \vec{v}_{cm}) + \sum_a m_a \left(\vec{r}_{cm} \times \frac{d\vec{r}_a}{dt} \right) + \sum_a m_a (\vec{r}_a \times \vec{v}_{cm}) + \sum_a (\vec{r}_a \times \vec{p}_a), \end{aligned} \quad (2.47)$$

donde hemos definido $\vec{p}_a = \vec{p}_a - \vec{p}_{cm}$ como el momento de la a -ésima partícula relativo al centro de masas. El segundo y el tercer término de (2.47) son cero, puesto que ambos contienen el factor

$\sum_a m_a \vec{r}_a$, que es el vector de posición del centro de masas relativo al centro de masas, o sea un vector nulo. El momento angular total se puede por lo tanto escribir como

$$\vec{L}_{tot} = (\vec{r}_{cm} \times M \vec{v}_{cm}) + \sum_a (\vec{r}_a \times \vec{p}_a) = \vec{L}_{cm} + \vec{L}_{int}. \quad (2.48)$$

Reconocemos en el primer término el momento angular del centro de masa y en el segundo el momento angular de las partículas individuales alrededor del centro de masas. El sistema no se comporta por lo tanto como una sola partícula con el momento angular concentrado en el centro de masas, sino también hay una contribución del momento angular interno de cada partícula.

Se define que el trabajo realizado por las fuerzas sobre un sistema de partículas como

$$W = \sum_a \int_x^y \vec{F}_a \cdot ds_a, \quad (2.49)$$

que, de modo similar al caso de una sola partícula se puede reescribir como (véase (2.10))

$$W = T(y) - T(x), \quad (2.50)$$

donde definimos la energía cinética del sistema como $T = \sum_a \frac{1}{2} m_a v_a^2$. Es fácil de demostrar que en terminos de la posición relativa al centro de masas, la energía cinética viene dado por (ejerc.) $T = \frac{1}{2} M v_{cm}^2 + \sum_a \frac{1}{2} m_a \tilde{v}_a^2$, es decir, la suma de la energía cinética del centro de masas y la energía cinética interna de cada partícula.

Si suponemos que tanto las fuerzas internas, como las externas son conservativas, podemos escribir (2.49) como (véase (2.14))

$$W = \sum_a \int_x^y \left(\vec{F}_a^{(ext)} + \sum_b \vec{F}_{ab}^{(int)} \right) \cdot ds_a = \sum_a \left(V_a(x) - V_a(y) \right) + \sum_{ab} \left(V_{ab}(x) - V_{ab}(y) \right). \quad (2.51)$$

A la cantidad V_{ab} se le llama la *energía potencial interna* y las identidades (2.50) y (2.51) indican que la energía mecánica total

$$E_{tot} = \frac{1}{2} M v_{cm}^2 + \sum_a \left(\frac{1}{2} m_a \tilde{v}_a^2 + V_a + \sum_b V_{ab} \right) \quad (2.52)$$

está conservada.

2.4. Cuerpos contínuos

Un cuerpo contínuo se puede considerar como una generalización de un sistema de N partículas individuales, en el límite donde $N \rightarrow \infty$. En lugar de hablar de una partícula individual con masa m_a , consideramos la masa dm contenida en un elemento de volumen dV situada en posición \vec{r} . La suma sobre todas las partículas se convierte entonces en un integral sobre el volumen del cuerpo. Por ejemplo, la masa del cuerpo está dado por

$$M = \int \rho(\vec{r}) dV, \quad (2.53)$$

con $\rho(\vec{r}) = dm(\vec{r})/dV$ la densidad del cuerpo en el punto \vec{r} . De la misma manera se define el centro de masas \vec{r}_{cm} , el momento \vec{p}_{cm} , el momento angular \vec{L}_{cm} y la energía cinética T_{cm} como

$$\vec{r}_{cm} = \frac{1}{M} \int \vec{r} dm = \frac{1}{M} \int \vec{r} \rho(\vec{r}) dV, \quad (2.54)$$

$$\vec{p}_{cm} = \int \vec{v}(\vec{r}) dm = \int \vec{v}(\vec{r}) \rho(\vec{r}) dV, \quad (2.55)$$

$$\vec{L}_{cm} = \int \vec{r} \times \vec{v}(\vec{r}) dm = \int \vec{r} \times \vec{v}(\vec{r}) \rho(\vec{r}) dV, \quad (2.56)$$

$$T_{cm} = \frac{1}{2} \int v^2(\vec{r}) dm = \frac{1}{2} \int v^2(\vec{r}) \rho(\vec{r}) dV. \quad (2.57)$$

Dejaremos al estudiante interesado derivar la dinámica de estos cuerpos continuos.

2.5. Problemas

1. Definimos el *tensor de inercia* I_{ij} de un sistema de N partículas como

$$I_{ij} = \sum_{a=1}^N m_a (r_a^2 \delta_{ij} - x_i^a x_j^a), \quad (2.58)$$

donde r_a es la norma del vector de posición \vec{r}_a de la a -ésima partícula y x_i^a son los componentes de \vec{r}_a . Considera ahora un sistema de tres partículas con masa m , donde la primera partícula está situado en el punto $(a, 0)$ del plano xy , la segunda en el punto $(-a, 0)$ y la tercera en el punto $(0, h)$ (con $a, h > 0$). Calcula primero el valor de I_{ij} y despues el valor en un sistema de referencia que está rotado con respecto al primero con un ángulo de 45 grados hacia la izquierda. Averigua algunas componentes, calculándolas explícitamente con la formula (2.58).

2. Demuestra que si un tensor T_{ij} es (anti-)simétrico en una base, lo es en cada base.
3. Demuestra que la delta de Kronecker (1.6) y el tensor de Levi-Civita (1.43) son tensores fundamentales, es decir, los componentes tienen el mismo valor en cada base.
4. Considera una base cartesiana en \mathbb{R}^3 , que consiste de los vectores $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$, mutuamente ortogonales y con norma 1. Sea $\vec{A} = A_i \vec{e}_i$ un vector arbitrario, demuestre que:

- a) $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$,
- b) $\vec{e}_i \times \vec{e}_j = \epsilon_{ijk} \vec{e}_k$,
- c) $\vec{A} \cdot \vec{e}_i = A_i$
- d) $\vec{A} \times \vec{A} = 0$

¿Cual es el significado de cada una de estas expresiones?

5. Definimos el operador vectorial *nabla* $\vec{\nabla}$ como $\vec{\nabla} = \vec{e}_i \partial_i$. Actuando sobre varios tipos de campos, se puede construir los siguientes operadores con el operador $\vec{\nabla}$:
 - el *gradiente* $\vec{\nabla} \phi = \partial_i \phi \vec{e}_i$ de un campo escalar ϕ ,
 - la *divergencia* $\vec{\nabla} \cdot \vec{A} = \partial_i A_i$ de un campo vectorial \vec{A} ,
 - el *rotacional* $\vec{\nabla} \times \vec{A} = \epsilon_{ijk} \partial_i A_j \vec{e}_k$ de un campo vectorial \vec{A} ,
 - el *laplaciano* $\Delta \phi = \partial_i \partial_i \phi$ de un campo escalar ϕ .

Demuestra los siguientes identidades vectoriales:

- a) $\vec{\nabla} \times \vec{\nabla} \phi = 0$
- b) $\vec{\nabla} \cdot \vec{\nabla} \times \vec{A} = 0$
- c) $\vec{\nabla} \cdot \vec{\nabla} \phi = \Delta \phi$,
- d) $\vec{\nabla} \cdot (\vec{A} \times \vec{B}) = \vec{B} \cdot \vec{\nabla} \times \vec{A} - \vec{A} \cdot \vec{\nabla} \times \vec{B}$,

e) $\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \Delta \vec{A}$, donde definimos el laplaciano actuando sobre un campo vectorial como actuando sobre cada componente: $\Delta \vec{A} = \Delta A_i \vec{e}_i$.

6. Varias partículas con masas m_a están moviendo libremente por el espacio y chocando elásticamente, conservando así la energía cinética y el momento. Instalamos ahora una lámpara y consideramos los movimientos de las sombras de las partículas, proyectadas sobre un plano. Atribuimos pues a cada sombra una masa que es proporcional a la masa de la partícula que causa la sombra. ¿Están todavía conservadas la energía cinética y el momento de las sombras en el plano?
7. Un vagón de tren abierto está rodando horizontalmente sin rozamiento por los carriles, mientras que llueve verticalmente. A lo largo del tiempo se va acumulando agua en el vagón, lo que aumenta considerablemente su masa. ¿Cómo afecta el agua acumulada la velocidad, el momento y la energía cinética del vagón?
Cuando deja de llover, se abre un desagüe y el agua acumulada desaparece poco a poco. ¿Cómo afecta la pérdida de agua a la velocidad, el momento y la energía cinética del vagón?
8. Escribe las ecuaciones de movimiento e resuélvelas para el caso de
 - a) el oscilador armónico
 - b) el caso de un campo gravitatorio constante.
9. Demuestra que la segunda ley de Newton es invariante bajo la transformación de Galilei:

$$x' = x + vt, \quad t' = t. \quad (2.59)$$

¿Cuál es el significado físico de esta invariancia?

10. Demuestra que en un campo de fuerza central el momento angular se conserva.
11. Explica con definición (2.7) la Ley de la palanca: el fuerza por la distancia a la fuerza es igual al peso por la distancia al peso (o según Arquímedes: dame una palanca y levantaré el mundo).
12. Calcula la altura que alcanza una piedra lanzada verticalmente con velocidad inicial \vec{v} . ¿Cómo variaría este resultado si hicieramos el experimento en la Luna en lugar de en la Tierra? ¿Con qué velocidad hay que lanzar la piedra para que no vuelva a caer a la tierra?

Capítulo 3

Formalismo lagrangiano

Hemos visto que la segunda ley de Newton (2.3), junto el principio de la relatividad, forma la base de la mecánica clásica. En la práctica, sin embargo, hay muchos casos donde no es fácil utilizar la segunda ley, puesto que algunas de las fuerzas no son conocidas. En este capítulo desarrollaremos un formalismo para resolver los problemas de la dinámica si las fuerzas desconocidas son del tipo que imponen ligaduras holónomas.

3.1. Ligaduras y coordenadas generalizadas

La segunda ley de Newton (2.3) relaciona las fuerzas que actúan sobre un sistema con los cambios de velocidad de ese sistema. Dado que un sistema de N partículas tiene en general la forma de un conjunto de $3N$ ecuaciones diferenciales parciales acopladas de segundo orden,

$$m_a \ddot{\vec{r}}_a = \vec{F}_a^{(tot)}, \quad (3.1)$$

el sistema queda completamente determinado con $6N$ constantes de integración, habitualmente tomadas las N posiciones iniciales y N velocidades iniciales. Esto es lo que llevó Laplace (1749-1827) a expresar su determinismo al decir que una vez conocida las posiciones y velocidades de todas las partículas en un momento $t = t_0$, toda la historia y todo el futuro del universo están determinados y son calculables.¹

Sin embargo hay muchos casos donde en la práctica es imposible aplicar la segunda ley para resolver un problema de dinámica, simplemente porque hay más variables que ecuaciones. Ilustremos esto con el ejemplo sencillo del péndulo matemático plano (véase Figura 3.1): una masa m cuelga de una cuerda inelástica y sin masa de longitud L en un campo gravitacional. La fuerza gravitatoria $\vec{F} = m\vec{g}$ no es la única fuerza actuando sobre la masa, puesto que la cuerda misma también ejerce una fuerza \vec{f} sobre m . (En realidad la fuerza \vec{f} de la cuerda es una fuerza efectiva que describe las interacciones de las partículas de la cuerda con la masa m , dado que el sistema de la masa y las partículas individuales de la cuerda es inmanejable). El problema ahora es que para determinar el movimiento de la masa m a través de la segunda ley, es necesario conocer también una expresión para \vec{f} , pero dado que \vec{f} es una fuerza efectiva que surge de las interacciones de la cuerda con la masa, no tenemos esa expresión.

El efecto de la fuerza desconocida \vec{f} es mantener la masa a distancia L del origen o . El movimiento de la masa por lo tanto está restringida en su movimiento debido a la fuerza \vec{f} . Decimos que la masa está sometida a una *ligadura* y a las fuerzas efectivas que restringen los movimientos de sistemas las llamamos *fuerzas de ligaduras*.

¹Y cuando Napoleón le preguntó dónde en su visión del mundo estaba Dios, contestó Laplace: “No necesito esa hipótesis.”

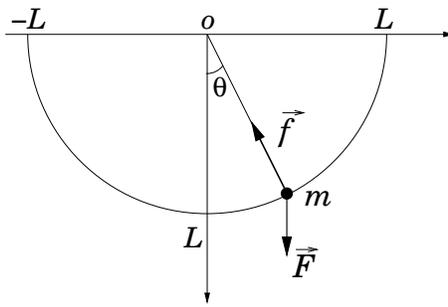


Figura 3.1: El péndulo matemático plano: una masa m cuelga de una cuerda con longitud L y está sometida a las fuerzas \vec{F} de la gravedad y \vec{f} de la tensión de la cuerda.

Hay varios tipos de ligaduras y varias maneras de clasificarlas. Una manera de distinguir las ligaduras es a base de su dependencia en el tiempo: las ligaduras que no dependen explícitamente de t se llaman *esclerónomas*, mientras las que sí dependen explícitamente se llaman *reónomas*. Ejemplos de un sistema con una ligadura esclerónoma son el péndulo plano mencionado arriba o un gas en el contenedor. Ejemplos de sistemas reónomas son una masa sobre una superficie que cambia con el tiempo o un gas en un contenedor deformable.

Otra manera de clasificar las ligaduras es según su expresión matemática. A las ligaduras que pueden escribirse como una ecuación en función de las coordenadas

$$S_n(\vec{r}_a, t) = 0 \quad (3.2)$$

se llaman *ligaduras holónomas*. Matemáticamente, la ligadura (3.2) define una superficie o una curva en el espacio \mathbb{R}^{3N} a la que las partículas están restringidas. Ejemplos de ligaduras holónomas son el péndulo plano mencionado arriba, una partícula que se mueve por una superficie o un cuerpo rígido. Ejemplos de sistemas no-holónomas son un gas en un contenedor, proyectiles que se caen a la tierra o una rueda que rueda sin resbalar. Las ligaduras en los dos primeros casos no-holónomas son de la forma $S_n(\vec{r}_a, t) \leq 0$, mientras el tercer ejemplo tiene una ligadura que conecta las coordenadas con las velocidades de una forma que no es integrable a una condición (3.2).

Los sistemas con ligaduras holónomas son mucho más fáciles de tratar, porque estas ligaduras imponen relaciones exactas entre las coordenadas que permiten escribir algunas en función de las otras. Esta propiedad hace que exista un sistema de coordenadas preferidas, las llamadas *coordenadas generalizadas*, y que en estas coordenadas las fuerzas efectivas se eliminen de la descripción, tal que el problema sea resoluble. En este curso nos restringiremos a sistemas con ligaduras holónomas.

El número de coordenadas generalizadas es igual al número de *grados de libertad*. Una partícula libre, sin ligaduras, tiene 3 grados de libertad, uno por cada dirección espacial en que se puede mover. Un sistema de N partículas sin ligaduras tiene $3N$ grados de libertad, pero si el sistema está sometida a k ligaduras holónomas, podemos utilizar estas ligaduras para eliminar k de los $3N$ grados de libertad. Un sistema de N partículas sometido a k ligaduras holónomas tiene $3N - k$ grados de libertad.

Existen por lo tanto $3N - k$ coordenadas generalizadas $q_\alpha(t)$ que están relacionadas con las coordenadas originales \vec{r}_a a través de

$$\vec{r}_a = \vec{r}_a(q_\alpha(t), t). \quad (3.3)$$

Una manera de pensar en esta relación consiste en verlo como un cambio de coordenadas \vec{r}_a a coordenadas q_α . Efectivamente el conjunto de las relaciones (3.3) y las ligaduras (3.2) es invertible.

En el caso del péndulo plano, tenemos dos ligaduras ($\sqrt{x^2 + y^2} = L$ y $z = 0$) y por lo tanto

sólo un grado de libertad. La coordenada generalizada por lo tanto es el ángulo θ que describe la desviación de la posición de equilibrio. Para una partícula en una esfera tenemos una ligadura ($\sqrt{x^2 + y^2 + z^2} = R$) y como coordenadas generalizadas podemos tomar los ángulos θ y ϕ que parametrizan la esfera. La relación (3.3) entre las coordenadas originales y las generalizadas es simplemente el cambio de coordenadas cartesianas a esféricas

$$x = R \sin \theta \cos \phi, \quad y = R \sin \theta \sin \phi, \quad z = R \cos \theta. \quad (3.4)$$

Claramente esta parametrización de x, y y z satisface la ligadura y las relaciones (3.4) junto con la ligadura son invertibles.

Entre las propiedades más importantes de las coordenadas generalizadas tenemos que, por un lado son independientes (por construcción, puesto que las ligaduras ya están tomado en cuenta) y por otro forman las coordenadas de una variedad diferenciable. La primera propiedad será muy útil a la hora de formular el problema en función de las coordenadas generalizadas y la segunda propiedad nos permite utilizar todas las herramientas conocidas de la geometría diferencial.

Con la herramienta de las coordenadas generalizadas podemos ahora reformular los problemas de la dinámica con ligaduras holónomas de tal forma que las fuerzas de ligaduras desaparecen de la descripción.

3.2. El principio de trabajo virtual y las ecuaciones de Lagrange

Una primera observación es que aunque no conocemos las fuerzas de ligadura cuantitativamente, sí podemos afirmar una propiedad importante: para las ligaduras holónomas, que restringen las partículas a una superficie (o una curva), las fuerzas de ligadura son perpendiculares a la superficie y por lo tanto proporcionales al gradiente de la función $S(\vec{r}_a, t)$:

$$\vec{f} = \sum_n \lambda \vec{\nabla} S_n(\vec{r}_a, t) = \sum_n \lambda \frac{\partial S_n(\vec{r}_a, t)}{\partial \vec{r}_a}. \quad (3.5)$$

Podemos definir un *desplazamiento virtual* $\delta \vec{r}_a$ como un desplazamiento infinitesimal desde un punto \vec{r}_a al punto $\vec{r}_a + \delta \vec{r}_a$ consistente con las fuerzas y las ligaduras presentes en el momento $t = t_0$. Fijaos que el desplazamiento virtual está definido por un momento fijo en el tiempo y difiere por lo tanto de lo que sería un desplazamiento actual en un intervalo dt . $\delta \vec{r}_a$ difiere de una derivada total por $d\vec{r}_a = \delta \vec{r}_a + \frac{\partial \vec{r}_a}{\partial t} dt$. Sólo para ligaduras esclerónomas (independientes de t), los dos desplazamientos coinciden.

Por definición tenemos, por lo tanto, que tanto en \vec{r}_a como en $\vec{r}_a + \delta \vec{r}_a$ las ecuaciones de las ligaduras se verifican

$$S_n(\vec{r}_a, t) = S_n(\vec{r}_a + \delta \vec{r}_a, t) = 0, \quad (3.6)$$

tal que, expandiendo la diferencia en una serie de Taylor y sumando sobre todas las ligaduras y todas las partículas, tenemos

$$\sum_{n,a} \frac{\partial S_n}{\partial \vec{r}_a} \cdot \delta \vec{r}_a = 0. \quad (3.7)$$

Comparando con la forma de las fuerzas de ligaduras (3.5), concluimos que estas fuerzas no realizan trabajo en un desplazamiento virtual

$$W_f = \int \vec{f} \cdot \delta \vec{r} = 0. \quad (3.8)$$

Este principio se llama el *principio del trabajo virtual* o el *principio de D'Alembert*. En lenguaje geométrico simplemente quiere decir que las fuerzas de ligadura y los desplazamientos virtuales

son ortogonales. El significado físico del principio del trabajo virtual consiste en que las únicas fuerzas que realizan trabajo sobre el sistema, son las fuerzas actuales, no las virtuales.

Considera ahora un sistema de N partículas, obedeciendo cada una a la segunda ley de Newton $\vec{F}_a^{(tot)} = \dot{\vec{p}}_a$.² Escribiendo la fuerza total actuando sobre la a -ésima partícula como la suma de las fuerzas actuales y las fuerzas de ligadura $\vec{F}_a^{(tot)} = \vec{F}_a + \vec{f}_a$, tenemos que

$$0 = \sum_a (\vec{F}_a + \vec{f}_a - \dot{\vec{p}}_a) \cdot \delta \vec{r}_a = \sum_a (\vec{F}_a - \dot{\vec{p}}_a) \cdot \delta \vec{r}_a, \quad (3.9)$$

donde en la última igualdad hemos utilizado el principio del trabajo virtual. Aunque de esta manera hemos podido eliminar las fuerzas de ligadura, no podemos igualar a cero los coeficientes $(\vec{F}_a - \dot{\vec{p}}_a)$, puesto que las $\delta \vec{r}_a$ no son independientes, sino que están relacionadas mutuamente por las ligaduras. Conviene por lo tanto reescribir (3.9) en función de las coordenadas generalizadas, que por construcción sí son independientes.

Derivando la relación (3.3) entre las coordenadas \vec{r}_a y las coordenadas q_α con respecto a t tenemos que (utilizando el convenio de sumación para las coordenadas generalizadas)

$$\vec{v}_a = \frac{d\vec{r}_a}{dt} = \frac{\partial \vec{r}_a}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \vec{r}_a}{\partial t}, \quad (3.10)$$

y del mismo modo, los desplazamientos virtuales se escriben como

$$\delta \vec{r}_a = \frac{\partial \vec{r}_a}{\partial q_\alpha} \delta q_\alpha. \quad (3.11)$$

(Fijaos que como los desplazamientos virtuales están tomados en $t = t_0$ no hay variación en t .) La ecuación (3.9) se puede por lo tanto reescribir como

$$\sum_a \vec{F}_a \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} \delta q_\alpha - \sum_a m_a \ddot{\vec{r}}_a \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} \delta q_\alpha = 0. \quad (3.12)$$

El coeficiente del último término se puede escribir como

$$\begin{aligned} m_a \ddot{\vec{r}}_a \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} &= \frac{d}{dt} \left(m_a \dot{\vec{r}}_a \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} \right) - m_a \dot{\vec{r}}_a \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_a}{\partial q_\alpha} \right) \\ &= \frac{d}{dt} \left(m_a \vec{v}_a \cdot \frac{\partial \vec{v}_a}{\partial \dot{q}_\alpha} \right) - m_a \vec{v}_a \cdot \left(\frac{\partial \vec{v}_a}{\partial q_\alpha} \right) \\ &= \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_\alpha} \left(\frac{1}{2} m_a v_a^2 \right) \right) - \frac{\partial}{\partial q_\alpha} \left(\frac{1}{2} m_a v_a^2 \right), \end{aligned} \quad (3.13)$$

donde en la segunda igualdad hemos utilizado que la derivada total y la derivada parcial con respecto a q_α son intercambiables y, por (3.10) que $\partial v_a / \partial \dot{q}_\alpha = \partial r_a / \partial q_\alpha$.

La cantidad $(\frac{1}{2} m_a v_a^2)$ en (3.13) es la energía cinética T_a de la a -ésima partícula, definido en (2.10). Si definimos la fuerza generalizada $Q_\alpha = \sum_a \vec{F}_a \cdot \partial \vec{r}_a / \partial q_\alpha$ veremos que (3.12) se convierte en

$$\left\{ Q_\alpha - \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) + \frac{\partial T}{\partial q_\alpha} \right\} \delta q_\alpha = 0, \quad (3.14)$$

donde T es la energía cinética total de las N partículas. Fijaos que en el caso de coordenadas cartesianas la energía cinética no depende de las coordenadas, sólo de las velocidades y el último término por lo tanto es cero en estas coordenadas. Sin embargo, en general ese término es no nulo y tiene que ver con la conexión no trivial de las coordenadas curvilineas.

²A partir de ahora adoptaremos la notación en que indicamos las derivadas (totales) con respecto a t con un punto: $\frac{d\Omega}{dt} = \dot{\Omega}$ y $\frac{d^2\Omega}{dt^2} = \ddot{\Omega}$.

Hasta ahora no hemos hecho más que un cambio de coordenadas de \vec{r}_a a q_α , utilizando el principio de trabajo virtual para eliminar las fuerzas de ligadura. Pero si ahora suponemos que las ligaduras son holónomas, sabemos que las q_α son coordenadas generalizadas y por lo tanto mutuamente independientes. En ese caso (3.14) tiene que ser válido para cada δq_α por separado, y así para ligaduras holónomas tenemos que

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} - Q_\alpha = 0. \quad (3.15)$$

Si además suponemos que las fuerzas actuales \vec{F}_a son conservativas (y por lo tanto derivables de un potencial), podemos escribir Q_α como

$$Q_\alpha = \sum_a \vec{F}_a \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} = \sum_a -\frac{\partial V}{\partial \vec{r}_a} \cdot \frac{\partial \vec{r}_a}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}, \quad (3.16)$$

tal que (3.15) se convierte en

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial (T - V)}{\partial q_\alpha} = 0, \quad (3.17)$$

o, para potenciales que son funciones sólo de las coordenadas generalizadas, no de las velocidades,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) - \frac{\partial L}{\partial q_\alpha} = 0. \quad (3.18)$$

Aquí hemos introducido la función $L = T - V$, llamada la *función de Lagrange* o simplemente el *Lagrangiano* y las ecuaciones (3.18) se llaman las *ecuaciones de Lagrange* o de *Euler-Lagrange*.

Con las ecuaciones de Lagrange hemos conseguido un formalismo para describir el comportamiento de un sistema sin que aparezcan las fuerzas de ligaduras. Es más, para la gran clase de problemas donde las fuerzas son conservativas, incluso el concepto de fuerzas desaparece del todo del formalismo. Sólo involucra a las funciones escalares T y V y es por lo tanto independiente de las coordenadas elegidas.

Fijaos que para un sistema dado, el lagrangiano no es único. Ya hemos visto que V esta determinado salvo una constante global, pero la degeneración va más lejos. Cada función L' relacionada con un lagrangiano L a través de la relación $L' = L + dG/dt$ con $G(q_\alpha, t)$ una función arbitraria, nos dará las mismas ecuaciones de movimiento. En otras palabras, dos lagrangianos que difieren por una derivada total son físicamente equivalentes y los consideraremos por lo tanto el mismo lagrangiano. Demostraremos esta propiedad en la siguiente sección.

La gran ventaja del formalismo lagrangiano es que, no sólo desaparecen las desconocidas fuerzas de ligaduras, sino que también es fácilmente generalizable a otros campos de la física, donde el concepto de fuerzas no tiene el sentido que tiene en la mecánica newtoniana, como en la mecánica de fluidos, teoría de la relatividad o la teoría de campos.

3.3. El principio de mínima acción

Antes de seguir con la discusión de las propiedades del lagrangiano, es útil comentar que hay otro método más fácil de conseguir las ecuaciones de Lagrange que a través del principio de trabajo virtual. Este método es el *principio de mínima acción*, propuesto por Maupertuis (1698 - 1759) y desarrollado por Euler, utilizando el cálculo variacional.

Según el principio de mínima acción, el movimiento de un sistema yendo de la posición $q_\alpha(t_1)$ en $t = t_1$ a $q_\alpha(t_2)$ en $t = t_2$ es tal que la integral S definida como

$$S(q_\alpha(t), \dot{q}_\alpha(t), t) = \int_{t_1}^{t_2} L(q_\alpha(t), \dot{q}_\alpha(t), t) dt, \quad (3.19)$$

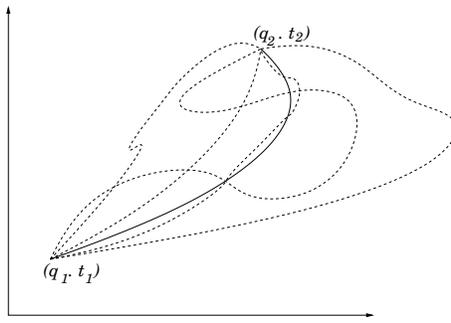


Figura 3.2: La trayectoria real (línea continua) y algunas trayectorias posibles (líneas discontinuas) de una partícula moviéndose de $q_\alpha(t_1)$ a $q_\alpha(t_2)$. La trayectoria real corresponde a la que tiene el valor mínimo para S .

llamada la *acción*, es mínima para la trayectoria $q_\alpha(t)$ seguida. Aquí $L = T - V$ es el lagrangiano del sistema, definido en la sección anterior.

Si la acción es mínima para la trayectoria real, cualquier trayectoria con las mismas condiciones iniciales y finales tiene que tener un valor superior para S y la trayectoria real corresponde a un valor estacionario de S . En otras palabras, la variación de la acción, debido a la variación de las trayectorias $q_\alpha(t) \rightarrow q_\alpha(t) + \delta q_\alpha(t)$, manteniendo $\delta q_\alpha(t_1) = \delta q_\alpha(t_2) = 0$, es cero. Demostraremos ahora que esto implica que las coordenadas $q_\alpha(t)$ satisfacen las ecuaciones de Lagrange (3.18).

La variación de la acción (3.19) viene dada por

$$\begin{aligned}
 \delta S &= \delta \int_{t_1}^{t_2} L(q_\alpha(t), \dot{q}_\alpha(t), t) dt \\
 &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) dt \\
 &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \right) \delta q_\alpha dt + \left[\frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \right]_{t_1}^{t_2}, \tag{3.20}
 \end{aligned}$$

donde en la última igualdad hemos utilizado que $\delta \dot{q}_\alpha = d(\delta q_\alpha)/dt$ y hemos integrado por partes. Tomando los puntos $q_\alpha(t_1)$ y $q_\alpha(t_2)$ fijos, está claro que el segundo término de (3.20) es cero.

El hecho de que la variación δS tiene que ser cero para cualquier variación δq_α arbitraria, implica por un lema conocido del cálculo variacional, que la función multiplicando δq_α es cero:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) - \frac{\partial L}{\partial q_\alpha} = 0, \tag{3.21}$$

lo que es justo la ecuación de Lagrange (3.18). Ahora también está claro lo que ya anunciamos en la sección anterior. Dos lagrangianos que difieren por una derivada total implican las mismas ecuaciones de movimiento, puesto que la variación de la derivada total es cero por las condiciones iniciales y finales.

