

Apuntes de Mecánica Clásica

Fernando O. Minotti

2do cuatrimestre de 2017

Índice general

1. Mecánica de Newton	3
1.1. Transformación de Galileo	4
1.2. Sistema de varias partículas	4
1.3. Teorema del Virial	8
2. Mecánica analítica	11
2.1. Definiciones básicas y notación	11
2.2. Principio de los trabajos virtuales (D 'Alembert)	13
2.3. Ecuaciones de Lagrange	15
2.3.1. Partículas en campos electromagnéticos	20
2.4. Principio de Hamilton	21
2.4.1. Principio de Maupertuis	24
2.5. Invarianza de las ecuaciones de Lagrange y simetrías	26
2.5.1. Teorema de Noether	29
2.6. Acción como función de las q 's	32
3. Ecuaciones canónicas de Hamilton	34
3.0.1. Transformación de Legendre y ecuaciones de Hamilton	34
3.0.2. Formalismo simpléctico	37
3.1. Transformaciones canónicas y corchetes de Poisson	38
3.1.1. Transformaciones canónicas	38
3.1.2. Transformaciones canónicas infinitesimales	42
3.1.3. Teorema de Liouville	44
3.1.4. Corchetes de Poisson	46
4. Ecuación de Hamilton-Jacobi	51
4.1. Variables ángulo-acción	53
4.2. Invariantes adiabáticos	56

5. Mecánica relativista	59
5.1. Cinemática relativista	59
5.1.1. Cuadriectores	63
5.2. Dinámica relativista	67
5.2.1. Leyes de conservación en sistemas de varias partículas .	72
5.2.2. Desintegración de partículas	75
5.2.3. Choque de partículas	77
6. Fuerzas centrales	79
6.1. Problema de Kepler	79
6.2. Choque elástico	86
6.3. Dispersión (Scattering)	90
7. Pequeñas oscilaciones	96
7.1. Modos normales	103
7.2. Oscilaciones de sistemas aislados (moléculas)	104
7.3. Oscilaciones forzadas y amortiguadas	105
7.4. Oscilaciones no lineales en una dimensión	111
8. Cuerpo rígido	114
8.1. Cinemática	114
8.1.1. Matrices de rotación	116
8.1.2. Ángulos de Euler	118
8.2. Dinámica	122
8.2.1. Energía cinética	122
8.2.2. Teorema de Steiner	125
8.2.3. Momento angular	126
8.2.4. Ejes principales del tensor de inercia	127
8.3. Ecuaciones de Euler	129
8.4. Movimiento del cuerpo sólido libre	131
8.4.1. Construcción de Poinsot	131
8.4.2. Otra representación geométrica	133
8.4.3. Estabilidad de la rotación alrededor de los ejes principales	133
8.4.4. Elipsoide con simetría de revolución	135
8.5. Movimiento de trompos y giróscopos	138
8.5.1. Trompo	138
8.5.2. Estabilidad del trompo vertical	144
8.5.3. Giróscopo	145

Capítulo 1

Mecánica de Newton

La mecánica de Newton (también llamada *mecánica vectorial*) se basa en sus tres leyes que adaptamos aquí a la nomenclatura de la materia:

i) Un punto material (o partícula) sobre el que no actúan fuerzas permanece continuamente en reposo o en movimiento rectilíneo y uniforme. (Esta ley establece el marco de referencia en el que son válidas las otras leyes; se entiende que no actúan fuerzas sobre el punto material si no existe ningún agente que las ejerza, tal como otra partícula o cuerpo extenso, hilos, resortes, cuerpos muy masivos que produzcan atracción gravitatoria, cuerpos cargados y corrientes eléctricas si el punto material tiene carga eléctrica, etc.. Si en tales condiciones el punto permanece en reposo o se mueve con velocidad constante, entonces se está en el marco adecuado, que es el que conocemos como sistema de referencia inercial).

ii) Si sobre la partícula actúan fuerzas la tasa de cambio de la cantidad de movimiento lineal de la partícula es igual a la fuerza total. (Se entiende por cantidad de movimiento lineal al producto de la masa de la partícula por su velocidad, y por tasa de cambio la variación por unidad de tiempo; nótese además que las fuerzas se dan por conocidas o determinables independientemente del cambio de estado de movimiento de la partícula)

iii) Cuando dos partículas interactúan la fuerza que la primera ejerce sobre la segunda es igual en intensidad y dirección, pero opuesta en sentido, a la que la segunda ejerce sobre la primera.

Para trabajar con estas leyes debemos definir los vectores involucrados; en particular, el vector posición de cada partícula respecto de algún origen conveniente. Llamamos \mathbf{x}_i al vector posición de la partícula i y escribimos la segunda ley para cada partícula i (de masa m_i y sobre la que actúa una fuerza \mathbf{F}_i) como (los puntos indican derivación temporal y el número de ellos el orden de la derivada)

$$m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i.$$

1.1. Transformación de Galileo

Si tenemos un sistema de referencia S en el que valen las leyes de Newton y elegimos describir la dinámica de un dado sistema de partículas desde otro sistema S' que se mueve respecto del primero con velocidad constante \mathbf{U} , las coordenadas espaciales y temporal de cualquier partícula en S' estarán relacionadas con las correspondientes en S por una *transformación de Galileo* expresada como (con notación evidente y suponiendo que en $t = 0$ los orígenes de ambos sistemas coinciden)

$$\begin{aligned}\mathbf{x} &= \mathbf{x}' + \mathbf{U}t', \\ t &= t'.$$

Esta transformación es puramente cinemática. La invarianza de las leyes de Newton frente a esta transformación; esto es, la validez de las mismas leyes, expresadas de igual manera, pero ahora en términos de objetos referidos al sistema S' , requiere que las masas y fuerzas sean invariantes ante la transformación. Esta última es una hipótesis básica de la dinámica de Newton que, de hecho, es válida para transformaciones genéricas, a sistemas de referencia acelerados de manera arbitraria.

1.2. Sistema de varias partículas

Llamando $\mathbf{p}_i \equiv m_i \dot{\mathbf{x}}_i$ a la cantidad de movimiento de la partícula i , la cantidad de movimiento total \mathbf{P} del sistema de N partículas es

$$\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i,$$

de la cual podemos decir que

$$\frac{d\mathbf{P}}{dt} = \sum_{i=1}^N \frac{d\mathbf{p}_i}{dt} = \sum_{i=1}^N \mathbf{F}_i^{ext} + \sum_{i=1}^N \sum_{j \neq i} \mathbf{f}_{ij}, \quad (1.1)$$

donde se ha diferenciado entre la fuerza externa neta sobre cada partícula, \mathbf{F}_i^{ext} , y la interna sufrida por la partícula i debida a la j , \mathbf{f}_{ij} . Por la tercera ley de Newton es $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$, por lo que la doble suma en (1.1) se anula y resulta entonces

$$\frac{d\mathbf{P}}{dt} = \sum_{i=1}^N \mathbf{F}_i^{ext} \equiv \mathbf{F}^{ext}. \quad (1.2)$$

Usando además que

$$\frac{d\mathbf{P}}{dt} = \frac{d}{dt} \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \equiv M \frac{d^2 \mathbf{X}_{CM}}{dt^2},$$

donde se ha introducido la masa total del sistema $M = \sum_{i=1}^N m_i$, y la posición del centro de masas $\mathbf{X}_{CM} = \left(\sum_{i=1}^N m_i \mathbf{x}_i \right) / M$, puede escribirse la ecuación de movimiento del centro de masas como

$$M \frac{d^2 \mathbf{X}_{CM}}{dt^2} = \mathbf{F}^{ext}.$$

Si definimos ahora la cantidad de movimiento angular (o momento angular) respecto de un punto fijo O ,

$$\mathbf{L}_o = \sum_{i=1}^N \mathbf{x}_{io} \times \mathbf{p}_i, \quad (1.3)$$

donde se entiende que \mathbf{x}_{io} es el vector posición de la partícula i referido al origen O . Tenemos entonces

$$\frac{d\mathbf{L}_o}{dt} = \sum_{i=1}^N \dot{\mathbf{x}}_{io} \times \mathbf{p}_i + \sum_{i=1}^N \mathbf{x}_{io} \times \dot{\mathbf{p}}_i.$$

Dado que $\dot{\mathbf{x}}_{io}$ es paralelo a \mathbf{p}_i cada término de la primera sumatoria es nulo, mientras que el segundo se escribe, usando la segunda ley de Newton

$$\frac{d\mathbf{L}_o}{dt} = \sum_{i=1}^N \mathbf{x}_{io} \times \mathbf{F}_i^{ext} + \sum_{i=1}^N \mathbf{x}_{io} \times \sum_{j \neq i} \mathbf{f}_{ij}.$$

Si escribimos que

$$\mathbf{x}_{jo} = \mathbf{x}_{io} + \Delta \mathbf{x}_{ij},$$

al efectuar la doble sumatoria se tendrán pares de términos de la forma

$$\begin{aligned} \mathbf{x}_{io} \times \mathbf{f}_{ij} + \mathbf{x}_{jo} \times \mathbf{f}_{ji} &= \mathbf{x}_{io} \times \mathbf{f}_{ij} + \mathbf{x}_{io} \times \mathbf{f}_{ji} + \Delta \mathbf{x}_{ij} \times \mathbf{f}_{ji} \\ &= \mathbf{x}_{io} \times (\mathbf{f}_{ij} + \mathbf{f}_{ji}) + \Delta \mathbf{x}_{ij} \times \mathbf{f}_{ji}. \end{aligned}$$

El paréntesis del segundo renglón se anula por la tercera ley de Newton, mientras que si la fuerza interna \mathbf{f}_{ij} tiene la dirección del vector que une las

partículas i, j (forma fuerte de la tercera ley) el último producto vectorial se anula también, con lo que se obtiene finalmente

$$\frac{d\mathbf{L}_o}{dt} = \sum_{i=1}^N \mathbf{x}_{io} \times \mathbf{F}_i^{ext}. \quad (1.4)$$

Tenemos entonces de (1.2) y (1.4) que si no actúan fuerzas externas sobre las partículas del sistema, entonces $\mathbf{P} = cte$ y $\mathbf{L}_o = cte$; la constancia de \mathbf{L}_o es válida cualquiera sea el punto fijo O considerado.

Si, conocido el valor de \mathbf{L}_o en un instante dado (no estamos pidiendo que se conserve), queremos calcular el momento angular respecto de otro punto fijo O' , basta escribir en (1.3) que $\mathbf{x}_{io} = \mathbf{x}_{io'} + \mathbf{X}_{o'} - \mathbf{X}_o$, con \mathbf{X}_o y $\mathbf{X}_{o'}$ los vectores posición de los puntos O y O' , respectivamente, para obtener

$$\begin{aligned} \mathbf{L}_o &= \sum_{i=1}^N \mathbf{x}_{io'} \times \mathbf{p}_i + (\mathbf{X}_{o'} - \mathbf{X}_o) \times \sum_{i=1}^N \mathbf{p}_i \\ &= \mathbf{L}_{o'} + (\mathbf{X}_{o'} - \mathbf{X}_o) \times \mathbf{P}, \end{aligned}$$

con lo que

$$\mathbf{L}_{o'} = \mathbf{L}_o - (\mathbf{X}_{o'} - \mathbf{X}_o) \times \mathbf{P}.$$

Conocido el valor del momento angular \mathbf{L}_o respecto del origen O de un sistema de referencia S , se plantea ahora calcular el momento angular $\mathbf{L}'_{o'}$ respecto del origen O' de un sistema de referencia S' que se mueve con velocidad arbitraria $\mathbf{U}(t)$ respecto de S . Si $\mathbf{X}_{o'}(t)$ es la posición de O' respecto de O (por supuesto, es $\mathbf{U}(t) = \dot{\mathbf{X}}_{o'}(t)$), tenemos

$$\mathbf{L}'_{o'} = \sum_{i=1}^N \mathbf{x}'_i \times m_i \dot{\mathbf{x}}'_i = \sum_{i=1}^N (\mathbf{x}_i - \mathbf{X}_{o'}) \times m_i (\dot{\mathbf{x}}_i - \mathbf{U}),$$

que, expandiendo los productos, se puede escribir como

$$\begin{aligned} \mathbf{L}'_{o'} &= \sum_{i=1}^N \mathbf{x}_i \times m_i \dot{\mathbf{x}}_i - \mathbf{X}_{o'} \times \sum_{i=1}^N m_i \dot{\mathbf{x}}_i - \\ &\quad \left(\sum_{i=1}^N m_i \dot{\mathbf{x}}_i \right) \times \mathbf{U} + \mathbf{X}_{o'} \times \left(\sum_{i=1}^N m_i \right) \mathbf{U} \\ &= \mathbf{L}_o - \mathbf{X}_{o'} \times \mathbf{P} - M \mathbf{X}_{CM} \times \mathbf{U} + M \mathbf{X}_{o'} \times \mathbf{U}. \end{aligned}$$

En particular, si se elige el CM como origen O' , la expresión se simplifica a

$$\mathbf{L}'_{CM} = \mathbf{L}_o - \mathbf{X}_{CM} \times \mathbf{P},$$

que escrito en la forma más cómoda

$$\mathbf{L}_o = \mathbf{L}'_{CM} + \mathbf{X}_{CM} \times \mathbf{P},$$

nos indica que el momento angular de un sistema respecto de un punto cualquiera puede descomponerse en el momento respecto al centro de masas (llamado momento angular propio o de spin) más el debido al movimiento del sistema de partículas como conjunto (momento angular orbital).

Consideremos finalmente la variación de energía cinética T de un sistema de partículas (el punto indica producto escalar de vectores)

$$\frac{dT}{dt} = \frac{d}{dt} \left(\frac{1}{2} \sum_{i=1}^N m_i |\dot{\mathbf{x}}_i|^2 \right) = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \ddot{\mathbf{x}}_i,$$

que por la segunda ley de Newton es igual a

$$\frac{dT}{dt} = \sum_{i=1}^N \dot{\mathbf{x}}_i \cdot \mathbf{F}_i^{ext} + \sum_{i=1}^N \dot{\mathbf{x}}_i \cdot \sum_{j \neq i} \mathbf{f}_{ij}.$$

Si las fuerzas entre pares de partículas derivan de un potencial (independiente del tiempo) $V_{ij} = V_{ji} = V_0(|\mathbf{x}_i - \mathbf{x}_j|)$ tenemos

$$\mathbf{f}_{ij} = -\frac{\partial V_{ij}}{\partial \mathbf{x}_i},$$

con lo que

$$\sum_{j \neq i} \mathbf{f}_{ij} = -\frac{\partial}{\partial \mathbf{x}_i} \sum_{j \neq i} V_{ij} = -\frac{\partial V^{(i)}}{\partial \mathbf{x}_i},$$

donde $V^{(i)} \equiv \sum_{j \neq i} V_{ij}$. Si se define adicionalmente $V \equiv \frac{1}{2} \sum_{i=1}^N V^{(i)}$ tenemos que, para cualquier i ,

$$\frac{\partial V}{\partial \mathbf{x}_i} = \frac{\partial V^{(i)}}{\partial \mathbf{x}_i},$$

con lo que podemos escribir

$$\sum_{i=1}^N \dot{\mathbf{x}}_i \cdot \sum_{j \neq i} \mathbf{f}_{ij} = -\sum_{i=1}^N \dot{\mathbf{x}}_i \cdot \frac{\partial V}{\partial \mathbf{x}_i} = -\frac{dV}{dt},$$

y, finalmente,

$$\frac{d}{dt} (T + V) = \sum_{i=1}^N \dot{\mathbf{x}}_i \cdot \mathbf{F}_i^{ext}.$$

Así, si no existen fuerzas externas, la energía mecánica $E \equiv T + V$, suma de las energías cinética T y potencial V , se conserva. Más aún, si la fuerza externa deriva también de un potencial independiente del tiempo

$$\mathbf{F}_i^{ext} = -\frac{\partial V^{ext}}{\partial \mathbf{x}_i},$$

podemos deducir de manera semejante

$$\frac{d}{dt} (T + V + V^{ext}) = 0.$$

1.3. Teorema del Virial

Si definimos la función virial de un sistema de partículas

$$\mathcal{V} \equiv \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \mathbf{x}_i,$$

tenemos que su variación temporal se escribe

$$\begin{aligned} \frac{d\mathcal{V}}{dt} &= \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \dot{\mathbf{x}}_i + \sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \mathbf{x}_i \\ &= 2T + \sum_{i=1}^N \mathbf{x}_i \cdot \left(\mathbf{F}_i^{ext} + \sum_{j \neq i} \mathbf{f}_{ij} \right). \end{aligned}$$

Si, como en el fin del punto anterior las fuerzas entre partículas son derivables de un potencial, y no existen fuerzas externas, se tiene

$$\frac{d\mathcal{V}}{dt} = 2T - \sum_{i=1}^N \mathbf{x}_i \cdot \frac{\partial V}{\partial \mathbf{x}_i}.$$

Si, además, el potencial V es una función homogénea de grado k ($k = -1$ para el potencial gravitatorio y $k = 2$ para el potencial elástico)

$$\sum_{i=1}^N \mathbf{x}_i \cdot \frac{\partial V}{\partial \mathbf{x}_i} = kV,$$

con lo cual

$$\frac{d\mathcal{V}}{dt} = 2T - kV. \quad (1.5)$$

Calculemos ahora el promedio temporal de esta variación, definido como

$$\left\langle \frac{d\mathcal{V}}{dt} \right\rangle \equiv \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} \frac{d\mathcal{V}}{dt} dt = \lim_{\tau \rightarrow \infty} \frac{\mathcal{V}(\tau) - \mathcal{V}(-\tau)}{2\tau}.$$

Si ninguna partícula se aleja al infinito, ni sus velocidades divergen, de la definición del virial, su valor será finito siempre y tendremos por lo tanto

$$\left\langle \frac{d\mathcal{V}}{dt} \right\rangle = 0$$

con lo cual resulta el teorema del virial:

$$2 \langle T \rangle = k \langle V \rangle.$$

Dado además que no hay fuerzas externas,

$$\langle T \rangle + \langle V \rangle = E,$$

con lo que se concluye

$$\begin{aligned} \langle T \rangle &= \frac{kE}{2+k}, \\ \langle V \rangle &= \frac{2E}{2+k}. \end{aligned}$$

En el caso gravitatorio ($k = -1$), dado que $T > 0$, debe ser entonces $E < 0$ y $\langle V \rangle = -2 \langle T \rangle$. En el caso elástico es $\langle V \rangle = \langle T \rangle = E/2$.

Consideremos ahora una función auxiliar w definida como

$$w \equiv \frac{1}{2} \sum_{i=1}^N m_i \mathbf{x}_i \cdot \mathbf{x}_i, \quad (1.6)$$

que por derivación directa da

$$\frac{dw}{dt} = \sum_{i=1}^N m_i \dot{\mathbf{x}}_i \cdot \mathbf{x}_i = \mathcal{V},$$

con lo que, volviendo a derivar, y usando (1.5),

$$\frac{d^2 w}{dt^2} = \frac{d\mathcal{V}}{dt} = 2T - kV.$$

En el caso gravitatorio tenemos entonces

$$\frac{d^2 w}{dt^2} = 2T + V = T + E \geq E,$$

desigualdad que, al integrar dos veces resulta en

$$w \geq \frac{1}{2}Et^2 + \dot{w}(0)t + w(0).$$

De esta ecuación se deduce que si la energía del sistema de partículas con interacción gravitatoria es positiva, entonces $w \rightarrow \infty$ cuando $t \rightarrow \pm\infty$, lo que significa, de la definición (1.6) de w , que al menos una de las partículas se aleja indefinidamente.

Capítulo 2

Mecánica analítica

2.1. Definiciones básicas y notación

En muchos casos los vectores posición no son los más convenientes para describir el movimiento de las partículas de un sistema mecánico. Claramente es más sencillo dar la ubicación de un planeta por su distancia al sol y un par de ángulos apropiados que por las tres componentes cartesianas de su vector posición. Llamemos entonces q_1, q_2, \dots, q_{3N} a las $3N$ variables que sirven para determinar en forma completa el estado de un sistema mecánico formado por N partículas. Determinar el estado de un sistema mecánico se entiende como dar las posiciones (en un dado instante) de todas las partículas dotadas de masa que lo conforman. Expresado en ecuaciones, para cada partícula i ($1 \leq i \leq N$) podremos expresar su vector posición como una función $\mathbf{x}_i(q_1, q_2, \dots, q_{3N})$.

Las q_k son llamadas *coordenadas generalizadas* y el espacio que determinan se denomina *espacio de configuración*. Al evolucionar el sistema en el tiempo las coordenadas generalizadas describen trayectorias $q_k(t)$ en el espacio de configuración, por lo que podemos definir las *velocidades generalizadas* $\dot{q}_k(t)$. Las ecuaciones dinámicas que se deducirán permiten la determinación (al menos en principio) de las trayectorias $q_k(t)$ e involucran funciones que dependen de las coordenadas y velocidades generalizadas y del tiempo.

Cuando existen vínculos que ligan algunas o todas las partículas entre sí o con cuerpos externos, el sistema de N partículas sigue siendo descrito por $3N$ coordenadas generalizadas, pero éstas no son independientes; se habla entonces de ligaduras o vínculos entre las variables.

Si existen m ligaduras independientes entre las n variables se denomina *número de grados de libertad*, o simplemente *grados de libertad*, al valor $n = 3N - m$. Este número es una propiedad del sistema y no de las variables que

se utilizan.

La pregunta es entonces si es posible trabajar con sólo n variables q_k . La respuesta es que esto depende de si se puede, al menos en principio, usar las ecuaciones de vínculo para escribir m de las variables en términos de las $3N - m$ restantes; cuando esto sucede se dice que los vínculos son *holónomos*. Los vínculos que no permiten esto se denominan *no holónomos*.

Un ejemplo extremo es el de un cuerpo rígido (sólido), que podemos pensar como constituido por un número enorme de partículas ligadas por la condición de que sus distancias relativas permanezcan constantes. Estas ligaduras reducen el número de variables desde el impensable número $3N$ a sólo 6; tres de las cuales son, por ejemplo, las coordenadas del centro de masa del cuerpo; las otras tres serán ángulos que determinan la orientación.

Si el sólido está además forzado a rotar alrededor de un eje fijo, basta con dar un solo ángulo para determinar su estado; las ecuaciones de ligadura serían que dos de los ángulos que determinan la orientación son fijos y que cada coordenada del centro de masa se puede escribir como función del tercer ángulo. Estas ligaduras son holónomas de manera que uno podría efectivamente hacer la reducción de variables hasta sólo una.

Si el sólido es una esfera que se mueve sobre una superficie horizontal la altura de su centro es fija (ligadura holónoma), por lo que sólo hacen falta dos coordenadas para determinarlo; si además rueda sin deslizar, su punto de contacto con la superficie está instantáneamente en reposo, lo que vincula la velocidad instantánea de su centro con su velocidad de rotación. Esto último corresponde a dos ligaduras no holónomas porque la ligadura de estas velocidades no alcanza para determinar el vínculo entre las coordenadas del centro y los ángulos de orientación mismos, y no hace posible la eliminación de unos en favor de otros (una esfera, partiendo del mismo lugar y con igual orientación, puede llegar a un mismo punto con distintas orientaciones, según el camino que haya seguido, aun respetando punto a punto la ligadura entre velocidades; por otro lado, un cilindro que rodara sin deslizar tendría una única relación entre coordenada de su centro y ángulo rotado, por lo que en este caso el vínculo dado por la rodadura es holónomo).

Así, para un sólido no ligado es $n = 6$, para uno que gira sobre un eje fijo es $n = 1$; para una esfera que se mueve sobre una superficie deslizando libremente es $n = 5$; si rota sin deslizar $n = 3$, pero se requieren 5 variables, ligadas por dos vínculos no holónomos, para describir el estado de la esfera.

Finalmente, los vínculos (holónomos o no) pueden depender explícitamente del tiempo; por ejemplo, el eje alrededor del que es forzado a rotar un sólido o la superficie sobre la que rueda una esfera pueden seguir un movimiento prefijado. Los vínculos dependientes del tiempo se denominan *reónomos* y los no dependientes del tiempo *esclerónomos*.

2.2. Principio de los trabajos virtuales (D'Alembert)

La introducción de ligaduras en el sistema mecánico lleva al concepto de *fuerza de vínculo*, que es justamente la que se ejerce sobre la partícula para forzar el cumplimiento de la ligadura. Esta fuerza de vínculo se diferencia de la denominada *fuerza aplicada* que es aquella determinada independientemente de cualquier otra fuerza, dando sólo las posiciones (y a veces también las velocidades) de las partículas. Así, si dos partículas están ligadas por un resorte la fuerza que ejerce el resorte sobre una de ellas es una fuerza aplicada, que depende de la posición de ambas partículas; también son fuerzas aplicadas el peso, la fuerza eléctrica sobre una partícula cargada, la fuerza magnética (que depende de la velocidad), etc. Por otro lado, la fuerza que ejerce un riel que guía el movimiento de una partícula es una fuerza de vínculo, que no puede ser determinada sin conocer las otras fuerzas que actúan.

Una restricción adicional que imponemos a las fuerzas de vínculo es que puedan ser tan grandes en magnitud como fuera necesario para imponer la ligadura, lo que es una idealización de los vínculos reales (los hilos se estiran, las varillas se doblan o se quiebran, pero trabajamos dentro de los límites en lo que esto no pasa o su efecto puede despreciarse).

Un problema con la condición anterior lo dan las fuerzas de rozamiento. Si las condiciones del problema son tales que el rozamiento es suficiente para impedir que haya deslizamiento (rozamiento estático), la fuerza de rozamiento se considera entonces de vínculo. Si pudiera haber deslizamiento (rozamiento dinámico) deberíamos considerar al rozamiento como una fuerza aplicada anómala (ya que no cumple con ser independiente de otras fuerzas dado que su magnitud depende de la fuerza de vínculo normal), pero ya no puede ser considerada fuerza de vínculo.

Otro concepto fundamental es el de *desplazamiento virtual*, que es un desplazamiento infinitesimal de la posición de una dada partícula, realizado *instantáneamente* (de aquí la condición de virtual, ya que no es posible realizarlo efectivamente); es decir, a velocidad infinita, sin que transcurra el tiempo durante el desplazamiento. Aparte de ser instantáneo, el desplazamiento es arbitrario, no relacionado con el movimiento real de la partícula en el instante considerado. Sin embargo, los desplazamientos virtuales más útiles son los que respetan los vínculos; esto es, no violan las condiciones de ligadura del sistema (no sacan la partícula del riel que la guía, no deforman los cuerpos rígidos, no estiran los hilos, etc.; sin embargo, aun así no corresponden necesariamente al movimiento real). Estos desplazamientos se denominan *compatibles con los vínculos*.

Es importante notar que si los vínculos fueran dependientes del tiempo, al ser instantáneo el desplazamiento virtual, los vínculos permanecen en el estado en que se encuentran en el instante del desplazamiento, por lo que los desplazamientos virtuales compatibles con los vínculos deben respetar la condición impuesta por éstos en ese dado instante.

El *trabajo virtual* de una fuerza es entonces el trabajo que ella realiza en el desplazamiento virtual.

Finalmente, el *principio de los trabajos virtuales* de D 'Alembert **postula** que la suma de los trabajos virtuales de **todas** las fuerzas de vínculo de un sistema es nula, para cualquier conjunto de desplazamientos virtuales, compatibles con los vínculos, de las partículas del sistema.

Es esencial incluir todas las fuerzas de vínculo en la suma, ya que el trabajo virtual de una dada fuerza de vínculo es en general no nulo; la cancelación se da entre todos los sumandos. Nótese que las fuerzas aplicadas (no de vínculo) producen un trabajo total no nulo en general; el principio se aplica sólo a las fuerzas de vínculo.

Algunos autores deducen este principio de la tercera ley de Newton, aunque esta ley más bien puede hacerlo plausible, verificando el principio en casos concretos, y no realmente probarlo de manera general.

Notamos \mathbf{R}_i a la fuerza de vínculo (total o neta) que actúa sobre la partícula i , y \mathbf{F}_i a la fuerza aplicada (también total o neta) que actúa sobre esta partícula. Si $\delta\mathbf{x}_i$ es el desplazamiento virtual de la partícula i , el principio de D 'Alembert asegura que (el punto simboliza el producto escalar de vectores)

$$\delta W_R = \sum_{i=1}^N \mathbf{R}_i \cdot \delta\mathbf{x}_i = 0,$$

para todos los $\delta\mathbf{x}_i$ compatibles con los vínculos.