Estrictamente hablando, las ecuaciones de Lagrange (3.21) no dan las condiciones necesarias para que la acción (3.19) sea mínima, sino extrema. En principio una solución de (3.21) podría corresponder con un máximo de S y habrá que averiguar la estabilidad de la solución. En la práctica estaremos contentos si logramos obtener una solución exacta para un sistema dado.

El principio de mínima acción da cierto rigor matemático a las ecuaciones de la mecánica newtoniana. Asocia un valor S a cada curva entre las posiciones iniciales y finales y selecciona la trayectoria real como la curva que tiene el valor mínimo de S .

Visto desde el punto de vista físico, la función $S = S(q_\alpha(t), \dot{q}_\alpha(t), t)$ es una manera compacta de resumir las ecuaciones de movimiento de un sistema, que vienen dadas por las ecuaciones de Lagrange. Muchas veces, en teoría de campos o en relatividad uno va tan lejos que identifica la acción (o el lagrangiano) con la teoría que utiliza, dado que S contiene codificada toda la dinámica del sistema.

3.4. Ejemplos

Antes de discutir en detalle el significado y las propiedades del lagrangiano y las ecuaciones de Lagrange, daremos primero unos ejemplos de sistemas sencillos. El objetivo es básicamente demostrar que el formalismo lagrangiano y equivalente a la mecánica newtoniana del capítulo 1, pero es más potente en la hora de tratar con sistemas con ligaduras.

- **Una partícula sin ligaduras en un potencial:** Un partícula sin ligaduras corresponde al caso más sencillo para analizar con el formalismo de Lagrange. En coordenadas cartesianas, el lagrangiano de una partícula con masa m en un potencial $V(x)$ viene dado por

$$L = \frac{1}{2}m\dot{x}_i^2 - V(x). \quad (3.22)$$

Dado que no hay ligaduras, podemos tomar las mismas coordenadas cartesianas como coordenadas generalizadas. Las ecuaciones de movimiento son por lo tanto

$$m\ddot{x}_i = -\frac{\partial V(x)}{\partial x_i}, \quad (3.23)$$

lo que efectivamente reconocemos como la segunda ley de Newton en el caso de fuerzas conservativas.

En coordenadas esféricas el lagrangiano tiene la forma

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2) - V(r, \theta, \varphi), \quad (3.24)$$

de forma que las ecuaciones de movimiento son

$$\begin{aligned} m\ddot{r} &= mr\dot{\theta}^2 + mr \sin^2 \theta \dot{\varphi}^2 - \frac{\partial V}{\partial r}, \\ \frac{d}{dt}(mr^2\dot{\theta}) &= \frac{1}{2}mr^2 \sin \theta \cos \theta \dot{\varphi}^2 - \frac{\partial V}{\partial \theta}, \\ \frac{d}{dt}(mr^2 \sin^2 \theta \dot{\varphi}) &= -\frac{\partial V}{\partial \varphi}. \end{aligned} \quad (3.25)$$

Estas ecuaciones son en general difíciles de resolver y no lo haremos ahora. Sólo haremos la observación de que las ecuaciones simplifican considerablemente si el potencial V es una función sólo de r , como por ejemplo en el caso de los planetas en el campo gravitatorio del sol o en problemas de choque. Estudiaremos estos casos en más detalle en el capítulo 5.

- **Péndulo matemático plano:** En la sección 3.1 hemos dado el ejemplo del péndulo plano matemático como sistema que es difícil de resolver a través de la segunda ley de Newton, por la presencia de fuerzas de ligadura, en ese caso la tensión de la cuerda. Demostraremos ahora que se pueden obtener y resolver las ecuaciones mucho más fácilmente en el formalismo lagrangiano.

El lagrangiano de una masa m tendido de una cuerda inelástica y sin masa con longitud L (vease Figura 3.1) es

$$L = \frac{1}{2}mL^2\dot{\theta}^2 + mgL \cos \theta, \quad (3.26)$$

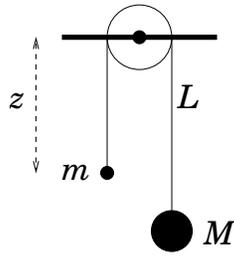


Figura 3.3: La máquina de Atwood: dos masas m y M atadas por una cuerda de longitud L cuelgan de una polea.

donde θ es el ángulo de desvío de la posición de equilibrio y g es la aceleración gravitacional que hemos encontrado en el capítulo 1. La ecuación de Lagrange para este sistema es

$$L\ddot{\theta} = -g \sin \theta. \quad (3.27)$$

Una primera observación es que la masa m no aparece en las ecuación del péndulo. El periodo y la amplitud del péndulo son por lo tanto independientes de la masa colgada.

Una segunda observación es que la ecuación (3.27) es muy difícil de resolver. De hecho, sólo se puede buscar una solución numérica. Sin embargo, es posible hacer un aproximación para ángulos pequeños: para $\theta \ll 1$ tenemos que $\sin \theta \approx \theta$ y podemos escribir (3.27) como

$$\ddot{\theta} = -\frac{g}{L}\theta, \quad (3.28)$$

lo que es la ecuación del oscilador armónico, con solución

$$\theta(t) = A \sin(\omega t + \theta_0), \quad (3.29)$$

donde la amplitud A y la fase θ_0 son constantes de integración, determinadas por las condiciones iniciales. $\omega = \sqrt{g/L}$ es la frecuencia, tal que el periodo de oscilación $T = 2\pi\sqrt{L/g}$ es proporcional a la raíz de la longitud de la cuerda.

Fijaos por último que el periodo T (ó equivalentemente la frecuencia ω) no depende de la amplitud θ . Este hecho es conocido como la *Ley del péndulo* de Galilei, quien lo descubrió empíricamente, según la leyenda midiendo el periodo de oscilación de un candelabro en la catedral de Pisa, con su propio pulso como cronómetro. En realidad la ley del péndulo sólo es válida en la aproximación de amplitudes pequeñas (3.28). Un análisis numérico indica que en general el periodo aumenta con la amplitud.

- **La máquina de Atwood:** La máquina de Atwood consiste de dos masas m y M conectadas a través de una cuerda inelástica y sin masa de longitud L que pasa por una polea. Las masas sólo están permitidas de moverse verticalmente (vease Figura 3.3).

El sistema claramente tiene sólo un grado de libertad (6 coordenadas y 5 ligaduras) y tomaremos como coordenadas generalizada la altura z de la masa m por debajo de la polea. El lagrangiano entonces es de la forma

$$L = \frac{1}{2}m\dot{z}^2 + \frac{1}{2}M\dot{z}^2 + mgz + Mg(L - z), \quad (3.30)$$

tal que la ecuación de Lagrange es

$$\ddot{z} = \frac{m - M}{M + m}g, \quad (3.31)$$

con la solución

$$z(t) = \frac{1}{2} \frac{m - M}{M + m}gt^2 + v_0t + z_0. \quad (3.32)$$

Las masas van por lo tanto acelerando en la dirección de la masa más grande. La posición y la velocidad inicial z_0 y v_0 son constante de integración. Fijaos que, igual que en el caso del péndulo, la tensión de la cuerda no aparece en el problema.

Con estos tres ejemplos sencillos hemos dado un primer indicio de la fuerza del formalismo lagrangiano, en comparación con la segunda ley de Newton. Veremos en la sección de ejercicios que también es capaz de resolver problemas más complicados.

3.5. Interpretación y propiedades del Lagrangiano

Miremos ahora la estructura de las ecuaciones de Lagrange (3.18). Ya hemos visto en el ejemplo de la partícula sin ligaduras en coordenadas cartesianas que recuperamos la segunda ley de Newton:

$$\dot{p}_i = -\frac{\partial V}{\partial x_i}, \quad (3.33)$$

donde p_i es la componente i del vector de momento \vec{p} . En general, para cualquier tipo de coordenadas generalizadas podemos definir el *momento generalizado* o *momento conjugado* p_α como³

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}. \quad (3.34)$$

En (3.16) ya hemos interpretado la variación de V respecto a q_α como la fuerza generalizada Q_α , así que la ecuación de Lagrange (3.18) se puede escribir de la forma

$$\dot{p}_\alpha - \frac{\partial T}{\partial q_\alpha} = -\frac{\partial V}{\partial q_\alpha}. \quad (3.35)$$

La generalización de (3.33) es obvia, sólo queda la interpretación del segundo término de la izquierda. Antes ya mencionamos su relación con la conexión de las coordenadas curvilíneas. Efectivamente, para el caso de ligaduras esclerónomas, la energía cinética T en coordenadas generalizadas es de la forma

$$T = \frac{1}{2} \sum_a m_a \dot{\vec{r}}_a^2 = \frac{1}{2} \sum_a m_a \frac{\partial \vec{r}_a}{\partial q_\alpha} \dot{q}_\alpha \cdot \frac{\partial \vec{r}_a}{\partial q_\beta} \dot{q}_\beta = \frac{1}{2} a_{\alpha\beta} \dot{q}_\alpha \dot{q}_\beta, \quad (3.36)$$

con

$$a_{\alpha\beta} = \sum_a m_a \frac{\partial \vec{r}_a}{\partial q_\alpha} \cdot \frac{\partial \vec{r}_a}{\partial q_\beta} \quad (3.37)$$

jugando el papel de una métrica en el espacio definido por los valores posibles de las coordenadas generalizadas. Efectivamente analizando en más detalle, se puede escribir (3.35) como

$$\ddot{q}_\alpha + \Gamma_{\beta\gamma}^\alpha \dot{q}_\beta \dot{q}_\gamma = -\frac{\partial V}{\partial q_\alpha}, \quad (3.38)$$

donde los $\Gamma_{\beta\gamma}^\alpha$ son los símbolos de Christoffel (la conexión) de la métrica que describe la superficie de los posibles valores de los q_α . En ausencia de fuerzas (V constante), reconocemos la ecuación de las geodésicas del curso de geometría diferencial, así que (3.38) es la ecuación de la trayectoria de una partícula sometida a fuerzas externas en un espacio curvo, donde la curvatura está causada por las ligaduras esclerónomas.

De la forma del lagrangiano también podemos deducir que en ciertas condiciones hay cantidades que se conservan a lo largo del movimiento. Ya hemos hablado en el capítulo 1 sobre la conservación de momento, energía y momento angular, pero el formalismo lagrangiano nos permite deducirlo

³Fijaos que los p_α no necesariamente forman un vector, puesto que los q_α tampoco necesariamente lo son.

con más rigor. También veremos que estas cantidades conservadas tienen una relación directa con las simetrías del espacio y el tiempo.

Una primera simetría tiene que ver con la invariancia bajo traslaciones en el tiempo. La derivada total del lagrangiano con respecto a t es

$$\frac{dL}{dt} = \frac{\partial L}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \ddot{q}_\alpha + \frac{\partial L}{\partial t} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \dot{q}_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \ddot{q}_\alpha + \frac{\partial L}{\partial t} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \dot{q}_\alpha \right) + \frac{\partial L}{\partial t}, \quad (3.39)$$

o en otras palabras, hay una cantidad

$$E = p_\alpha \dot{q}_\alpha - L, \quad (3.40)$$

cuya derivada total está directamente relacionada con la dependencia de L en t :

$$\frac{dE}{dt} = - \frac{\partial L}{\partial t}. \quad (3.41)$$

En general, la dependencia del lagrangiano de t suele venir de ligaduras reónomas o de variaciones del potencial V con el tiempo, pero en los casos más interesantes para la física fundamental, el lagrangiano no depende explícitamente de t , sino sólo a través de q_α y \dot{q}_α . Esto corresponde con la homogeneidad del tiempo (las leyes de la física no varían de momento en momento) y por lo tanto (3.41) dice que en estos casos la cantidad E es constante.

La interpretación de la función E no siempre está clara, pero bajo ciertas circunstancias (bastante comunes), resulta que E es la energía mecánica del sistema. En el caso de que la relación (3.3) entre las coordenadas generalizadas y las coordenadas originales no dependen de t , la energía cinética es una función que viene dada por la expresión (3.36) y por lo tanto

$$\frac{\partial T}{\partial \dot{q}_\alpha} \dot{q}_\alpha = 2T \quad (3.42)$$

(por el teorema de Euler para funciones homogéneas, o por directamente calcularlo). Si además suponemos que el potencial V sólo depende de las coordenadas q_α , y no de las velocidades \dot{q}_α , tenemos que

$$E = \frac{\partial T}{\partial \dot{q}_\alpha} \dot{q}_\alpha - (T - V) = T + V. \quad (3.43)$$

La función E , dada en (3.40) es (por lo menos en cuanto forma) idéntica al Hamiltoniano H que desarrollaremos en el capítulo 4, a través de una transformación de Legendre.

Pasemos ahora a la homogeneidad del espacio. Si un lagrangiano no depende de una coordenada q_α , aunque puede depender de la velocidad \dot{q}_α , decimos que esta coordenada es *cíclica*. En este caso tenemos inmediatamente por (3.18) que

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0, \quad (3.44)$$

o sea, que el momento conjugado p_α en la dirección q_α está conservado. En particular, para una partícula libre sin ligaduras, el potencial es cero y el lagrangiano depende sólo de las velocidades. En este caso las coordenadas generalizadas son las coordenadas cartesianas y el momento conservado es el momento mecánico total \vec{p} , como ya sabíamos por la segunda ley de Newton.

Sin embargo el momento conjugado p_α no siempre tiene que coincidir con el momento mecánico $p_i = mv_i$. Un ejemplo típico es el caso de una partícula con carga e en un campo electromagnético, especificado por los potenciales electromagnéticos ϕ y \vec{A} . El lagrangiano en este caso está dado por

$$L = \frac{1}{2} m \dot{x}_i^2 - e\phi(x) + e \dot{x}_i A_i(x). \quad (3.45)$$

Está claro que en este caso por la definición (3.34), el momento conjugado p_i viene dado por

$$p_i = m\dot{x}_i + eA_i(x_i). \quad (3.46)$$

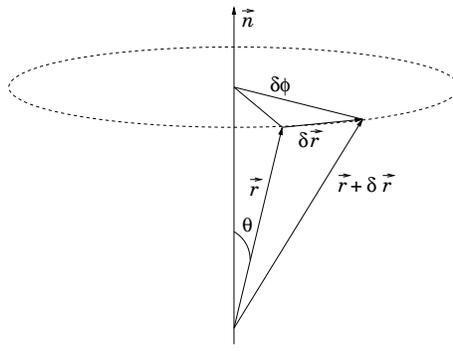


Figura 3.4: El cambio de un vector \vec{r} bajo una rotación infinitesimal $\delta\phi$.

Efectivamente, en el caso de que los potenciales electromagnéticos no dependan de alguna dirección x_k , no es el momento mecánico $m\dot{x}_i$ el que se conserva, sino la combinación (3.46)

La conservación de momento conjugado está por lo tanto relacionada con la simetría de traslaciones del lagrangiano. En general en presencia de un potencial, el momento no se conserva, puesto que el potencial rompe la invariancia bajo traslaciones.

Finalmente miramos el caso de la invariancia bajo rotaciones, debido a la isotropía del espacio. Si no hay una dirección preferida en el espacio, no puede haber dependencia explícita en las coordenadas angulares. Si ahora actuamos con una rotación infinitesimal $\delta\phi$ alrededor de un eje, tenemos que el vector de posición \vec{r} está transformado en el vector $\vec{r} + \delta\vec{r}$, donde la norma de $|\delta\vec{r}|$ está dada por $|\delta\vec{r}| = r \sin\theta \delta\phi$, y θ es el ángulo entre el vector \vec{r} y el eje de rotación (véase Figura 3.4). Dada la forma de la norma y el hecho de que $\delta\vec{r}$ es ortogonal tanto a \vec{r} y el eje de rotación, podemos escribir

$$\delta\vec{r} = \delta\phi (\vec{n} \times \vec{r}), \quad (3.47)$$

donde \vec{n} es un vector de unidad paralelo al eje de rotación. De (3.47) derivamos directamente que bajo la rotación infinitesimal $\delta\phi$, la velocidad se transforma como

$$\delta\vec{v} = \delta\phi (\vec{n} \times \vec{v}). \quad (3.48)$$

El cambio infinitesimal δL de lagrangiano, producida por la rotación $\delta\phi$ es por lo tanto

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial \vec{r}} \cdot \delta\vec{r} + \frac{\partial L}{\partial \dot{\vec{r}}} \cdot \delta\dot{\vec{r}} \\ &= \frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{r}}} \cdot (\delta\phi \vec{n} \times \vec{r}) + \frac{\partial L}{\partial \dot{\vec{r}}} \cdot (\delta\phi \vec{n} \times \vec{v}) \\ &= \vec{p} \cdot (\delta\phi \vec{n} \times \vec{r}) + \vec{p} \cdot (\delta\phi \vec{n} \times \vec{v}) \\ &= \delta\phi \vec{n} \cdot \left(\vec{r} \times \vec{p} + \dot{\vec{r}} \times \vec{p} \right) \\ &= \delta\phi \vec{n} \cdot \frac{d}{dt} (\vec{r} \times \vec{p}), \end{aligned} \quad (3.49)$$

donde en la cuarta igualdad hemos utilizado la propiedad cíclica (??) del producto mezclado. Si el lagrangiano es invariante bajo rotaciones, la variación δL tiene que ser cero para cualquier rotación $\delta\phi$ arbitraria, tal que de (3.49) concluimos que la cantidad $\vec{L} = \vec{r} \times \vec{p}$ se conserva. Reconocemos claramente el momento angular, definido en (2.6).

En esta sección hemos visto que la conservación de energía, momento y momento angular están relacionados con la invariancia del lagrangiano bajo respectivamente traslaciones en el tiempo, en el espacio y bajo rotaciones. En el fondo, hemos demostrado tres casos particulares del *teorema*

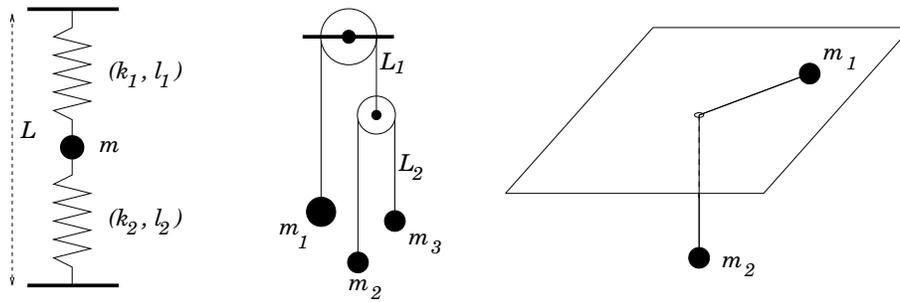


Figura 3.5: El problema de los muelles (ejerc 4), la máquina de Atwood doble (ejerc 6) y el problema de las masas en la mesa (ejerc 9).

de Noether que dice que por cada simetría continua global de un sistema existe una cantidad (carga) conservada y vice versa, cada cantidad conservada está relacionada con una simetría global continua del sistema. Otro ejemplo de una carga de Noether es la carga eléctrica, debido al hecho de que la teoría de Maxwell es invariante bajo transformaciones gauge actuando en los potenciales electromagnéticos.

3.6. Problemas

1. Calcula por sustitución explícita que el lagrangiano $L' = L + dF/dt$ tiene las mismas ecuaciones de movimiento que L , para cualquier función $F = F(q_\alpha, t)$ diferenciable.
2. Demuestra que la ecuación de Lagrange para una partícula libre sin ligaduras es igual que la ecuación de una geodésica ($\ddot{q}_\alpha + \Gamma_{\beta\gamma}^\alpha \dot{q}_\beta \dot{q}_\gamma = 0$) en el espacio plano para el caso de coordenadas cilíndricas y coordenadas esféricas.
3. Considera una partícula restringida a moverse sobre la superficie de una esfera de radio R en un campo gravitacional constante. Escribe las ecuaciones de movimiento y demuestra que las orbitas con θ (el ángulo azimut) constante sólo son posibles en el hemisferio sur de la esfera. Comenta como varía el ángulo de equilibrio θ_0 en función del momento angular.
4. Una masa puntual está colgada de dos muelles entre dos paredes de distancia L (vease Figura 3.5). Los muelles tienen tamaño l_1 y l_2 y constantes de muelle k_1 y k_2 respectivamente. Calcula el movimiento de la masa si sólo los desplazamientos verticales están permitidos.
5. Una masa se mueve a lo largo de una hélice dado por las ecuaciones $\rho = az$ y $\varphi = -bz$ en un campo gravitatorio. Demuestra que la ecuación de movimiento está dada por

$$(a^2 b^2 z^2 + a^2 + 1)\ddot{z} + a^2 b^2 z \dot{z}^2 + g = 0. \quad (3.50)$$

6. Calcula las ecuaciones de movimiento de tres masas sujetadas en una máquina de Atwood doble (vease Figura 3.5). Demuestra que para $m_2 = m_3$ el problema se reduce a la máquina de Atwood simple. ¿Cuál es el significado físico de esta observación? Resuelve el caso general donde las tres masas son distintas.
7. Da un ejemplo de un sistema donde la cantidad E definida en (3.40)
 - es la energía mecánica y está conservada.
 - no es la energía mecánica pero sí está conservada.
 - sí es la energía mecánica pero no está conservada.

- no es la energía mecánica y no está conservada.

8. El lagrangiano de un sistema particular está dado por

$$L = \frac{1}{2}m(a\dot{x}^2 + 2b\dot{x}\dot{y} + c\dot{y}^2) - \frac{1}{2}k(ax^2 + 2bxy + cy^2) \quad (3.51)$$

con a, b y c constantes tal que $ac \neq b^2$. Calcula y resuelve las ecuaciones de movimiento. ¿Qué sistema describe L y qué tiene de especial la situación $ac = b^2$?

9. Dos masas están conectadas a través de una cuerda de longitud L . Una masa está colocada en una mesa donde puede moverse sin rozamiento y la otra cuelga de la cuerda que pasa por un agujero en la mesa y sólo puede moverse verticalmente (vease Figura 3.5). Calcula las ecuaciones de Lagrange y determina las constantes de movimiento. ¿Cuál es su interpretación? Resuelve el problema en el caso de que una de las cantidades conservadas sea cero.
10. El lagrangiano de una partícula relativista en un campo eléctrico \vec{E} y el campo magnético \vec{B} viene dado por

$$L = -m_0c^2\sqrt{1 - v^2/c^2} - e(\phi(x, t) - \dot{x}_i A_i(x, t)), \quad (3.52)$$

donde m_0 es la masa en reposo, $v^2 = \dot{x}_i \dot{x}_i$ la velocidad de la partícula al cuadrado y c es la velocidad de la luz, una constante universal. (Fijaos que para el límite newtoniano $v \ll c$, el lagrangiano se reduce a la versión no-relativista (3.45), salvo un factor constante famoso.) Dado que los campos electromagnéticos \vec{E} y \vec{B} están relacionado con los potenciales electromagnéticos ϕ y \vec{A} a través de

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial\vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A}, \quad (3.53)$$

demuestra que la ecuación de movimiento de una partícula relativista sometida a la fuerza de Lorentz viene dada por

$$\frac{d}{dt} \left(\frac{m_0 \dot{x}_i}{\sqrt{1 - v^2/c^2}} \right) = e [E_i + (\vec{v} \times \vec{B})_i]. \quad (3.54)$$

Calcula la energía conservada E en ausencia de campos electromagnéticos ($\phi = \vec{A} = 0$). ¿A qué corresponde esta fórmula en el sistema de reposo de la partícula?

Capítulo 4

Formalismo hamiltoniano y transformaciones canónicas

Pasaremos ahora a otro formalismo matemático para derivar las leyes de la mecánica, desarrollado por el matemático irlandés William Rowan Hamilton. El formalismo hamiltoniano es físicamente equivalente al formalismo lagrangiano, aunque el punto de partida matemático es muy diferente. Aunque el formalismo hamiltoniano no aporta ninguna física nueva, ni hace más fácil la resolución de las ecuaciones de movimiento, la gran ventaja del formalismo está en los conceptos matemáticos nuevos que introduce y que es formalismo permite extenderse a nuevas áreas de la física moderna.

4.1. El Hamiltoniano como transformada de Legendre

En el formalismo lagrangiano, un estado físico de un sistema a un momento $t = t_0$ está caracterizado por los valores de las coordenadas generalizadas q_α y sus derivadas \dot{q}_α . Los n coordenadas generalizadas q_α describen un punto en el *espacio de configuraciones* n -dimensional y la evolución del sistema viene dado por la trayectoria $q_\alpha(t)$ por el espacio de configuraciones. Matemáticamente, las ecuaciones de Lagrange son un problema de n variables y n ecuaciones de segundo orden, con condiciones iniciales $q_\alpha(t_0)$ y $\dot{q}_\alpha(t_0)$. Observa que, aunque matemáticamente a la hora de derivar las ecuaciones de Lagrange hemos tratado las velocidades generalizadas \dot{q}_α como independientes de las posiciones q_α , físicamente las \dot{q}_α no aparecen como variables independientes en el espacio de configuraciones, sino sólo como las derivadas de q_α y como condiciones iniciales.

En el formalismo hamiltoniano las ecuaciones no aparecen en función de las velocidades generalizadas, sino de los momentos generalizados p_α , definidos en (3.34). Pero la gran diferencia con el formalismo lagrangiano es que ahora se le considera nos n momentos conjugados como variables independientes de las n coordenadas generalizadas. El estado de un sistema ahora está caracterizado por $2n$ variables q_α y p_α en un espacio $2n$ -dimensional, llamado el *espacio de las fases*. Veremos que matemáticamente el formalismo hamiltoniano es un problema de $2n$ variables y $2n$ ecuaciones de primer orden.

La *función de Hamilton* o *hamiltoniano* está definido como la transformada de Legendre del lagrangiano. El diferencial total de un lagrangiano $L(q_\alpha, \dot{q}_\alpha, t)$ viene dado por

$$\begin{aligned} dL(q_\alpha, \dot{q}_\alpha, t) &= \frac{\partial L}{\partial q_\alpha} dq_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} d\dot{q}_\alpha + \frac{\partial L}{\partial t} dt \\ &= \dot{p}_\alpha dq_\alpha + p_\alpha d\dot{q}_\alpha + \frac{\partial L}{\partial t} dt \end{aligned}$$

$$= \dot{p}_\alpha dq_\alpha + d(p_\alpha \dot{q}_\alpha) - \dot{q}_\alpha dp_\alpha + \frac{\partial L}{\partial t} dt, \quad (4.1)$$

donde en la segunda igualdad hemos utilizado la definición (3.34) del momento conjugado y la ecuación de Lagrange, mientras en la tercera igualdad hemos escrito el término $p_\alpha d\dot{q}_\alpha$ como un diferencial total menos la parte que faltaba para el diferencial total.

Si ahora agrupamos los términos de (4.1) que son diferenciales totales, obtenemos

$$dH(q_\alpha, p_\alpha, t) = \dot{q}_\alpha dp_\alpha - \dot{p}_\alpha dq_\alpha - \frac{\partial L}{\partial t} dt, \quad (4.2)$$

donde hemos definido el hamiltoniano $H(q_\alpha, p_\alpha, t)$ como

$$H(q_\alpha, p_\alpha, t) = p_\alpha \dot{q}_\alpha(p_\alpha) - L(q_\alpha, \dot{q}_\alpha(p_\alpha), t). \quad (4.3)$$

¿Cómo hay que leer esta definición? Pues ya hemos visto que los momentos conjugados p_α están definidos por (3.34) como

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}, \quad (4.4)$$

lo que nos da en general una expresión de p_α en función de q_α , \dot{q}_α y t . En la mayoría de los casos relevantes, esta expresión es invertible, tal que podemos resolver \dot{q}_α en función de q_α , p_α y t . Rellenando esta expresión para las \dot{q}_α en la definición (4.3) del hamiltoniano, obtenemos una expresión para H en función de las coordenadas generalizadas, sus momentos conjugados y el tiempo.

En general, se puede definir la *transformación de Legendre* de una función $f(x, y)$ con diferencial

$$df(x, y) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \quad (4.5)$$

a una función $h(v, y) = vx - f(x, y)$ con diferencial

$$dh(v, y) = \frac{\partial h}{\partial v} dv - \frac{\partial h}{\partial y} dy \quad (4.6)$$

si tomamos que

$$v = \frac{\partial f}{\partial x} \quad y \quad x = \frac{\partial h}{\partial v}. \quad (4.7)$$

Fijaos que lo que tomamos por las variables de una función son las derivadas parciales de la transformada de Legendre. Las transformaciones de Legendre son por lo tanto una herramienta muy útil si queremos describir la misma física en términos de otras variables fundamentales. Es una técnica muy común en la termodinámica. Fijaos que las transformaciones de Legendre forman un \mathbb{Z}_2 : si $h(v, y)$ es la transformada de Legendre de $f(x, y)$, también $f(x, y)$ es la transformada de $h(v, y)$.

Vemos pues que el lagrangiano y el hamiltoniano son la transformada de Legendre el uno del otro, donde las variables conjugadas son las velocidades generalizadas \dot{q}_α y los momentos p_α . Ambos contienen la misma información física, pero codificada de otra manera. Las ecuaciones de movimiento o las *ecuaciones de Hamilton* se pueden derivar fácilmente de (4.2):

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha}, \quad \dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (4.8)$$

Tanto físicamente como matemáticamente son equivalentes a las ecuaciones de Lagrange (3.18). Desde un punto de vista matemático, (4.8) nos da $2n$ ecuaciones diferenciales de primer orden, mientras de (3.18) obtenemos n ecuaciones de segundo orden.