Nota histórica: El principio de los trabajos virtuales como es presentado aquí es el resultado de muchas contribuciones a lo largo del tiempo. Inicialmente se aplicó en forma elemental y medio velada a problemas de estática comenzando por Aristóteles (384-322 AC) y pasando por Stevinus (1598-1620) y Galileo (1564-1642). En su forma más explícita y general (siempre en el caso estático) fue dado por Juan Bernoulli (1667-1748) alrededor de 1717. Fue D 'Alembert (1717-1785) quien formuló su principio, que dice en realidad que un problema dinámico puede reducirse a uno estático en el que se han agregado a las fuerzas reales (de vínculo y aplicadas) las fuerzas de inercia, de expresión $-m_i\ddot{\mathbf{x}}_i$, permitiendo así el uso del principio en casos dinámicos. Denominar principio de D 'Alembert al presentado aquí es entonces no del todo correcto, aunque usual.

2.3. Ecuaciones de Lagrange

Si escribimos la ecuación de movimiento de la partícula i denotando ambos tipos de fuerzas actuantes

$$m_i \ddot{\mathbf{x}}_i = \mathbf{R}_i + \mathbf{F}_i, \quad (2.1)$$

multiplicamos escalarmente esta ecuación por el desplazamiento virtual $\delta \mathbf{x}_i$ de la partícula y sumamos para todas las partículas, el principio de D'Alembert nos dice que podemos escribir (pasando todo al lado izquierdo)

$$\sum_{i=1}^N m_i \ddot{\mathbf{x}}_i \cdot \delta \mathbf{x}_i - \sum_{i=1}^N \mathbf{F}_i \cdot \delta \mathbf{x}_i = 0 \quad (2.2)$$

Como existen vínculos los desplazamientos de las distintas partículas no son independientes entre sí (recuérdese que aquí los desplazamientos virtuales deben respetar los vínculos). Supongamos tener m vínculos holónomos entre las partículas, que pueden escribirse como m relaciones entre las posiciones de las partículas

$$G_r(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N, t) = 0, \quad 1 \leq r \leq m, \quad (2.3)$$

donde se ha puesto de manifiesto que las relaciones de vínculo pueden depender explícitamente del tiempo, como se discutió más arriba. La condición de que los desplazamientos respeten los vínculos se escribe entonces

$$\sum_{i=1}^N \nabla_i G_r \cdot \delta \mathbf{x}_i = 0, \quad 1 \leq r \leq m, \quad (2.4)$$

donde ∇_i representa el gradiente respecto de las coordenadas de partícula i . Si los vínculos no son holónomos, de cualquier manera pueden expresarse en general de forma diferencial como en (2.4), sólo que en lugar de $\nabla_i G_r$ aparecerá una cantidad vectorial $\mathbf{A}_{ir}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N, t)$ que no puede ser expresada como el gradiente de una función respecto de las coordenadas \mathbf{x}_i (se deduce más adelante su forma general). Escribimos entonces en general para vínculos holónomos o no las m condiciones sobre los desplazamientos como ($1 \leq r \leq m$)

$$\sum_{i=1}^N \mathbf{A}_{ir} \cdot \delta \mathbf{x}_i = 0. \quad (2.5)$$

La idea es multiplicar cada una de las relaciones (2.5) por una función escalar desconocida $\lambda_r(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N, t)$ (*multiplicador de Lagrange*) y sumar

todas ellas para escribir (cambiando el orden de las sumatorias)

$$\sum_{i=1}^N \left(\sum_{r=1}^m \lambda_r \mathbf{A}_{ir} \right) \cdot \delta \mathbf{x}_i = 0,$$

que a su vez podemos sumar a (2.2) para escribir (agrupando todo)

$$\sum_{i=1}^N \left(m_i \ddot{\mathbf{x}}_i - \mathbf{F}_i + \sum_{r=1}^m \lambda_r \mathbf{A}_{ir} \right) \cdot \delta \mathbf{x}_i = 0. \quad (2.6)$$

Como existen las m relaciones (2.5) entre las $3N$ componentes de los desplazamientos $\delta \mathbf{x}_i$, podemos elegir $3N - m$ de éstas en forma arbitraria, y las m restantes estarán dadas como función de las anteriores (que son independientes). La idea entonces es elegir los λ_r para que se satisfagan las ecuaciones

$$m_i \ddot{\mathbf{x}}_i - \mathbf{F}_i + \sum_{r=1}^m \lambda_r \mathbf{A}_{ir} = 0, \quad (2.7)$$

para los m componentes no independientes de los desplazamientos $\delta \mathbf{x}_i$. De esta manera, en (2.6) sólo sobreviven las $3N - m$ componentes independientes que, son por ser arbitrarias, indican que debe ser nulo cada uno de los factores que las multiplica. Así, para **todas** las partículas se debe satisfacer la ecuación (2.7) que reescribimos

$$m_i \ddot{\mathbf{x}}_i = \mathbf{F}_i - \sum_{r=1}^m \lambda_r \mathbf{A}_{ir}, \quad 1 \leq i \leq N, \quad (2.8)$$

denominadas ecuaciones de Lagrange de primera especie.

En particular, si comparamos (2.8) con la ecuación (2.1) vemos que las fuerzas de vínculo están dadas por

$$\mathbf{R}_i = - \sum_{r=1}^m \lambda_r \mathbf{A}_{ir}.$$

Por supuesto, como se tienen m funciones incógnita λ_r adicionales a las N coordenadas \mathbf{x}_i , se requieren m ecuaciones adicionales. Si los m vínculos son holónomos, las ecuaciones adicionales son las (2.3). Si algunos vínculos son no holónomos éstos están dados por una expresión de la forma (2.5) que no es posible integrar para obtener una expresión que relacione a las \mathbf{x}_i ; sin embargo, las ligaduras no holónomas son debidas en general a que se tiene una relación entre las velocidades de la forma genérica

$$\sum_{i=1}^N \mathbf{A}_{ir} \cdot \dot{\mathbf{x}}_i = a_r, \quad (2.9)$$

que sirven entonces como las ecuaciones auxiliares. Las (2.5) se deducen de estas expresiones al multiplicar por δt y luego tomar $\delta t = 0$ considerando que, por ser virtual el desplazamiento, es $|\dot{\mathbf{x}}_i| \rightarrow \infty$, de manera que $\delta \mathbf{x}_i = \dot{\mathbf{x}}_i \delta t$ es no nulo, pero $a_r \delta t = 0$.

Una forma mucho más útil de las ecuaciones de Lagrange se obtiene si se usan coordenadas generalizadas y si, además, los m vínculos son holónomos, de manera que pueden emplearse las relaciones de ligadura para eliminar m variables y trabajar con $n = 3N - m$ variables independientes. Así, es posible expresar las posiciones de cada partícula por una función de n coordenadas generalizadas *independientes*

$$\mathbf{x}_i = \mathbf{x}_i(q_1, q_2, \dots, q_n, t), \quad (2.10)$$

donde hemos incluido la posibilidad de que la relación dependa explícitamente del tiempo. Así, para los desplazamientos virtuales se tiene (el tiempo se mantiene fijo, $\delta t = 0$)

$$\delta \mathbf{x}_i = \sum_{k=1}^n \frac{\partial \mathbf{x}_i}{\partial q_k} \delta q_k, \quad (2.11)$$

que al reemplazar en (2.2) da

$$\sum_{i=1}^N \sum_{k=1}^n m_i \ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \delta q_k - \sum_{i=1}^N \sum_{k=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \delta q_k = 0. \quad (2.12)$$

Escribimos ahora

$$\ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} = \frac{d}{dt} \left(\dot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right) - \dot{\mathbf{x}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_k} \right), \quad (2.13)$$

y notamos que

$$\dot{\mathbf{x}}_i = \frac{\partial \mathbf{x}_i}{\partial t} + \sum_{k=1}^n \frac{\partial \mathbf{x}_i}{\partial q_k} \dot{q}_k,$$

de donde vemos que $\dot{\mathbf{x}}_i$ es una función lineal de las velocidades generalizadas \dot{q}_k que cumple

$$\frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_k} = \frac{\partial \mathbf{x}_i}{\partial q_k}. \quad (2.14)$$

Por otro lado,

$$\frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_k} \right) = \frac{\partial^2 \mathbf{x}_i}{\partial t \partial q_k} + \sum_{p=1}^n \frac{\partial \mathbf{x}_i}{\partial q_p \partial q_k} \dot{q}_p = \frac{\partial \dot{\mathbf{x}}_i}{\partial q_k}. \quad (2.15)$$

Usando (2.14) y (2.15) escribimos entonces (2.13) como

$$\begin{aligned}\ddot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} &= \frac{d}{dt} \left(\dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_k} \right) - \dot{\mathbf{x}}_i \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial q_k} \\ &= \frac{d}{dt} \left(\frac{1}{2} \frac{\partial |\dot{\mathbf{x}}_i|^2}{\partial \dot{q}_k} \right) - \frac{1}{2} \frac{\partial |\dot{\mathbf{x}}_i|^2}{\partial q_k}.\end{aligned}$$

Usando que la energía cinética del sistema de partículas es

$$T = \sum_{i=1}^N \frac{1}{2} m_i |\dot{\mathbf{x}}_i|^2,$$

y definiendo la fuerza generalizada

$$Q_k \equiv \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k},$$

podemos escribir la ecuación (2.12) como

$$\sum_{k=1}^n \left[\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - Q_k \right] \delta q_k = 0. \quad (2.16)$$

Notemos que al ser las variaciones δq_k arbitrarias e independientes debe anularse el corchete para cada k , con lo que resulta la forma útil de las ecuaciones de Lagrange

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - Q_k = 0. \quad (2.17)$$

Si las fuerzas aplicadas se derivan de un potencial independiente de las velocidades, $\mathbf{F}_i = -\nabla_i V$, entonces

$$Q_k = - \sum_{i=1}^N \nabla_i V \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} = - \frac{\partial V}{\partial q_k}, \quad (2.18)$$

por lo que, usando que $\partial V / \partial \dot{q}_k = 0$, podemos reescribir (2.17) como

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0, \quad (2.19)$$

donde se ha definido el *lagrangiano* $L = T - V$. Las (2.19) se denominan ecuaciones de Lagrange de segunda especie, o simplemente ecuaciones de Lagrange.

Si algunos de los vínculos son no holónomos no es posible reducir el número de coordenadas generalizadas hasta el número de grados de libertad n . Supongamos que se han usado algunas de las ligaduras holónomas (pero no necesariamente todas) para reducir en algo el número de coordenadas q_k y se tiene entonces s de ellas, con $3N \geq s > n$. Todas las deducciones que llevan desde (2.10) hasta (2.16) siguen siendo válidas con s en lugar de n , sólo que ahora las variaciones δq_k no son independientes entre sí, sino que están ligadas por las $s - n$ ligaduras no usadas para reducir el número de variables (sea porque no se ha querido usarlas o porque las ligaduras son no holónomas y no es posible hacerlo). Las relaciones entre las ligaduras estarán dadas por $s - n$ expresiones del tipo

$$\sum_{k=1}^s B_{kr} \delta q_k = 0, \quad 1 \leq r \leq s - n, \quad (2.20)$$

donde las B_{kr} son de la forma $\partial G_r / \partial q_k$ sólo si la ligadura r es holónoma. Usando el método de los multiplicadores de Lagrange multiplicamos estas expresiones por funciones desconocidas λ_r y sumamos los $s - n$ expresiones a la (2.16) (en la que se ha reemplazado n por s) para obtener

$$\sum_{k=1}^s \left[\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} - Q_k + \sum_{r=1}^{s-n} \lambda_r B_{kr} \right] \delta q_k = 0.$$

Procediendo como se hizo con (2.6) se obtiene finalmente

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_k} \right) - \frac{\partial T}{\partial q_k} = Q_k - \sum_{r=1}^{s-n} \lambda_r B_{kr}, \quad (2.21)$$

que deben completarse con las ecuaciones de vínculo no usadas

$$G_r(q_1, q_2, \dots, q_s, t) = 0,$$

en caso de vínculos holónomos. En el caso de vínculos no holónomos se tiene ecuaciones análogas a las (2.9) vistas arriba

$$\sum_{k=1}^s B_{kr} \dot{q}_k = b_r.$$

Si las fuerzas son derivables de un potencial (ver (2.18)), entonces podemos reescribir (2.21) como

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = - \sum_{r=1}^{s-n} \lambda_r B_{kr}. \quad (2.22)$$

Nota histórica: Fue Lagrange (1736-1813) quien en su obra *Mécanique Analytique* de 1788 desarrolló la mecánica analítica, que denominó así por el uso que hace en ella del análisis matemático, junto con el método de multiplicadores y técnicas variacionales (a ver en el punto siguiente).

2.3.1. Partículas en campos electromagnéticos

El caso de una partícula cargada en un campo electromagnético es un ejemplo importante de existencia del lagrangiano para fuerzas dependientes de la velocidad. La fuerza electromagnética sobre la partícula de carga e y velocidad \mathbf{u} es (en unidades del SI)

$$\mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B}),$$

donde \mathbf{E} es el campo eléctrico y \mathbf{B} el magnético en la posición de la partícula. En términos de los potenciales escalar y vector es

$$\mathbf{F} = e \left[-\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} + \mathbf{u} \times (\nabla \times \mathbf{A}) \right].$$

Si desarrollamos explícitamente, se tiene para la componente cartesiana x

$$F_x = e \left[-\frac{\partial\phi}{\partial x} - \frac{\partial A_x}{\partial t} + u_y \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - u_z \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \right].$$

Si a esta expresión le sumamos y restamos $u_x \partial A_x / \partial x$, podemos escribirla como

$$F_x = e \left[-\frac{\partial\phi}{\partial x} - \frac{dA_x}{dt} + \frac{\partial}{\partial x} (\mathbf{u} \cdot \mathbf{A}) \right],$$

donde

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + u_x \frac{\partial A_x}{\partial x} + u_y \frac{\partial A_x}{\partial y} + u_z \frac{\partial A_x}{\partial z}.$$

Tenemos así que podemos escribir en general

$$\mathbf{F} = -e \left[\nabla (\phi - \mathbf{u} \cdot \mathbf{A}) + \frac{d\mathbf{A}}{dt} \right].$$

Esta expresión nos lleva a proponer el “potencial”

$$V = e(\phi - \mathbf{u} \cdot \mathbf{A}),$$

de donde tenemos inmediatamente (con expresiones análogas para las otras componentes)

$$\frac{d}{dt} \left(\frac{\partial V}{\partial u_x} \right) - \frac{\partial V}{\partial x} = F_x.$$

Para un conjunto de N partículas tenemos

$$V = \sum_{i=1}^N e_i (\phi_i - \dot{\mathbf{x}}_i \cdot \mathbf{A}_i), \quad (2.23)$$

donde

$$\phi_i \equiv \phi(\mathbf{x}_i, t), \quad \mathbf{A}_i \equiv \mathbf{A}(\mathbf{x}_i, t).$$

Así,

$$\mathbf{F}_i = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{\mathbf{x}}_i} \right) - \frac{\partial V}{\partial \mathbf{x}_i},$$

y

$$\begin{aligned} Q_k &= \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} = \sum_{i=1}^N \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{\mathbf{x}}_i} \right) \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} - \sum_{i=1}^N \frac{\partial V}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \\ &= \frac{d}{dt} \sum_{i=1}^N \left(\frac{\partial V}{\partial \dot{\mathbf{x}}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right) - \sum_{i=1}^N \frac{\partial V}{\partial \dot{\mathbf{x}}_i} \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{x}_i}{\partial q_k} \right) - \sum_{i=1}^N \frac{\partial V}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \\ &= \frac{d}{dt} \sum_{i=1}^N \left(\frac{\partial V}{\partial \dot{\mathbf{x}}_i} \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial \dot{q}_k} \right) - \sum_{i=1}^N \frac{\partial V}{\partial \dot{\mathbf{x}}_i} \cdot \frac{\partial \dot{\mathbf{x}}_i}{\partial q_k} - \sum_{i=1}^N \frac{\partial V}{\partial \mathbf{x}_i} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \\ &= \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{q}_k} \right) - \frac{\partial V}{\partial q_k}, \end{aligned}$$

y es entonces posible definir $L = T - V$, con el potencial (2.23).