Para concretar un poco el formalismo y demostrar que también desde el punto de vista físico las ecuaciones de Hamilton son equivalentes a las de Lagrange, miraremos el ejemplo de una partícula

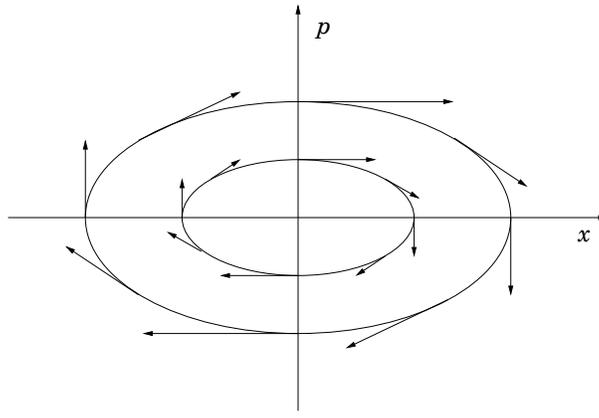


Figura 4.1: La evolución del oscilador armónico como curvas integrales del campo vectorial $(\dot{q}_\alpha, \dot{p}_\alpha)$ en el espacio de las fases.

en un potencial. En coordenadas cartesianas, el lagrangiano de una partícula con masa m en un potencial $V(x)$ viene dado por

$$L = \frac{1}{2}m\dot{x}_i\dot{x}_i - V(x). \quad (4.9)$$

Por la definición (3.34) tenemos que el momento conjugado $p_i = m\dot{x}_i$, o invirtiendo esta relación $\dot{x}_i = p_i/m$. Para el hamiltoniano de una partícula en un potencial obtenemos por lo tanto a través de la definición (4.3)

$$H = \frac{1}{2m}p_i p_i + V(x) \quad (4.10)$$

Las ecuaciones de Hamilton (4.8) en este caso son

$$\dot{x}_i = \frac{\partial H}{\partial p_i} = \frac{p_i}{m}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} = -\frac{\partial V(x)}{\partial x_i}. \quad (4.11)$$

Reconocemos en la primera ecuación simplemente la relación conocida entre las velocidades y los momentos conjugados y en la segunda ecuación la ley de Newton en el caso de fuerzas conservativas. Combinando las dos ecuaciones obtenemos la ecuación (3.23) deducida en el formalismo lagrangiano. Dejaremos al estudiante interesado averiguar la equivalencia del formalismo hamiltoniano y el lagrangiano para los otros ejemplos tratados en Capítulo 3.

Un concepto útil en el formalismo hamiltoniano son las trayectorias en el espacio de fase. Las ecuaciones de Hamilton (4.8) definen un campo vectorial $(\dot{q}_\alpha, \dot{p}_\alpha)$ en el espacio de las fases $2n$ -dimensional y la evolución del sistema está caracterizada por las curvas integrales de este campo vectorial. Miremos el ejemplo sencillo de un oscilador armónico unidimensional. El hamiltoniano viene dado por

$$H = \frac{1}{2m}p^2 + \frac{1}{2}kx^2 \quad (4.12)$$

y las ecuaciones de movimiento

$$\dot{x} = \frac{1}{m}p, \quad \dot{p} = -kx \quad (4.13)$$

definen un campo vectorial en el plano (x, p) (Véase Figura 4.1.) El eje x representa la desviación del punto de equilibrio y el eje p la norma del momento. Vemos por lo tanto para el oscilador armónico que cuando la desviación es máxima, el momento es mínimo, y viceversa. La evolución del sistema viene dado por las curvas integrales del campo vectorial, donde las condiciones iniciales $(q(t_0), p(t_0))$ determinan la curva específica. Para el caso del oscilador armónico, cada curva corresponde a un valor de la energía total $E = T + V$, que se conserva a lo largo de la curva, aunque en general esto no es verdad. (Por ejemplo en el caso del oscilador armónico amortiguado.)

4.2. Interpretación y cantidades conservadas

De las mismas ecuaciones de Hamilton (4.8) se ve que la derivada total del hamiltoniano con respecto al tiempo viene dada por

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial H}{\partial p_\alpha} \dot{p}_\alpha + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}, \quad (4.14)$$

donde los primeros dos términos se cancelan mutuamente, debido a (4.8). Vemos por lo tanto que si el hamiltoniano no depende explícitamente de t , H es una cantidad conservada. En realidad, esto ya lo sabíamos de la discusión de las cantidades conservadas en el formalismo lagrangiano, en sección 3.5. La definición (4.3) del hamiltoniano obtenido a través de la transformación de Legendre coincide con la definición de la cantidad conservada E , obtenida en (3.40) y la relación (3.41) es lo mismo que la tercera ecuación de Hamilton (4.8).

Ya hemos visto en la sección 3.5 que para el caso que las coordenadas generalizadas (3.3) no son funciones del tiempo y que además el potencial no depende de las velocidades generalizadas \dot{q}_α , la cantidad conservada E es la energía mecánica $T + V$ del sistema. Efectivamente, en este caso el lagrangiano viene dado por

$$L = \frac{1}{2} a_{\alpha\beta} \dot{q}_\alpha \dot{q}_\beta - V(q_\alpha) \quad (4.15)$$

y es fácil de ver, utilizando la definición (4.3) y la inversa de (4.4), que el hamiltoniano es por lo tanto de la forma

$$H = \frac{1}{2} a_{\alpha\beta} \dot{q}_\alpha \dot{q}_\beta + V(q_\alpha) = T + V. \quad (4.16)$$

Hay casos en los que HS no se conserva y/o que no es la energía mecánica, pero para un estudio más sistemático referimos a la literatura.

El hecho de que el hamiltoniano es, dadas ciertas circunstancias, la energía mecánica del sistema, es una de las ventajas del formalismo hamiltoniano. Sabiendo que los mínimos del potencial corresponden a soluciones estáticas estables, sólo tenemos que buscar los mínimos de H para encontrarlas. También por la misma propiedad, el formalismo hamiltoniano es ideal para calcular el espectro energético de un sistema en la mecánica cuántica. Sin embargo, en teorías relativistas, la energía del sistema no es una función escalar, como la acción, sino una componente del cuadrivector de energía-momento. El hamiltoniano por lo tanto no es invariante bajo cambios de coordenadas (relativistas) y aunque existen formulaciones hamiltonianas de la mecánica relativista, son bastante enredadas y menos naturales que el formalismo lagrangiano, que resulta ser covariante de manera natural.

Los mismos teoremas de conservación del formalismo lagrangiano también se aplican al caso hamiltoniano, quizá incluso más claramente. Si una coordenada q_α es cíclica (i.e. no aparece explícitamente, sino a través de su velocidad) en el lagrangiano, también lo es en el hamiltoniano. Pero de (4.8) se ve directamente que en este caso el momento conjugado es constante $p_\alpha = c_\alpha$. Como en el formalismo hamiltoniano las posiciones y los momentos son variables independientes, es posible eliminar dos grados de libertad del sistema: la coordenada cíclica simplemente porque no aparece y su momento conjugado, puesto que en el hamiltoniano apareciera como un parámetro constante.¹

Muchas veces buscar la solución de las $2n$ ecuaciones acopladas de Hamilton no resulta más fácil que la de las n ecuaciones de Lagrange, aunque estas últimas son de segunda orden. A la hora de resolver un problema, por lo tanto, el formalismo hamiltoniano no proporciona ninguna ventaja grande, aparte de los momentos conservados. El *formalismo de Routh*, desarrollado por el matemático Edward Routh (1831-1907), escoge lo mejor de los dos mundos. Consiste en no aplicar una transformación de Legendre a todas las coordenadas q_α , sino sólo a las l coordenadas cíclicas

¹Fijaos que aquí no hemos dicho nada sobre si la coordenada cíclica era una coordenada angular o no. Los teoremas de conservación y las cantidades conservadas se aplican por lo tanto a momentos lineales y a momentos angulares.

$q_{\bar{\alpha}}$ (con $0 \leq \bar{\alpha} \leq l$). El *routhiano*, la función

$$R(q_a, \dot{q}_A, p_{\bar{\alpha}}, t) = \sum_{\bar{\alpha}} p_{\bar{\alpha}} \dot{q}_{\bar{\alpha}} - L(q_a, \dot{q}_A, \dot{q}_{\bar{\alpha}}, t) \quad (4.17)$$

obtenida de esta forma, depende las $n - l$ coordenadas y las velocidades no-cíclicas q_A y \dot{q}_A (con $0 \leq A \leq n - l$) y los l momentos $p_{\bar{\alpha}}$. Pero dado que los momentos conjugados de las coordenadas son constantes, sólo aparecen como parámetros en el routhiano $R = R(q_a, \dot{q}_A, c_{\bar{\alpha}}, t)$. En la práctica el routhiano va ser una función de las coordenadas no-cíclicas y sus velocidades, a las cuales se aplica el formalismo lagrangiano.

4.3. Transformaciones canónicas

Dado que las ecuaciones de Hamilton (4.8) son tan fáciles en el caso de coordenadas cíclicas, es conveniente estudiar bajo qué condiciones el hamiltoniano toma la forma más sencilla. Un hamiltoniano podría tener alguna coordenada cíclica en un sistema de coordenadas (q_{α}, p_{α}) , sin que esto fuera obvio en otro sistema de coordenadas. Por ejemplo, del hamiltoniano

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(\sqrt{x^2 + y^2 + z^2}) \quad (4.18)$$

en coordenadas cartesianas no está claro, pero pasando a coordenadas esféricas se ve que la coordenada $\phi = \arctg \frac{\sqrt{x^2 + y^2}}{z}$ es cíclica y el momento angular asociado se conserva.

A las ecuaciones de Hamilton en la forma (4.8)

$$\dot{q}_{\alpha} = \frac{\partial H}{\partial p_{\alpha}}, \quad \dot{p}_{\alpha} = -\frac{\partial H}{\partial q_{\alpha}}, \quad (4.19)$$

se le llama las *ecuaciones canónicas* y cualquier transformación de coordenadas y momentos

$$Q_{\alpha} = Q_{\alpha}(q_{\beta}, p_{\beta}, t), \quad P_{\alpha} = P_{\alpha}(q_{\beta}, p_{\beta}, t), \quad (4.20)$$

se llama una *transformación canónica* si existe una función $\tilde{H}(Q_{\alpha}, P_{\alpha}, t)$ tal que se satisfacen las ecuaciones

$$\dot{Q}_{\alpha} = \frac{\partial \tilde{H}}{\partial P_{\alpha}}, \quad \dot{P}_{\alpha} = -\frac{\partial \tilde{H}}{\partial Q_{\alpha}}. \quad (4.21)$$

Resulta que el producto de dos transformaciones canónicas es una transformación canónica y que el inverso de una transformación canónica y la identidad también son canónicas. Por lo tanto, el conjunto de transformaciones canónicas forma un grupo bajo la composición de funciones.

Para determinar la forma del hamiltoniano nuevo $\tilde{H}(Q_{\alpha}, P_{\alpha}, t)$ y su relación con $H(q_{\alpha}, p_{\alpha}, t)$, hay que darse cuenta que tanto (4.19) como (4.21) surgen de un principio variacional

$$\delta \int_{t_0}^{t_1} L(q_{\alpha}, \dot{q}_{\alpha}, t) dt = \delta \int_{t_0}^{t_1} [p_{\alpha} \dot{q}_{\alpha} - H(q_{\alpha}, p_{\alpha}, t)] dt = 0, \quad (4.22)$$

$$\delta \int_{t_0}^{t_1} \tilde{L}(Q_{\alpha}, \dot{Q}_{\alpha}, t) dt = \delta \int_{t_0}^{t_1} [P_{\alpha} \dot{Q}_{\alpha} - \tilde{H}(Q_{\alpha}, P_{\alpha}, t)] dt = 0, \quad (4.23)$$

Dado la relación (4.20) entre (q_{α}, p_{α}) y (Q_{α}, P_{α}) , esto implica que los integrandos de las dos ecuaciones están relacionados a través de la identidad

$$C(p_{\alpha} \dot{q}_{\alpha} - H) = P_{\alpha} \dot{Q}_{\alpha} - \tilde{H} + \frac{d}{dt} F(q_{\alpha}, p_{\alpha}, t), \quad (4.24)$$

donde C es una constante arbitraria y $F(q_\alpha, p_\alpha, t)$ una función arbitraria de las variables canónicas y el tiempo. Sin pérdida de generalidad se puede coger $C = 1$.² Del mismo modo, F se puede escribir como función de las variables canónicas (q_α, p_α) , de (Q_α, P_α) o incluso de combinaciones de éstas, utilizando la relación (4.20) entre un par de variables canónicas y otro. En la práctica resultará más conveniente elegir la mitad de las variables pertenecientes a las variables antiguas y la otra mitad a las nuevas.

Si escribimos por ejemplo F en función de las coordenadas q_α y Q_α , obtenemos para (4.24)

$$p_\alpha \dot{q}_\alpha - H = P_\alpha \dot{Q}_\alpha - \tilde{H} + \frac{\partial F}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial F}{\partial Q_\alpha} \dot{Q}_\alpha + \frac{\partial F}{\partial t}, \quad (4.25)$$

o, reagrupando los términos con \dot{q}_α y \dot{Q}_α

$$\left(p_\alpha - \frac{\partial F}{\partial q_\alpha}\right) \dot{q}_\alpha - H = \left(P_\alpha + \frac{\partial F}{\partial Q_\alpha}\right) \dot{Q}_\alpha - \tilde{H} + \frac{\partial F}{\partial t}. \quad (4.26)$$

Dado que tanto H , \tilde{H} como F son independientes de \dot{q}_α y \dot{Q}_α , esto último sólo puede ser verdad si

$$p_\alpha = \frac{\partial F}{\partial q_\alpha}, \quad P_\alpha = -\frac{\partial F}{\partial Q_\alpha}, \quad \tilde{H} = H + \frac{\partial F}{\partial t}. \quad (4.27)$$

Vemos por lo tanto que la función $F(q_\alpha, Q_\alpha, t)$ actúa como una función generadora de la transformación canónica: dada una transformación canónica (4.20), se pueden expresar p_α y P_α en función de q_α , Q_α y t . Las primeras dos ecuaciones (4.27) proporcionan un sistema de ecuaciones diferenciales que nos permiten encontrar la función $F(q_\alpha, Q_\alpha, t)$ que relaciona el hamiltoniano antiguo H con el nuevo \tilde{H} a través de la tercera ecuación de (4.27).

También es posible invertir el procedimiento: dada una función generadora $F(q_\alpha, Q_\alpha, t)$, de las dos primeras ecuaciones de (4.27) podemos obtener una expresión para el cambio de variables (4.20) de Q_α y P_α en función de q_α y p_α , mientras la tercera ecuación de (4.27) nos da el hamiltoniano del cual Q_α y P_α son las variables canónicas.

A veces es más conveniente escribir la función generadora F en términos de otras variables, como por ejemplo q_α y P_α (por ejemplo porque en la transformación canónica (4.20) los P_α no dependen de los q_α). En este caso la derivación es análoga, salvo la sutileza de que la función generadora que aparece en (4.25) tiene que ser

$$\tilde{F}(q_\alpha, P_\alpha, t) = F(q_\alpha, P_\alpha, t) - P_\alpha Q_\alpha. \quad (4.28)$$

Sustituyendo esta expresión en (4.25) y agrupando de términos con \dot{q}_α y \dot{P}_α , obtenemos las ecuaciones

$$p_\alpha = \frac{\partial F}{\partial q_\alpha}, \quad Q_\alpha = \frac{\partial F}{\partial P_\alpha}, \quad \tilde{H} = H + \frac{\partial F}{\partial t}. \quad (4.29)$$

Del mismo modo se obtienen las fórmulas análogas para los casos en que F se expresa en función de p_α y Q_α ó de p_α y P_α .

Veremos ahora unos ejemplos de transformaciones canónicas. Consideramos primero la función generadora $F = q_\alpha P_\alpha$. De las ecuaciones (4.29) tenemos que $Q_\alpha = q_\alpha$ y $P_\alpha = p_\alpha$. En otras palabras, la función $F = q_\alpha P_\alpha$ genera una transformación trivial, donde las nuevas variables canónicas son las mismas que las antiguas y el hamiltoniano es invariante. La importancia de este ejemplo trivial está en demostrar que la identidad es una transformación canónica, una propiedad necesaria para la estructura de grupo de las transformaciones canónicas.

Consideramos ahora la transformación

$$\begin{aligned} Q_\alpha &= a q_\alpha + b p_\alpha, \\ P_\alpha &= c q_\alpha + d p_\alpha. \end{aligned} \quad (4.30)$$

²Para una $C \neq 1$ siempre se puede hacer una redefinición de las variables canónicas $Q'_\alpha = A Q_\alpha$ y $P'_\alpha = B P_\alpha$ y del hamiltoniano $\tilde{H}' = A B \tilde{H}$ con $A B = C$, tal que en la relación (4.24) C ya no aparece.

Queremos saber bajo qué condiciones esta transformación es canónica, cuál es su función generadora F y de qué hamiltoniano \tilde{H} son Q_α y P_α las variables canónicas. Lo primero es por lo tanto expresar (por ejemplo) p_α y P_α en función de las coordenadas q_α y Q_α :

$$p_\alpha = b^{-1}(Q_\alpha - aq_\alpha), \quad P_\alpha = (c - adb^{-1})q_\alpha + db^{-1}Q_\alpha. \quad (4.31)$$

Podremos obtener una expresión para F resolviendo el sistema de ecuaciones diferenciales (4.27)

$$\frac{\partial F}{\partial q_\alpha} = b^{-1}(Q_\alpha - aq_\alpha), \quad (4.32)$$

$$\frac{\partial F}{\partial Q_\alpha} = (c - adb^{-1})q_\alpha + db^{-1}Q_\alpha. \quad (4.33)$$

No es difícil de ver que la solución es

$$F = -b^{-1}\left(\frac{1}{2}aq_\alpha^2 + \frac{1}{2}dQ_\alpha^2 - q_\alpha Q_\alpha\right), \quad (4.34)$$

con las condiciones $ad - bc = 1$ y $b \neq 0$. Vemos por lo tanto que podemos actuar con cualquier transformación $SL(2, \mathbb{R})$, mezclando coordenadas y momentos sin cambiar nada físico. Esto indica que la distinción entre coordenadas y momentos es completamente arbitraria en el formalismo hamiltoniano. Efectivamente, en el caso espacial donde $a = d = 0$ y $b = -c = 1$ tenemos que la transformación canónica intercambia las posiciones de las variables antiguas con los momentos de las nuevas y viceversa. Fijaos por último que la identidad no se puede escribir como una transformación canónica de la forma (4.30), por violar la condición $b \neq 0$. Ya hemos visto arriba que la función generadora no es de la forma (4.34).

También hay otra manera de saber si una transformación (4.20) es canónica, sin tener que recurrir a la función generadora, por lo menos en el caso en que la transformación no depende del tiempo. Consideramos la transformación canónica

$$Q_\alpha = Q_\alpha(q_\beta, p_\beta), \quad P_\alpha = P_\alpha(q_\beta, p_\beta), \quad (4.35)$$

y su inverso

$$q_\alpha = q_\alpha(Q_\beta, P_\beta), \quad p_\alpha = p_\alpha(Q_\beta, P_\beta). \quad (4.36)$$

De (4.35) sabemos que

$$\dot{Q}_\alpha = \frac{\partial Q_\alpha}{\partial q_\beta} \dot{q}_\beta + \frac{\partial Q_\alpha}{\partial p_\beta} \dot{p}_\beta = \frac{\partial Q_\alpha}{\partial q_\beta} \frac{\partial H}{\partial p_\beta} - \frac{\partial Q_\alpha}{\partial p_\beta} \frac{\partial H}{\partial q_\beta}, \quad (4.37)$$

donde en la segunda igualdad hemos utilizado las ecuaciones de Hamilton (4.8). Por otro lado, dado que Q_α y P_α también son variables canónicas (por construcción), sabemos de (4.36) que

$$\dot{Q}_\alpha = \frac{\partial H}{\partial P_\alpha} = \frac{\partial H}{\partial q_\beta} \frac{\partial q_\beta}{\partial P_\alpha} + \frac{\partial H}{\partial p_\beta} \frac{\partial p_\beta}{\partial P_\alpha}. \quad (4.38)$$

Fijaos en este punto que los hamiltonianos que aparecen en (4.37) y (4.38) son el mismo, puesto que la transformación (4.35) no depende del tiempo. Podemos por lo tanto comparar directamente las dos expresiones (4.37) y (4.38) para \dot{Q}_α . Las dos expresiones sólo pueden ser válidas a la vez, es decir las transformaciones (4.35) y (4.35) sólo pueden ser canónicas si

$$-\frac{\partial Q_\alpha}{\partial p_\beta} = \frac{\partial q_\beta}{\partial P_\alpha}, \quad \frac{\partial Q_\alpha}{\partial q_\beta} = \frac{\partial p_\beta}{\partial P_\alpha}. \quad (4.39)$$

Del mismo modo comparando las expresiones para \dot{P}_α obtenemos las condiciones

$$\frac{\partial P_\alpha}{\partial p_\beta} = \frac{\partial q_\beta}{\partial Q_\alpha}, \quad -\frac{\partial P_\alpha}{\partial q_\beta} = \frac{\partial p_\beta}{\partial Q_\alpha}. \quad (4.40)$$

Estas expresiones, aparte de darnos las condiciones necesarias para que una transformación sea canónica, nos serán útiles para demostrar la invariancia de los corchetes de Poisson bajo transformaciones canónicas.

Finalmente miramos a un ejemplo (un poco rebuscado) donde el hamiltoniano toma una forma más sencilla debido a una transformación canónica: el caso del oscilador armónico unidimensional. La forma tradicional de escribir el hamiltoniano es

$$H = \frac{1}{2m}p^2 + \frac{k}{2}x^2 = \frac{1}{2m}(p^2 + m^2\omega^2x^2), \quad (4.41)$$

donde $\omega = \sqrt{k/m}$ es la frecuencia del oscilador. Si pensamos en p y $m\omega x$ como las componentes de un vector en el espacio de las fases, vemos que la segunda expresión para el hamiltoniano sugiere una simetría “rotacional” para H , puesto que el hamiltoniano sólo depende de la “norma” $p^2 + m^2\omega^2x^2$. En coordenadas polares, la coordenada angular sería cíclica. Intentamos por lo tanto el cambio de coordenadas

$$p = f(P) \cos Q, \quad m\omega x = f(P) \sin Q, \quad (4.42)$$

donde $f(P)$ es una función cuya forma tenemos que determinar para que el cambio de coordenadas sea una transformación canónica. Igual que en el ejemplo anterior tenemos que resolver el sistema de ecuaciones diferenciales (4.27) para encontrar la expresión para F . Del cambio de coordenadas (4.42) tenemos que

$$p = m\omega x \cotg Q, \quad (4.43)$$

tal que una solución particular de (4.27) sería

$$F = \frac{1}{2}m\omega x^2 \cotg Q. \quad (4.44)$$

De la segunda ecuación de (4.27) deducimos por lo tanto la expresión para P

$$P = \frac{1}{2}m\omega x^2 \sin^{-2} Q, \quad (4.45)$$

o, depejando x

$$x = \sqrt{\frac{2P}{m\omega}} \sin Q. \quad (4.46)$$

Comparando con el cambio de coordenadas (4.42) deducimos que $f(P) = \sqrt{2m\omega P}$. En las coordenadas P y Q , el hamiltoniano (4.41) es de la forma

$$H = \omega P, \quad (4.47)$$

y es obvio que la coordenada Q es cíclica. Las ecuaciones de movimiento son

$$\dot{P} = 0 \quad \dot{Q} = \omega \quad (4.48)$$

lo que se puede intergar directamente

$$Q = \omega t + \phi_0, \quad P = \frac{1}{2}m\omega A^2 \quad (4.49)$$

o, en las coordenadas antiguas

$$x = A \sin(\omega t + \phi_0). \quad (4.50)$$

4.4. Corchetes de Poisson

Existe otro formalismo para describir la dinámica de un sistema, que está íntimamente relacionado con el formalismo hamiltoniano y con las transformaciones canónicas: el formalismo de

los corchetes de Poisson. Igual que el formalismo hamiltoniano no aporta física nueva, ya que es equivalente a este último (y por lo tanto también al formalismo lagrangiano), pero sí proporciona un enfoque nuevo para tratar sistemas dinámicos, relacionándolo con teoría de álgebras de Lie y proporcionando un lenguaje que es directamente aplicable a la formulación de la mecánica cuántica.

En el formalismo hamiltoniano, la derivada total con respecto al tiempo de cualquier cantidad $f(q_\alpha, p_\alpha, t)$ viene dada por

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial f}{\partial p_\alpha} \dot{p}_\alpha + \frac{\partial f}{\partial t} \\ &= \frac{\partial f}{\partial q_\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial H}{\partial q_\alpha} + \frac{\partial f}{\partial t}.\end{aligned}\quad (4.51)$$

Si ahora definimos los *corchetes de Poisson* $\{f, g\}$ de dos funciones $f(q_\alpha, p_\alpha, t)$ y $g(q_\alpha, p_\alpha, t)$ como

$$\{f, g\} = \frac{\partial f}{\partial q_\alpha} \frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial g}{\partial q_\alpha}, \quad (4.52)$$

entonces vemos que se puede escribir (4.51) como

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t}. \quad (4.53)$$

Esta última ecuación describe la evolución de la cantidad $f(q_\alpha, p_\alpha, t)$ y por lo tanto se puede considerar como su ecuación de movimiento. No es difícil demostrar que de la ecuación (4.53) se puede recuperar las ecuaciones de Hamilton (4.8). Tomando para la función $f(q_\alpha, p_\alpha, t)$ la propia coordenada canónica q_α , se obtiene

$$\dot{q}_\alpha = \frac{\partial q_\alpha}{\partial q_\beta} \frac{\partial H}{\partial p_\beta} - \frac{\partial q_\alpha}{\partial p_\beta} \frac{\partial H}{\partial q_\beta} + \frac{\partial q_\alpha}{\partial t} = \delta_{\alpha\beta} \frac{\partial H}{\partial p_\beta} = \frac{\partial H}{\partial p_\alpha}, \quad (4.54)$$

donde en segundo y el tercer término de la segunda igualdad son cero porque q_α es independiente tanto de p_α como de t . Similarmente, cogiendo para $f(q_\alpha, p_\alpha, t)$ el momento canónico p_α y el propio hamiltoniano H tenemos respectivamente

$$\dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}, \quad \frac{dH}{dt} = \frac{\partial H}{\partial t}. \quad (4.55)$$

En particular la ecuación de Poisson (4.53) toma una forma sencilla si la función $f(q_\alpha, p_\alpha, t)$ es una constante de movimiento. En este caso la ecuación se reduce a

$$\{f, H\} + \frac{\partial f}{\partial t} = 0. \quad (4.56)$$

El formalismo de los corchetes de Poisson resulta especialmente útil para estudiar los conjuntos de las cantidades conservadas. Pero antes de entrar en este tema, discutiremos algunas de sus propiedades.

Se puede comprobar (ejerc.), utilizando directamente la definición (4.52), que los corchetes de Poisson tienen las siguientes propiedades

$$\{f, g\} = -\{g, f\}, \quad (4.57)$$

$$\{af + bg, h\} = a\{f, h\} + b\{g, h\}, \quad (4.58)$$

$$\frac{\partial}{\partial \eta} \{f, g\} = \left\{ \frac{\partial f}{\partial \eta}, g \right\} + \left\{ f, \frac{\partial g}{\partial \eta} \right\}, \quad (4.59)$$

$$\{fg, h\} = f\{g, h\} + \{f, h\}g, \quad (4.60)$$

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0, \quad (4.61)$$

para $a, b \in \mathbb{R}$ (ó \mathbb{C} si es necesario) y η cualquier variable de las funciones consideradas. Las ecuaciones (4.57)-(4.59) expresan la antisimetría y la linealidad de los corchetes de Poisson, la ecuación (4.60) la decomposición del producto y la ecuación (4.61) es la *identidad de Poisson* o la *identidad de Jacobi*.

Aparte de estas propiedades, los corchetes de Poisson son invariantes bajo transformaciones canónicas. Para demostrar esto, consideramos primero los llamados *corchetes fundamentales de Poisson* de las variables canónicas q_α y p_α :

$$\{q_\alpha, q_\beta\} = \{p_\alpha, p_\beta\} = 0, \quad \{q_\alpha, p_\beta\} = \delta_{\alpha\beta}. \quad (4.62)$$

Otro conjunto de variables canónicas Q_α y P_α , generado a través de una transformación canónica (4.35) satisface las mismas relaciones. Aplicando la regla de la cadena y las condiciones para canonicidad (4.39)-(4.40) vemos que

$$\{Q_\alpha, Q_\beta\} = \frac{\partial Q_\alpha}{\partial q_\gamma} \frac{\partial Q_\beta}{\partial p_\gamma} - \frac{\partial Q_\alpha}{\partial p_\gamma} \frac{\partial Q_\beta}{\partial q_\gamma} = -\frac{\partial Q_\alpha}{\partial q_\gamma} \frac{\partial q_\gamma}{\partial P_\beta} - \frac{\partial Q_\alpha}{\partial p_\gamma} \frac{\partial p_\gamma}{\partial P_\beta} = -\frac{\partial Q_\alpha}{\partial P_\beta} = 0. \quad (4.63)$$

Del mismo modo es fácil de demostrar que

$$\{P_\alpha, P_\beta\} = 0, \quad \{Q_\alpha, P_\beta\} = \delta_{\alpha\beta}. \quad (4.64)$$

En general, los corchetes de Poisson de dos funciones f y g transformada bajo una transformación canónica (4.35) como

$$\begin{aligned} \{f, g\} &= \frac{\partial f}{\partial q_\alpha} \frac{\partial g}{\partial p_\alpha} - \frac{\partial f}{\partial p_\alpha} \frac{\partial g}{\partial q_\alpha} \\ &= \frac{\partial f}{\partial Q_\beta} \frac{\partial g}{\partial Q_\gamma} \{Q_\beta, Q_\gamma\} + \frac{\partial f}{\partial Q_\beta} \frac{\partial g}{\partial P_\gamma} \{Q_\beta, P_\gamma\} \\ &\quad + \frac{\partial f}{\partial P_\beta} \frac{\partial g}{\partial Q_\gamma} \{P_\beta, Q_\gamma\} + \frac{\partial f}{\partial P_\beta} \frac{\partial g}{\partial P_\gamma} \{P_\beta, P_\gamma\} \\ &= \frac{\partial f}{\partial Q_\beta} \frac{\partial g}{\partial P_\beta} - \frac{\partial f}{\partial P_\beta} \frac{\partial g}{\partial Q_\beta}, \end{aligned} \quad (4.65)$$

donde en la segunda igualdad hemos aplicado la regla de la cadena y las condiciones para canonicidad (4.39)-(4.40) y en la tercera igualdad la invariancia canónica de los corchetes fundamentales de Poisson (4.63)-(4.64). Por lo tanto vemos que no sólo los corchetes fundamentales son invariantes bajo transformaciones canónicas, sino también los corchetes de cualesquieres dos cantidades f y g . Esto implica que la dinámica descrita por la ecuación de Poisson (4.53) es independiente de las variables canónicas que utilizamos.

Pasaremos ahora al estudio de las constantes del movimiento. Ya hemos visto que si una función $f(q_\alpha, p_\alpha, t)$ es una constante durante el movimiento del sistema, esta función satisface la ecuación

$$\{f, H\} + \frac{\partial f}{\partial t} = 0. \quad (4.66)$$

Muchas veces a una función que satisface esta condición se le denomina (*primera*) *integral*. En particular, si f no depende del tiempo explícitamente, entonces (4.66) se reduce a $\{f, H\} = 0$.

Estará claro que cuanto más constantes del movimiento conozcamos, más simple será en análisis del sistema, puesto que se reduce el número de ecuaciones diferenciales que tenemos que resolver. El formalismo de los corchetes de Poisson nos proporciona una manera elegante de estudiar las constantes del movimiento: el *teorema de Poisson* (o el *teorema de Poisson-Jacobi*) dice que si dos funciones $f(q_\alpha, p_\alpha, t)$ y $g(q_\alpha, p_\alpha, t)$ son constantes del movimiento, entonces el corchete de Poisson de f y g también es una constante del movimiento. Efectivamente, tenemos que

$$\frac{\partial}{\partial t} \{f, g\} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} = \{ \{f, H\}, g \} + \{ f, \{g, H\} \} = \{ \{f, g\}, H \}, \quad (4.67)$$

donde en la segunda identidad hemos utilizado que f y g son constantes del movimiento y en la última igualdad la identidad de Jacobi (4.61). De esta manera, conociendo dos o más constantes del movimiento, se puede encontrar más aplicando los corchetes de Poisson a los distintos pares de constantes. Sin embargo, no siempre es posible obtener de este modo todas las constantes del sistema, ni todas las constantes obtenidas son independientes. Es posible que el resultado obtenido sea una combinación (constante) de las constantes ya conocidas y también puede ocurrir que de un conjunto de constantes de movimiento genera sólo un subconjunto de todas las constantes del sistema.

En este punto debería estar clara la relación entre los corchetes de Poisson y las álgebras de Lie. Efectivamente, las constantes del movimiento juegan el papel de los elementos del álgebra T_i , mientras los corchetes de Poisson corresponden a los conmutadores de dos elementos del álgebra $[T_i, T_j]$. Las propiedades (4.57)-(4.61) son propiedades conocidas de los conmutadores (salvo (4.59)) y (4.67) corresponde al hecho de que un álgebra cierra bajo conmutación:

$$[T_i, T_j] = f_{ij}^k T_k, \quad (4.68)$$

donde f_{ij}^k son las constantes de estructura del álgebra. Ahora también es obvio porque no siempre es posible obtener todas las constantes de movimiento a partir de unas pocas. Es posible que las pocas constantes conocidas pertenezcan a una subálgebra, que por definición es cerrada.

Esta relación entre las constantes del movimiento y los elementos de un álgebra no es sorprendente. Ya hemos visto en sección 3.5 que según el teorema de Noether a cada cantidad conservada (es decir, a cada constante del movimiento) corresponde una invariancia del sistema bajo una simetría continua. Las invariancias del sistema están caracterizadas por elementos del grupo de simetría, del cual los generadores forman un álgebra. En cierto modo, las primeras integrales forman los generadores de las simetrías del sistema.

Veremos ahora un ejemplo concreto, el de una partícula libre, no sometida a fuerzas ni ligaduras. Está claro que en este caso el potencial es cero y el hamiltoniano sólo consiste en la energía cinética y depende por lo tanto, en coordenadas cartesianas, sólo de los momentos $p_i = m\dot{x}_i$ ($i = 1, 2, 3$). Por el análisis hecho en las secciones 3.5 y 4.2 sabemos que aparte de los p_i , también los momentos angulares $L_i = \epsilon_{ijk} x_j p_k$ se conservan. Efectivamente, es fácil comprobar que los corchetes de Poisson $\{p_i, H\}$ y $\{L_i, H\}$ son cero, con lo cual satisfacen la condición (4.66), dado que ni p_i ni L_i dependen del tiempo. Calculando los distintos corchetes de Poisson de las distintas combinaciones de cantidades conservadas obtenemos (ejerc.)

$$\{p_i, p_j\} = 0, \quad \{p_i, L_j\} = \epsilon_{ijk} p_k, \quad \{L_i, L_j\} = \epsilon_{ijk} L_k. \quad (4.69)$$

Estas 6 constantes de movimiento describen completamente el sistema de 3 grados de libertad. Pero es fácil de ver que si sólo conociéramos L_1 y L_2 , podríamos obtener L_3 a través de los corchetes de Poisson, aunque no sería posible obtener los p_i .

El conjunto de constantes p_i y L_i describen la simetría bajo traslaciones y rotaciones de \mathbb{R}^3 . En el lenguaje del álgebra, esto corresponde a los operadores $\hat{P}_i = \partial_i$ y $\hat{L}_i = \epsilon_{ijk} x_j \partial_k$, que satisfacen el álgebra

$$[\hat{P}_i, \hat{P}_j] = 0, \quad [\hat{P}_i, \hat{L}_j] = \epsilon_{ijk} \hat{P}_k, \quad [\hat{L}_i, \hat{L}_j] = \epsilon_{ijk} \hat{L}_k. \quad (4.70)$$

En el último conmutador reconocemos el álgebra de $SO(3)$. De igual manera la energía E es una cantidad conservada y corresponde al generador de las traslaciones en el tiempo. El operador correspondiente es el mismo hamiltoniano $\hat{H} = \partial_t$, que determina la evolución del sistema.

4.5. Unas palabras cuánticas

La estrecha relación entre los corchetes de Poisson y los álgebras de Lie es crucial en la formulación de la mecánica cuántica, la parte de la mecánica que describe el mundo (sub)atómico.

Allí una cantidad física \mathcal{A} está asociada a un operador hermítico \hat{A} , llamado *observable*. Los posibles valores que puede tomar la cantidad \mathcal{A} corresponden a los autovalores a_n del operador \hat{A} . En general los autovalores de operadores toman valores discretos, y debido a esto ciertas cantidades físicas aparecen *cuantizadas*. Ejemplos son los niveles de energía de los electrones en un átomo de hidrógeno o la cuantización del momento angular.

Pero un aspecto más contraria a la intuición aún es el hecho de que si dos operadores no conmutan, no son simultáneamente diagonalizables y no podemos conocer simultáneamente los autovalores de cada uno de ellos. Esto implica que no podemos conocer simultáneamente los valores de las cantidades físicas asociadas a estos operadores. Por ejemplo en (4.70) hemos visto que las componentes del generador del momento angular no conmutan. Siempre podremos elegir una base donde uno de estas componentes es diagonal, digamos \hat{L}_3 , lo que nos permite afirmar ciertas cosas sobre los posibles valores de la componente z de \vec{L} , pero no podemos decir nada sobre las componentes x ó y . Según la (interpretación tradicional de la) mecánica cuántica, los valores de estas componentes simplemente no están determinadas y carece de sentido preguntar cuánto mide por ejemplo L_x . La única manera de saber esto sería diagonalizar \hat{L}_1 , pero entonces volveríamos a perder la información sobre la componente z .

El ejemplo más famoso de esta *relación de incertidumbre* de Heisenberg viene directamente de los corchetes fundamentales de Poisson (4.62). Si \hat{X} es el operador correspondiente a la posición x de una partícula y \hat{P} el operador correspondiente a su momento, entonces el equivalente de los corchetes fundamentales de Poisson (4.62) es

$$[\hat{X}, \hat{P}] = \frac{i\hbar}{2\pi} \mathbb{1}, \quad (4.71)$$

donde $\mathbb{1}$ es el operador de identidad y $\hbar = 6,626068 \cdot 10^{-34}$ kg m²/s es la constante de Planck, el cuanto mínimo de acción. En la mecánica cuántica todas las cantidades físicas aparecen cuantizadas en múltiplos de \hbar . El conmutador no trivial (4.71) entre el operador de posición y el operador de momento impone una incertidumbre fundamental a la hora de intentar medir la posición y el momento de una partícula. Heisenberg demostró que si la incertidumbre en la posición viene dada por $\Delta\mathcal{X}$ y la incertidumbre en la posición por $\Delta\mathcal{P}$, tenemos inevitablemente en cada medición que

$$\Delta\mathcal{X}\Delta\mathcal{P} \geq \frac{\hbar}{4\pi}. \quad (4.72)$$

Matemáticamente esta incertidumbre entre posición y momento viene del hecho de que x y p son variables conjugadas en el análisis de Fourier. Se puede describir la información de un sistema con una función $\psi(x)$, o equivalentemente con su transformada de Fourier $\phi(p) = \int \psi(x)e^{ipx} dx$. Si $\psi(x)$ es una gaussiana con una amplitud alta y una variancia baja (la posición está bien determinada), entonces $\phi(p)$ es otra gaussiana, pero con un máximo baja y una variancia muy grande (muchacha indeterminación en el momento), y viceversa.

4.6. Problemas

1. Dado el lagrangiano

$$L = \frac{1}{2}a_{\alpha\beta}\dot{q}_\alpha\dot{q}_\beta - V(q_\alpha), \quad (4.73)$$

con $a_{\alpha\beta}$ la métrica (3.37) de la variedad impuesta por las ligaduras reónomas.

- a) Demuestra que el hamiltoniano correspondiente viene dado por

$$H = \frac{1}{2}a_{\alpha\beta}^{-1}p_\alpha p_\beta + V(q_\alpha), \quad (4.74)$$

donde $a_{\alpha\beta}^{-1}$ es la inversa de $a_{\alpha\beta}$.

b) Aplica una transformación de Legendre al hamiltonino (4.74), eliminando los momentos p_α a través de (la mitad de) las ecuaciones de Hamilton y demuestra que se recupera el lagrangiano original (4.73).

c) Aplica ahora una transformación de Legendre al hamiltonina (4.74), eliminando las coordenadas q_α a través de (la otra mitad de) las ecuaciones de Hamilton, para obtener el lagrangiano del espacio de momentos $K = K(p_\alpha, \dot{p}_\alpha)$. ¿Cuál es la relación entre L y K ?

2. Considera el lagrangiano

$$L(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x) + Ax^n\dot{x}, \quad (4.75)$$

con A y n constantes arbitrarios. Calcula el hamiltoniano correspondiente $H = H(x, p)$ y demuestra que se reduce al hamiltoniano $\tilde{H} = H(x, \tilde{p})$, correspondiente al lagrangiano

$$\tilde{L}(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x), \quad (4.76)$$

bajo la redefinición de los momentos $p = \tilde{p} + Ax^n$. Explica porqué. (Pista: compara las ecuaciones de Lagrange de L y \tilde{L} .)

3. Considera la partícula restringida a moverse sobre la superficie de una esfera de radio R en un campo gravitacional constante del ejercicio 3 del capítulo 3. Calcula el hamiltoniano y demuestra que las ecuaciones de Hamilton son equivalentes a las de Lagrange obtenidas en el ejercicio 3.
4. Considera una partícula con masa m en un campo gravitatorio constante, restringida a moverse sobre un espiral con ecuación $R = \text{cte}$ y $z = k\theta$ con K una constante. Calcula el hamiltoniano y resuelve las ecuaciones de Hamilton.
5. Dos masas m_1 y m_2 están conectadas a través de un muelle sin masa con constante k y tamaño en reposo L . El sistema se mueve libremente en el plano \mathbb{R}^2 . Escribe el hamiltoniano del sistema y demuestra que hay tres coordenadas cíclicas. ¿Con qué cantidades conservadas corresponden estas simetrías? ¿Cuál es cuarta cantidad conservada? (Pista: describe el sistema en términos de la posición del centro de masas y las posiciones relativas, como en la sección 5.2).
6. Considera lagrangiano de un oscilador armónico amortiguado

$$L = \frac{1}{2}m\dot{q}^2 e^{2\gamma t} - \frac{1}{2}k^2 q^2 e^{2\gamma t}, \quad (4.77)$$

con γ una constante positiva. Deriva las ecuaciones de movimiento y interpreta cada término. Calcula la forma del hamiltoniano. Dada la transformación canónica

$$Q = q e^{\gamma t}, \quad P = p e^{-\gamma t}, \quad (4.78)$$

busca la función generadora $F(p, P)$ que genera esta transformación y calcula el hamiltoniano \tilde{H} para el cual Q y P son las variables canónicas. Resuelva las ecuaciones de movimiento de \tilde{H} , distinguiendo entre los casos que sea γ mayor, igual o menor que $\sqrt{k/m}$. ¿Cuál es la interpretación de cada uno de los casos en las variables antiguas q y p ?

7. Considera el hamiltoniano $H_0(p, q)$ que no depende explícitamente del tiempo y que describe algún sistema de un grado de libertad. Sobre el sistema H_0 actúa una perturbación tal que el sistema nuevo está descrito por el hamiltoniano perturbado $H = H_0 - \epsilon q \sin \omega t$. ¿Cómo se modifican las ecuaciones de movimiento debido a la perturbación? Dada la función generadora

$$F(q, P) = qP - \frac{\epsilon q}{\omega} \cos \omega t, \quad (4.79)$$

demuestra que el hamiltoniano \tilde{H} , generado de la transformación canónica inducido por F , es otra vez de la forma $\tilde{H} = H_0(Q, P - \frac{\epsilon}{\omega} \cos \omega t)$. En otras palabras, la transformación canónica absorbe la perturbación en el hamiltoniano. Demuestra a través de las ecuaciones de movimiento que Q y P son efectivamente las variables canónicas de $\tilde{H} = H_0$.

8. Un sistema está descrito por el lagrangiano

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}A(x^2 + y^2) - Bxy, \quad (4.80)$$

con A y B constantes positivas. Aplica el cambio de coordenadas

$$\theta = \frac{1}{\sqrt{2}}(x + y), \quad \phi = \frac{1}{\sqrt{2}}(x - y), \quad (4.81)$$

y interpreta el sistema a través de la nueva forma del lagrangiano. Demuestra que el hamiltoniano se descompone en dos partes H_θ y H_ϕ que dependen sólo de θ y de ϕ respectivamente. Utiliza los corchetes de Poisson para demostrar que ambas son cantidades conservadas. Define la cantidad $\ell = m(\theta p_\phi - \phi p_\theta)$ y demuestra que ℓ se conserva en el caso de $B = 0$. Interpreta las cantidades H_θ , H_ϕ y ℓ y comenta por qué están conservadas.

Capítulo 5

Potenciales centrales

En este capítulo estudiaremos una primera aplicación del formalismo lagrangiano: una partícula moviéndose en un potencial que sólo depende de la distancia r del origen. Hay muchos ejemplos de este tipo de sistema en la naturaleza, como un planeta en un campo gravitatorio, un sistema de estrellas dobles, la electrostática o problemas de dispersión.

5.1. Repaso de coordenadas esféricas

Pasaremos ahora rápidamente al caso de coordenadas esféricas. El cambio de coordenadas cartesianas $\{x, y, z\}$ a esféricas $\{r, \theta, \phi\}$ y vice versa está dado por

$$\begin{aligned}x &= r \sin \theta \cos \phi, & r &= \sqrt{x^2 + y^2 + z^2}, \\y &= r \sin \theta \sin \phi, & \theta &= \arctan \frac{\sqrt{x^2 + y^2}}{z}, \\z &= r \cos \theta, & \phi &= \arctan \frac{y}{x}.\end{aligned}\tag{5.1}$$

La base ortonormal $\{\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi\}$ está relacionada con la base cartesiana $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ como

$$\vec{e}_r = \sin \theta \cos \phi \vec{e}_x + \sin \theta \sin \phi \vec{e}_y + \cos \theta \vec{e}_z,\tag{5.2}$$

$$r\vec{e}_\theta = r(\cos \theta \cos \phi \vec{e}_x + \cos \theta \sin \phi \vec{e}_y - \sin \theta \vec{e}_z),\tag{5.3}$$

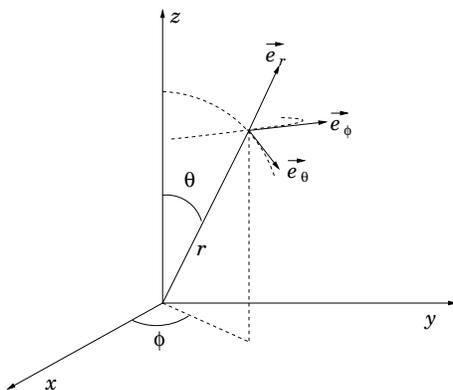


Figura 5.1: El cambio de coordenadas cartesianas $\{x, y, z\}$ a coordenadas esféricas $\{r, \theta, \phi\}$.

$$r \sin \theta \vec{e}_\phi = r \sin \theta \left(-\sin \phi \vec{e}_x + \cos \phi \vec{e}_y \right), \quad (5.4)$$

y la métrica en coordenadas esféricas es

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \quad (5.5)$$

En coordenadas esféricas la posición de una partícula está dado por el vector de posición $\vec{r}(t) = r(t)\vec{e}_r$ y la velocidad es

$$\vec{v}(t) = \frac{d\vec{r}(t)}{dt} = \dot{r}\vec{e}_r + r\frac{d\vec{e}_r}{dt} \quad (5.6)$$

$$= \dot{r}\vec{e}_r + r\dot{\theta}\vec{e}_\theta + r\sin\theta\dot{\phi}\vec{e}_\phi, \quad (5.7)$$

El momento angular \vec{L} en estas coordenadas está dado por

$$\vec{L} = \vec{r} \times \vec{p} = mr^2 \sin^2 \theta \dot{\phi} \vec{e}_\theta + mr^2 \dot{\theta} \vec{e}_\phi. \quad (5.8)$$

También es importante darse cuenta que el componente z en coordenadas cartesianas es igual que el componente θ en coordenadas esféricas:

$$L_z = m(x\dot{y} - y\dot{x}) = mr^2 \sin^2 \theta \dot{\phi} = L_\theta, \quad (5.9)$$

donde en la segunda igualdad hemos utilizado el cambio de coordenadas (5.1).

Finalmente, para completitud damos las expresiones para el gradiente, la divergencia, el rotacional y la divergencia en coordenadas esféricas:

$$\vec{\nabla}\psi = \partial_r \psi \vec{e}_r + \frac{1}{r} \partial_\theta \psi \vec{e}_\theta + \frac{1}{r \sin \theta} \partial_\phi \psi \vec{e}_\phi, \quad (5.10)$$

$$\vec{\nabla} \cdot \vec{A} = \frac{1}{r^2} \partial_r (r^2 A_r) + \frac{1}{r \sin \theta} \partial_\theta (\sin \theta A_\theta) + \frac{1}{r \sin \theta} \partial_\phi A_\phi, \quad (5.11)$$

$$\begin{aligned} \vec{\nabla} \times \vec{A} = & \frac{1}{r \sin \theta} \left(\partial_\theta (\sin \theta A_\phi) - \partial_\phi A_\theta \right) \vec{e}_r + \left(\frac{1}{r \sin \theta} \partial_\phi A_r - \frac{1}{r} \partial_r (r A_\phi) \right) \vec{e}_\theta \\ & + \frac{1}{r} \left(\partial_r (r A_\theta) - \partial_\theta A_r \right) \vec{e}_\phi, \end{aligned} \quad (5.12)$$

$$\Delta \psi = \frac{1}{r^2} \partial_r (r^2 \partial_r \psi) + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \partial_\theta \psi) + \frac{1}{r^2 \sin^2 \theta} \partial_\phi^2 \psi. \quad (5.13)$$

5.2. Reducción del problema de dos cuerpos

Una primera observación importante es que bajo ciertas circunstancias, un sistema de dos partículas con interacciones se puede reducir a un problema de una sola partícula en un potencial. Al ser así, muchos problemas de dos cuerpos se reducen a un problema de una partícula.

Considera dos partículas con masas m_1 y m_2 y posiciones \vec{r}_1 y \vec{r}_2 que no están sometidas a fuerzas externas, ni a ligaduras y que interactúan mutuamente a través un potencial de interacción que sólo depende de la posición relativa $\vec{r} = \vec{r}_2 - \vec{r}_1$, y posiblemente de la velocidad relativa $\vec{v} = \vec{v}_2 - \vec{v}_1$. Ejemplos de sistemas de este tipo son dos masas conectadas a través de un muelle o dos masas moviéndose en su propio campo gravitatorio. El lagrangiano de este sistema está dado por

$$L = \frac{1}{2}(m_1 + m_2)(\dot{\vec{r}}_{cm})^2 + \frac{1}{2}m_1(\dot{\vec{r}}_1)^2 + \frac{1}{2}m_2(\dot{\vec{r}}_2)^2 - V(\vec{r}, \vec{v}), \quad (5.14)$$

donde en el término de la energía cinética reconocemos la energía cinética del centro de masas y la energía cinética de cada uno de los componentes, relativo al centro de masas (vease sección 3.1).

Los vectores de posición relativo al centro de masas \vec{r}_1 y \vec{r}_2 están relacionados con el vector de posición relativo $\vec{r} = \vec{r}_2 - \vec{r}_1$ a través de (vease figura 5.2) (ejerc.):

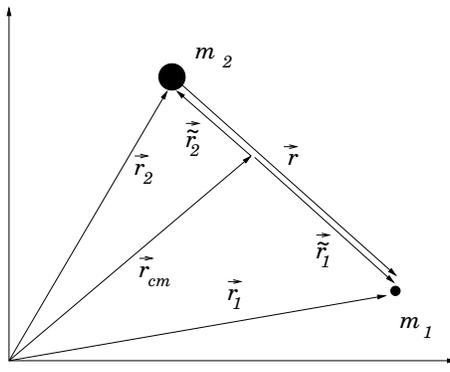


Figura 5.2: Un sistema de dos masas m_1 y m_2 : el vector de posición de centro de masas \vec{r}_{cm} , los vectores de posición relativo al centro de masas \vec{r}_1 y \vec{r}_2 y el vector de posición relativa $\vec{r} = \vec{r}_2 - \vec{r}_1$.

$$\vec{r}_1 = -\frac{m_2}{m_1 + m_2}\vec{r}, \quad \vec{r}_2 = \frac{m_1}{m_1 + m_2}\vec{r}, \quad (5.15)$$

tal que en términos del vector de posición relativo \vec{r} y el vector de posición de centro de masas \vec{r}_{cm} , podemos escribir el lagrangiano (5.14) como

$$L = \frac{1}{2}(m_1 + m_2)(\dot{\vec{r}}_{cm})^2 + \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}(\dot{\vec{r}})^2 - V(\vec{r}, \vec{v}). \quad (5.16)$$

Una primera observación es que \vec{r}_{cm} no sólo es una coordenada cíclica, sino que también se desacopla completamente del resto del sistema. Su ecuación de movimiento $\ddot{\vec{r}}_{cm} = 0$ dice que el centro de masas se mueve con una velocidad constante, independientemente de como se comportan las masas m_1 y m_2 . (Esto era de esperar, dado que no hay fuerzas externas ni ligaduras que actúan sobre el sistema.) Podemos por lo tanto, meternos en el sistema de referencia que se mueve con la misma velocidad que el centro de masas y el lagrangiano se reducirá a

$$L = \frac{1}{2}M(\dot{\vec{r}})^2 - V(\vec{r}, \vec{v}). \quad (5.17)$$

La segunda observación es que la expresión (5.17) es en realidad el lagrangiano de una sola partícula con masa

$$M = \frac{m_1 m_2}{m_1 + m_2} \quad (5.18)$$

moviéndose en el potencial $V(\vec{r}, \vec{v})$. Lo que hemos hecho, en otras palabras es meternos en el sistema de la masa m_2 , que ve la otra masa moverse alrededor de ella, pero comportándose como tuviera una masa efectiva M . La masa efectiva M se llama usualmente la *masa reducida*.

Fijaos que si $m_1 \ll m_2$ tenemos que $M \approx m_1$ y $\vec{r}_1 = -\vec{r}$, mientras $\vec{r}_2 = 0$. En este caso tenemos que el centro de masas (casi) coincide con la posición de m_2 , que a penas está influenciado por m_1 . Este es el caso de una estrella masiva con un planeta orbitándola.

Hemos pues conseguido reducir el problema de dos cuerpos al problema más sencillo de una sola partícula. Este resultado ya era conocido por Newton. Sin embargo el tratamiento general de N masas moviéndose en el potencial generado por estas mismas es mucho más complicado. De hecho, el matemático francés Jules Henri Poincaré (1854-1912) demostró en 1889 que el problema de 3 partículas no tiene solución expresable con funciones uniformes para condiciones iniciales generales. Algunas soluciones específicas sí son conocidas, correspondiendo al caso de que una o dos de las masas sean muy pequeñas en comparación con la otra o al caso de que las condiciones iniciales sean muy especiales. Un ejemplo del primer caso es de un planeta cerca de un sistema de una estrella doble, lo que se resuelve relativamente fácilmente con teoría de perturbaciones. Ejemplos del segundo caso fueron ya dados por Lagrange, que demostró que tres masas pueden recorrer orbitas

elípticas (y por lo tanto periódicas) alrededor del centro de masas si uno elige adecuadamente las condiciones iniciales. También demostró que tres cuerpos posicionados en las vértices de un triángulo equilátero o en posiciones colineales tienen un movimiento metaestable que mantiene estas posiciones. Son los llamados *puntos de Lagrange* y ejemplos de estos casos se encuentran en el sistema solar y la navegación espacial: el grupo de asteroideos los Troyanos se encuentran en los puntos de Lagrange del sistema Jupiter-Sol y se utilizan los puntos de Lagrange del sistema Tierra-Sol y Tierra-Luna para colocar satélites en órbitas estables y económicas. Finalmente es interesante observar que las soluciones periódicas de sistemas de N partículas son inestables y que el sistema corresponde a un sistema caótico: pequeñas variaciones en las condiciones iniciales llevan a desviaciones arbitrariamente grandes en escalas de tiempo finitas. Esto en particular también se aplica a nuestro sistema solar.

5.3. El lagrangiano y las ecuaciones de movimiento

Consideramos ahora el lagrangiano de una partícula en un potencial $V(r)$ en coordenadas esféricas:

$$L = \frac{1}{2}m (\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2) - V(r). \quad (5.19)$$

Ya hemos visto en la sección anterior que el lagrangiano también corresponde a un sistema de dos partículas con interacciones mutuas, donde m juega el papel de masas reducida. Por el resto de este capítulo consideraremos el caso de una sola partícula, pero es útil saber que en cualquier momento se puede volver a traducir los resultados al caso de dos masas.

Las ecuaciones de movimiento de este sistema son

$$m\ddot{r} = mr\dot{\theta}^2 + mr \sin^2 \theta \dot{\phi}^2 - \frac{\partial V}{\partial r}, \quad (5.20)$$

$$\frac{d}{dt}(mr^2\dot{\theta}) = \frac{1}{2}mr^2 \sin \theta \cos \theta \dot{\phi}^2, \quad (5.21)$$

$$\frac{d}{dt}(mr^2 \sin^2 \theta \dot{\phi}) = 0. \quad (5.22)$$

Debido a la simetría esférica del sistema, la única componente de la fuerza es en la dirección radial. (Fijaos que no decimos si apunta hacia fuera o hacia dentro. Esto depende de la forma específica de $V(r)$.) De hecho, ya hemos visto en el capítulo anterior que las ecuaciones de θ y ϕ son las ecuaciones geodésicas en coordenadas esféricas.

La consecuencia de esta simetría esférica es que el momento angular \vec{L} se conserva, componente por componente. Esto impone restricciones sobre la posición y el momento de la partícula: el momento angular $\vec{L} = \vec{r} \times \vec{p}$ sólo se conserva si tanto \vec{r} como \vec{p} están restringidos al plano perpendicular a \vec{L} . El movimiento en un potencial central por lo tanto siempre es un movimiento plano.

Con esta observación se puede simplificar bastante las ecuaciones de movimiento (5.20)-(5.22). Sin pérdida de generalidad podemos elegir el origen del sistema de coordenadas en el plano perpendicular a \vec{L} y la dirección $\theta = 0$ tal que coincide con la orientación de \vec{L} (o equivalentemente, en coordenadas cartesianas, elegimos la dirección z perpendicular al plano de movimiento). En este sistema de coordenadas, todo el movimiento toma lugar en el plano $\theta = \pi/2$ y los únicos grados de libertad son por lo tanto las coordenadas polares planas r y ϕ . Las ecuaciones de movimiento se reducen a

$$m\ddot{r} = mr\dot{\phi}^2 - \frac{\partial V}{\partial r}, \quad (5.23)$$

$$\frac{d}{dt}(mr^2\dot{\phi}) = 0, \quad (5.24)$$

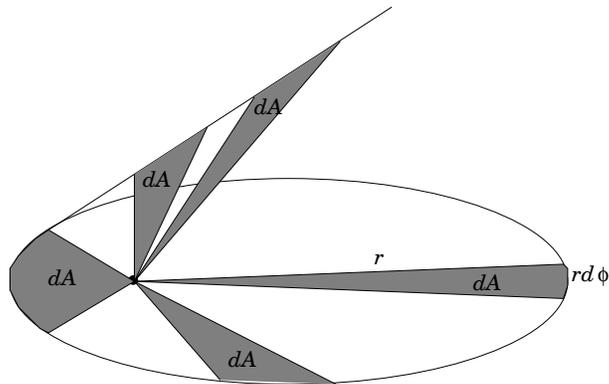


Figura 5.3: La segunda Ley de Kepler: en intervalos iguales de tiempo dt el vector de posición recorre superficies iguales $dA = r^2 d\phi$. Es independiente de la forma del potencial central y (por lo tanto) sigue válida para una partícula libre, que se mueve en línea recta.

lo que corresponde a las ecuaciones de movimiento del lagrangiano

$$L = \frac{1}{2}m (\dot{r}^2 + r^2\dot{\phi}^2) - V(r). \quad (5.25)$$

Discutamos ahora una por una las ecuaciones de movimiento. La más sencilla es la de ϕ , ya que ϕ es una coordenada cíclica. La ecuación (5.24) dice que la cantidad $\ell = mr^2\dot{\phi}$ se conserva. Por (5.8) sabemos que ℓ corresponde a la componente θ (y por lo tanto también z) del momento angular y por nuestra elección de la orientación de la base también es la norma del momento angular total $|\vec{L}| = \ell$.