2.4. Principio de Hamilton

Un punto extremadamente importante es que la forma de las ecuaciones (2.19) es la misma que las denominadas ecuaciones de Euler-Lagrange del cálculo de variaciones.

Un ejemplo característico del cálculo de variaciones es encontrar la función $y = f(x)$ que hace que la integral

$$I = \int_{x_1}^{x_2} F(y, y', x) dx, \quad (2.24)$$

sea extrema (máxima o mínima). En (2.24) F es una función conocida de sus tres argumentos, la prima indica derivación respecto de x , y se considera que en los límites de integración la función y toma valores dados y_1, y_2 . Si para una dada $f(x)$ la integral es un extremo, cuando $f(x)$ se reemplaza por una

función ligeramente modificada $f(x) + \varepsilon\eta(x)$ (ε es una constante infinitesimal y $\eta(x)$ una función arbitraria que vale cero en x_1, x_2 para cumplir que el valor de y en los extremos está fijado) la integral no debe variar a primer orden en ε . Así

$$\delta I = \int_{x_1}^{x_2} F(f + \varepsilon\eta, f' + \varepsilon\eta', x) dx - \int_{x_1}^{x_2} F(f, f', x) dx = 0. \quad (2.25)$$

Si escribimos desarrollando a primer orden en serie de Taylor para ε

$$F(f + \varepsilon\eta, f' + \varepsilon\eta', x) = F(f, f', x) + \frac{\partial F}{\partial y}\varepsilon\eta + \frac{\partial F}{\partial y'}\varepsilon\eta',$$

(donde se entiende que las derivadas de F se evalúan en $y = f(x), y' = f'(x)$) podemos escribir (2.25) como

$$\delta I = \varepsilon \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' \right) dx = 0. \quad (2.26)$$

Dado que

$$\frac{\partial F}{\partial y'}\eta' = \frac{d}{dx} \left(\frac{\partial F}{\partial y'}\eta \right) - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \eta,$$

al integrar por partes el segundo término del integrando de (2.26) y usar que $\eta(x_1) = \eta(x_2) = 0$ podemos escribir

$$\delta I = \varepsilon \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right] \eta dx = 0,$$

pero como $\eta(x)$ es una función arbitraria (podría en particular elegirse igual a la función que multiplica, con lo cual se tiene la integral de una magnitud definida positiva en todo punto) debe anularse el corchete y se tiene entonces la ecuación de Euler

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0.$$

Si F es función de $2n$ variables $y_k, y'_k: F(y_1, y'_1, y_2, y'_2, \dots, y_n, y'_n, x)$ y deben hallarse las n funciones $y_k = f_k(x)$ que extreman la integral

$$I = \int_{x_1}^{x_2} F(y_1, y'_1, y_2, y'_2, \dots, y_n, y'_n, x) dx,$$

con valores de las y_k fijos en los límites de integración, se tiene que al variar las n funciones: $f_k(x) \rightarrow f_k(x) + \varepsilon_k \eta_k(x)$, la integral no varía, por lo que un

cálculo análogo al anterior (desarrollando ahora en Taylor en cada uno de los ε_k) lleva a

$$\delta I = \int_{x_1}^{x_2} \sum_{k=1}^n \varepsilon_k \left[\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'_k} \right) \right] \eta_k dx = 0. \quad (2.27)$$

Si no hay ligaduras entre las funciones y_k , los η_k son independientes (y arbitrarios) por lo que puede elegirse a todos salvo uno igual a cero y, así como antes, ver que debe ser para cada y_k

$$\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'_k} \right) = 0. \quad (2.28)$$

Si hubiera m ligaduras de la forma

$$G_r(y_1, y_2, \dots, y_n, x) = 0, \quad 1 \leq r \leq m, \quad (2.29)$$

como tanto las f_k como sus variaciones deben respetar las ligaduras, tenemos

$$G_r(f_1, \dots, f_n, x) = G_r(f_1 + \varepsilon_1 \eta_1, \dots, f_n + \varepsilon_n \eta_n, x) = 0, \quad 1 \leq r \leq m,$$

Al desarrollar la segunda serie de ecuaciones en serie de Taylor a primer orden en los ε_i (y usar también la primera serie de condiciones) se tiene

$$\sum_{k=1}^n \varepsilon_k \frac{\partial G_r}{\partial y_k} \eta_k = 0, \quad 1 \leq r \leq m.$$

Usando el método de los multiplicadores de Lagrange multiplicamos estas últimas ecuaciones por funciones a determinar $\lambda_r(y_1, y_2, \dots, y_n, x)$, y sumamos las m ecuaciones a la (2.27) para obtener

$$\int_{x_1}^{x_2} \sum_{k=1}^n \varepsilon_k \left[\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'_k} \right) + \sum_{r=1}^m \lambda_r \frac{\partial G_r}{\partial y_k} \right] \eta_k dx = 0.$$

Como siempre, elegimos los m multiplicadores λ_r para que se anulen m términos de la sumatoria en k , por lo que sobreviven $n - m$ términos que podemos tomar entonces como independientes, con lo que el resultado final es que, para todas las y_k , se tiene

$$\frac{\partial F}{\partial y_k} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'_k} \right) + \sum_{r=1}^m \lambda_r \frac{\partial G_r}{\partial y_k} = 0, \quad (2.30)$$

que deben completarse con las (2.29).

Es claro entonces que las ecuaciones (2.19) y (2.28) (o las (2.22) y (2.30)); si las ligaduras fueran no holónomas la deducción de (2.30) es equivalente a la dada, sólo que en lugar de $\partial G_r / \partial y_k$ se tendrían funciones B_{kr} no reducibles a derivadas) son totalmente equivalentes si se identifican

$$\begin{aligned} L &\leftrightarrow F \\ q_k &\leftrightarrow y_k \\ t &\leftrightarrow x \end{aligned}$$

En otras palabras, las trayectorias del sistema en el espacio de configuración (las $q_k(t)$) son las que extreman la integral

$$S = \int_{t_1}^{t_2} L(q_1, \dot{q}_1, \dots, q_n, \dot{q}_n, t) dt, \quad (2.31)$$

sujeta a la condición que las q_k tomen valores dados en los instantes t_1 y t_2 . Éste es el *principio variacional de Hamilton*, denominado también *principio de mínima acción*, y la integral S es denominada la *acción* del sistema.

Notemos que S es una función de t_1 , t_2 , y de los valores de las coordenadas en estos instantes; por otro lado S depende de la forma de las trayectorias $q_k(t)$ en $t_1 \leq t \leq t_2$, esto es, distintas *funciones* $q_k(t)$ dan distintos valores de S . Se dice entonces que S es una *funcional* de las $q_k(t)$.

2.4.1. Principio de Maupertuis

Para sistemas mecánicos en los que se conserva la energía mecánica $E = T + V$ es posible escribir un principio variacional que determina directamente la forma geométrica de las trayectorias del sistema, sin considerar cómo son recorridas al transcurrir el tiempo (el tiempo desaparece de la formulación). Para ver esto usemos que $L = T - V$, junto con $E = T + V$, para escribir la acción (2.31) como

$$S = \int_{t_1}^{t_2} (2T - E) dt = \int_{t_1}^{t_2} 2T dt - E(t_2 - t_1),$$

y consideramos que T tiene la forma genérica (las a_{ij} dependen sólo de las q 's)

$$T = \frac{1}{2} \sum_{i,j} a_{ij} \dot{q}_i \dot{q}_j = \frac{1}{2dt^2} \sum_{i,j} a_{ij} dq_i dq_j,$$

de donde podemos despejar el diferencial de tiempo

$$dt = \sqrt{\frac{\sum_{i,j} a_{ij} dq_i dq_j}{2T}},$$

y reescribir la acción como

$$S = \int_{t_1}^{t_2} \sqrt{2T \sum_{i,j} a_{ij} dq_i dq_j} - E(t_2 - t_1).$$

Como ya impusimos la constancia de E , el término $E(t_2 - t_1)$ no varía al variar S , por lo que podemos descartarlo y considerar sólo la primera integral que denominamos *acción reducida* S_0

$$S_0 = \int_{t_1}^{t_2} \sqrt{2T \sum_{i,j} a_{ij} dq_i dq_j},$$

y que podemos escribir en términos de sólo las q 's al reemplazar T por $E - V$ y tomar a una de las q 's, digamos la q_1 , como coordenada independiente (que integramos entre valores fijos arbitrarios q_1 y q_1')

$$S_0 = \int_{q_1}^{q_1'} \sqrt{2(E - V) \sum_{i,j} a_{ij} \frac{dq_i}{dq_1} \frac{dq_j}{dq_1}} dq_1.$$

Al tomar entonces variaciones arbitrarias de las q_i (con $i \neq 1$ y con valores fijos en los límites de integración) e imponer $\delta S_0 = 0$ obtenemos las ecuaciones de la forma geométrica de las trayectorias. Éste es el principio de Maupertuis.

Nota histórica: Los principios variacionales se usaron desde tiempos remotos en forma más o menos elemental, como por ejemplo en el problema de hallar el área máxima encerrada por una curva de perímetro dado (problema de Dido) conocido por los geómetras de la Grecia antigua. Su desarrollo moderno comenzó con el planteo de Juan Bernoulli en 1696 del problema de cuál es la curva de descenso más rápido (braquistocrona) entre dos puntos dados en un plano vertical, por la que desciende una partícula sin rozamiento con el peso como única fuerza aplicada; problema que fue resuelto por Newton y Bernoulli en forma casi simultánea. Maupertuis en 1746 postuló su principio de mínima acción en forma elemental (que formularon luego correctamente Euler y Lagrange), al que consideraba como una expresión de la “economía” de la Naturaleza. Lagrange en su *Mécanique Analytique* de 1788 desarrolló la metodología más general de resolver problemas variacionales y utilizó el principio de mínima acción en la forma vista aquí. Previamente Euler (1707-1783) dedujo las ecuaciones que llevan su nombre y también consideró un principio de mínima acción basado en lo que llamaríamos hoy día energía potencial. Finalmente Hamilton en 1834 desarrolló su dinámica (ver más adelante) considerando variaciones más generales de la acción, en las que varían también los valores en los extremos de integración.

2.5. Invarianza de las ecuaciones de Lagrange y simetrías

Un punto importante de las ecuaciones de Lagrange es que la forma de las ecuaciones es la misma en cualquier sistema de referencia. En efecto, si hacemos una transformación de las coordenadas generalizadas, incluso dependiente del tiempo, y escribimos las coordenadas q_k en términos de otras coordenadas Q_k (no confundir con las fuerzas generalizadas, que no aparecerán de ahora en más en estos apuntes)

$$q_k = f_k(Q_1, \dots, Q_n, t), \quad 1 \leq k \leq n, \quad (2.32)$$

al reemplazar las variables originales por las variables nuevas en el lagrangiano se obtiene éste en función de las nuevas variables; por supuesto, reemplazando también las velocidades generalizadas

$$\dot{q}_k = \frac{\partial f_k}{\partial t} + \sum_{j=1}^n \frac{\partial f_k}{\partial Q_j} \dot{Q}_j \equiv g_k(Q_1, \dot{Q}_1, \dots, Q_n, \dot{Q}_n, t),$$

para obtener $\tilde{L}(Q_1, \dot{Q}_1, \dots, Q_n, \dot{Q}_n, t)$. Transformaciones del tipo (2.32) son denominadas *de punto* o *de contacto*. Sin embargo, el *valor* de L y de \tilde{L} es el mismo en cada instante dado (sólo que expresado en coordenadas transformadas unas de otras), por lo que la acción (2.31) se expresa de igual manera en las nuevas coordenadas

$$S = \int_{t_1}^{t_2} \tilde{L}(Q_1, \dot{Q}_1, \dots, Q_n, \dot{Q}_n, t) dt,$$

y las ecuaciones de Lagrange correspondientes obtenidas al extremar la acción tienen igual *forma*

$$\frac{d}{dt} \left(\frac{\partial \tilde{L}}{\partial \dot{Q}_k} \right) - \frac{\partial \tilde{L}}{\partial Q_k} = 0. \quad (2.33)$$

Nótese que las nuevas coordenadas podrían ser, por ejemplo, coordenadas del sistema mecánico en un sistema de referencia no inercial y, sin embargo, no hay necesidad de incluir ninguna fuerza de inercia o algo equivalente (aquí se ve la ventaja de la formulación variacional; si hubiésemos usado el principio de D'Alembert tendríamos que haber calculado la aceleración de cada partícula y el trabajo virtual de las fuerzas de inercia para deducir (2.33) en un sistema no inercial).

Por otro lado, la covarianza de las ecuaciones de Lagrange nos permite usar las coordenadas mejor adaptadas al problema, siempre y cuando podamos expresar la función lagrangiana L en términos de ellas.

Otro punto importante es que si a L se le suma una derivada total del tiempo

$$L \rightarrow L + \frac{d}{dt}M(q_1, \dots, q_n, t), \quad (2.34)$$

la acción cambia a

$$S \rightarrow S + M|_{t=t_2} - M|_{t=t_1},$$

pero como en la variación de S se mantienen fijos los valores de las coordenadas en los extremos, las ecuaciones obtenidas son exactamente las mismas; ambos lagrangianos llevan a las mismas ecuaciones de movimiento. Así, por ejemplo, términos como $C_1 \dot{q}_1 = d(C_1 q_1)/dt$, o $C_1 q_1 \dot{q}_1 = d(C_1 q_1^2/2)/dt$, con C_1 constante, sumados a L no aportan a las ecuaciones de movimiento. Transformaciones del tipo dado por (2.34) son denominadas de *gauge*.

De enorme importancia son las transformaciones continuas a las variables q'_k , que son aquellas que pueden escribirse de la forma

$$q_k = f_k(q'_1, \dots, q'_n, s, t), \quad (2.35)$$

donde s es un parámetro que puede variarse continuamente desde, digamos, s_0 tal que cuando $s = s_0$ es $f_k = q'_k$, de manera que $q'_k = q_k$ en $s = s_0$. De esta manera podemos pensar que las coordenadas generalizadas son transformadas desde sus valores originales en forma continua al variar s .

De las (2.35) deducimos la transformación de las velocidades generalizadas (para un s dado)

$$\dot{q}_k = \dot{f}_k = \frac{\partial f_k}{\partial t} + \sum_{j=1}^n \frac{\partial f_k}{\partial q'_j} \dot{q}'_j,$$

por lo que al reemplazar en el lagrangiano $L(q_1, \dot{q}_1, \dots, q_n, \dot{q}_n, t)$ obtenemos un $\tilde{L}(q'_1, \dot{q}'_1, \dots, q'_n, \dot{q}'_n, t)$, que en $s = s_0$ coincide con el lagrangiano original. Consideremos ahora una variación infinitesimal δs en el entorno de $s = s_0$, que (a valores fijos de las q'_k) induce una variación infinitesimal en el lagrangiano que podemos escribir como

$$\delta L = \sum_{k=1}^n \left(\frac{\partial L}{\partial q_k} \frac{\partial f_k}{\partial s} \Big|_{s=s_0} + \frac{\partial L}{\partial \dot{q}_k} \frac{\partial \dot{f}_k}{\partial s} \Big|_{s=s_0} \right) \delta s, \quad (2.36)$$

donde notamos que en $s = s_0$ es $\tilde{L} = L$, $q'_k = q_k$ y $\dot{q}'_k = \dot{q}_k$. Nótese que el lagrangiano \tilde{L} varía en general al variar s porque las q_k varían como lo especifica (2.35) aun manteniendo fijas las q'_k .

Si usamos que

$$\left. \frac{\partial L}{\partial \dot{q}_k} \frac{\partial f_k}{\partial s} \right|_{s=s_0} = \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \dot{q}_k} \frac{\partial f_k}{\partial s} \right|_{s=s_0} \right) - \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \dot{q}_k} \right) \frac{\partial f_k}{\partial s} \right|_{s=s_0}, \quad (2.37)$$

y tenemos en cuenta las ecuaciones de Lagrange, podemos escribir a δL como

$$\delta L = \delta s \frac{d}{dt} \left(\sum_{k=1}^n \left. \frac{\partial L}{\partial \dot{q}_k} \frac{\partial f_k}{\partial s} \right|_{s=s_0} \right).$$

De esta expresión obtenemos un resultado de mucha importancia: si para alguna transformación continua de la forma (2.35) el lagrangiano es *invariante* ($\delta L = 0$), entonces la magnitud

$$C = \sum_{k=1}^n \left. \frac{\partial L}{\partial \dot{q}_k} \frac{\partial f_k}{\partial s} \right|_{s=s_0}, \quad (2.38)$$

es una constante de movimiento. Este resultado es una versión simplificada del *teorema de Noether* (Emmy Noether (1882-1935), publicado en 1918); al final de este punto veremos el teorema general. Cuando el lagrangiano es invariante ante una transformación continua se dice que posee una *simetría continua*.

Supongamos, por ejemplo, que el lagrangiano no varía si alguna coordenada q_r se incrementa en una cantidad arbitraria $q_r \rightarrow q_r + s = q'_r$, por lo que podemos escribir la transformación como $f_r = q'_r - s$ (aquí es $s_0 = 0$). La cantidad conservada será entonces

$$\frac{\partial L}{\partial \dot{q}_r} = cte.$$

Por supuesto, que L no varíe ante cambios arbitrarios de q_r significa que es independiente de esta coordenada, $\partial L / \partial q_r = 0$, por lo que las ecuaciones de Lagrange nos dan inmediatamente la misma información (el teorema de Noether es, por supuesto, mucho más general).

La derivada de L respecto de una velocidad generalizada $\partial L / \partial \dot{q}_r$ se denomina *impulso generalizado* p_r , y si el lagrangiano es independiente de una coordenada generalizada se dice que ésta es *cíclica*. Así, el impulso generalizado correspondiente a una coordenada cíclica es constante durante la evolución del sistema.

Finalmente calculemos la derivada temporal *total* de L :

$$\frac{dL}{dt} \equiv \frac{\partial L}{\partial t} + \sum_{k=1}^n \left(\frac{\partial L}{\partial q_k} \dot{q}_k + \frac{\partial L}{\partial \dot{q}_k} \frac{d\dot{q}_k}{dt} \right),$$

que, usando

$$\frac{\partial L}{\partial \dot{q}_k} \frac{d\dot{q}_k}{dt} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \dot{q}_k \right) - \dot{q}_k \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right),$$

reescribimos como

$$\frac{dL}{dt} \equiv \frac{\partial L}{\partial t} + \left[\sum_{k=1}^n \frac{\partial L}{\partial q_k} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \dot{q}_k + \frac{d}{dt} \left(\sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k \right);$$

y, usando las ecuaciones de Lagrange, (y escribiendo p_k por $\partial L/\partial \dot{q}_k$) tenemos

$$\frac{dL}{dt} \equiv \frac{\partial L}{\partial t} + \frac{d}{dt} \left(\sum_{k=1}^n p_k \dot{q}_k \right),$$

por lo que que si L no depende *explícitamente* del tiempo, $\partial L/\partial t = 0$, resulta

$$\frac{d}{dt} \left(\sum_{k=1}^n p_k \dot{q}_k - L \right) = 0.$$

Así, la magnitud entre paréntesis es constante durante la evolución del sistema, y se identifica con la energía de éste,

$$E = \sum_{k=1}^n p_k \dot{q}_k - L. \quad (2.39)$$

Notemos que si para tal sistema es $V(q_1, \dots, q_n)$ y

$$T = \frac{1}{2} \sum_{i,j} f_{ij}(q_1, \dots, q_n) \dot{q}_i \dot{q}_j,$$

entonces es inmediato ver que (2.39) corresponde a $E = T + V$, la energía mecánica; pero si, por ejemplo, T no fuese bilineal en las velocidades generalizadas, la energía conservada no corresponde a la suma de la energías cinética y potencial.

2.5.1. Teorema de Noether

Veamos ahora la forma más general del teorema de Noether. El punto es que la invarianza de L es a veces una condición muy restrictiva. De hecho, un sistema mecánico posee tal o cual simetría si las ecuaciones de movimiento la tienen, para lo que basta que la acción sea invariante, aunque el lagrangiano no lo sea (la inversa no es cierta por la posibilidad de las transformaciones de

gauge vistas arriba (ver final de este punto)). Además, para generalizar aún más el tipo de transformación, permitamos que se efectúe una transformación continua a la variable de integración de la acción, esto es, al tiempo t , de la forma

$$t = f_0(q'_1, \dots, q'_n, s, t'),$$

tal que para $s = s_0$ es $f_0 = t'$, y escribimos la transformación de las coordenadas en forma análoga a las (2.35)

$$q_k = f_k(q'_1, \dots, q'_n, s, t').$$

Como hicimos antes veamos cómo se altera la acción por un cambio infinitesimal δs en el entorno de $s = s_0$, de manera que

$$t = t' + \theta_0 \delta s, \quad q_k = q'_k + \theta_k \delta s,$$

donde se han definido

$$\theta_0 \equiv \left. \frac{\partial f_0}{\partial s} \right|_{s=s_0}, \quad \theta_k \equiv \left. \frac{\partial f_k}{\partial s} \right|_{s=s_0},$$

en las que, al estar trabajando a primer orden en δs , sus argumentos se han tomado como los q_k y t (en lugar de los q'_k y t'), por lo que podemos escribir en forma totalmente explícita la transformación inversa (ésta es una ventaja de las transformaciones infinitesimales)

$$t' = t - \theta_0(q_1, \dots, q_n, t) \delta s, \quad q'_k = q_k - \theta_k(q_1, \dots, q_n, t) \delta s, \quad (2.40)$$

Debe tenerse cuidado con las velocidades transformadas, ya que éstas son derivadas de las q'_k respecto de t' (y no de t), por lo que podemos calcular, usando las (2.40) (evaluando todo a primer orden en δs),

$$\begin{aligned} \dot{q}'_k &= \frac{dq'_k}{dt'} = \frac{dq'_k}{dt} \frac{1}{dt'/dt} = \frac{\dot{q}_k - \dot{\theta}_k \delta s}{1 - \dot{\theta}_0 \delta s} \\ &\simeq (\dot{q}_k - \dot{\theta}_k \delta s) (1 + \dot{\theta}_0 \delta s) \simeq \dot{q}_k + (\dot{q}_k \dot{\theta}_0 - \dot{\theta}_k) \delta s. \end{aligned} \quad (2.41)$$

Por otro lado, la acción transformada se escribe en general

$$S' = \int_{t'_1}^{t'_2} \tilde{L}(q'_1, \dot{q}'_1, \dots, q'_n, \dot{q}'_n, t') dt',$$

en la que los límites de integración t'_1 y t'_2 son los que corresponden a t_1 y t_2 en las variables originales. De esta manera, en el cambio infinitesimal de coordenadas se induce un cambio infinitesimal de la acción dado por

$$\delta S = S' - S = \int_{t'_1}^{t'_2} \tilde{L}(q'_1, \dot{q}'_1, \dots, q'_n, \dot{q}'_n, t') dt' - \int_{t_1}^{t_2} L(q_1, \dot{q}_1, \dots, q_n, \dot{q}_n, t) dt. \quad (2.42)$$

Usando las (2.40) y (2.41) podemos escribir, desarrollando \tilde{L} a orden uno en δs en el entorno de $s = s_0$,

$$\tilde{L} = L + \frac{\partial L}{\partial t} (-\theta_0 \delta s) + \sum_{k=1}^n \left[\frac{\partial L}{\partial q_k} (-\theta_k \delta s) + \frac{\partial L}{\partial \dot{q}_k} (\dot{q}_k \dot{\theta}_0 - \dot{\theta}_k) \delta s \right] \equiv L + \delta L, \quad (2.43)$$

y, además,

$$dt' = (1 - \dot{\theta}_0 \delta s) dt,$$

con lo que (2.42) se reescribe a primer orden en δs (los límites de integración han vuelto a ser t_1 y t_2 porque la integración es en la variable t)

$$\delta S = \int_{t_1}^{t_2} (L + \delta L) (1 - \dot{\theta}_0 \delta s) dt - \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} (\delta L - L \dot{\theta}_0 \delta s) dt, \quad (2.44)$$

con el δL definido en (2.43). En este punto integramos por partes el segundo término del integrando

$$\delta s \int_{t_1}^{t_2} L \dot{\theta}_0 dt = \delta s L \theta_0 \Big|_{t_1}^{t_2} - \delta s \int_{t_1}^{t_2} \theta_0 \frac{dL}{dt} dt,$$

donde, por supuesto, es

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \sum_{k=1}^n \left[\frac{\partial L}{\partial q_k} \dot{q}_k + \frac{\partial L}{\partial \dot{q}_k} \frac{d}{dt} \dot{q}_k \right].$$

Con esto, los pasos a seguir son directos y se dejan como ejercicio; se requiere usar la expresión de δL dada en (2.43), integrar por partes algunos términos y usar las ecuaciones de Lagrange, para poder finalmente escribir a δS en términos de sólo expresiones evaluadas en los extremos de integración

$$\delta S = \delta s \left[\left(\sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L \right) \theta_0 - \sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \theta_k \right] \Big|_{t_1}^{t_2}. \quad (2.45)$$

Si tenemos entonces la *información adicional* que la acción es invariante ante la transformación, $\delta S = 0$, el término entre corchetes debe valer lo mismo en t_1 y en t_2 , o sea, debe ser constante ya que t_1 y t_2 son arbitrarios. Así obtenemos una constante de movimiento dada por

$$C = \left(\sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L \right) \theta_0 - \sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \theta_k. \quad (2.46)$$

Nótese que si no hacemos ningún cambio en t , con lo que la invarianza de S equivale a la de L , es $\theta_0 = 0$ y se reobtiene la (2.38). Si no hacemos cambios de las coordenadas (o sea $\theta_k = 0$) y la acción no varía ante una translación temporal $t' = t + s$ (lo que requiere que L no dependa de t), es $\theta_0 = -1$ y se recupera la constancia de la energía (2.39). Sin embargo, más generalmente la acción puede ser invariante ante transformaciones combinadas de las coordenadas y del tiempo, en cuyo caso la utilidad de (2.46) es clara.