Del hecho que ℓ es constante podemos deducir una ley de la mecánica celeste. La cantidad

$$\frac{dA}{dt} = \frac{1}{2}r^2\dot{\phi} = \frac{\ell}{2m} \quad (5.26)$$

es obviamente constante. Aquí $dA = \frac{1}{2}r^2d\phi$ es la superficie recorrida por el vector de posición \vec{r} en el intervalo dt , tal que (5.26) dice que la velocidad sectorial dA/dt es constante. Es la *segunda ley de Kepler* o la *ley de las áreas*: en intervalos de tiempo iguales el vector de posición recorre áreas iguales. Esto implica que la masa se moverá más rápido cerca del centro del potencial y más lento lejos (véase Figura 5.3). Johannes Kepler (1571-1630) formuló esta ley (junto con dos más) para el movimiento de los planetas en un potencial gravitatorio $V(r) \sim 1/r$, basado en las observaciones que había hecho el astrónomo danés Tycho Brahe (1546-1601), sin la ayuda de un telescopio. Sin embargo a base de la derivación que acabamos de hacer, está claro que la segunda ley es independiente de la forma del potencial y es solamente una manifestación de la conservación de momento angular. Efectivamente, incluso es válido en ausencia de potenciales, con la masa moviéndose en línea recta (Véase Figura 5.3).

Pasamos ahora a la ecuación de r . Substituyendo $\dot{\phi}$ por su expresión equivalente en función del momento angular ℓ , obtenemos para (5.23)

$$m\ddot{r} = \frac{\ell^2}{mr^3} - \frac{\partial V}{\partial r}. \quad (5.27)$$

Esta es una ecuación diferencial ordinaria de segunda orden y el problema se ha reducido al de una partícula unidimensional en un potencial efectivo

$$U(r) = V(r) + \frac{\ell^2}{2mr^2}, \quad (5.28)$$

donde el término debido al momento angular se llama *energía centrífuga* y representa la influencia de ϕ en el movimiento en la dirección r . Veremos que este término hará que la masa no se pueda acercarse arbitrariamente cerca al centro de fuerzas.

Aparte del momento angular, hay otra cantidad conservada: la energía, dado que el lagrangiano no depende explícitamente del tiempo y las fuerzas son conservativas. La energía total es la suma de la energía cinética y la energía potencial

$$\begin{aligned} E &= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + V(r) \\ &= \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r) \\ &= \frac{1}{2}m\dot{r}^2 + U(r). \end{aligned} \quad (5.29)$$

También aquí vemos que la energía se reduce a la energía de una partícula unidimensional en un potencial $U(r)$.

En lugar de intentar resolver la ecuación (5.27), es más fácil (y equivalente¹) utilizar las expresiones de las cantidades conservadas E y ℓ , ya que estas consisten de dos ecuaciones diferenciales de primera orden:

$$mr^2\dot{\phi} = \ell, \quad (5.30)$$

$$\frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r) = E. \quad (5.31)$$

Estas dos ecuaciones se pueden integrar y resolver el problema (en principio). De la ecuación para r sacamos que

$$\dot{r} = \sqrt{\frac{2}{m}\left(E - V(r) - \frac{\ell^2}{2mr^2}\right)}, \quad (5.32)$$

así que integrando obtenemos la siguiente expresión para t en función de la distancia del centro:

$$t = \int_{r_0}^r \frac{dr'}{\sqrt{\frac{2}{m}\left(E - V(r') - \frac{\ell^2}{2mr'^2}\right)}}. \quad (5.33)$$

En realidad estamos interesados en la expresión de r en función de t , pero en principio la función obtenido al calcular la integral (5.33) es invertible. Del mismo modo podemos deducir de (5.30) la expresión para ϕ :

$$\phi = \int_{t_0}^t \frac{\ell}{mr^2(t')} dt', \quad (5.34)$$

donde la expresión para $r(t)$ se obtiene de (5.33).

En la práctica, estamos más interesados en una expresión para r en función de ϕ que en la expresión de r y ϕ en función de t , puesto que la función $r(\phi)$ nos da la ecuación de la trayectoria. Esta expresión se puede derivar fácilmente de (5.30) y (5.32). Tenemos que

$$\frac{d\phi}{dr} = \frac{d\phi}{dt} \frac{dt}{dr} = \frac{\ell}{mr^2} \frac{1}{\sqrt{\frac{2}{m}\left(E - V(r) - \frac{\ell^2}{2mr^2}\right)}} \quad (5.35)$$

tal que integrando, obtenemos

$$\phi - \phi_0 = \int_{r_0}^r \frac{\ell dr'}{r'^2 \sqrt{2mE - 2mV(r') - \frac{\ell^2}{r'^2}}}. \quad (5.36)$$

¹Se puede derivar la expresión para la energía total (5.29) de la ecuación de movimiento de r . Multiplicando (5.27) con \dot{r} se puede reescribir la ecuación como una derivada total con respecto al tiempo. Al integrar, la cantidad conservada que se encuentra es justamente (5.29).

Igual que (5.33), esta expresión es en principio invertible, tal que se puede derivar la expresión inversa para $r(\theta)$. Con la expresión (5.36) (o equivalentemente (5.33) y (5.34)) hemos resuelto el problema de las trayectorias de una masa en un potencial arbitrario $V(r)$, por lo menos formalmente.

En la práctica sin embargo, resulta muy difícil integrar las expresiones (5.33), (5.34) y (5.36), puesto sólo para unos pocos potenciales $V(r)$ podemos expresar los integrales como funciones conocidas. Una primera observación es que la integral (5.36) se simplifica bastante al hacer el cambio de variables $u = r^{-1}$. Tenemos entonces que

$$\phi - \phi_0 = \int_{u_0}^u \frac{\ell du'}{\sqrt{2mE - 2mV(u') - \ell^2 u'^2}}. \quad (5.37)$$

Una clase de potenciales importantes son las potenciales de la forma

$$V = ar^p = au^{-p}, \quad (5.38)$$

desde el punto de vista físico porque aparecen con cierta frecuencia en la naturaleza y desde el punto de vista matemático porque hacen la expresión (5.37) tratable para ciertos valores de p .

El caso más sencillo es cuando $V(r) = 0$, es decir, una partícula libre. En este caso la integral (5.37) se reduce a

$$\phi - \phi_0 = \int \frac{dw}{\sqrt{1 - w^2}}, \quad (5.39)$$

donde hemos hecho el cambio de variables $w = \ell u / \sqrt{2mE}$. Integrando obtenemos pues para la trayectoria

$$r \sin \phi = \frac{\ell}{\sqrt{2mE}}, \quad (5.40)$$

o sea, la partícula moviéndose con velocidad constante a lo largo de la recta que pasa a una distancia $\ell / \sqrt{2mE}$ del origen (lo que era de esperar, dado que se trata de una partícula libre).

Para el potencial del oscilador armónico $V = \frac{1}{2}ar^2$ podemos escribir (5.37) como

$$\phi - \phi_0 = \int \frac{\ell dw}{\sqrt{2mEw - ma - \ell^2 w^2}}, \quad (5.41)$$

donde hemos aplicado el cambio $u^2 = w$. También aquí hemos podido reducir la integral en una forma manejable, donde la solución es

$$r^{-2} = \frac{mE}{\ell^2} \left(\sqrt{1 - \frac{a\ell^2}{mE^2}} \cos(\phi - \phi_0) + 1 \right). \quad (5.42)$$

En general se puede resolver la integral (5.37) con funciones trigonométricas para $p = -1$ y $p = -2$ y con integrales elípticas para $p = -3$ y $p = -4$. El caso $p = -1$ corresponde al potencial de Newton, que estudiaremos en más detalle en sección 5.5. Un caso físicamente interesante es el potencial $V = a/r^2 - b/r$, que representa una atracción a distancias grandes, pero una repulsión fuerte a distancias cortas. Potenciales de este tipo se encuentran en física molecular.

Hay muchos tipos de potenciales que no son de la forma polinómica (5.38) y que sin embargo son relevantes en la física, como el potencial de Yukawa $V(r) = -e^{-ar}/r$, que aparece en física nuclear, o el potencial logarítmico $V(r) = \log(ar)$ de una fuente lineal. Para estos casos es mucho más difícil resolver el integral (5.37) y en estos casos habrá que recorrer a aproximaciones o cálculo numérico. Sin embargo veremos en la siguiente sección que a pesar de que no tenemos una expresión exacta para la trayectoria, sí se puede deducir muchas propiedades cualitativas de las órbitas.

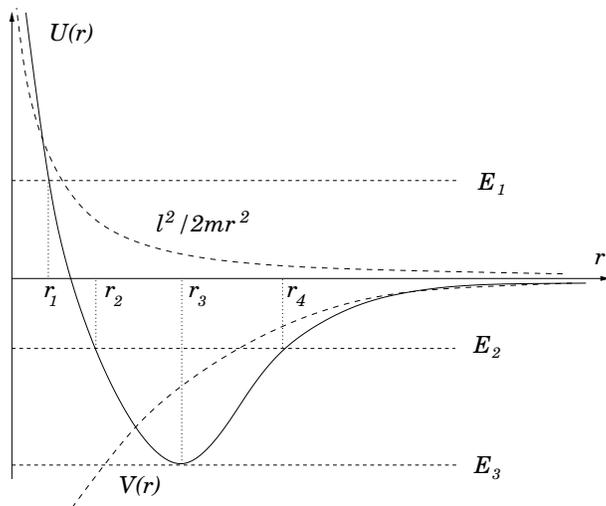


Figura 5.4: El potencial efectivo $U(r)$ es la suma del potencial $V(r)$ y el término centrífuga $\ell^2/2mr^2$. El movimiento de una partícula en un potencial $V(r)$ está restringido a las zonas donde $E \geq V(r) + \ell^2/2mr^2$. Dependiendo de las condiciones iniciales, la partícula puede acercarse o alejarse más o menos al centro del potencial.

5.4. Estudio cualitativo de las trayectorias

Hemos visto en la sección anterior que es posible reducir el problema de una partícula en un potencial central $V(r)$ a la ecuación (5.27) de una partícula unidimensional en un potencial

$$U(r) = V(r) + \frac{\ell^2}{2mr^2}. \quad (5.43)$$

Del estudio del movimiento en el potencial efectivo $U(r)$ es posible sacar mucha información sobre las trayectorias en el potencial original $V(r)$, incluso si no es posible resolver la ecuación (5.27) de manera exacta.

En el problema bi-dimensional, la energía y el momento angular vienen dados por las condiciones iniciales.² Sin embargo, para el problema uni-dimensional el momento angular ℓ no juega ningún papel, dado que es proporcional a $\dot{\phi}$. Por lo tanto, dado un valor para ℓ , queda completamente determinado el potencial uni-dimensional.

Para diferentes valores de E , la ley de conservación de energía restringe el movimiento en el potencial efectivo $U(r)$. De (5.29) tenemos para la velocidad radial

$$\dot{r} = \sqrt{\frac{2}{m} (E - U(r))}. \quad (5.44)$$

Vemos por lo tanto que la partícula con energía E sólo se puede mover por las zonas donde $E \geq V(r) + \ell^2/2mr^2$, si no la velocidad radial se volvería imaginaria.

Por ejemplo, consideremos partículas con distintas energías moviéndose en el potencial atractivo $V(r)$ que da lugar al potencial efectivo $U(r)$, como en la Figura 5.4. Por convenio hemos elegido que el potencial vale cero en el infinito. Una partícula con energía $E_1 > 0$ puede venir de lejos, acercarse al centro del potencial no más que r_1 , que es la distancia por la cual $E = U(r)$. Es

²Estará claro que la posición y la velocidad inicial implican un cierto valor para la energía y el momento angular. En muchas ocasiones es incluso más fácil dar las condiciones iniciales en función de E y ℓ que en función de posición y velocidad inicial.

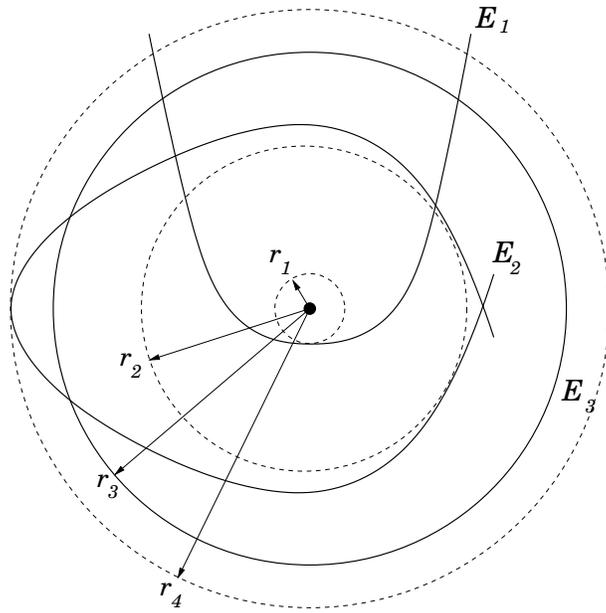


Figura 5.5: Las trayectorias de partículas con energía E_1 , E_2 y E_3 en el potencial de Figura 5.4.

el punto donde la partícula nota el efecto de la fuerza centrífuga, debido al momento angular. Observa que aunque en este punto $\dot{r} = 0$, la velocidad de la partícula no es cero, puesto que tiene todavía una velocidad angular $\dot{\phi}$ no nula. Por otro lado, para esta energía no hay cota superior para la distancia radial: la partícula puede alejarse del centro hasta el infinito (Véase Figura 5.5). Esto es el caso de un cometa que entre en el sistema solar desde el espacio interestelar, se aproxima al sol en una trayectoria parabólica, llega a una distancia mínima y sale del sistema solar para no volver nunca más.

La trayectoria de una partícula con energía $E_2 < 0$ es muy diferente, ya que ahora \dot{r} se hace zero en dos puntos distintos, r_2 y r_4 . La trayectoria por lo tanto tiene un punto de máxima aproximación y uno de máximo alejamiento, correspondiendo a r_2 y r_4 respectivamente. Este es el caso de un cometa capturado por el campo gravitatorio o la luna dando vueltas alrededor de la tierra. El hecho de que la órbita está acotada por arriba y por abajo no implica necesariamente que la órbita está cerrada.

El mínimo de energía que una partícula puede tener en el potencial $V(r)$ de Figura 5.4 es E_3 , correspondiendo al mínimo del potencial efectivo $U(r)$. La trayectoria es circular y el radio r_3 corresponde a la distancia donde la fuerza atractiva del potencial cancela justamente la fuerza centrífuga del momento angular. El ejemplo en el sistema solar que más se parece a este caso es el del planeta Venus, que sigue una trayectoria elíptica con tan baja excentricidad ($e = 0,007$) que parece casi circular.

Esta clasificación de órbitas (abiertas, acotadas y circulares) es válida para un grupo grande de potenciales atractivos, con las propiedades que tiene una singularidad menos fuerte que r^{-2} en $r = 0$ y que tiende a cero asintóticamente para r grande. En esta clase encontramos en potencial de Newton en tres dimensiones y el potencial de Yukawa. El potencial molecular $V = a/r^2 - b/r$ también cae en esta clase, para valores adecuados para a , b y ℓ .

Entre los potenciales que no satisfacen estas condiciones tenemos el potencial $V(r) = a/r^n$, con $n \geq 2$, que corresponde al potencial de Newton en dimensiones mayores que tres. En este caso el comportamiento del potencial efectivo cerca de cero está dominado por $V(r)$ y el efecto de la energía centrífuga $\ell^2/2mr^2$ es despreciable (vease Figura 5.6). Para una partícula con energía

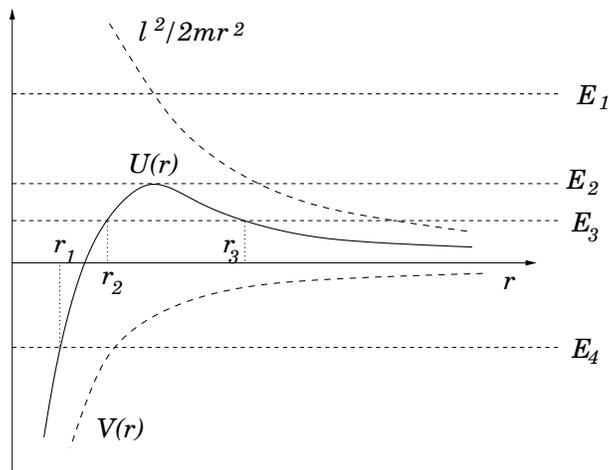


Figura 5.6: Los movimientos de partícula con distintas energías en en potencial $V(r) = ar^{-n}$, con $n \geq 2$.

E_1 no existe ninguna restricción: la partícula puede recorrer todas las distancias entre el centro y infinito. Para una partícula con energía E_3 , su movimiento depende de su posición inicial: si está cerca del centro del potencial, puede moverse entre el centro y r_2 , mientras si está lejos, puede salir hasta el infinito, pero no acercarse más que una distancia r_3 . Una partícula con energía E_4 sólo puede moverse en la región entre el centro y r_1 . Finalmente fijaos que una órbita circular aquí también es posible, si la partícula tiene una energía E_2 , pero al contrario con el potencial de Figura 5.4, este movimiento es inestable frente a pequeñas perturbaciones y decaerá al caso con energía E_1 o uno de los casos de energía E_3 .

Finalmente, potenciales que no tienden a un valor asintótico, sino tienden a infinito para r grande³ siempre tienen órbitas acotadas. Ejemplos de estos potenciales son el oscilador armónico y el potencial logarítmico.

5.5. El problema de Kepler

A Newton y sus contemporáneos les costó mucho adivinar como la fuerza de la gravedad varía con la distancia del sol. Históricamente, el primero en sugerir que la fuerza varía inversamente proporcional al cuadrado de la distancia fue Robert Hooke (1635-1703), pero más o menos al mismo tiempo fue propuesto también por gente como Newton y Christopher Wren (1632-1723). El mérito de Newton era derivar las leyes de Kepler de la fórmula de la fuerza gravitacional y darse cuenta que la gravedad es una fuerza universal entre dos masas cualesquiera.

Hoy en día es más fácil derivar la forma del potencial gravitacional, gracias al teorema de Gauss. Una masa situada en un punto o es la fuente de un campo vectorial de fuerzas saliendo de (o dirigido hacia) este punto. El flujo del campo vectorial a través de esferas concéntricas con centro en o tiene que ser constante, dado que no hay otras fuentes de campo. Pero como la superficie de las esferas crece con r^2 , la fuerza tiene que variar con el inverso de la distancia para tener el mismo flujo para todas las esferas. Esta es la razón por la cual las potenciales $V(r) \sim 1/r$ aparecen en tantas áreas distintas de la física, como la ley de Newton para gravedad y la ley de Coulomb para electromagnetismo: son simplemente aplicaciones del teorema de Gauss con distintos campos vectoriales. Fijaos que la forma de $V(r) \sim 1/r$ es típica para tres dimensiones espaciales. En

³Potenciales que tienden a $-\infty$ para r grandes no son físicos, puesto que no están acotados por debajo. una partícula ganaría más y más energía, simplemente alejándose del centro del potencial.

espacios D -dimensionales la superficie de una hipersfera es proporcional a r^{D-1} , tal que la ley de Gauss predice potenciales de la forma $V(r) \sim 1/r^{D-1}$ para estos casos.

En esta sección vamos a prestar especial atención al potencial de Newton (o el de Coulomb) $V(r) = k/r$ y a las trayectorias de partículas en este potencial. En particular queremos resolver el conocido *problema de Kepler*, es decir, derivar las tres leyes de Kepler de las ecuaciones de movimiento de una partícula en este potencial. Recordamos que las leyes de Kepler son:

1. Las planetas se mueven alrededor del sol en órbitas elípticas, donde el sol se encuentra en uno de los focos del elipse.
2. El vector de posición del planeta recorre superficies iguales en intervalos iguales de tiempo.
3. El cuadrado del periodo de un planeta es proporcional al cubo del tamaño del eje largo de la órbita, con la constante de proporcionalidad K , la *constante de Kepler*, igual para todos los planetas.

Ya hemos visto que la segunda ley es simplemente un consecuencia de la conservación de momento angular, válida en cualquier potencial central. Las otras dos leyes son específicas del potencial $V(r) = -kr^{-1}$.

Rellenando $V(u) = -ku$ en la ecuación de la órbita (5.37), podemos resolver la integral directamente

$$\phi = \int \frac{\ell du}{\sqrt{2mE + 2mku - \ell^2 u^2}} = \arccos \frac{-2mk + 2\ell^2 u}{2mk\sqrt{1 + 2\ell^2 E/mk^2}}, \quad (5.45)$$

lo que en términos de r se reduce a

$$r^{-1} = \frac{mk}{\ell^2} \left(1 + \sqrt{1 + \frac{2\ell^2 E}{mk^2}} \cos \phi \right). \quad (5.46)$$

Reconocemos en (5.46) la ecuación de una sección cónica

$$r = \frac{p}{1 + e \cos \phi}, \quad (5.47)$$

si identificamos la excentricidad e y la constante p como

$$e = \sqrt{1 + \frac{2\ell^2 E}{mk^2}}, \quad p = \frac{\ell^2}{mk}. \quad (5.48)$$

De las propiedades de secciones cónicas sabemos para la ecuación (5.47) representa una hipérbola para $e > 1$ ($E > 0$), una parábola para $e = 1$ ($E = 0$), un elipse para $e < 1$ ($E < 0$) y un círculo para $e = 0$ ($E = -mk^2/2\ell^2$).

Para un potencial atractivo ($k > 0$), como en el caso del potencial gravitatorio, este análisis coincide con la clasificación de órbitas que hemos hecho en la sección anterior. Los planetas tienen trayectorias acotadas, con energía $E < 0$. Ahora vemos que sus órbitas corresponden a elipses, tal como dice la primera ley de Kepler.

Para un potencial repulsivo ($k < 0$), como el potencial de Coulomb entre dos partículas con el mismo signo de carga eléctrica, sabemos por el análisis de sección 5.4 que la energía siempre es positiva, tal que las únicas trayectorias posibles son hipérbolas.

Nos concentraremos ahora en el caso de una partícula en una órbita acotada en un potencial atractivo, para derivar la tercera ley de Kepler. De las propiedades de elipses sabemos que el semi-eje mayor a y el semi-eje menor b están relacionados con e y p como

$$a = \frac{p}{1 - e^2} = -\frac{k}{2E}, \quad b = a\sqrt{1 - e^2} = \frac{\ell}{\sqrt{-2mE}}. \quad (5.49)$$

Vice versa, la excentricidad e expresada en función de a es

$$e = \sqrt{1 - \frac{p}{a}} = \sqrt{1 - \frac{\ell^2}{mka}}. \quad (5.50)$$

La manera más fácil de derivar la tercera ley de Kepler es saliendo de la segunda ley. Integrando (5.26) sobre el intervalo de un periodo $[0, T]$ obtenemos que

$$T = \frac{2m}{\ell} \pi ab = 2\pi \sqrt{\frac{m}{k}} a^{3/2}, \quad (5.51)$$

donde en la primera igualdad hemos utilizado que el área de un elipse es $A = \pi ab$ y en la segunda las expresiones (5.49) y (5.50) de b y e en función de a .

De (5.51) vemos que efectivamente el cuadrado del periodo es proporcional al cubo del eje grande del elipse, pero la constante de Kepler

$$K = \frac{4\pi^2 m}{k} \quad (5.52)$$

depende de m y no puede ser igual por todos los planetas. Por lo tanto la tercera ley de Kepler sólo se cumple de forma débil para los potenciales $V(r) \sim 1/r$, en el sentido que $T^2 \sim a^3$, pero la constante no es universal. Por ejemplo para el potencial de Coulomb, $k = cqQ$, donde Q y q son las cargas eléctricas de la partículas y $c = 8,98 \cdot 10^9 \text{ kg m}^3 / \text{C}^2 \text{ s}^2$ la constante eléctrica. Por lo tanto vemos que la constante de Kepler en electromagnetismo depende tanto de la masa como de la carga de las partículas.

Para el potencial gravitatorio tenemos que $k = Gm_1 m_2$, donde m_1 y m_2 son (en el caso del sistema solar) las masas del sol y del planeta. Sin embargo la masa m que aparece en (5.52) es la masa reducida (5.18). En el caso del sol y los planetas sabemos que la masa del planeta m_1 es mucho menor que la masa del sol m_2 , tal que la masa reducida m es primera aproximación es igual a m_1 (vease la discusión de sección 5.2). Por lo tanto, la expresión para la constante de Kepler se reduce a

$$K = \frac{4\pi^2 m}{Gm_1 m_2} \approx \frac{4\pi^2}{Gm_2}. \quad (5.53)$$

Efectivamente, en esta aproximación la constante de Kepler sólo depende de la masa m_2 del sol, o en general, de la masa central que da origen al potencial.

En el caso de que las dos masas m_1 y m_2 sean comparables (por ejemplo en una estrella doble), esta aproximación ya no es válida y hay que volver a traducir muchos de los resultados al problema de dos partículas. Por ejemplo para $m_2 = 2m_1$ tenemos por (5.15) que $\vec{r}_1 = -2\vec{r}_2$ y si \vec{r} sigue una trayectoria elíptica, también lo harán \vec{r}_1 y \vec{r}_2 (vease Figura 5.7). Aunque la versión débil de la tercera ley de Kepler es válida (con $K = 4\pi^2 / 3Gm_1$) en la formulación equivalente de una partícula de masa reducida, no es válida para cada una de las masas m_1 y m_2 por separada. Es fácil de ver que, para que el centro de masas siga situado en el origen el periodo de m_1 tiene que ser el mismo que el de m_2 , a pesar de tener una órbita más pequeña.

También la segunda ley de Kepler es más difícil, puesto el momento angular total es la suma del los momentos angulares de cada una de las partículas. Y aunque el momento angular total se conserva (y por lo tanto en el sistema de la masa reducida vale la ley de las áreas), esto no es verdad para cada momento angular por separado.

El problema de Kepler es importante desde el punto de vista histórico, puesto que era la primera vez donde se pone una teoría física a prueba con experimentos (en este caso observaciones): por un lado las leyes de Kepler eran puramente empíricas, sacadas de datos observacionales y describían el movimiento de los planetas sin explicar el porqué. Por otro lado la teoría de Newton era la primera teoría fundamental, basado en deducciones a base de principios. La importancia del problema de Kepler está justamente en que Newton era capaz de demostrar que sus ideas podían explicar el movimiento de los planetas.

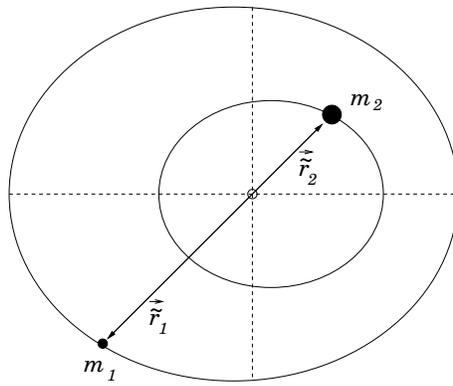


Figura 5.7: Un sistema de estrellas dobles con masas $m_2 = 2m_1$.

5.6. Problemas

1. Comprueba que la definición del producto escalar (??) es consistente con el cambio de coordenadas (??)-(??) y la definición (??).
2. Averigua que las $\Gamma_{\mu\nu}^\rho$ definidas en (1.65) son simétricas en μ y ν . Comprueba también la fórmula (1.67). Comprueba que $\nabla_\mu A_\nu$ transforma como un tensor de rango dos, dado que A_μ es un vector.
3. Calcula la expresión (5.5) para la métrica plana en coordenadas esféricas, aplicando el cambio de coordenadas (5.1) a la métrica en coordenadas cartesianas $ds^2 = dx^2 + dy^2 + dz^2$. Calcula también las componentes de la conexión de Levi-Civita para coordenadas esféricas.
4. Calcula los componentes del momento angular en coordenadas cartesianas explícitamente y a través del cambio de coordenadas (5.1). Calcula la norma de \vec{L} en cartesianas y en esféricas y compara las dos expresiones.
5. Se puede demostrar (a través del teorema de Bertrand) que sólo hay dos potenciales con la propiedad que todas las trayectorias acotadas son cerradas. Uno es el conocido potencial de Newton, el otro es el potencial del oscilador armónico $V(r) = \frac{1}{2}kr^2$. Deriva para este último potencial el equivalente de las leyes de Kepler. En otras palabras:
 - a) ¿Cuál es la forma de las órbitas?
 - b) ¿Existe algún equivalente de la ley de las áreas?
 - c) ¿Cuál es la relación entre el periodo y el tamaño de la órbita?

Pista: En lugar de intentar resolver las integrales de sección 5.3, es más fácil plantarse el problema en coordenadas cartesianas. Fijaos que la ecuación de un elipse en estas coordenadas viene dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (5.54)$$

6. Considera las dos masas en una mesa con agujero, introducido en el problema 9 de sección 3. Haz un estudio cualitativo del movimiento del sistema y determina la condición para que la masa m_2 permanece a la misma altura. Dado una energía E y un momento angular ℓ , ¿cuál es el valor mínimo y máximo de r ?
7. Demuestra que ecuación (5.40) corresponde efectivamente a una partícula con energía E y momento angular ℓ .

8. El potencial de interacción (clásica) entre dos moléculas neutras, separadas por una distancia r , viene dado por

$$V(r) = -\frac{2a}{r^6} + \frac{b}{r^{12}}. \quad (5.55)$$

Calcula la distancia de equilibrio y la energía que corresponde con este estado. ¿Cuál es la frecuencia de oscilación alrededor de la posición de equilibrio si el sistema tiene una energía ligeramente más alta que la del equilibrio?

9. Excavamos un túnel penetrando de tierra, conectando dos puntos antípodas. Si suponemos que la densidad de la tierra es constante y paramos la rotación diaria alrededor de su eje, describe el movimiento de una piedra que cae por el túnel. ¿Qué pasaría si no despreciamos la rotación de la tierra?

Capítulo 6

Teoría de Maxwell

La teoría de Maxwell es una teoría de campos, es decir una teoría con infinitos grados de libertad. El concepto físico de un campo fue introducido por Micheal Faraday (1791-1869), un autodidacta que aprendió física leyendo los libros que tenía que encuadernar en la imprenta donde trabajaba. Faraday había observado que el serrín metálico se agrupaba según filamentos en la proximidad de cargas y corrientes eléctricas y creía que estas cargas y corrientes causaban campos que llenan el espacio, actuando sobre otras cargas y corrientes. En este sentido un campo eléctrico es un intermediario a través del cual interaccionan las partículas y resuelve el problema conceptual que había introducido Newton al postular la acción a distancia de la gravedad.

Las teorías de campos forman la base de la física moderna de altas energías. Toda la física de partículas está basada en la teoría cuántica de campos, mientras en cierto modo la relatividad general no es más que formular la gravedad en lenguaje de una teoría de campos.

6.1. El límite a sistemas continuos

Antes de tratar a fondo el electromagnetismo como teoría de campos, intentaremos hacer la conexión con la mecánica analítica que hemos estudiado en los capítulos anteriores. Desde el punto de vista mecánico, una teoría de campos no es nada más que una teoría con un número infinito (continuo) de grados de libertad. Son aplicables por lo tanto las mismas técnicas que ya conocemos de (por ejemplo) el formalismo lagrangiano, sólomente tomando en cuenta la sutileza de tomar de manera adecuada el límite continuo. En esta sección revisaremos cómo tomar este límite.

Considera un sistema que consiste de una serie infinita (pero contable) de masas iguales alineadas a lo largo del eje x y conectadas por muelles identicos de tamaño L_0 y constante elástica k . Supondremos además que las masas sólo se pueden mover en la dirección x (Véase Figura 6.1). Tomamos como coordenadas generalizadas $q_\alpha(t)$ la posición de la α -ésima masa m

$$q_\alpha(t) = \alpha L_0 + \phi_\alpha(t), \quad (6.1)$$

donde el indice α corre de $-\infty$ a ∞ y $\phi_\alpha(t)$ mide la desviación de la posición de equilibrio. Las velocidades generalizadas por lo tanto vienen dadas por $\dot{q}_\alpha = \dot{\phi}_\alpha$. La energía potencial es proporcional al cuadrado de la desviación de los muelles del tamaño de equilibrio y viene dado por

$$V = \frac{1}{2} \sum_{\alpha} k(q_{\alpha+1} - q_\alpha - L_0)^2 = \frac{1}{2} \sum_{\alpha} k(\phi_{\alpha+1} - \phi_\alpha)^2. \quad (6.2)$$

Podemos por lo tanto escribir el lagrangiano como

$$L = \frac{1}{2} \sum_{\alpha} m \dot{\phi}_\alpha^2 - \frac{1}{2} \sum_{\alpha} k(\phi_{\alpha+1} - \phi_\alpha)^2$$

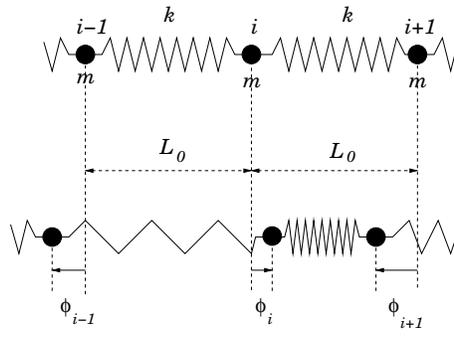


Figura 6.1: Una serie de masas m están conectadas a través de muelles con constante k y posición de equilibrio L_0 . En la situación de equilibrio (arriba) las masas están separadas por la distancia L_0 . En la situación general (abajo) el desplazamiento de la masa m de la posición de equilibrio está caracterizado por ϕ_α .

$$= \frac{1}{2} \sum_{\alpha} L_0 \left[\frac{m}{L_0} \dot{\phi}_{\alpha}^2 - k L_0 \left(\frac{\phi_{\alpha+1} - \phi_{\alpha}}{L_0} \right)^2 \right], \quad (6.3)$$

donde en la última igualdad hemos sacado un factor L_0 por razones que se harán claras un poco más adelante. Las ecuaciones de movimiento vienen dadas por

$$\frac{m}{L_0} \ddot{\phi}_{\alpha} = k L_0 \left(\frac{\phi_{\alpha+1} - \phi_{\alpha}}{L_0^2} \right) - k L_0 \left(\frac{\phi_{\alpha} - \phi_{\alpha-1}}{L_0^2} \right). \quad (6.4)$$

Aquí no estamos interesados en intentar resolver estas ecuaciones, sino queremos saber qué ocurre con el sistema y la ecuaciones de movimiento en el límite donde la posición de equilibrio L_0 tiende a cero.

Una primera observación es que en este caso la distribución de masas se hace continua y el sistema de masas conectadas por muelles tiende a una varilla elástica con cierta densidad de masa ρ . El índice discreto α que hace referencia a la α -ésima masa, se convierte en una variable continua x , indicando los puntos a lo largo de la varilla. Concretamente, la coordenada generalizada $\phi_{\alpha}(t)$ que indica el desplazamiento de la masa α se convierte en una función $\phi(x, t)$ del desplazamiento de una cantidad infinitesimal de masa dm a lo largo de la recta. De este modo hemos pasado de una teoría con un número discreto de grados de libertad a una teoría con un número continuo de grados de libertad. En otras palabras, una teoría de campos. Fijaos que, a pesar de que $\phi(x, t)$ es una función de x y de t sin que ninguna de los dos variables juegue un rol más fundamental que el otro, el origen de cada uno de las variables es muy distinto. Merece la pena enfatizar que en el límite de teoría de campos, las cantidades físicas fundamentales, las que juegan el papel de coordenada generalizada, son los campos $\phi(x, t)$ y no las posiciones x . Las posiciones x no son en cierto modo más que una etiqueta continua para distinguir los diferentes grados de libertad ϕ .

Miremos ahora como se transforman el lagrangiano (3.19) y las ecuaciones de movimiento (6.4) en el límite $L_0 \rightarrow 0$. Está claro que en este caso las cantidades $\dot{\phi}_{\alpha}(t)$ y $\ddot{\phi}_{\alpha}(t)$ se convierten en derivadas parciales con respecto al tiempo:

$$\dot{\phi}_{\alpha}(t) \longrightarrow \frac{\partial}{\partial t} \phi(x, t), \quad \ddot{\phi}_{\alpha}(t) \longrightarrow \frac{\partial^2}{\partial t^2} \phi(x, t). \quad (6.5)$$

Además, el segundo término de (6.3) es (casi) por definición e la derivada parcial con respecto a x

$$\lim_{L_0 \rightarrow 0} \frac{\phi_{\alpha+1}(t) - \phi_{\alpha}(t)}{L_0} = \lim_{L_0 \rightarrow 0} \frac{\phi(x + L_0, t) - \phi(x, t)}{L_0} = \frac{\partial}{\partial x} \phi(x, t) \quad (6.6)$$

y consecuentemente los términos del lado derecho de (6.4) se convierte en una derivada segunda:

$$\lim_{L_0 \rightarrow 0} \left[\left(\frac{\phi_{\alpha+1} - \phi_\alpha}{L_0^2} \right) - \left(\frac{\phi_\alpha - \phi_{\alpha-1}}{L_0^2} \right) \right] = \lim_{L_0 \rightarrow 0} \frac{1}{L_0} \left[\frac{\partial \phi(x + L_0, t)}{\partial x} - \frac{\partial \phi(x, t)}{\partial x} \right] = \frac{\partial^2 \phi(x, t)}{\partial x^2}. \quad (6.7)$$

Queda por lo tanto identificar el límite de los parámetros que aparecen en el lagrangiano y la ecuación de movimiento. La cantidad m/L_0 tiene dimensión de masa por unidad de longitud, y en el límite continuo se convierte en la densidad ρ de la varilla. La cantidad ka se identifica de la siguiente manera: la fuerza necesaria para extender una de las muelles la distancia $(\phi_{\alpha+1} - \phi_\alpha)$ viene dado por

$$F = k(\phi_{\alpha+1} - \phi_\alpha) = ka \cdot \frac{(\phi_{\alpha+1} - \phi_\alpha)}{a} = Y \Delta L_0, \quad (6.8)$$

donde $\Delta L_0 = (\phi_{\alpha+1} - \phi_\alpha)/a$ es la extensión por unidad de longitud y Y es el módulo de Young, la constante de proporcionalidad entre la fuerza aplicada y la extensión por unidad en materiales elásticos, una constante que depende del material considerado.

Con esta identificación de la densidad y el módulo de Young podemos escribir la acción (6.3) como

$$L = \int dx \left[\frac{1}{2} \rho (\partial_t \phi)^2 - Y (\partial_x \phi)^2 \right], \quad (6.9)$$

donde la suma sobre los desplazamientos $\sum_\alpha L_0$ se ha convertido en una integral $\int dx$ sobre la variable continua x . Del mismo modo, la ecuación de movimiento (6.4) se convierte en una ecuación de ondas

$$\rho \partial_t^2 \phi(x, t) - Y \partial_x^2 \phi(x, t) = 0, \quad (6.10)$$

que describe ondas (longitudinales) de densidad en el material de la varilla que se propagan con una velocidad $v = \sqrt{Y/\rho}$.

Podemos escribir esta integral como $L = \int dx \mathcal{L}$, donde el integrando $\mathcal{L}(\phi, \partial_x \phi, \partial_t \phi)$ es la *densidad lagrangiana*, que contiene toda la física del sistema. A su vez se define la acción S como en (3.19)

$$S(\phi, \partial_x \phi, \partial_t \phi) = \int dt L(\phi, \partial_x \phi, \partial_t \phi) = \int dt dx \mathcal{L}(\phi, \partial_x \phi, \partial_t \phi). \quad (6.11)$$

Fijaos que hemos tenido mucho cuidado en derivar el límite continuo de tanto el lagrangiano (6.10) como las ecuaciones de movimiento (6.9), sin decir todavía nada sobre cómo derivar las ecuaciones de movimiento del lagrangiano. En principio no es muy diferente de la derivación que hemos hecho en la sección 3.3, salvo que ahora hay que variar con respecto a los campos $\phi(x, t)$. La herramienta matemática necesaria para esta operación es la *derivada funcional* definida como

$$\frac{\delta \phi(x, t)}{\delta \phi(y, t')} = \delta(x - y) \delta(t - t'), \quad \frac{\delta F(\phi(x, t))}{\delta \phi(y, t')} = \frac{\partial F(\phi(x, t))}{\partial \phi(x, t)} \frac{\delta \phi(x, t)}{\delta \phi(y, t')}, \quad (6.12)$$

para cualquier función $F(\phi(x, t))$ y donde $\delta(x - y)$ es la *delta de Dirac*, la generalización continua de la delta de Kronecker. Repitiendo los mismos pasos que en el caso de una acción un número discreto de grados de libertad y tomando como condiciones de contorno que

$$\delta \phi(x, t_1) = \delta \phi(x, t_2) = \delta \phi(x_1, t) = \delta \phi(x_2, t) = 0, \quad (6.13)$$

es decir, tomando la variación igual a cero tanto en los puntos iniciales y finales como en los contornos¹, llegamos a la ecuación de Lagrange para el caso de una teoría de campos:

$$\frac{1}{c^2} \frac{\partial}{\partial t} \left(\frac{\delta \mathcal{L}}{\delta (\partial_t \phi)} \right) + \frac{\partial}{\partial x} \left(\frac{\delta \mathcal{L}}{\delta (\partial_x \phi)} \right) - \frac{\delta \mathcal{L}}{\delta \phi} = 0, \quad (6.14)$$

¹En el caso de la varilla elástica hemos tomado $x_1 = -\infty$ y $x_2 = \infty$, pero en general la integración se puede hacer tanto en intervalos finitos como infinitos.

donde c es una constante con unidad de velocidad que aparece por razones dimensionales. Concretamente para el caso de la acción (6.9) tenemos que

$$\frac{\delta\mathcal{L}\left(\phi(y, t'), \partial_y\phi(y, t'), \partial_{t'}\phi(y, t')\right)}{\delta(\partial_t\phi(x, t))} = \rho \partial_{t'}\phi(y, t')\delta(x-y)\delta(t-t') = \rho \partial_t\phi(x, t), \quad (6.15)$$

$$\frac{\delta\mathcal{L}\left(\phi(y, t'), \partial_y\phi(y, t'), \partial_{t'}\phi(y, t')\right)}{\delta(\partial_y\phi(x, t))} = Y \partial_x\phi(y, t')\delta(x-y)\delta(t-t') = Y \partial_x\phi(x, t), \quad (6.16)$$

tal que (6.14) aplicado al lagrangiano (6.9) implica la ecuación de movimiento (6.10), como hemos derivado antes calculando el límite explícitamente.

En la práctica, las derivadas funcionales, por muy complicadas que parecen, se aplican como si fueran derivadas ordinarias, olvidándose que uno en realidad está derivando con respecto a funciones. Podemos por lo tanto reducir (6.15) y (6.16) a

$$\frac{\delta\mathcal{L}}{\delta(\partial_t\phi)} = \rho \partial_t\phi, \quad \frac{\delta\mathcal{L}}{\delta(\partial_x\phi)} = \rho \partial_x\phi. \quad (6.17)$$

La generalización a dos y tres dimensiones debería ser obvia. En lugar de ser funciones de sólo x y t , los campos $\phi(x_i, t)$ van a depender en general de x_i y t y el lagrangiano será una función de $\phi(x_i, t)$ y sus derivadas $\partial_i\phi(x_i, t)$ y $\partial_t\phi(x_i, t)$. La ecuación de Lagrange (6.14) será por lo tanto de la forma

$$\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{L}}{\delta(\partial_t\phi)}\right) + \frac{\partial}{\partial x_i}\left(\frac{\delta\mathcal{L}}{\delta(\partial_i\phi)}\right) - \frac{\delta\mathcal{L}}{\delta\phi} = 0. \quad (6.18)$$

En este punto es conveniente considerar las tres coordenadas espaciales x_i y el tiempo t como las coordenadas de un espacio cuatro-dimensional.² Anotamos el vector de posición en este espacio como x_μ , donde el índice μ corre de 0 a 3, siendo $x_0 = t$. La ecuación de Lagrange simplifica en esta notación a

$$\partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\right) - \frac{\delta\mathcal{L}}{\delta\phi} = 0. \quad (6.19)$$

Esta notación resultará especialmente útil para estudiar la teoría de Maxwell en el formalismo lagrangiano, dado que las leyes de Maxwell son invariantes bajo las transformaciones de Lorentz. Sin embargo, antes de pasar al formalismo lagrangiano de la teoría de Maxwell, pasaremos primero unas secciones estudiando la interpretación y la estructura matemática de las leyes del electromagnetismo. Dejaremos la formulación covariante de la teoría de Maxwell para el capítulo 9.

6.2. Las leyes de Maxwell

El físico escocés James Clerk Maxwell (1831-1879) publicó sus cuatro leyes de Maxwell en 1865, aunque la mayoría de estas leyes ya habían sido descubiertos por Coulomb (1736-1806), Ørsted (1777-1851), Ampère (1775-1836), Biot (1774-1862), Savart (1791-1841) y Faraday (1791-1867) a base de investigación experimental. De hecho Maxwell añadió sólo un término nuevo a las ecuaciones que ahora llevan su nombre. Pero el gran logro de Maxwell fue unificar el conjunto de leyes empíricas sueltas sobre electrostática, corrientes eléctricas e inducción magnética en una sólida teoría que describe todos los fenómenos relacionados con el electromagnetismo. Y como extra resultó que su teoría era capaz de dar un fundamento teórico a la óptica, una parte de la física que hasta entonces parecía completamente desconexa de los fenómenos electromagnéticos.

²Esta simplificación viene inspirada por la estructura del espacio-tiempo tal como surge desde la teoría de la relatividad especial, pero en este punto no es más que un truco matemático para simplificar las ecuaciones. De momento no tiene ningún significado físico (todavía).

Dada una densidad de carga eléctrica $\rho(\vec{x}, t)$ y una corriente $\vec{j}(\vec{x}, t)$, las leyes de Maxwell para los campos eléctricos $\vec{E}(\vec{x}, t)$ y magnéticos $\vec{B}(\vec{x}, t)$ vienen dados en la siguiente forma

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}, \quad (6.20)$$

$$\vec{\nabla} \times \vec{E} = -\partial_t \vec{B}, \quad (6.21)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad (6.22)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \epsilon_0 \mu_0 \partial_t \vec{E}. \quad (6.23)$$

Los constantes ϵ_0 y μ_0 son respectivamente la permitividad eléctrica y la permeabilidad magnética del vacío, que en unidades adecuadas se puede igualar a 1,³ lo que haremos a partir de ahora. La única cosa relevante será que el producto $\epsilon_0 \mu_0$ tiene la dimensión de velocidad ($[\epsilon_0 \mu_0] = LT^{-1}$) y el valor de $\epsilon_0 \mu_0 = 1/c^2$, donde c es la velocidad de la luz.

Las ecuaciones de Maxwell en forma diferencial (6.20)-(6.23) consiste en 8 ecuaciones diferenciales lineales acopladas, donde en general se toma como condiciones de contorno que los campos \vec{E} y \vec{B} tienden a cero en el infinito (para sistemas finitos). Aunque la forma diferencial (6.20)-(6.23) es más conveniente para buscar soluciones de las ecuaciones, la física detrás de estas ecuaciones se ve mejor en la formulación integral. Utilizando los teoremas de Stokes sobre las integrales de la divergencia y el rotacional de un campo vectorial \vec{a}

$$\iiint_V \vec{\nabla} \cdot \vec{a} \, d^3x = \oiint_{\partial V} \vec{a} \cdot \vec{n} \, d^2x, \quad \iint_S (\vec{\nabla} \times \vec{a}) \cdot \vec{n} \, d^2x = \oint_{\partial S} \vec{a} \cdot d\vec{r}, \quad (6.24)$$

podemos reescribir las ecuaciones de Maxwell (6.20)-(6.23) en su forma integral:

$$\oiint_{\partial V} \vec{E} \cdot \vec{n} \, d^2x = \iiint_V \rho \, d^3x, \quad (6.25)$$

$$\oint_{\partial S} \vec{E} \cdot d\vec{r} = -\frac{d}{dt} \iint_S \vec{B} \cdot \vec{n} \, d^2x, \quad (6.26)$$

$$\oiint_{\partial V} \vec{B} \cdot \vec{n} \, d^2x = 0, \quad (6.27)$$

$$\oint_{\partial S} \vec{B} \cdot d\vec{r} = \iint_S \vec{j} \cdot \vec{n} \, d^2x + \frac{d}{dt} \iint_S \vec{E} \cdot \vec{n} \, d^2x. \quad (6.28)$$

La primera ley de Maxwell o la *Ley de Gauss* (6.25) nos dice que el flujo del campo eléctrico a través de una superficie cerrada es igual a la carga total encerrada dentro de la superficie. No es nada más que una versión del teorema de Gauss sobre flujos de campos vectoriales, aplicada a \vec{E} , identificando la carga eléctrica $Q = \int \rho d^3x$ como fuente del campo eléctrico. En particular, si tomamos el caso de una carga puntual e en el origen, podemos calcular el flujo de \vec{E} a través de una esfera de radio R alrededor del origen. Aprovechándonos de la simetría esférica del problema es fácil de ver que (ejerc.)

$$\vec{E} = \frac{e}{4\pi r^2} \vec{e}_r, \quad (6.29)$$

con lo que hemos recuperado la ley de Coulomb.

La tercera ley de Maxwell (6.27) nos intenta dar el equivalente magnético de la ley de Gauss. Sin embargo aquí vemos que el lado derecho de la ecuación, que correspondería al término de la

³La teoría de Maxwell no es sólo una teoría para los campos electromagnéticos en el vacío, sino también describe el comportamiento de estos campos en materiales dieléctricos, ferromagnéticos y paramagnéticos. Cada material tiene sus propias características electromagnéticas, descritas por la permitividad eléctrica ϵ_{mat} y la permeabilidad magnética μ_{mat} del material. Históricamente, Maxwell formuló sus ecuaciones para el caso de campos electromagnéticos en materiales. La primera persona que escribió las leyes de Maxwell para el vacío en su forma fundamental (6.20)-(6.23) fue el físico holandés Hendrik Anton Lorentz (1853-1928).

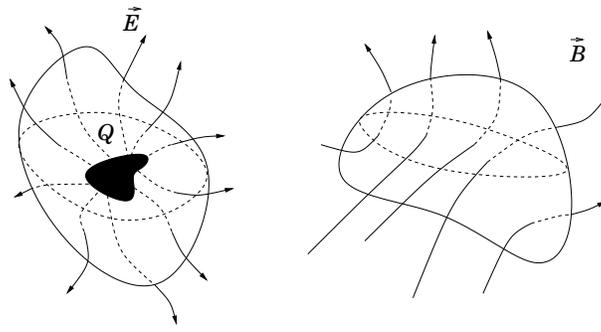


Figura 6.2: La primera y la tercera ley de Maxwell: el flujo del campo eléctrico \vec{E} a través de una superficie cerrada es proporcional a la carga total encerrada dentro de la superficie, mientras que el flujo del campo magnético \vec{B} a través de cualquier superficie cerrada es cero.

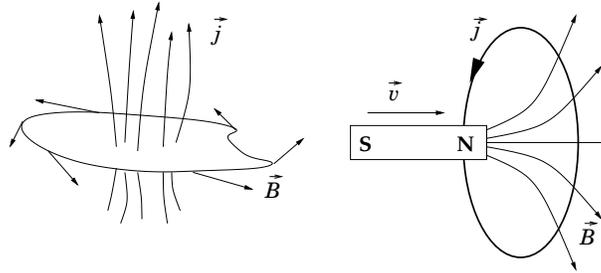


Figura 6.3: La Ley de Ampère y la Ley de Faraday: Una corriente eléctrica a través de una superficie genera un campo magnético a lo largo de la curva que borde la superficie, mientras que la variación de flujo magnético a través de una superficie induce una corriente eléctrica a lo largo de la curva.

fuente del campo magnético, es cero. La ecuación (6.27) dice que el flujo magnético a través de cualquier superficie cerrada siempre es cero. En otras palabras, no existe un equivalente magnético de una carga eléctrica. Mientras las cargas eléctricas positivas y negativas pueden existir separadas e independientes unas de otras, los polos norte y sur de un imán siempre aparecen en parejas y las líneas del campo magnético siempre se cierran o se extienden al infinito. Aunque en la física moderna existen soluciones de monopolos magnéticos con propiedades muy interesantes, no hay (de momento) ninguna indicación que estas soluciones corresponden a objetos reales en la naturaleza. Discutiremos un ejemplo de una solución de monopolo en la sección 6.5.

La Ley de Ampère, el primer término de la derecha de la ecuación (6.28), nos cuenta cuál es la fuente del campo magnético: la corriente eléctrica \vec{j} . La integral de \vec{B} a lo largo de una curva cerrada C es igual al flujo de corriente eléctrica a través de la superficie cuyo contorno es la curva C (Véase Figura 6.3 a). En particular podemos distinguir dos casos de interés. El primero es el caso del conductor lineal infinito con una corriente constante $I = \int \vec{j} \cdot \vec{n} d^2x$ (Figura 6.4 a). A través de la ley de Ampère y las simetrías del sistema no es muy difícil de ver (ejerc.) que el campo magnético generado por esta corriente I tiene sólo una componente en la dirección θ , el ángulo polar en coordenadas cilíndricas:

$$\vec{B} = \frac{I}{2\pi r} \vec{e}_\theta. \quad (6.30)$$

El segundo caso interesante es el del solenoide infinito, donde una corriente I pasa por una hélice conductor (Figura 6.4 b). De manera similar al caso del conductor lineal se puede demostrar (ejerc.) que el campo magnético fuera del solenoide es cero, mientras que dentro hay un campo

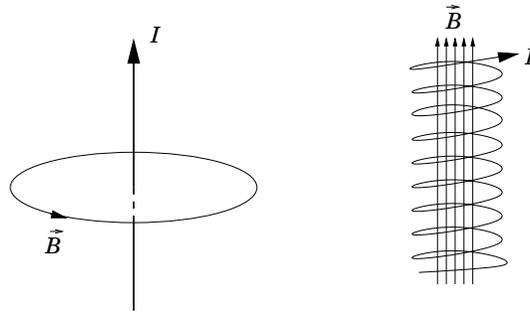


Figura 6.4: Ejemplos típicos de la Ley de Ampère: un campo magnético rotacional alrededor de un conductor lineal infinito y un campo magnético constante en el interior de un solenoide.

constante $\vec{B} = I\vec{e}_z$ a lo largo del eje del solenoide. Una consecuencia de esto es que un metal se convierte en un imán si lo metemos dentro de un solenoide. Dependiendo del tipo de material es posible que las propiedades magnéticas perduran incluso si sacamos el metal del solenoide. Los casos aquí comentados son ejemplos sencillos de la ley de Biot-Savart sobre el campo magnético generada por corrientes eléctricas. Vemos por lo tanto que la ley de Biot-Savart, y en general la de Ampère, relacionan claramente los fenómenos eléctricos interaccionan con los fenómenos magnéticos.

La Ley de Faraday o la ley de la inducción magnética (6.26) describe en cierto sentido la situación contraria a la ley de Ampère con el solenoide. Si una corriente circular causa un campo magnético, ¿también un imán metido en un anillo conductor causa una corriente eléctrica en el anillo? Esto claramente no es verdad, puesto que violaría de manera brutal la conservación de energía. Pero Faraday descubrió que sí se genera una corriente en el momento en que el imán se acerca o se aleja del anillo, es decir cuando el flujo magnético a través del anillo cambia. La ley de Faraday dice que un cambio en el flujo magnético a través de una superficie S induce un campo eléctrico rotacional alrededor de la curva que bordea S (Véase Figura 6.3 b). Este campo eléctrico rotacional general una corriente \vec{j} que a su vez causa a través de la ley de Ampère un campo magnético. El sentido de la corriente \vec{j} es tal que el campo magnético causado intenta compensar el cambio de flujo del campo original. Las corrientes eléctricas inducidas por cambio de flujos magnéticos a través de una superficie es el principio básico detrás de una dinamo y es la base del motor eléctrico.

Las cuatro leyes de electricidad y magnetismo que acabamos de comentar arriba, son las que eran conocidas en los tiempos de Maxwell. En ese momento eran leyes empíricas que describían bien los experimentos realizados hasta entonces. Sin embargo Maxwell se dió cuenta que en particular la ley de Ampère (las ecuaciones (6.23) y (6.28) sin el segundo término de la derecha) no podía ser válida siempre. Si aplicáramos la ley de Biot-Savart (6.30) al caso donde una corriente \vec{j} está cargando un condensador (un acumulador de carga), la corriente genera claramente un campo magnético \vec{B} alrededor del conductor. Sin embargo, el flujo de corriente a través de una superficie es cero, si dejamos que la superficie pase entre las dos placas del condensador (Véase figura 6.5), por lo que la ley de Ampère no es válida. Maxwell se dió cuenta que mientras corría la corriente eléctrica, iba aumentando la carga acumulada en el condensador y por lo tanto aumentando el flujo de campo eléctrico entre las dos placas y la superficie. De allí dedujo Maxwell que ese cambio de flujo eléctrico inducía el campo magnético \vec{B} alrededor del conductor, lo que resulta en el segundo término de (6.23) y (6.28), el término de Maxwell. Al conjunto de la ley de Ampère y el término de Maxwell, se le suele llamar la ley de Ampère-Maxwell.

Maxwell introdujo su término a base de razones puramente teóricas, ya que en este momento no había ninguna indicación experimental para creer en la existencia de una contribución de este tipo. Fijaos que al volver a introducir las constantes ϵ_0 y μ_0 explícitamente, el término de

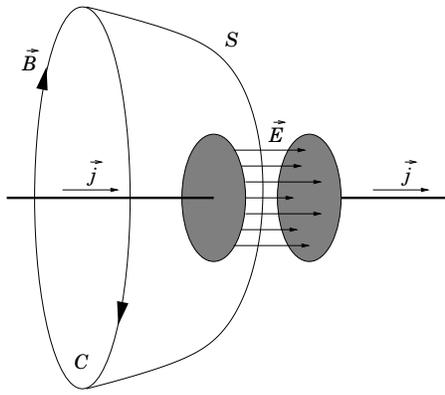


Figura 6.5: El término de Maxwell como corrección a la ley de Ampère: el campo magnético \vec{B} a lo largo de una curva C alrededor del conductor con una corriente \vec{j} no es igual al flujo de la corriente pasando por la superficie S , si S pasa entre las dos placas de un condensador. La carga que se acumula en el condensador crea un cambio en el campo eléctrico, que es el responsable del campo magnético \vec{B} .

Maxwell va multiplicado por un factor $1/c^2 \approx 10^{-17} \text{ (m/s)}^2$, un factor muy pequeño. Sin embargo el término resulta ser de crucial importancia para la consistencia de las ecuaciones de Maxwell y da lugar directamente a la existencia de las soluciones de ondas electromagnéticas, un hecho experimentalmente encontrado por Heinrich Hertz (1847-1894) alrededor de 1890.

Las leyes de Maxwell describen cómo cargas y corrientes eléctricas generan campos eléctricos y magnéticos, pero no hemos dicho todavía cómo una partícula cargada interacciona con estos campos electromagnéticos. Para esto recurrimos a la *fuerza de Lorentz*. La fuerza ejercida por un campo \vec{E} y un campo \vec{B} sobre una partícula con carga q viene dada por

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \quad (6.31)$$

donde \vec{v} es la velocidad de la partícula. Vemos que esta fuerza consiste en dos partes: un término eléctrico que básicamente resume la ley de Coulomb entre dos cargas y un término que dice la carga notará el campo magnético si se mueve con una velocidad \vec{v} . En particular en un campo magnético constante, una carga q con velocidad \vec{v} perpendicular a \vec{B} se moverá en círculos de radio $r = mv/qB$ (ejerc.). En el capítulo 2 ya hemos visto que la fuerza de Lorentz es conservativa en el caso de cargas estáticas, pero no en general. Se puede ver fácilmente de las leyes de Maxwell que $\vec{\nabla} \times \vec{F} \neq 0$. Sin embargo, el segundo término no realiza ningún trabajo sobre la partícula. Aunque cambia la dirección de \vec{v} , no cambia la norma y por lo tanto la energía cinética sigue igual. El cambio de energía cinética T de una partícula en un campo electromagnético solo es debido al campo eléctrico. Por (2.19) y (6.31) vemos que

$$\frac{dT}{dt} = \vec{v} \cdot \vec{F} = q\vec{v} \cdot \vec{E}. \quad (6.32)$$

6.3. Cantidades conservadas

Las mismas leyes de Maxwell indican por su propia estructura que existen ciertas cantidades conservadas. Tomando la divergencia de la Ley de Ampère-Maxwell (6.23) tenemos que

$$0 = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) = \vec{\nabla} \cdot \vec{j} + \vec{\nabla} \cdot \partial_t \vec{E} = \vec{\nabla} \cdot \vec{j} + \partial_t \rho, \quad (6.33)$$

donde en la última igualdad hemos intercambiado las derivadas parciales ∂_t y ∂_i y utilizado la ley de Gauss (6.20). La expresión (6.33) es la forma en que en una teoría de campos se expresa una

ley de conservación, en particular aquí tenemos la *ley de conservación de carga eléctrica*. En la formulación integral se lee

$$\oint_{\partial V} \vec{j} d^2x = -\frac{d}{dt} \iiint_V \rho d^3x, \quad (6.34)$$

en otras palabras, el cambio de carga en un volumen V es igual al flujo de corriente que pasa por la superficie que bordea el volumen V . Si tomamos para el volumen V el universo entero, vemos que la carga total $Q = \int \rho d^3x$ está conservada. Pero la ley de conservación en su forma (6.33) ó (6.34) es más fuerte por su carácter local. No sólo nos dice que la carga eléctrica está conservada, sino también nos dice cómo está conservada, en el sentido que indica a donde ha ido la parte que falta: la posible pérdida de carga es debido al flujo de corriente eléctrica. Fijaos que para poder derivar esta ley de conservación de carga hemos utilizado explícitamente el término de Maxwell en (6.23). Así vemos que conservación de carga también podría haber sido el argumento para corregir la ley de Ampère.