Una versión más completa del teorema de Noether involucra directamente la invarianza de las ecuaciones de movimiento mismas, lo que implica que la acción S' difiere de la S por sólo una transformación de gauge. El cambio infinitesimal de la acción será entonces, $\delta S = \varepsilon g(q_1, \dots, q_n, t)|_{t_1}^{t_2}$, con g una función explícita de las coordenadas y el tiempo por lo que, usando (2.45), es (hacemos ahora $\varepsilon = \delta s$, ya que cualquier diferencia puede absorberse en la definición de la función $g(q_1, \dots, q_n, t)$)

$$\left(\sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L \right) \theta_0 - \sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \theta_k - g = cte. \quad (2.47)$$

2.6. Acción como función de las q 's

Mencionemos finalmente una interpretación adicional de la acción que será útil más adelante. Sabemos que la acción definida por (2.31) fue definida entre tiempos fijos, con lo que la única dependencia libre es entonces respecto de las funciones $q_i(t)$, y decimos por esto que es una funcional de las q_i . Podemos alternativamente definirla como

$$S = \int_{t_0}^t L(q_1, \dot{q}_1, \dots, q_n, \dot{q}_n, t') dt',$$

donde t_0 es un tiempo fijo en el que las q_i tienen también valores fijos. Además imponemos que las $q_i(t')$ son las soluciones de las ecuaciones de movimiento, con “valor inicial” $q_i(t_0)$ y “valor final” $q_i(t)$. De esta manera S será una función de t y de las $q_i(t)$ (éstas no están fijadas porque en t_0 sólo se dan los valores de las q_i , pero no los de sus derivadas; la trayectoria real está entonces definida por los posibles valores finales $q_i(t)$). De esta manera, si mantenemos fijo t y variamos los valores finales $q_i(t)$ (con lo que variaremos

las trayectorias) estaremos haciendo variaciones de S de la forma

$$\begin{aligned}
 \delta S &= \int_{t_0}^t \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt' \\
 &= \int_{t_0}^t \sum_i \left[\frac{\partial L}{\partial q_i} \delta q_i + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \right] dt' \\
 &= \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i = \sum_i p_i \delta q_i,
 \end{aligned}$$

donde para pasar al último renglón se usó que las q_i satisfacen las ecuaciones de movimiento, y que las variaciones de las q_i son nulas en t_0 , donde están fijados sus valores. Deducimos así que

$$\frac{\partial S}{\partial q_i} = p_i. \quad (2.48)$$

Por otro lado, si derivamos S respecto de t

$$\frac{dS}{dt} = L = \frac{\partial S}{\partial t} + \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i,$$

de donde, usando (2.48) y (2.39),

$$\frac{\partial S}{\partial t} = L - \sum_i p_i \dot{q}_i = -E. \quad (2.49)$$

Capítulo 3

Ecuaciones canónicas de Hamilton

3.0.1. Transformación de Legendre y ecuaciones de Hamilton

Las ecuaciones de Lagrange pueden ser transformadas en un sistema de ecuaciones de primer orden en las derivadas temporales. Para esto analicemos una transformación muy importante descubierta por Legendre; comencemos con una función de n variables independientes

$$F = F(u_1, u_2, \dots, u_n),$$

y consideremos la transformación desde las variables u_k a nuevas variables v_k definidas por las derivadas de F

$$v_k = \frac{\partial F}{\partial u_k}. \quad (3.1)$$

Sabemos del análisis matemático que para que las n variables v_k sean independientes entre sí el *Hessiano* de la transformación debe ser distinto de cero; el Hessiano es el determinante de la matriz de $n \times n$ formada por las derivadas $\partial v_i / \partial u_j$; o sea

$$\det \left(\frac{\partial^2 F}{\partial u_i \partial u_j} \right) \neq 0. \quad (3.2)$$

Si se cumple (3.2) las (3.1) pueden ser resueltas para hallar las u 's en función de las v 's, con lo cual podemos definir una nueva función G de las n variables v_k como

$$G \equiv \sum_{k=1}^n u_k v_k - F. \quad (3.3)$$

Si calculamos el diferencial de G tenemos

$$dG = \sum_{k=1}^n \left(u_k dv_k + v_k du_k - \frac{\partial F}{\partial u_k} du_k \right) = \sum_{k=1}^n u_k dv_k,$$

donde se usó la propia definición (3.1) para cancelar los últimos dos términos de la sumatoria. Vemos entonces que

$$u_k = \frac{\partial G}{\partial v_k}, \quad (3.4)$$

que al comparar con (3.1) nos indica la notable simetría (o dualidad) de la transformación; dadas las variables u_k y la F podemos pasar a las v_k , y de éstas a las u_k se puede volver con una transformación equivalente a través de la G .

Si en la F tuviéramos otras m variables w_j que no entran en la transformación (variables pasivas)

$$F = F(u_1, u_2, \dots, u_n, w_1, w_2, \dots, w_m),$$

las v_k y la G se definen como antes, pero al calcular dG ,

$$\begin{aligned} dG &= \sum_{k=1}^n \left(u_k dv_k + v_k du_k - \frac{\partial F}{\partial u_k} du_k \right) - \sum_{j=1}^m \frac{\partial F}{\partial w_j} dw_j \\ &= \sum_{k=1}^n u_k dv_k - \sum_{j=1}^m \frac{\partial F}{\partial w_j} dw_j, \end{aligned}$$

de donde resulta, además de (3.4),

$$\frac{\partial G}{\partial w_j} = -\frac{\partial F}{\partial w_j}. \quad (3.5)$$

La idea es aplicar una transformación de Legendre tomando como variables u_k a las velocidades generalizadas \dot{q}_k , como variables pasivas a las q_k y al tiempo t , y como función F al lagrangiano L . Así, se definen las equivalentes de las v_k , de (3.1),

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

que son los impulsos p_k . Como función G , de (3.3), se define a la *función hamiltoniana* H (o hamiltoniano)

$$H = \sum_{k=1}^n \dot{q}_k p_k - L, \quad (3.6)$$

de la que tenemos entonces, de (3.4),

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad (3.7)$$

mientras que de las (3.5) es

$$\frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k}, \quad (3.8)$$

y

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (3.9)$$

Si usamos las ecuaciones de Lagrange podemos escribir las (3.8) como

$$\frac{\partial H}{\partial q_k} = -\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) = -\dot{p}_k,$$

que, junto con las (3.7), nos da un sistema de ecuaciones de primer orden para las $2n$ variables q_k y p_k , conocido como sistema de *ecuaciones canónicas* o de *Hamilton*, que reescribimos

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad (3.10)$$

en las que el hamiltoniano se calcula del lagrangiano a través de la (3.6), luego de reemplazar las \dot{q}_k en términos de las p_k (y de las q_k), lo que es siempre posible si se cumple la condición (3.2), que reescribimos en términos de las variables de interés,

$$\det \left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \right) \neq 0.$$

Vemos de (3.9) que si el lagrangiano no depende explícitamente de t tampoco lo hace H . Recordemos que en tal caso se conserva justamente la cantidad que hemos definido como H , que identificamos en tal caso con la energía del sistema E .

Para concluir notemos que la acción S puede ser escrita, de (3.6), como

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \left(\sum_{k=1}^n p_k \dot{q}_k - H \right) dt, \quad (3.11)$$

en la que si hacemos una variación arbitraria de las funciones $p_k(t)$ obtenemos

$$\delta S = \int_{t_1}^{t_2} \left(\sum_{k=1}^n \dot{q}_k \delta p_k - \frac{\partial H}{\partial p_k} \delta p_k \right) dt = 0,$$

en virtud de las ecuaciones de la transformación (3.7). Vemos entonces que una variación arbitraria de las p_k no tiene efecto sobre S . De esta manera, podemos determinar las ecuaciones de Hamilton (3.10) al extremar la acción S (escrita como en la segunda igualdad de (3.11)) con variaciones *arbitrarias e independientes* de las $q_k(t)$ y las $p_k(t)$. Si hacemos esto tenemos

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^n \left(p_k \delta \dot{q}_k + \dot{q}_k \delta p_k - \frac{\partial H}{\partial q_k} \delta q_k - \frac{\partial H}{\partial p_k} \delta p_k \right) dt,$$

que al integrar por partes, de la manera usual, el primer término entre paréntesis (y usar que las variaciones son cero en los extremos) da

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^n \left[\left(-\dot{p}_k - \frac{\partial H}{\partial q_k} \right) \delta q_k + \left(\dot{q}_k - \frac{\partial H}{\partial p_k} \right) \delta p_k \right] dt;$$

de donde las ecuaciones de Hamilton resultan inmediatamente al pedir $\delta S = 0$.

Nomenclatura: El espacio generado por las $2n$ variables q_k, p_k se denomina *espacio de fases* del sistema, y se dice, para cada valor de k , que q_k y p_k son *variables conjugadas* entre sí.

Nota histórica: Hamilton (1805-1865) publicó su trabajo *On a General Method in Dynamics* en 1834, en el que deduce las ecuaciones presentadas aquí siguiendo un método diferente, a través de lo que él denominó función principal (ver último punto de estos apuntes).

3.0.2. Formalismo simpléctico

Un formalismo muy conveniente es el denominado simpléctico, en el cual se trabaja en el espacio de vectores

$$\eta = \begin{pmatrix} q_1 \\ \dots \\ q_n \\ p_1 \\ \dots \\ p_n \end{pmatrix},$$

con lo que las ecuaciones de Hamilton pueden escribirse en forma matricial compacta como

$$\dot{\eta} = \mathbf{J} \frac{\partial H}{\partial \eta}, \quad (3.12)$$

donde la *matriz simpléctica* \mathbf{J} de $(2n \times 2n)$ está definida por

$$\mathbf{J} = \begin{pmatrix} \mathbf{0}_{(n \times n)} & \mathbf{I}_{(n \times n)} \\ -\mathbf{I}_{(n \times n)} & \mathbf{0}_{(n \times n)} \end{pmatrix},$$

donde los bloques de $(n \times n)$ son ceros o la matriz identidad \mathbf{I} , con el signo asignado.

Propiedades útiles inmediatas de la misma definición de \mathbf{J} son

$$\mathbf{J}^T = -\mathbf{J}, \quad (3.13)$$

y, dado que

$$\mathbf{J}^T \mathbf{J} = \begin{pmatrix} \mathbf{I}_{(n \times n)} & \mathbf{0}_{(n \times n)} \\ \mathbf{0}_{(n \times n)} & \mathbf{I}_{(n \times n)} \end{pmatrix} = \mathbf{I}_{(2n \times 2n)},$$

es

$$\mathbf{J}^T = \mathbf{J}^{-1}, \quad (3.14)$$

y

$$\mathbf{J}^2 = -\mathbf{I}_{(2n \times 2n)}. \quad (3.15)$$

3.1. Transformaciones canónicas y corchetes de Poisson

3.1.1. Transformaciones canónicas

Vimos que las transformaciones de punto no cambian la forma de las ecuaciones de Lagrange. Queremos ver ahora qué tipo de transformaciones hace lo propio con las ecuaciones de Hamilton. Si de las variables (q_k, p_k) , en las que el hamiltoniano es H , cambiamos a las variables (Q_k, P_k) , en las que el hamiltoniano pasa a ser \tilde{H} , la cuestión es cuáles transformaciones llevan a que en las nuevas variables las ecuaciones sean

$$\dot{Q}_k = \frac{\partial \tilde{H}}{\partial P_k}, \quad \dot{P}_k = -\frac{\partial \tilde{H}}{\partial Q_k}. \quad (3.16)$$

Una forma conveniente de proceder es a través del principio variacional de Hamilton en la forma (3.11). Es claro que se obtendrán ecuaciones de la forma (3.16) si al transformar a las nuevas variables se tiene

$$\sum_{k=1}^n p_k \dot{q}_k - H = \sum_{k=1}^n P_k \dot{Q}_k - \tilde{H} + \frac{dM}{dt}, \quad (3.17)$$

donde hemos tenido en cuenta que el agregado de una derivada total del tiempo dM/dt no modifica la variación de la acción.

Si multiplicamos (3.17) por dt y reordenamos tenemos

$$dM = \sum_{k=1}^n p_k dq_k - \sum_{k=1}^n P_k dQ_k + (\tilde{H} - H) dt, \quad (3.18)$$

que nos dice que

$$p_k = \frac{\partial M}{\partial q_k}, \quad P_k = -\frac{\partial M}{\partial Q_k}, \quad \tilde{H} = H + \frac{\partial M}{\partial t}. \quad (3.19)$$

De esta manera, para una función *arbitraria* $M(q_1, \dots, q_n, Q_1, \dots, Q_n, t)$ el primer grupo de ecuaciones en (3.19) nos permite despejar la expresión de las Q 's en función de las q 's y las p 's, que al ser reemplazadas en el segundo grupo nos da la expresión de las P 's. La función M es denominada *función generatriz*.

Podemos fácilmente determinar relaciones como las (3.19) para funciones generatrices que dependan de otro par de grupos de variables efectuando una transformación de Legendre (vista en el punto anterior) de la siguiente manera. A ambos lados de (3.18) sumamos (o restamos) un diferencial exacto; por ejemplo,

$$\begin{aligned} dM + d\left(\sum_{k=1}^n P_k Q_k\right) &= dM + \sum_{k=1}^n P_k dQ_k + \sum_{k=1}^n Q_k dP_k \\ &= \sum_{k=1}^n p_k dq_k + \sum_{k=1}^n Q_k dP_k + (\tilde{H} - H) dt, \end{aligned}$$

de donde, llamando

$$M' = M + \sum_{k=1}^n P_k Q_k,$$

tenemos

$$p_k = \frac{\partial M'}{\partial q_k}, \quad Q_k = \frac{\partial M'}{\partial P_k}, \quad \tilde{H} = H + \frac{\partial M'}{\partial t}, \quad (3.20)$$

que define la transformación en términos de una función generatriz nueva $M'(q_1, \dots, q_n, P_1, \dots, P_n)$.

Si a ambos lados de (3.18) restamos $d(\sum_{k=1}^n p_k q_k)$ obtenemos de manera similar

$$q_k = -\frac{\partial M''}{\partial p_k}, \quad P_k = -\frac{\partial M''}{\partial Q_k}, \quad \tilde{H} = H + \frac{\partial M''}{\partial t}, \quad (3.21)$$

con

$$M'' = M - \sum_{k=1}^n p_k q_k.$$

Finalmente si a ambos lados de (3.18) sumamos $d[\sum_{k=1}^n (P_k Q_k - p_k q_k)]$ tenemos

$$q_k = -\frac{\partial M'''}{\partial p_k}, \quad Q_k = \frac{\partial M'''}{\partial P_k}, \quad \tilde{H} = H + \frac{\partial M'''}{\partial t}, \quad (3.22)$$

con

$$M''' = M + \sum_{k=1}^n P_k Q_k - \sum_{k=1}^n p_k q_k.$$

De esta manera, a través de las funciones generatrices obtenemos las posibles transformaciones que preservan la forma de las ecuaciones de Hamilton. Estas transformaciones se denominan *transformaciones canónicas*. Esta metodología nos permite generar transformaciones canónicas al precio que debemos resolver ecuaciones algebraicas, en general complicadas, para obtener la forma explícita de la transformación.