Por (6.32) sabemos que el campo eléctrico puede cambiar la energía cinética de una partícula cargada. Si creemos en la ley de conservación de la energía, debemos por lo tanto aceptar que el campo electromagnético también tiene energía, que pierde o gana al aumentar o disminuir la energía de la partícula. Esto tampoco resulta tan raro, puesto que si consideramos el campo electromagnético como una realidad física, no es de extrañar que posea energía y momento.

Podemos derivar la expresión para la energía del campo electromagnético de las propias leyes de Maxwell. Si tomamos el producto escalar de (6.23) con \vec{E} y de (6.21) con \vec{B} tenemos

$$\begin{aligned} \vec{E} \cdot (\vec{\nabla} \times \vec{B}) &= \vec{j} \cdot \vec{E} + \vec{E} \cdot \partial_t \vec{E}, \\ \vec{B} \cdot (\vec{\nabla} \times \vec{E}) &= -\vec{B} \cdot \partial_t \vec{B} \end{aligned} \quad (6.35)$$

Restando estas dos expresiones y utilizando la propiedad (d) del ejercicio 5 en capítulo 2, vemos que

$$\begin{aligned} -\vec{\nabla} \cdot (\vec{E} \times \vec{B}) &= \partial_t(E^2 + B^2) + \vec{j} \cdot \vec{E} \\ &= \partial_t(E^2 + B^2) + \sum_a q_a \vec{v}_a \cdot \vec{E} \\ &= \partial_t(E^2 + B^2) + \partial_t T \end{aligned} \quad (6.36)$$

donde en la segunda igualdad hemos escrito la corriente \vec{j} como el conjunto de cargas en movimiento y en la última igualdad hemos utilizado (6.32). La expresión tiene la forma de una ley de conservación: si identificamos la cantidad $E^2 + B^2$ con la energía del campo electromagnético y el vector $\vec{S} = \vec{E} \times \vec{B}$ con la corriente de energía, vemos que la ley de conservación de energía dice que el cambio de energía cinética de las partículas cargadas más el cambio de energía de campo electromagnético en un volumen es igual al flujo de energía a través de la superficie. El vector $\vec{S} = \vec{E} \times \vec{B}$ se llama el *vector de Poynting*.

Para obtener una expresión para el momento podemos hacer una derivación parecida. Sólo que ahora el momento es un vector y por lo tanto el flujo de densidad de momento será un tensor de rango dos. Lo más cómodo por lo tanto es hacer la derivación en componentes. La derivada parcial del vector de Poynting con respecto del tiempo viene dado por

$$\begin{aligned} \partial_t S_i &= \epsilon_{ijk} E_j \partial_t B_k + \epsilon_{ijk} \partial_t E_j B_k \\ &= -\epsilon_{ijk} E_j \epsilon_{klm} \partial_t E_m + \epsilon_{ijk} (\epsilon_{jlm} \partial_t B_m - j_j) B_k \\ &= -E_j \partial_i E_j + E_j \partial_j E_i + B_k \partial_k B_i - B_k \partial_i B_k - \epsilon_{ijk} j_j B_k \\ &= -\frac{1}{2} \partial_i (E^2 + B^2) + \partial_j (E_i E_j + B_i B_j) - E_i \partial_j E_j - B_i \partial_j B_j - \epsilon_{ijk} j_j B_k \\ &= -\frac{1}{2} \partial_i (E^2 + B^2) + \partial_j (E_i E_j + B_i B_j) - \rho E_i - \epsilon_{ijk} j_j B_k, \end{aligned} \quad (6.37)$$

donde en la segunda igualdad hemos utilizado las ecuaciones de Maxwell para $\vec{\nabla} \times \vec{E}$ y $\vec{\nabla} \times \vec{B}$ y en la última las ecuaciones para $\vec{\nabla} \cdot \vec{E}$ y $\vec{\nabla} \cdot \vec{B}$. En los últimos dos términos reconocemos la expresión

para la fuerza de Lorentz (6.31) de un conjunto de partículas cargadas y podemos por lo tanto escribirlo como $F_i = \frac{d}{dt} \sum_a (p_a)_i$, la derivada del momento de las partículas. Si ahora definimos el *tensor de momento*

$$T_{ij} = \frac{1}{2}(E^2 + B^2)\delta_{ij} + E_i E_j + B_i B_j \quad (6.38)$$

la ecuación (6.37) la podemos escribir como

$$\partial_t \left(S_i + \sum_a (p_a)_i \right) = -\partial_i T_{ij}. \quad (6.39)$$

Vemos por lo tanto que podemos interpretar el vector de Poynting como el momento del campo electromagnético y la ecuación (6.39) nos dice que la pérdida de momento del campo electromagnético y la pérdida de momento de las partículas es igual al flujo del tensor de momento T_{ij} .

Fijaos que el vector de Poynting aparece dos veces en estas derivaciones: una vez como el flujo de energía y otra vez como el momento conservado. Por la teoría de la relatividad especial sabemos que la energía y el momento están relacionados como los componentes del cuadrivector de energía-momento. La teoría de Maxwell es explícitamente invariante Lorentz (aunque no sea visible en esta notación) y el hecho de que el vector de Poynting aparece en estas dos formas es una manifestación de la covariancia de la teoría.

6.4. Potenciales electromagnéticos e invariancia gauge

Las leyes de Maxwell en su forma diferencial (6.20)-(6.23) son más cómodas de resolver que en su forma integral (6.25)-(6.28), pero aún así no son fáciles por la aparición del operador diferencial $\vec{\nabla}$ en forma de divergencias y rotacionales. Resulta que las dos leyes de Maxwell homogéneas, la ley de Faraday (6.21) y la de la divergencia de \vec{B} (6.22) nos permiten escribir el conjunto en una forma más accesible.

El hecho de que la divergencia de \vec{B} siempre es cero, implica, por las propiedades de divergencias y rotacionales demostradas en el capítulo 2, que podemos escribir \vec{B} como el rotacional de algún campo \vec{A} :⁴

$$\vec{B} = \vec{\nabla} \times \vec{A}. \quad (6.40)$$

Sustituyendo este resultado en la ley de Faraday (6.21), podemos escribirla como

$$\vec{\nabla} \times \left(\vec{E} + \partial_t \vec{A} \right) = 0, \quad (6.41)$$

intercambiando las derivadas parciales ∂_t y ∂_i . Por las mismas propiedades de rotacionales y gradientes, se ve que la combinación $\vec{E} + \partial_t \vec{A}$ es el gradiente de un campo $-\phi$, donde el signo menos se introduce por futura conveniencia. En otras palabras, podemos escribir el campo eléctrico como

$$\vec{E} = -\vec{\nabla} \phi - \partial_t \vec{A}. \quad (6.42)$$

Los campos auxiliares ϕ y \vec{A} se llaman los *potenciales electromagnéticos*. Dado ϕ y \vec{A} , los campos electromagnéticos \vec{E} y \vec{B} están completamente determinados. Pero lo contrario no es verdad: ni ϕ , ni \vec{A} están determinados unívocamente. Si calculamos los campos \vec{E}' y \vec{B}' a partir de campos ϕ' y \vec{A}' que están relacionados con ϕ y \vec{A} mediante

$$\phi' = \phi + \partial_t \Lambda, \quad \vec{A}' = \vec{A} - \vec{\nabla} \Lambda, \quad (6.43)$$

con $\Lambda = \Lambda(\vec{x}, t)$ una función arbitraria, tenemos que

$$\vec{E}' = -\vec{\nabla} \phi' - \partial_t \vec{A}' = -\vec{\nabla} \left(\phi + \partial_t \Lambda \right) - \partial_t \left(\vec{A} - \vec{\nabla} \Lambda \right) = -\vec{\nabla} \phi - \partial_t \vec{A} = \vec{E}, \quad (6.44)$$

$$\vec{B}' = \vec{\nabla} \times \vec{A}' = \vec{\nabla} \times \left(\vec{A} - \vec{\nabla} \Lambda \right) = \vec{\nabla} \times \vec{A} - \vec{\nabla} \times \vec{\nabla} \Lambda = \vec{B}. \quad (6.45)$$

⁴Aquí suponemos que el espacio es simplemente conexo, lo que es el caso para \mathbb{R}^3 . En la sección 6.5 estudiaremos soluciones en espacios que son topológicamente no triviales.

Hay por lo tanto infinitos campos ϕ y \vec{A} que dan todos los mismos \vec{E} y \vec{B} . ¿Es esto un problema? ¿Cómo sabemos ahora cuál coger? En realidad, ϕ y \vec{A} son sólo potenciales, campos auxiliares que hemos introducido por mayor facilidad matemática y que no corresponden a ninguna realidad física. La verdadera física está en las leyes de Maxwell y allí sólo aparecen \vec{E} y \vec{B} . No hay experimentos que permitan medir los potenciales electromagnéticos.⁵ La elección de ϕ y \vec{A} no tiene por lo tanto ninguna relevancia física y podemos elegir lo que más nos convenga. Dentro de muy poco veremos que esta libertad de elección se convierte en una herramienta muy poderosa.

Habiendo sacado de las ecuaciones homogéneas las expresiones (6.40) y (6.42) para \vec{E} y \vec{B} en función de ϕ y \vec{A} , nos queda sustituir estas expresiones en las ecuaciones inhomogéneas restantes. Utilizando las propiedades de divergencias y rotacionales, no es difícil de encontrar que obtenemos

$$-\Delta\phi - \partial_t(\vec{\nabla} \cdot \vec{A}) = \rho, \quad (6.46)$$

$$\partial_t^2 \vec{A} - \Delta \vec{A} - \vec{\nabla}(\partial_t \phi + \vec{\nabla} \cdot \vec{A}) = \vec{j}, \quad (6.47)$$

para la ley de Gauss (6.20) y la ley de Ampère-Maxwell (6.23) respectivamente. A primera vista estas relaciones no parecen mucho más simples que las ecuaciones de Maxwell de donde las hemos derivado, sobre todo debido a los últimos términos en el lado izquierdo. Si no fuera por este término, la ecuación (6.47) se convertiría en una ecuación de ondas inhomogénea en tres dimensiones.

Sin embargo recordaos que tenemos la libertad de cambiar los potenciales ϕ y \vec{A} como (6.43) y de este modo elegir los potenciales ϕ y \vec{A} según más nos convenga. Por lo tanto siempre podemos elegir los potenciales de forma que satisfagan la condición

$$\vec{\nabla} \cdot \vec{A} + \partial_t \phi = 0. \quad (6.48)$$

Con esta elección de los potenciales las ecuaciones (6.46) y (6.47) se reducen a

$$\partial_t^2 \phi - \Delta \phi = \rho, \quad \partial_t^2 \vec{A} - \Delta \vec{A} = \vec{j}, \quad (6.49)$$

o sea la ecuación inhomogénea de ondas en tres dimensiones.

Aquí hemos visto un primer ejemplo de cómo la ambigüedad de los potenciales electromagnéticos es una gran ventaja más que un problema. Eligiendo bien una condición para ϕ y \vec{A} puede simplificarse mucho un problema dado. La invariancia de las leyes de Maxwell bajo la transformación (6.43) se llama *invariancia gauge* y la elección de unos ϕ y \vec{A} específicos se llama *elección de gauge*. En particular la elección (6.48) se llama el *gauge de Lorenz*.⁶ Otras elecciones de gauge comunes son el gauge de Coulomb $\vec{\nabla} \times \vec{A} = 0$, el gauge temporal $\partial_t \phi = \partial_t \vec{A} = 0$ y el gauge axial $\vec{A} \cdot \vec{n} = 0$, donde \vec{n} es un vector de unidad en una dirección espacial específica.

La invariancia gauge de las leyes de Maxwell tiene otra gran ventaja. Dado que los potenciales no corresponden a ninguna realidad física, los verdaderos resultados físicos, como la energía, el momento o los campos electromagnéticos no pueden depender de la elección de ϕ y \vec{A} . Diferentes elecciones deberían dar el mismo resultado. No sólo nos da esto una manera de comprobar los resultados, también pone restricciones fuertes sobre la forma de las expresiones de cantidades físicas.

Resumiendo podemos escribir las ecuaciones de Maxwell como

$$\vec{E} = -\vec{\nabla}\phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}, \quad (6.50)$$

$$\partial_t^2 \phi - \Delta \phi = \rho, \quad \partial_t^2 \vec{A} - \Delta \vec{A} = \vec{j}, \quad (6.51)$$

⁵También aquí suponemos que estamos en un espacio simplemente conexo. Veremos un contraejemplo en el capítulo 9.

⁶El gauge de Lorenz recibe su nombre por Ludwig Lorenz, el físico danés que lo introdujo en 1867 y no por H.A. Lorentz, el físico holandés, conocido por la fuerza de Lorentz y el grupo de Lorentz, el grupo de invariancia de la relatividad especial. La jugada graciosa de la Historia está en que el interés del gauge de Lorenz está en el hecho que es invariante Lorentz.

siempre y cuando asumimos que ϕ y \vec{A} están relacionados a través de (6.48). Las ecuaciones (6.51) nos dicen cómo los potenciales se comportan frente a las cargas y las corrientes, mientras (6.50) nos da la relación entre los potenciales y los campos físicos \vec{E} y \vec{B} . Veremos en la sección 9.1 que la invariancia gauge está íntimamente ligado con la teoría matemática de formas y que las ecuaciones homogéneas (6.50) son en realidad una identidad de Bianchi, que dice que el tensor electromagnético es (localmente) exacto.

6.5. Soluciones de las ecuaciones de Maxwell

Con las ecuaciones (6.49) para los potenciales no resulta difícil encontrar soluciones de las ecuaciones de Maxwell. Una gran ventaja de la estructura de las ecuaciones es que son lineales, así que cualquier combinación lineal de dos soluciones también es una solución.⁷

En la sección 6.2 ya hemos mencionado algunas soluciones, como la de la carga puntual, el conductor lineal y el solenoide infinito. En esta sección repasaremos alguna soluciones más que son interesantes por su relevancia física.

6.5.1. Ondas electromagnéticas

Una de las propiedades sorprendentes de las ecuaciones de Maxwell es que existen soluciones no triviales en ausencia de cargas y corrientes. Las llamadas ecuaciones del vacío son⁸

$$\vec{\nabla} \cdot \vec{E} = 0, \quad \vec{\nabla} \cdot \vec{B} = 0, \quad (6.52)$$

$$\vec{\nabla} \times \vec{E} = -\partial_t \vec{B}, \quad \vec{\nabla} \times \vec{B} = \epsilon_0 \mu_0 \partial_t \vec{E}. \quad (6.53)$$

Una primera observación es la remarcable simetría entre campos eléctricos y magnéticos en la ausencia de cargas. En la presencia de cargas esta simetría esta rota debido al hecho de que existen cargas eléctricas pero no hay monopolos magnéticos. Volveremos a este tema un poco más adelante.

Una segunda observación es que la estructura de las ecuaciones del vacío permite ondas de campo viajando por el espacio: en la ecuación (6.53) vemos que un cambio en el campo magnético induce un campo eléctrico por la ley de Faraday, y a su vez un cambio de campo eléctrico induce por el término de Maxwell un campo magnético. Un cambio en el campo electromagnético genera por lo tanto más cambios, resultando en paquetes de campo propagándose por el espacio independientemente de cargas y corrientes. Estos paquetes de campo se llaman *ondas electromagnéticas*. Estudiamos ahora estas soluciones un poco más en detalle.

El caso más sencillo es cuando los campos \vec{E} y \vec{B} sólo dependen de una coordenada, digamos x y del tiempo. Entonces las ecuaciones (6.52)-(6.53) se reducen a

$$\partial_x E_x = \partial_x B_x = \partial_t E_x = \partial_t B_x = 0, \quad (6.54)$$

$$\partial_x E_z = \partial_t B_y, \quad \partial_x B_y = \epsilon_0 \mu_0 \partial_t E_z, \quad (6.55)$$

$$\partial_x E_y = -\partial_t B_z, \quad \partial_x B_z = -\epsilon_0 \mu_0 \partial_t E_y. \quad (6.56)$$

La solución de (6.54) es $E_x = c_1$ y $B_x = c_2$, con c_1 y c_2 constantes que tomaremos cero, puesto que no estamos interesados de momento en campos constantes. Las dos ecuaciones de (6.55) forman, igual que las dos de (6.56), un conjunto de ecuaciones que se desacopla de las otras. Sustituyendo la segunda ecuacion en en la primera vemos que podemos escribir (6.55) como

$$\frac{1}{c^2} \partial_t^2 E_z - \partial_x^2 E_z = 0, \quad \frac{1}{c^2} \partial_t^2 B_y - \partial_x^2 B_y = 0, \quad (6.57)$$

⁷Esta propiedad ya no es verdad por ejemplo en relatividad general, donde la ecuación de Einstein es no lineal, por lo que encontrar soluciones de las ecuaciones de Einstein es un problema notoriamente difícil.

⁸En esta subsección y en la siguiente volveremos a poner explícitamente la permitividad y ϵ_0 y la permeabilidad μ_0 , por razones que se harán obvios más adelante.

y (6.56) como

$$\frac{1}{z^2} \partial_t^2 E_y - \partial_x^2 E_y = 0, \quad \frac{1}{z^2} \partial_t^2 B_z - \partial_x^2 B_z = 0, \quad (6.58)$$

donde hemos sustituido también el producto de las constantes ϵ_0 y μ_0 por $\epsilon_0 \mu_0 = 1/c^2$. Vemos por lo tanto que cada componente satisface una ecuación de onda unidimensional, que tiene como solución más general

$$\begin{aligned} E_y &= f_1(x+ct) + g_1(x-ct), & E_z &= f_2(x+ct) + g_2(x-ct), \\ B_y &= c^{-1} f_1(x+ct) - c^{-1} g_1(x-ct), & B_z &= -c^{-1} f_2(x+ct) + c^{-1} g_2(x-ct), \end{aligned} \quad (6.59)$$

donde f_1, f_2, g_1 y g_2 son funciones arbitrarias de su argumento. Las funciones $f(x+ct)$ representan una onda que se propaga con velocidad c , la velocidad de la luz, en la dirección del eje x negativo, mientras las funciones $g(x-ct)$ una onda en la dirección del eje x positivo. Dado que el campo no depende de las direcciones y ó z , el campo electromagnético toma los mismos valores en los planos $x = \text{constante}$. Las soluciones que tienen esta propiedad se llaman *ondas planas*. Fijaos que la amplitud de la onda es perpendicular a la dirección de propagación. Las ondas electromagnéticas son por lo tanto ondas transversales.

Una clase importante de ondas planas son las llamadas *ondas monocromáticas*, donde las funciones f y g son senos y cosenos. La solución (6.59) es entonces de la forma

$$E_i = A_i \cos(kx \pm \omega t + \phi_i), \quad B_i = c^{-1} A_i \cos(kx \pm \omega t + \psi_i), \quad (6.60)$$

donde A_i es la amplitud, k el número de onda, ω la frecuencia angular y ϕ_i y ψ_i las fases. Para que las ondas monocromáticas sean de la forma (6.59), el número de onda y la frecuencia tienen que satisfacer la relación

$$\omega = ck. \quad (6.61)$$

Desde el punto de vista físico, las ondas monocromáticas no corresponden a una realidad física, puesto que se extienden por el espacio entero y tendrían una energía infinita. Sin embargo desde el punto de vista matemático es un concepto muy útil, ya que son fáciles de manejar. Además sabemos por el análisis de Fourier que cualquier solución (6.59) se puede escribir como una superposición de ondas monocromáticas de distintas frecuencias.

La longitud de onda λ viene dada por el número de onda y la frecuencia ν está relacionada con la frecuencia angular ω mediante

$$\lambda = \frac{2\pi}{k}, \quad \nu = \frac{\omega}{2\pi}. \quad (6.62)$$

Fijaos que la conocida fórmula para la velocidad de propagación de ondas nos da efectivamente la velocidad de la luz:

$$v = \lambda \nu = \frac{\omega}{k} = c. \quad (6.63)$$

Al darse cuenta que existen soluciones de paquetes de campo electromagnético que se propagan con la velocidad de la luz, Maxwell sacó la conclusión que la luz en realidad eran ondas electromagnéticas. Efectivamente la luz visible corresponde a ondas electromagnéticas con una longitud de onda entre 400 y 720 nm ($1\text{nm} = 10^{-9}$ m). Otras longitudes de onda corresponden con otros tipos de radiación electromagnética, como la radiación infraroja y ondas de radio con longitudes de ondas más largas o rayos ultravioletas, rayos X o rayos gamma con longitud de onda muy cortas.

La conexión entre el electromagnetismo y la óptica fue uno de los grandes éxitos de la teoría de Maxwell. No sólo era la primera vez que alguien venía con una teoría fundamental para explicar la naturaleza de la luz, sino esa teoría también daba una base sólida a la óptica y permitía derivar las leyes empíricas de la óptica y calcular índices de reflexión y refracción de materiales. Pero aparte de la teoría de Maxwell es capaz de dar una explicación de cómo surgen estas ondas electromagnéticas. Esto lo veremos en la siguiente subsección.

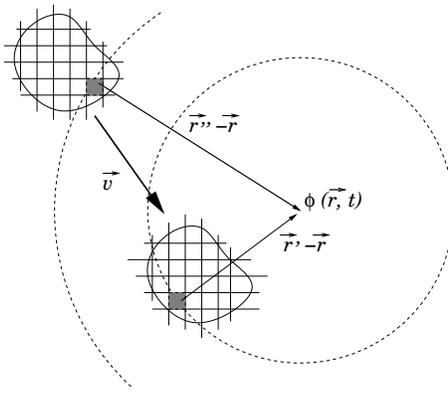


Figura 6.6: Los potenciales retardados: los potenciales electromagnéticos en el punto \vec{r} en $t = t_0$ vienen dado por la suma de todas las contribuciones de cargas y corrientes en los puntos \vec{r}' en el momento $t - |\vec{r} - \vec{r}'|/c$.

6.5.2. Potenciales retardados

Resolvemos ahora las ecuaciones generales de Maxwell. Ya hemos visto que, eligiendo el gauge de Lorenz, se puede escribir las ecuaciones de Maxwell en función de las potenciales electromagnéticos como

$$\frac{1}{c^2} \partial_t^2 \phi - \Delta \phi = \frac{\rho}{\epsilon_0}, \quad \frac{1}{c^2} \partial_t^2 \vec{A} - \Delta \vec{A} = \mu_0 \vec{j}. \quad (6.64)$$

Por la teoría de ecuaciones diferenciales sabemos que la solución general de unas ecuaciones lineales inhomogéneas consiste de la suma de las soluciones de las ecuaciones homogéneas y una solución particular. Las soluciones de las ecuaciones homogéneas son las ondas electromagnéticas que hemos visto en la subsección anterior.

La derivación de la solución particular en principio no es muy diferente al caso de la electrostática o las corrientes constantes. Para calcular el valor de los potenciales en un punto \vec{r} en el momento t tenemos que sumar todas las contribuciones de cargas y corrientes en todos los elementos de volumen d^3x . La diferencia con el caso estático es que para obtener la contribución de un elemento de volumen en posición \vec{r}' no hay que tomar la densidad de carga y corriente en el momento t , sino en el momento $t - |\vec{r} - \vec{r}'|/c$, dado que el tiempo $|\vec{r} - \vec{r}'|/c$ es el que necesita una onda electromagnética (y por lo tanto también la contribución a los potenciales) para viajar de \vec{r}' a \vec{r} . Los potenciales que reflejan esta propiedad se llaman los *potenciales retardados* y tienen la forma

$$\phi(\vec{r}, t) = \frac{1}{4\pi\epsilon_0} \iiint \frac{\rho(\vec{r}', t - \frac{|\vec{r} - \vec{r}'|}{c})}{|\vec{r} - \vec{r}'|} d\vec{r}', \quad (6.65)$$

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi} \iiint \frac{\vec{j}(\vec{r}', t - \frac{|\vec{r} - \vec{r}'|}{c})}{|\vec{r} - \vec{r}'|} d\vec{r}'. \quad (6.66)$$

Fijaos que en el caso estático recuperemos el potencial de Coulomb y el potencial de Biot-Savart.

Con la expresión de los potenciales retardados vemos lo que ya habíamos mencionado en la introducción de este capítulo: los campos electromagnéticos (o aquí los potenciales) juegan el rol de intermediarios que transmiten la interacción entre partículas. Aunque el viejo modelo newtoniano de interacciones a distancia ya parecía poco físico y causaban muchas críticas desde los primeros momentos⁹, estos inconvenientes fueron olvidados rápidamente al ver el éxito de la

⁹Newton nunca pretendía tener una explicación física para la gravedad, sólo quería dar la expresión correcta de la fuerza. De allí su respuesta: “Hypotheses non fingo” (No invento hipótesis).

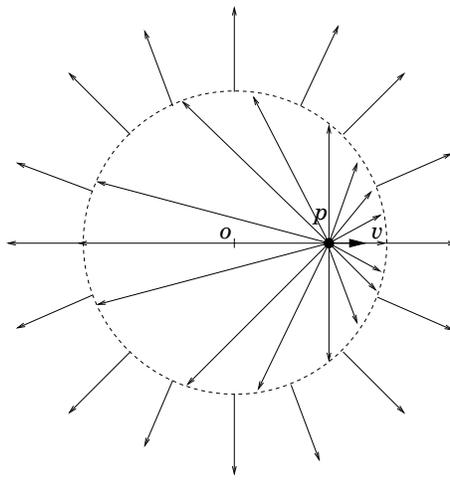


Figura 6.7: *Ondas electromagnéticas debidas a cargas aceleradas: una carga originalmente en reposo en el punto o acelera y se mueve al punto p. A distancias grandes las líneas de campo todavía están como si la carga estuviera en o, mientras cerca de de carga las líneas ya se han adaptado. Entre medias hay una discontinuidad que se expande con la velocidad de la luz.*

ley de la gravedad a la hora de resolver problemas concretos. En el contexto de teoría de campos, una partícula no interacciona directamente con otra, sino con los campos presentes, que hacen propagar la influencia de la partícula con una velocidad finita (la velocidad de la luz) hasta llegar a otras partículas.

En este sentido las ondas electromagnéticas en la teoría de Maxwell son precisamente esto: distorsiones del campo electromagnético debido a la aceleración de cargas que se propagan con la velocidad de la luz. Consideremos un carga inicialmente en reposo (o lo que es físicamente equivalente, moviéndose con velocidad constante). En el momento $t = t_0$ la carga acelera a una velocidad \vec{v} durante un intervalo Δt . La configuración de las líneas de campo en un momento $t > t_0 + \Delta t$ consiste de dos partes: a distancias $r > ct$ las líneas de campo todavía están como estaban antes de que la carga se moviese, dado que la señal no ha tenido tiempo para llegar hasta allí. La contribución a los potenciales retardados en los puntos con $r > ct$ es como si la carga todavía estuviera en el sitio original. A distancias $r < ct$ sin embargo las líneas de campo apuntan directamente a la carga por la ley de Coulomb. Entre las dos zonas hay una discontinuidad esférica de grosor $c\Delta t$ que, debido a la estructura de las ecuaciones del vacío, se propaga con la velocidad de la luz, formando una onda electromagnética.

La propiedad de que la información se propague con velocidad menor o igual a la de la luz es típica de teorías relativistas, dado que la velocidad de la luz como velocidad máxima es justo uno de los postulados de la teoría de la relatividad especial. La teoría de Maxwell ya es relativista en su forma (6.20)-(6.23), aunque fue formulada 40 años antes de la relatividad especial. Históricamente era justo la incompatibilidad de las leyes de Newton con las de Maxwell y el principio de la relatividad de Galilei, lo que inspiró Einstein a reformular la mecánica newtoniana.

6.5.3. El efecto Aharonov-Bohm

Hemos mencionado en varias ocasiones que los campos físicos son los campos electromagnéticos \vec{E} y \vec{B} , mientras los potenciales ϕ y \vec{A} sólo son construcciones matemáticas para manejar mejor las ecuaciones de Maxwell. Sin embargo esto sólo es verdad en la física clásica. En la mecánica cuántica sí puede haber efectos físicos debido a los potenciales. El ejemplo más conocido es el

En la mecánica cuántica, la posición de un electrón (o cualquier otra partícula) no está completamente determinada, sino que está descrita por la *función de onda* $\Psi(\vec{r}, t)$, una función compleja cuya norma al cuadrado $|\Psi(\vec{r}, t)|^2$ da la probabilidad de encontrar el electrón en la posición \vec{r} . Por lo tanto la función de onda no está determinada unívocamente: dada una función de onda $\Psi(\vec{r}, t)$, siempre se puede añadir una fase tal que $\Psi(\vec{r}, t)$ y $\Psi'(\vec{r}, t) = e^{i\alpha}\Psi(\vec{r}, t)$ tienen la misma física.

El hecho que tanto los potenciales como la fase de la función de onda no representan conceptos físicos, sugiere que una cosa está relacionada con la otra. Efectivamente, a través del formalismo hamiltoniano se puede demostrar que una transformación gauge de los potenciales induce un cambio de fase en la función de onda.

El hamiltoniano de una partícula en un campo electromagnético

$$H = \frac{1}{2m}(p_i - eA_i)^2 + V(x_i) \quad (6.67)$$

en la mecánica cuántica se convierte, reemplazando los momentos p_i por el operador $-i\hbar\partial_i$, en

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\partial_i + \frac{ie}{\hbar}A_i\right)^2 + V(x_i) \quad (6.68)$$

donde $\hbar = h/2\pi$ y h es la constante de Planck. En la mecánica cuántica el hamiltoniano es el operador de evolución de la función de onda $\Psi(\vec{r}, t)$ y la ecuación de movimiento es la *ecuación de Schrödinger*

$$\hat{H}\Psi(\vec{r}, t) = i\hbar \partial_t\Psi(\vec{r}, t). \quad (6.69)$$

Al aplicar una transformación gauge $\vec{A} \rightarrow \vec{A} + \vec{\nabla}\Lambda$, cambiará el hamiltoniano (6.68). Sin embargo, la transformación gauge no puede tener consecuencias físicas, por lo tanto la única manera de conseguir que la función de onda siga satisfaciendo la ecuación de Schrödinger es añadirle un factor de fase $e^{-ie\Lambda/\hbar}$ (ejerc.)

Es justo este cambio de fase la que utiliza el efecto Aharonov-Bohm. El experimento es el siguiente: considera una fuente de electrones colocada frente a una pantalla (detector). Entre la fuente de electrones y la pantalla colocamos un solenoide por el cual pasa una corriente (vease figura 6.8). Ya hemos visto en la sección 6.2 que el solenoide genera un campo magnético $\vec{B} = 0$ fuera y $\vec{B} = I\vec{e}_z$ dentro del solenoide. Tomando por simplicidad el radio del solenoide a cero, manteniendo el flujo magnético $\Phi = \iint \vec{B} \cdot \vec{n} d^2x$ constante, el potencial que genera este campo magnético es en coordenadas cilíndricas (con el eje z a lo largo del solenoide)

$$\vec{A} = \frac{\Phi}{2\pi r}\vec{e}_\phi. \quad (6.70)$$

Efectivamente, no es difícil de comprobar que con esta expresión \vec{B} es cero en todo el espacio, menos en el origen $r = 0$.¹⁰

Clásicamente los electrones no notan la presencia del flujo magnético, puesto que $\vec{B} = 0$ a lo largo de toda su trayectoria (y por lo tanto la fuerza de Lorentz es cero), sin embargo en la mecánica cuántica, al conectar el flujo, la función de ondas de los electrones acoge un factor de fase

$$\Psi^{(0)}(\vec{r}) \rightarrow \Psi^{(\vec{A})} = \exp\left(-\frac{ie}{\hbar} \int_{\infty}^{\vec{r}} \vec{A}(\vec{r}') \cdot d\vec{r}'\right) \Psi^{(0)}(\vec{r}). \quad (6.71)$$

¹⁰La expresión para el rotacional en coordenadas cilíndricas viene dado por

$$\vec{\nabla} \times \vec{A} = \left(r^{-1}\partial_\phi A_z - \partial_z A_\phi\right)\vec{e}_r + \left(\partial_z A_r - \partial_r A_z\right)\vec{e}_\phi + r^{-1}\left(\partial_r(rA_\phi) - \partial_\phi A_r\right)\vec{e}_z.$$

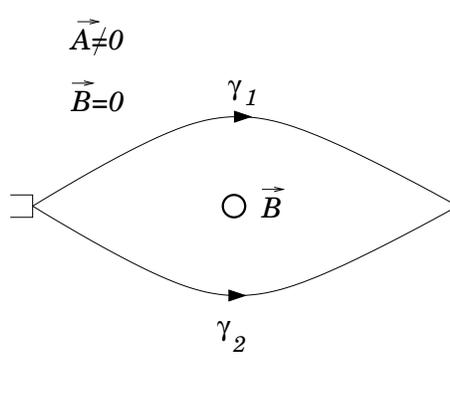


Figura 6.8: El efecto Aharonov-Bohm: electrones emitidos por una fuente, pasan alrededor de un solenoide. El campo magnético \vec{B} fuera del solenoide es cero, sin embargo el potencial \vec{A} no. Clásicamente los electrones no están influenciados por el flujo magnético dentro del solenoide. Cuánticamente la función de ondas acoge una fase, que da lugar a un patrón de interferencias en una pantalla (detector).

Debido a esta fase, el comportamiento en la pantalla será diferente. La probabilidad de encontrar el electrón en la posición \vec{r} viene dado por el cuadrado de la norma de la suma de las dos contribuciones, una por cada camino γ_1 y γ_2 , por encima y por debajo del solenoide respectivamente (véase figura 6.8):

$$P \sim \left| \Psi_{\gamma_1}(\vec{r}) + \Psi_{\gamma_2}(\vec{r}) \right|^2. \quad (6.72)$$

En general la contribución de los dos caminos causará un patrón de interferencias en la pantalla, debido a que en algunos puntos las dos contribuciones se anulan, mientras en otros puntos se suman.

Es fácil de ver que, aunque los electrones no notan en ningún momento el campo magnético, la probabilidad P será distinta para $\Phi = 0$ que para $\Phi \neq 0$. En general tenemos que

$$\begin{aligned} P &\sim \left| \exp\left(-\frac{ie}{\hbar} \int_{\gamma_1} \vec{A}(\vec{r}') \cdot d\vec{r}'\right) \Psi_{\gamma_1}^{(0)}(\vec{r}) + \exp\left(-\frac{ie}{\hbar} \int_{\gamma_2} \vec{A}(\vec{r}') \cdot d\vec{r}'\right) \Psi_{\gamma_2}^{(0)}(\vec{r}) \right|^2 \\ &= \left| \exp\left(-\frac{ie}{\hbar} \int_{\gamma_2} \vec{A}(\vec{r}') \cdot d\vec{r}'\right) \right|^2 \cdot \left| \exp\left(-\frac{ie}{\hbar} \int_{\gamma_1 - \gamma_2} \vec{A}(\vec{r}') \cdot d\vec{r}'\right) \Psi_{\gamma_1}^{(0)}(\vec{r}) + \Psi_{\gamma_2}^{(0)}(\vec{r}) \right|^2 \\ &= \left| e^{-ie\Phi/\hbar} \Psi_{\gamma_1}^{(0)}(\vec{r}) + \Psi_{\gamma_2}^{(0)}(\vec{r}) \right|^2, \end{aligned} \quad (6.73)$$

donde hemos utilizado que la integral de contorno de \vec{A} a lo largo de una trayectoria cerrada $\gamma = \gamma_1 - \gamma_2$ alrededor del solenoide es igual a Φ :

$$\oint_{\gamma} \vec{A} \cdot d\vec{r} = \iint (\vec{\nabla} \times \vec{A}) \cdot \vec{n} d^2x = \iint \vec{B} \cdot \vec{n} d^2x = \Phi. \quad (6.74)$$

De (6.73) está claro que la medición en la pantalla será diferente para $\Phi = 0$ y $\Phi \neq 0$. El factor de fase extra, en el caso de que el flujo magnético sea distinto de cero, causa que el patrón de interferencias generado se desplace con respecto al caso sin flujo. Esto es francamente sorprendente, puesto que el desplazamiento del patrón de interferencias es un efecto físico, mientras que los electrones sólo han podido interactuar con \vec{A} , puesto que $\vec{B} = 0$. ¿Significa esto que el potencial \vec{A} sí tiene significado físico? Analicemos la situación en más detalle.

Aunque el campo magnético fuera del solenoide es cero, hay claramente un efecto físico debido a la presencia de \vec{A} . Sin embargo, como se puede ver en (6.74) este efecto es proporcional a Φ , el flujo de \vec{B} a través del solenoide. Mirando bien la formula (6.74), hay algo raro que nos debería llamar la atención. Dado que $\vec{B} = 0$, el potencial \vec{A} tiene que ser un *gauge puro*, es decir $\vec{A} = \vec{\nabla}\Lambda$. Efectivamente, a partir de (6.70) sacamos que¹¹

$$\Lambda = \frac{1}{2}\Phi\phi. \quad (6.75)$$

Sin embargo, la integral sobre una curva cerrada de un gradiente (o equivalentemente, el rotacional de un gradiente) es cero, mientras la contribución en (6.74) nos da un resultado finito Φ , justo el efecto físico.

La razón por la que la integral (6.74) no es cero, a pesar de que \vec{A} es un gradiente, es porque la curva γ es topológicamente no trivial. El potencial diverge en $r = 0$, donde está colocado el solenoide. No podemos por lo tanto contraer la curva γ a un punto sin que pase por la singularidad. El espacio de configuraciones en que estamos trabajando es por lo tanto $\mathbb{R}^2 \setminus \{0\}$, el plano quitando el origen, y este espacio no es simplemente conexo: no todas las curvas son contraíbles. Es justo debido a esta estructura topológica que la integral (6.74) (y por lo tanto el efecto físico Φ) no es cero.

Desde el punto de vista matemático, esta estructura es debida a que no sólo el espacio de configuraciones es no simplemente conexa, sino también el grupo de gauge. Por un lado el grupo de gauge de la teoría de Maxwell es $U(1)$ (ya que las transformaciones gauge son abelianas), cuyo espacio de grupo es el círculo S^1 . Por otro lado, el espacio de configuraciones $\mathbb{R}^2 \setminus \{0\}$, que es topológicamente igual a $\mathbb{R} \times S^1$. Pero dado que (6.75) no depende de r , la parte (topológicamente) relevante del espacio de configuraciones es S^1 . El parámetro Λ de una transformación gauge es por lo tanto una función del espacio de grupo S^1 al espacio (relevante) de configuraciones S^1 . Diferentes funciones Λ se distinguen por el número de veces que la primera S^1 recorre (enrolla) a la segunda. Funciones con distinto número de enrollamiento no son continuamente deformables unas en otras y funciones con número de enrollamiento distinto de cero no son contraíbles. El número de enrollamiento es por lo tanto un invariante topológico y la clasificación de las distintas funciones $\Lambda : S^1 \rightarrow S^1$ se hace a través del (*primer*) grupo de homotopía π_1 . Para el caso del grupo $U(1)$, el grupo de homotopía $\pi_1(U(1)) = \pi_1(S^1) = \mathbb{Z}$.

Por último, ¿cuál es al final el estatus físico del potencial en el problema de Aharonov-Bohm? ¿Corresponde a un campo “real”, con efectos físicos? Y si fuera así, ¿qué campo \vec{A} cogeríamos, dado que hay una ambigüedad debido a la invariancia gauge? La respuesta es que todavía podemos hacer las transformaciones gauge $\vec{A} \rightarrow \vec{A} + \vec{\nabla}\Lambda$ que queremos (y que por lo tanto \vec{A} no tiene significado físico real), siempre y cuando respetemos las condiciones de contorno, es decir la estructura topológica del espacio de configuraciones. El verdadero efecto físico está en el flujo magnético $\Phi = \iint \vec{B} \cdot \vec{n} d^2x$.

6.5.4. El monopolo de Dirac

Ya hemos comentado en la subsección 6.5.1 que las ecuaciones del vacío (6.52)-(6.53)

$$\vec{\nabla} \cdot \vec{E} = 0, \quad \vec{\nabla} \cdot \vec{B} = 0, \quad (6.76)$$

$$\vec{\nabla} \times \vec{E} = -\partial_t \vec{B}, \quad \vec{\nabla} \times \vec{B} = \partial_t \vec{E}. \quad (6.77)$$

son remarcablemente simétricas bajo intercambio de \vec{E} y \vec{B} . Efectivamente en ausencia de cargas las ecuaciones (6.77) son invariantes bajo la transformación $SO(2)$

$$\vec{E}' = \cos \alpha \vec{E} + \sin \alpha \vec{B}, \quad (6.78)$$

$$\vec{B}' = -\sin \alpha \vec{E} + \cos \alpha \vec{B}. \quad (6.79)$$

¹¹El gradiente en coordenadas cilíndricas viene dado por $\vec{\nabla}\Lambda = \partial_r\Lambda\vec{e}_r + r^{-1}\partial_\phi\Lambda\vec{e}_\phi + \partial_z\Lambda\vec{e}_z$.

Fijaos que esta transformación intercambia completamente \vec{E} y \vec{B} , de modo que los campos eléctricos y magnéticos en el vacío son en realidad indistinguibles.

Esta simetría se rompe en presencia de cargas, debido al hecho que existen cargas eléctricas, pero no hay cargas magnéticas. El físico-matemático inglés Paul Dirac (1902-1984) descubrió una configuración de campos que describe un monopolo magnético.

En analogía con la expresión (6.29) de una carga puntual, un monopolo magnético con carga magnética g en el origen produce un campo magnético \vec{B} de la forma

$$\vec{B} = \frac{g}{4\pi r^2} \vec{e}_r, \quad (6.80)$$

tal que se puede introducir en las ecuaciones de Maxwell una densidad de carga magnética ρ_m y una corriente magnética \vec{j}_m que restaurarían la simetría $SO(2)$ de las ecuaciones del vacío:

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \rho_e, & \vec{\nabla} \cdot \vec{B} &= \rho_m, \\ \vec{\nabla} \times \vec{E} &= -\vec{j}_m - \partial_t \vec{B}, & \vec{\nabla} \times \vec{B} &= \vec{j}_e + \partial_t \vec{E}. \end{aligned} \quad (6.81)$$

Un potencial \vec{A} que da lugar al campo (6.80) viene dado por

$$\vec{A}^{(n)} = \frac{g(1 - \cos \theta)}{4\pi r \sin \theta} \vec{e}_\phi. \quad (6.82)$$

Efectivamente, utilizando la expresión (5.12) para el rotacional en coordenadas esféricas, es fácil de ver que $\vec{\nabla} \times \vec{A}^{(n)} = \vec{B}$, salvo en $r = 0$. Fijaos que la expresión (6.82) es singular para $\theta = \pi$, es decir el eje z negativo. Esta singularidad, llamada la *cuerda de Dirac*, es una consecuencia del hecho que hemos modificado la expresión para la divergencia de \vec{B} : sólo a través de una singularidad en \vec{A} podemos escribir

$$\vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) \neq 0. \quad (6.83)$$

Fijaos que esta situación es muy parecida a la de (6.74) en el caso del efecto Aharanov-Bohm, donde también los teoremas del análisis vectorial son circunvaladas por divergencias en el potencial. Esto sugiere que también aquí habrá un fundamento topológico en la base de esta solución. Veremos que éste efectivamente es el caso.

A nivel clásico la singularidad no nos debe preocupar, puesto que sólo es una singularidad en el potencial y no corresponde a nada físico. Sin embargo hemos visto en la subsección anterior que a nivel cuántico sí puede haber efectos físicos. Por lo tanto si queremos tomar el monopolo de Dirac en serio, tenemos que encontrar una manera de deshacernos de la cuerda de Dirac.

Otra expresión para \vec{A} que da el mismo campo magnético (6.80) es

$$\vec{A}^{(s)} = \frac{-g(1 + \cos \theta)}{4\pi r \sin \theta} \vec{e}_\phi, \quad (6.84)$$

que evita la singularidad en $\theta = \pi$, pero que es singular en $\theta = 0$, el eje z positivo. El hecho de que la singularidad cambia de sitio al elegir otro potencial, indica que la cuerda de Dirac no es algo físico, sino un artefacto de la elección de gauge.¹² Efectivamente, $\vec{A}^{(n)}$ y $\vec{A}^{(s)}$ están relacionados a través de una transformación gauge (puesto que dan el mismo \vec{B}):

$$\vec{A}^{(n)} = \vec{A}^{(s)} + \frac{g}{2\pi r \sin \theta} \vec{e}_\phi = \vec{A}^{(s)} - \vec{\nabla} \left(-\frac{g\phi}{2\pi} \right), \quad (6.85)$$

donde en la última igualdad hemos utilizado la expresión (5.10) para el gradiente en coordenadas esféricas.

¹²Incluso se puede demostrar relativamente fácil con la ley de Gauss que la cuerda no tiene porqué ser recta. Puede tener la forma de una curva semi-infinita, mientras que sea continua.

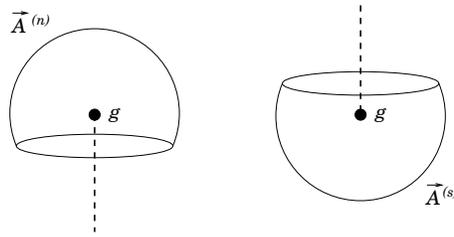


Figura 6.9: El monopolo de Dirac según Wu y Yang: dividimos el espacio en dos parches $\theta \in [0, \frac{\pi}{2} + \epsilon]$ y $\theta \in [\frac{\pi}{2} - \epsilon, \pi]$ y utilizamos en cada parche los potenciales $\vec{A}^{(n)}$ y $\vec{A}^{(s)}$ respectivamente, que son regulares en su respectivos parches. La cuerda de Dirac (la línea interrumpida) no aparece en la descripción, ya que es un artefacto de la elección de gauge.

El truco desarrollado por los físicos chinos Chien-Shiung Wu (1912-1997) y Chen Ning Yang (1922) consiste en dividir el espacio en parches y utilizar un potencial en cada parche, tal que la singularidad del potencial cae fuera del parche en que lo utilizas (véase figura 6.9). En particular cogemos como primer parche el hemisferio norte $\theta \in [0, \frac{\pi}{2} + \epsilon]$, donde utilizamos el potencial $\vec{A}^{(n)}$ y como segundo parche el hemisferio sur $\theta \in [\frac{\pi}{2} - \epsilon, \pi]$ con el potencial $\vec{A}^{(s)}$. En la zona de solapamiento $\theta \in [\frac{\pi}{2} - \epsilon, \frac{\pi}{2} + \epsilon]$ los dos potenciales están relacionados a través de la transformación gauge (6.85). Fijaos que esta transformación no está bien definida en $\theta = 0$ y $\theta = \pi$, los puntos de la cuerda de Dirac, pero dado que sólo estamos interesados en aplicar la transformación gauge en la zona de solapamiento, no nos tiene que preocupar.

El flujo magnético total a través de una esfera alrededor del monopolo viene dado por

$$\begin{aligned}
 \oiint \vec{B} \cdot \vec{n} d^2x &= \iint_N (\vec{\nabla} \times \vec{A}^{(n)}) \cdot \vec{n} d^2x + \iint_S (\vec{\nabla} \times \vec{A}^{(s)}) \cdot \vec{n} d^2x \\
 &= \oint (\vec{A}^{(n)} - \vec{A}^{(s)}) \cdot d\vec{r} \\
 &= \oint \vec{\nabla} \left(\frac{g\phi}{2\pi} \right) \cdot d\vec{r} = g,
 \end{aligned} \tag{6.86}$$

donde en la primera y la segunda igualdad hemos utilizado los teoremas de Stokes (6.24). Las integrales de superficie están tomadas respectivamente en el hemisferio norte y sur y la integral de contorno en la segunda igualdad está tomado a lo largo del ecuador, donde el signo relativo es debido al hecho de que recorremos la curva en sentido opuesto. Finalmente en la última igualdad hemos utilizado las expresiones (6.85) y (5.10) para el gradiente. La ecuación (6.86) es por lo tanto la ley de Gauss para cargas magnéticas: el flujo total a través de una superficie cerrada es igual a la carga encerrada dentro de la superficie, en acuerdo con la solución (6.80) propuesta al principio.

Con esta construcción la singularidad de la cuerda de Dirac no aparece en la descripción, demostrando que no es más que un artefacto de la elección de gauge. Clásicamente el procedimiento de utilizar potenciales distintos en regiones distintas del espacio es directo. Cuánticamente es más sutil, puesto que en la subsección anterior hemos visto que al cambiar de gauge, la función de onda de un electrón coge una fase. Por lo tanto, las funciones de onda en el hemisferio norte y sur están relacionadas mediante

$$\Psi^{(n)}(\vec{r}) = \exp\left(-\frac{ieg}{2\pi}\phi\right)\Psi^{(s)}(\vec{r}). \tag{6.87}$$

Sin embargo, en el ecuador $\theta = \pi/2$, la función de onda tiene que ser univaluada al dar una vuelta alrededor de la esfera. Vemos por lo tanto que la fase ha de tener el mismo valor para $\phi = 0$ y $\phi = 2\pi$, lo que implica que $eg/\hbar = 2n\pi$, donde n es un número entero. Escrito de otra forma,

llegamos a la llamada *condición de cuantización de carga de Dirac*:¹³

$$e = \frac{2n\pi\hbar}{g}. \quad (6.88)$$

En otras palabras, con que haya un sólo monopolio magnético en alguna parte del universo, todas las cargas eléctricas están cuantizadas, es decir aparecen como múltiplos de una carga mínima $e = 2\pi\hbar/g$. Esto es curioso, aunque nunca se ha visto (con seguridad) un monopolio magnético, la cuantización de carga es precisamente lo que ocurre en la naturaleza: todas las partículas conocidas tienen una carga que es un múltiplo entero de la carga del electrón.¹⁴ Por ejemplo, aunque el electrón y el protón son dos partículas muy distintos, la cota experimental de la diferencia de sus cargas es

$$\frac{|q_{e^-} + q_p|}{e} < 1,0 \cdot 10^{-21}. \quad (6.89)$$

Obviamente esto no puede servir como prueba de la existencia de monopolos, pero por otro lado no se conoce ningún otro mecanismo convincente que pueda explicar la cuantización de la carga. Es más, con otros intentos, embebiendo el grupo gauge $U(1)$ dentro de un grupo G más grande, uno consigue cuantizar la carga por métodos estándar de la mecánica cuántica, pero entonces resulta que los llamados *monopolos de 't Hooft-Polyakov* aparecen por la puerta trasera como defectos topológicos al romper la simetría del grupo G a $U(1)$. En otras palabras, parece que (de momento) no hay manera de cuantizar la carga eléctrica sin de algún modo introducir monopolos magnéticos también.

Igual que en el caso del efecto Aharonov-Bohm, la base matemática del monopolio de Dirac es la estructura topológica no trivial. El espacio de configuraciones es $\mathbb{R}^3 \setminus \{0\}$, puesto que el potencial es divergente en $\vec{r} = 0$. (El resto de la cuerda de Dirac la hemos eliminado por el truco de Wu y Yang, pero la singularidad en el origen aparece en ambos potenciales). Estamos por lo tanto trabajando con un fibrado $U(1)$ sobre el espacio $\mathbb{R}^3 \setminus \{0\}$, que tiene la misma homotopía que S^2 . La S^2 está descrita por dos cartas (parches), el hemisferio norte y el sur, y la función de transición, que nos cuenta cómo encajan los parches, es justo la transformación gauge $\exp(-ieg\phi/2\pi)$, definida en el ecuador, donde se solapan los dos parches. Igual que en el caso del efecto Aharonov-Bohm tenemos por lo tanto una función $\Lambda : S^1 \rightarrow S^1$, clasificada por el grupo de homotopía $\pi(S^1) = \mathbb{Z}$. El número entero n que aparece en la condición de cuantización de carga (6.88) es básicamente la clase de homotopía de Λ .

Los monopolos magnéticos, a pesar de no tener ninguna evidencia de su existencia, son de gran interés en la física moderna. Dado que en la teoría cuántica de campos, al carga eléctrica e es la constante de acoplamiento de la teoría, la simetría entre \vec{E} y \vec{B} en las leyes de Maxwell, indicarían una descripción dual, en función de unos potenciales duales y una constante de acoplo g . Además, la condición de cuantización de carga (6.88) implica que si la descripción eléctrica es débilmente acoplada, la descripción dual sería fuertemente acoplada. Las teorías fuertemente acopladas son en general muy difíciles de manejar, puesto que no se puede aplicar técnicas perturbativas. Pero dado la simetría de las leyes de Maxwell, el comportamiento con acoplamientos fuertes sería completamente simétrico al acoplamiento débil. La dualidad electromagnética y la existencia de monopolos, sería una oportunidad enorme de extender nuestros conocimientos más allá que el régimen perturbativo.

¹³La misma fórmula se puede derivar también utilizando un sólo potencial en el espacio entero. Para evitar un efecto Aharonov-Bohm visible alrededor de la cuerda de Dirac, hay que ajustar la fase tal que es un múltiplo de 2π .

¹⁴Existen partículas llamadas *quarks*, los constituyentes de por ejemplo los protones y neutrones, que tienen una carga $q = \pm e/3$ ó $q = \pm 2e/3$, por lo que parecen violar la condición de cuantización de la carga. Sin embargo, por una propiedad de la teoría gauge que describe su comportamiento, llamada *libertad asintótica*, estos quarks no pueden aparecer libremente. Debido a este *confinamiento de quarks*, estos se manifiestan siempre en conjuntos de dos o tres, tal que la carga total del conjunto es un múltiplo de la carga del electrón.

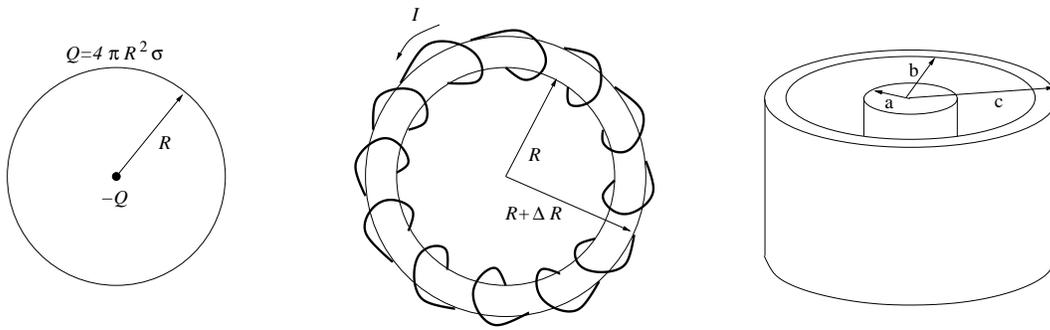


Figura 6.10: El problema de las cargas concéntricas (ejercicio 1), el problema del solenoide toroidal (ejercicio 3) y el problema de los conductores concéntricos (ejercicio 4).

6.6. Problemas

1. Considera la configuración de Figura 6.10a: una esfera con radio R tiene en su superficie una carga positiva $Q = 4\pi R^2\sigma$ uniformemente distribuida, donde la constante σ es la densidad de carga por unidad de superficie. En el centro de la esfera hay otra carga negativa $-Q$ puntual. Calcula el campo eléctrico \vec{E} dentro y fuera de la esfera, utilizando la ley de Gauss.
2. Calcula el campo eléctrico \vec{E} generado por una cuerda recta infinita con radio a , que lleva una densidad de carga constante de $\rho = \lambda/(\pi a^2)$, donde λ es la carga por unidad de longitud.
3. Considera un toro con radio interior R y radio exterior $R + \Delta R$, alrededor del cual está enroscado N veces un conductor, de modo que se forma un solenoide toroidal (vease Figura 6.10b). Calcula el campo magnético \vec{B} para $r < R$, para $R < r < R + \Delta R$ y para $R + \Delta R < r$, cuando pasa una corriente I por el conductor.
4. Un conductor infinito recto hueco, con radio interior b y radio exterior c está colocado coaxialmente con respecto a otro conductor infinito recto con radio a (vease Figura 6.10c). A través del conductor hueco pasa una densidad de corriente $\vec{j}_1 = I\vec{e}_z/\pi(c^2 - b^2)$ y a través del conductor interno una densidad de corriente $\vec{j}_2 = -I\vec{e}_z/\pi a^2$. Calcula el campo magnético para $r < a$, para $a < r < b$, para $b < r < c$ y para $c < r$. Compara los resultados con los campos generados por cada uno de los conductores independientes. ¿Qué conclusión se puede sacar de esta comparación?
5. Considera una densidad de corriente $\vec{j}(r)$ esféricamente simétrica, saliendo de una distribución de carga esféricamente simétrica (vease Figura 6.11). Motiva, usando argumentos de simetría, que esta corriente esféricamente simétrica no puede generar un campo magnético. Después, demuestra explícitamente que el campo magnético es cero, utilizando la conservación de carga y el término de Maxwell en la ley de Ampère-Maxwell.
6. El lagrangiano de electromagnetismo viene dado por

$$\mathcal{L} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2) - \rho\phi + \vec{j} \cdot \vec{A}. \quad (6.90)$$

- a) Escribe el lagrangiano en términos de los potenciales ϕ y \vec{A} y comprueba que es invariante gauge si se cumple la ley de conservación de carga.
- b) Demuestra que las ecuaciones de movimiento de ϕ y \vec{A} corresponden con las ecuaciones inhomogéneas de Maxwell. ¿Qué ocurre con las ecuaciones homogéneas?
- c) Considera el lagrangiano (6.90) en ausencia de cargas y corrientes ($\rho = \vec{j} = 0$) e impone

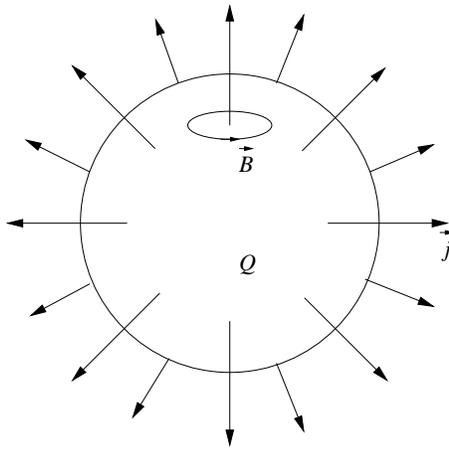


Figura 6.11: *El problema de la corriente radialmente simétrica (ejercicio 5)*

sobre los potenciales la condición del gauge axial $\phi = 0$. Calcula el hamiltoniano correspondiente a \mathcal{L} , si definimos el momento conjugado a A_i como

$$\Pi_i = \frac{\delta \mathcal{L}}{\delta(\partial_t A_i)}. \quad (6.91)$$

Escribe el hamiltoniano otra vez en términos de \vec{E} y \vec{B} y interpretalo. ¿Es el hamiltoniano invariante gauge?

Capítulo 7

Pequeñas oscilaciones

Capítulo 8

Introducción a la mecánica cuántica

Capítulo 9

Electromagnetismo avanzado y teorías gauge

- 9.1. Teoría de Maxwell en forma covariante
- 9.2. Soluciones topológicas
- 9.3. Teorías gauge