Usando el formalismo simpléctico, para el caso que las variables η' resultan de una transformación canónica independiente del tiempo desde las variables η , $\eta'(\eta)$, podemos escribir que para una variable genérica η'_k , con $1 \leq k \leq 2n$, es

$$\dot{\eta}'_k = \sum_{j=1}^{2n} \frac{\partial \eta'_k}{\partial \eta_j} \dot{\eta}_j. \quad (3.23)$$

Así, definiendo la matriz \mathbf{M} de $(2n \times 2n)$ con componentes

$$M_{kj} = \frac{\partial \eta'_k}{\partial \eta_j},$$

podemos reescribir (3.23) en forma matricial como

$$\dot{\eta}' = \mathbf{M}\dot{\eta},$$

que usando (3.12) se escribe

$$\dot{\eta}' = \mathbf{M}\mathbf{J} \frac{\partial H}{\partial \eta}.$$

Como a su vez

$$\frac{\partial H}{\partial \eta_k} = \sum_{j=1}^{2n} \frac{\partial H}{\partial \eta'_j} \frac{\partial \eta'_j}{\partial \eta_k},$$

que matricialmente se escribe

$$\frac{\partial H}{\partial \eta} = \mathbf{M}^T \frac{\partial H}{\partial \eta'} = \mathbf{M}^T \frac{\partial \tilde{H}}{\partial \eta'},$$

donde la última igualdad resulta por ser la transformación canónica independiente de t , con lo que tenemos finalmente

$$\dot{\eta}' = \mathbf{M} \mathbf{J} \mathbf{M}^T \frac{\partial \tilde{H}}{\partial \eta'},$$

y la condición que la transformación sea canónica corresponde a

$$\mathbf{M} \mathbf{J} \mathbf{M}^T = \mathbf{J}. \quad (3.24)$$

De (3.24) tenemos, postmultiplicando por $(\mathbf{M}^T)^{-1}$,

$$\mathbf{M} \mathbf{J} = \mathbf{J} (\mathbf{M}^T)^{-1},$$

a la que, al premultiplicar por \mathbf{J} y postmultiplicar por \mathbf{J}^T , conduce inmediatamente a, usando (3.13), (3.14) y (3.15),

$$\mathbf{J} \mathbf{M} = (\mathbf{M}^T)^{-1} \mathbf{J},$$

que luego de premultiplicar por \mathbf{M}^T da finalmente una forma equivalente de la (3.24),

$$\mathbf{M}^T \mathbf{J} \mathbf{M} = \mathbf{J}.$$

Veamos finalmente que (3.24) es válida en general, aún cuando la transformación canónica depende del tiempo, $\eta'(\eta, t)$, y por la tanto

$$\dot{\eta}'_k = \frac{\partial \eta'_k}{\partial t} + \sum_{j=1}^{2n} \frac{\partial \eta'_k}{\partial \eta_j} \dot{\eta}_j,$$

por lo que, procediendo como antes es

$$\dot{\eta}' = \frac{\partial \eta'}{\partial t} + \mathbf{M} \mathbf{J} \mathbf{M}^T \frac{\partial H}{\partial \eta'}, \quad (3.25)$$

donde ahora $H \neq \tilde{H}$. Consideremos entonces que la transformación canónica se deriva de una función generatriz \mathcal{M} , de tal manera que

$$\begin{aligned} q'_k &= \frac{\partial \mathcal{M}}{\partial p'_k}, \\ p'_k &= -\frac{\partial \mathcal{M}}{\partial q'_k}. \end{aligned}$$

Notemos que en la primera línea \mathcal{M} corresponde a una dependencia tipo M' , y en la otra a una tipo M (ambas representan sin embargo la misma transformación, sólo que expresada en términos de funciones de diferentes variables primarias; además, en ambos casos las derivadas son a variables originales (q_1, \dots, q_n) constantes). Con el formalismo simpléctico podemos entonces escribir

$$\eta' = \mathbf{J} \frac{\partial \mathcal{M}}{\partial \eta'},$$

que, junto a

$$H = \tilde{H} - \frac{\partial \mathcal{M}}{\partial t},$$

nos permite escribir (3.25) como

$$\dot{\eta}' = \mathbf{J} \frac{\partial^2 \mathcal{M}}{\partial \eta' \partial t} + \mathbf{M} \mathbf{J} \mathbf{M}^T \frac{\partial \tilde{H}}{\partial \eta'} - \mathbf{M} \mathbf{J} \mathbf{M}^T \frac{\partial^2 \mathcal{M}}{\partial t \partial \eta'},$$

con lo que (3.24) nos dice inmediatamente que

$$\dot{\eta}' = \mathbf{J} \frac{\partial \tilde{H}}{\partial \eta'},$$

y vemos así que (3.24) es condición suficiente, también cuando la transformación canónica depende del tiempo.

3.1.2. Transformaciones canónicas infinitesimales

Consideremos ahora transformaciones canónicas infinitesimales genéricas

$$Q_i = q_i + \xi_i(q_1, \dots, q_n, p_1, \dots, p_n, t), \quad P_i = p_i + \eta_i(q_1, \dots, q_n, p_1, \dots, p_n, t), \quad (3.26)$$

donde ξ_i y η_i son funciones infinitesimales. En relación con esto la función generatriz de la forma (3.20) es especialmente importante pues, entre todas las posibles, contiene la transformación identidad $Q_i = q_i$, $P_i = p_i$, que es generada por

$$M' = \sum_{k=1}^n q_k P_k.$$

La transformación canónica infinitesimal (3.26) será entonces generada por una función de la forma

$$M' = \sum_{k=1}^n q_k P_k + W(q_1, \dots, q_n, P_1, \dots, P_n, t),$$

con W una función infinitesimal que, por (3.20), nos da

$$Q_i = q_i + \frac{\partial W}{\partial P_i}, \quad p_i = P_i + \frac{\partial W}{\partial q_i},$$

de donde, comparando con (3.26), podemos identificar

$$\xi_i = \frac{\partial W}{\partial P_i}, \quad \eta_i = -\frac{\partial W}{\partial q_i}.$$

Por otro lado, como P_i difiere infinitesimalmente de p_i y W es ya infinitesimal, puede reemplazarse en W a P_i por p_i y escribir entonces la transformación infinitesimal canónica general en forma *explícita* como

$$Q_i = q_i + \frac{\partial W}{\partial p_i}, \quad P_i = p_i - \frac{\partial W}{\partial q_i}. \quad (3.27)$$

Un punto importante es que si se toma $W = Hdt$, entonces

$$\begin{aligned} Q_i &= q_i + \frac{\partial H}{\partial p_i} dt = q_i + \dot{q}_i dt, \\ P_i &= p_i - \frac{\partial H}{\partial q_i} dt = p_i + \dot{p}_i dt, \end{aligned}$$

que nos dice que las variables Q_i y P_i son en este caso los valores de q_i y p_i evolucionados en un dt . De esta manera, podemos interpretar la evolución temporal como una transformación canónica generada por el hamiltoniano mismo.

La forma simpléctica de las transformaciones infinitesimales es

$$\eta' = \eta + \mathbf{J} \frac{\partial W}{\partial \eta}.$$

Que una transformación de este tipo es canónica se prueba al construir la matriz \mathbf{M} correspondiente, que resulta simplemente

$$\mathbf{M} = \mathbf{I}_{(2n \times 2n)} + \mathbf{J} \frac{\partial^2 W}{\partial \eta \partial \eta}, \quad (3.28)$$

de donde (usando las propiedades conocidas de \mathbf{J})

$$\mathbf{M}^T = \mathbf{I}_{(2n \times 2n)} + \frac{\partial^2 W}{\partial \eta \partial \eta} \mathbf{J}^T = \mathbf{I}_{(2n \times 2n)} - \frac{\partial^2 W}{\partial \eta \partial \eta} \mathbf{J},$$

con lo que

$$\mathbf{M} \mathbf{J} \mathbf{M}^T = \left(\mathbf{J} + \mathbf{J} \frac{\partial^2 W}{\partial \eta \partial \eta} \mathbf{J} \right) \left(\mathbf{I}_{(2n \times 2n)} - \frac{\partial^2 W}{\partial \eta \partial \eta} \mathbf{J} \right) = \mathbf{J},$$

al orden uno en W , que prueba la condición canónica.

Una consecuencia importante es que en esta prueba no se usó la condición que la transformación canónica infinitesimal sea independiente del tiempo, por lo que tenemos la prueba de que la condición canónica (3.24) es también necesaria en general (se puede pasar de una transformación infinitesimal a una finita integrando en el parámetro infinitesimal).

3.1.3. Teorema de Liouville

Consideremos un dado volumen en el espacio de fases

$$V = \int_U dq_1 \dots dq_n dp_1 \dots dp_n,$$

determinado por el conjunto de puntos contenidos en una región U del mismo. Si tomamos cada conjunto de puntos $(q_1, \dots, q_n, p_1, \dots, p_n)$ dentro de U como las condiciones iniciales del movimiento de un sistema mecánico de hamiltoniano H , podemos preguntarnos cómo evoluciona este volumen al transcurrir el tiempo. En un dt los puntos dentro de U evolucionan y llenan otra región U' de volumen

$$V' = \int_{U'} dq'_1 \dots dq'_n dp'_1 \dots dp'_n,$$

donde las coordenadas de cada punto han evolucionado de acuerdo a

$$q'_i = q_i + \dot{q}_i dt = q_i + \frac{\partial H}{\partial p_i} dt, \quad (3.29a)$$

$$p'_i = p_i + \dot{p}_i dt = p_i - \frac{\partial H}{\partial q_i} dt, \quad (3.29b)$$

que nos permite determinar el volumen V' a través del jacobiano de la transformación (3.29)

$$V' = \int_{U'} dq'_1 \dots dq'_n dp'_1 \dots dp'_n = \int_U \Gamma dq_1 \dots dq_n dp_1 \dots dp_n, \quad (3.30)$$

con el jacobiano dado por

$$\Gamma = \frac{\partial (q'_1 \dots q'_n p'_1 \dots p'_n)}{\partial (q_1 \dots q_n p_1 \dots p_n)} = \det \left(\frac{\partial X'_i}{\partial X_j} \right),$$

donde hemos denotado por X'_i y X_i al conjunto (ordenado) de las coordenadas $(q'_1, \dots, q'_n, p'_1, \dots, p'_n)$ y $(q_1, \dots, q_n, p_1, \dots, p_n)$ respectivamente, con $1 \leq i \leq 2n$. Con esta notación podemos reescribir las (3.29) como

$$X'_i = X_i + h_i dt,$$

donde h_i es el conjunto ordenado de valores

$$\left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n} \right),$$

con lo cual

$$\Gamma = \det \left(\delta_{ij} + \frac{\partial h_i}{\partial X_j} dt \right),$$

donde δ_{ij} es la delta de Kronecker, que vale uno si $i = j$, y cero si $i \neq j$. Si usamos ahora la relación general (ε es un factor infinitesimal y A_{ij} una matriz cualquiera)

$$\det(\delta_{ij} + \varepsilon A_{ij}) = 1 + \varepsilon \text{Traza}(A_{ij}) + O(\varepsilon^2),$$

podemos escribir, a orden uno en dt ,

$$\Gamma = 1 + dt \text{Traza} \left(\frac{\partial h_i}{\partial X_j} \right) = 1 + dt \sum_{i=1}^{2n} \frac{\partial h_i}{\partial X_i}.$$

Pero es claro de las definiciones de las h_i y X_i que $\sum \partial h_i / \partial X_i = 0$, con lo cual $\Gamma = 1$ y resulta entonces de (3.30) que $V' = V$. En otras palabras, el volumen de una región cualquiera del espacio de fases se conserva durante la evolución de un sistema mecánico regido por una función de Hamilton. Éste es el *teorema de Liouville*, de gran importancia en mecánica estadística y teoría de Caos.

Notemos que de igual manera podemos deducir que una transformación canónica infinitesimal arbitraria (3.27) preserva el volumen de una región cualquiera del espacio de fases (usando W en lugar de Hdt). Si pensamos además en una sucesión de transformaciones infinitesimales podemos ver que el volumen es preservado también por transformaciones canónicas finitas, que es una propiedad importante de éstas.

El teorema de Liouville tiene una prueba muy simple usando el formalismo simpléctico, dado que, de la definición misma de \mathbf{M} , es

$$\Gamma = |\det(\mathbf{M})|,$$

pero por (3.24) es

$$\det(\mathbf{M}) \det(\mathbf{J}) \det(\mathbf{M}^T) = [\det(\mathbf{M})]^2 \det(\mathbf{J}) = \det(\mathbf{J}),$$

con lo que

$$[\det(\mathbf{M})]^2 = 1,$$

y $\Gamma = 1$.

3.1.4. Corchetes de Poisson

Pasemos ahora a considerar los denominados *corchetes de Poisson*. Se define un corchete de Poisson de dos magnitudes $f(q_1, \dots, q_n, p_1, \dots, p_n, t)$ y $g(q_1, \dots, q_n, p_1, \dots, p_n, t)$ como

$$\{f, g\} \equiv \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) = -\{g, f\}. \quad (3.31)$$

Veamos algunas propiedades importantes de estos corchetes. Es inmediato de la definición que se tiene

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij}, \quad (3.32)$$

Por otro lado, también es fácil ver que

$$\{q_i, f\} = \frac{\partial f}{\partial p_i}, \quad \{p_i, f\} = -\frac{\partial f}{\partial q_i}. \quad (3.33)$$

Otras relaciones muy útiles que también se deducen de la sola definición (3.31) son

$$\begin{aligned} \{f_1 + f_2, g\} &= \{f_1, g\} + \{f_2, g\}, \\ \{f_1 f_2, g\} &= f_1 \{f_2, g\} + f_2 \{f_1, g\}, \\ \frac{\partial}{\partial t} \{f, g\} &= \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\}, \end{aligned}$$

cuya deducción es inmediata, y la importante *identidad de Jacobi* entre tres funciones cualesquiera

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad (3.34)$$

cuya comprobación puede hacerse por el siguiente argumento. Es fácil convencerse de que si se desarrolla (3.34) usando la definición de los corchetes se tiene una suma de términos, *todos* los cuales contienen alguna derivada segunda de alguna de las funciones f, g, h . Sin embargo, si se consideran los términos que contienen derivadas segundas de sólo una dada función, digamos f , se comprueba fácilmente que la suma de todos éstos se anula. Como lo mismo pasa obviamente con cualquiera de las otras dos funciones, la suma de *todos* los términos debe anularse, con lo que se comprueba la identidad. Veamos entonces los términos que contienen derivadas segundas de f ; éstos

proviene claramente de sólo los dos últimos términos de (3.34), los cuales podemos desarrollar usando la propiedad de distribución de la suma como

$$\begin{aligned}\{g, \{h, f\}\} &= \sum_{k=1}^n \left\{ g, \frac{\partial h}{\partial q_k} \frac{\partial f}{\partial p_k} \right\} - \sum_{k=1}^n \left\{ g, \frac{\partial h}{\partial p_k} \frac{\partial f}{\partial q_k} \right\}, \\ \{h, \{f, g\}\} &= \sum_{k=1}^n \left\{ h, \frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} \right\} - \sum_{k=1}^n \left\{ h, \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right\},\end{aligned}$$

de los cuales los que contienen derivadas segundas de f son sólo (por propiedad de distribución del producto)

$$\begin{aligned}\sum_{k=1}^n \left\{ g, \frac{\partial f}{\partial p_k} \right\} \frac{\partial h}{\partial q_k} - \sum_{k=1}^n \left\{ g, \frac{\partial f}{\partial q_k} \right\} \frac{\partial h}{\partial p_k}, \\ \sum_{k=1}^n \left\{ h, \frac{\partial f}{\partial q_k} \right\} \frac{\partial g}{\partial p_k} - \sum_{k=1}^n \left\{ h, \frac{\partial f}{\partial p_k} \right\} \frac{\partial g}{\partial q_k},\end{aligned}$$

la suma de los cuales se ve fácilmente que se anula, con lo que probamos la identidad.

Consideremos ahora la derivada temporal de una función cualquiera f , que se expresa como

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial t} + \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \dot{q}_k + \frac{\partial f}{\partial p_k} \dot{p}_k \right) \\ &= \frac{\partial f}{\partial t} + \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial H}{\partial q_k} \right) \\ &= \frac{\partial f}{\partial t} + \{f, H\},\end{aligned}\tag{3.35}$$

donde para pasar al segundo renglón se han usado las ecuaciones de Hamilton.

De esta manera, si la magnitud f es constante de movimiento del sistema ($df/dt = 0$) y, además, no depende explícitamente del tiempo ($\partial f/\partial t = 0$), entonces su corchete de Poisson con H es nulo

$$\{f, H\} = 0,$$

que nos proporciona una manera de verificar si una función no dependiente en forma explícita del tiempo es constante de movimiento. Además, un teorema debido a Poisson nos dice que si dos magnitudes f y g son constantes de movimiento su corchete de Poisson $\{f, g\}$ también lo es; lo que a veces permite encontrar nuevas constantes de movimiento. La demostración para el caso

en que f y g no dependen explícitamente del tiempo es sencilla; usando la identidad de Jacobi (3.34), con el hamiltoniano H en lugar de h , se tiene

$$\{f, \{g, H\}\} + \{g, \{H, f\}\} + \{H, \{f, g\}\} = 0;$$

dado que por hipótesis $\{f, H\} = \{g, H\} = 0$, resulta inmediatamente que $\{H, \{f, g\}\} = 0$.

Si f y/o g dependen explícitamente del tiempo basta calcular la derivada total de $\{f, g\}$:

$$\begin{aligned} \frac{d}{dt} \{f, g\} &= \frac{\partial}{\partial t} \{f, g\} + \{\{f, g\}, H\} \\ &= \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} + \{\{f, H\}, g\} + \{f, \{g, H\}\} \\ &= \left\{ \frac{df}{dt}, g \right\} + \left\{ f, \frac{dg}{dt} \right\}; \end{aligned}$$

donde para pasar al segundo renglón se usó la identidad de Jacobi (3.34) y propiedades anteriores. Como por hipótesis $df/dt = dg/dt = 0$, se prueba el teorema.

Finalmente, veamos que los corchetes de Poisson no dependen de las variables usadas para calcularlos, siempre que estas variables correspondan a transformaciones canónicas de unas a otras; en otras palabras, los corchetes son invariantes ante transformaciones canónicas. Si llamamos,

$$\begin{aligned} \{f, g\}_{q,p} &= \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right), \\ \{f, g\}_{Q,P} &= \sum_{k=1}^n \left(\frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial P_k} - \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial Q_k} \right), \end{aligned}$$

entonces $\{f, g\}_{q,p} = \{f, g\}_{Q,P}$ si la transformación $Q_k(q_1, \dots, q_n, p_1, \dots, p_n, t)$, $P_k(q_1, \dots, q_n, p_1, \dots, p_n, t)$ es canónica. Para ver esto calculemos

$$\{Q_i, g\}_{q,p} = \sum_{k=1}^n \left(\frac{\partial Q_i}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial g}{\partial q_k} \right).$$

Si consideramos la transformación canónica generada por una $M'(q, P)$ tenemos

$$\frac{\partial Q_i}{\partial q_k} = \frac{\partial}{\partial q_k} \frac{\partial M'}{\partial P_i} = \frac{\partial}{\partial P_i} \frac{\partial M'}{\partial q_k} = \frac{\partial p_k}{\partial P_i},$$

mientras que si la consideramos generada por una $M'''(p, P)$ es

$$\frac{\partial Q_i}{\partial p_k} = \frac{\partial}{\partial p_k} \frac{\partial M'''}{\partial P_i} = \frac{\partial}{\partial P_i} \frac{\partial M'''}{\partial p_k} = -\frac{\partial q_k}{\partial P_i},$$

con lo que obtenemos

$$\begin{aligned} \{Q_i, g\}_{q,p} &= \sum_{k=1}^n \left(\frac{\partial p_k}{\partial P_i} \frac{\partial g}{\partial p_k} + \frac{\partial q_k}{\partial P_i} \frac{\partial g}{\partial q_k} \right) \\ &= \frac{\partial g}{\partial P_i} = \{Q_i, g\}_{Q,P}, \end{aligned}$$

donde para la última igualdad se usó (3.33). Análogamente se prueba que

$$\{P_i, g\}_{q,p} = \{P_i, g\}_{Q,P}.$$

Ahora bien, en el entorno de cualquier punto $(Q_1^0, \dots, Q_n^0, P_1^0, \dots, P_n^0)$ del espacio de fases podemos desarrollar a una función f de estas variables a primer orden (no hace falta más pues sólo se requieren derivadas primeras en los corchetes) como

$$f(Q, P) = f(Q^0, P^0) + \sum_{k=1}^n \frac{\partial f}{\partial Q_k} (Q_k - Q_k^0) + \sum_{k=1}^n \frac{\partial f}{\partial P_k} (P_k - P_k^0).$$

Evaluando $\{f, g\}_{q,p}$ en $(Q_1^0, \dots, Q_n^0, P_1^0, \dots, P_n^0)$ con este desarrollo de f , y usando propiedades básicas de los corchetes, y las relaciones recién deducidas, obtenemos finalmente

$$\begin{aligned} \{f, g\}_{q,p} &= \sum_{k=1}^n \frac{\partial f}{\partial Q_k} \{Q_k, g\}_{q,p} + \sum_{k=1}^n \frac{\partial f}{\partial P_k} \{P_k, g\}_{q,p} \\ &= \sum_{k=1}^n \frac{\partial f}{\partial Q_k} \{Q_k, g\}_{Q,P} + \sum_{k=1}^n \frac{\partial f}{\partial P_k} \{P_k, g\}_{Q,P} \\ &= \sum_{k=1}^n \frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial P_k} - \sum_{k=1}^n \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial Q_k} = \{f, g\}_{Q,P}, \end{aligned} \quad (3.36)$$

que es lo que queríamos probar.

Procediendo a la inversa, si se verifica (3.36) para cualquier par de funciones, basta reemplazar g con H y f con Q_k o P_k para concluir que la transformación de las q_k, p_k a las Q_k, P_k es canónica. De hecho, basta con verificarlo para sólo las Q_k y P_k , de las que se deduce (como se hizo para la (3.36)) que vale para cualquier f y g . Así, de las (3.32), si se cumple

$$\{Q_i, Q_j\}_{q,p} = \{P_i, P_j\}_{q,p} = 0, \quad \{Q_i, P_j\}_{q,p} = \delta_{ij},$$

la transformación es canónica, que es un método práctico de determinar si una transformación dada tiene esta propiedad.

Con formalismo simpléctico los corchetes de Poisson se escriben en forma matricial como

$$\{f, g\} = \left(\frac{\partial f}{\partial \eta} \right)^T \mathbf{J} \frac{\partial g}{\partial \eta},$$

y la prueba de que su valor es independiente de las variables usadas (siempre que sean canónicas) es mucho más simple que la vista. En efecto, dado que

$$\frac{\partial f}{\partial \eta_k} = \sum_{j=1}^{2n} \frac{\partial f}{\partial \eta'_j} \frac{\partial \eta'_j}{\partial \eta_k} = \sum_{j=1}^{2n} \frac{\partial f}{\partial \eta'_j} M_{jk},$$

que en forma matricial es

$$\frac{\partial f}{\partial \eta} = \mathbf{M}^T \frac{\partial f}{\partial \eta'},$$

se tiene

$$\begin{aligned} \{f, g\}_\eta &= \left(\frac{\partial f}{\partial \eta} \right)^T \mathbf{M} \mathbf{J} \mathbf{M}^T \frac{\partial g}{\partial \eta'} \\ &= \left(\frac{\partial f}{\partial \eta'} \right)^T \mathbf{J} \frac{\partial g}{\partial \eta'} = \{f, g\}_{\eta'}, \end{aligned}$$

donde para pasar a la segunda línea se usó la condición de transformación canónica (3.24).

Finalmente, consideremos el cambio de una función ante una transformación canónica infinitesimal (ver (3.27))

$$\begin{aligned} \delta f &= \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \delta q_k + \frac{\partial f}{\partial p_k} \delta p_k \right) \\ &= \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \frac{\partial W}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial W}{\partial q_k} \right) = \{f, W\}. \end{aligned}$$

En particular, si se toma a f como el hamiltoniano,

$$\delta H = \{H, W\}.$$

Vemos entonces que si el hamiltoniano es invariante frente a la transformación canónica infinitesimal debe ser $\{H, W\} = 0$; pero si W no depende explícitamente de t esto significa que W es constante de movimiento. En otras palabras, analizando sólo las propiedades de simetría de H pueden encontrarse las constantes de movimiento del sistema considerado.

Capítulo 4

Ecuación de Hamilton-Jacobi

Las ecuaciones de Lagrange para un sistema descrito por n coordenadas generalizadas son n ecuaciones de segundo orden de derivación en el tiempo. El mismo sistema puede ser estudiado de acuerdo a Hamilton con $2n$ ecuaciones de primer orden. Existe una notable alternativa, la descripción de Hamilton y Jacobi, que se reduce a una única ecuación en derivadas parciales. Esto puede verse como sigue.

Si se pudiera encontrar una transformación canónica tal que el hamiltoniano \tilde{H} expresado en las nuevas variables Q_k, P_k fuese una constante (que conviene, sin pérdida de generalidad, tomar igual a cero), las ecuaciones de Hamilton serían simplemente

$$\dot{Q}_k = \frac{\partial \tilde{H}}{\partial P_k} = 0, \quad \dot{P}_k = -\frac{\partial \tilde{H}}{\partial Q_k} = 0,$$

con la solución inmediata

$$Q_k = \alpha_k = cte, \quad P_k = \beta_k = cte. \quad (4.1)$$

De las relaciones (3.19) tenemos que la función generatriz $M(q_1, \dots, q_n, Q_1, \dots, Q_n, t)$ que genera la transformación buscada debe cumplir (tomamos $\tilde{H} = 0$)

$$H(q_1, \dots, q_n, p_1, \dots, p_n) + \frac{\partial M}{\partial t} = 0.$$

Antes de continuar veamos que, de la ecuación (3.17), teniendo en cuenta que $\tilde{H} = 0$ y $\dot{Q}_k = 0$, obtenemos

$$\frac{dM}{dt} = \sum_{k=1}^n p_k \dot{q}_k - H,$$

de donde (consideramos, sin pérdida de generalidad, que $M(t_0) = 0$, con t_0 fijo pero arbitrario)

$$M(t) = \int_{t_0}^t \left(\sum_{k=1}^n p_k \dot{q}_k - H \right) dt, \quad (4.2)$$

que, al comparar con (3.11), nos indica que M no es otra cosa que la acción S del sistema, considerada como función del límite superior de integración, y en la que las $q_k(t)$ y $p_k(t)$ usadas en su evaluación son las trayectorias reales del sistema en el espacio de fases. De esta manera llamamos de aquí en más S a la función generatriz buscada. Con esto, si usamos además (también de (3.19) o (2.48)) que

$$p_k = \frac{\partial S}{\partial q_k}, \quad (4.3)$$

la (4.2) se reescribe como (véase la (2.49))

$$H \left(q_1, \dots, q_n, \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n} \right) + \frac{\partial S}{\partial t} = 0. \quad (4.4)$$

Esta ecuación, denominada de *Hamilton-Jacobi*, resulta entonces ser una ecuación diferencial para la función generatriz $S(q_1, \dots, q_n, Q_1, \dots, Q_n, t)$ que, por (4.1), es función de sólo las q_k y del tiempo t , con las $Q_k = \alpha_k$ sólo constantes de integración.

Para resolver un problema de esta manera se procede entonces así:

1) Se encuentra una solución completa de la ecuación (4.4); esto es, una solución que contenga n constantes de integración α_k : $S(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n, t)$.

2) Usando, de (3.19) y (4.1), que $\partial S / \partial Q_k = -P_k = -\beta_k$, derivamos parcialmente la solución encontrada respecto de las constantes α_k y las igualamos a las constantes β_k :

$$\frac{\partial S}{\partial \alpha_k} = -\beta_k. \quad (4.5)$$

3) Resolvemos las n ecuaciones (4.5) para hallar las q_k en términos de las constantes α_k , β_k y del tiempo t :

$$q_k(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n, t).$$

Con lo que se obtiene la solución dinámica completa del problema ya que se encuentran las n coordenadas generalizadas como funciones explícitas del tiempo y de $2n$ constantes de integración que permiten ajustar condiciones iniciales genéricas.

No existen métodos generales para obtener soluciones completas de la ecuación de Hamilton-Jacobi; sin embargo, existen casos importantes en los

que es posible obtener una solución completa de ésta por el método de *separación de variables*. Si consideramos hamiltonianos que no dependen explícitamente del tiempo, sabemos que $H = E$, por lo que se propone

$$S(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n, t) = S_0(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n) - E(\alpha_1, \dots, \alpha_n)t, \quad (4.6)$$

donde hemos definido la *acción reducida* S_0 , que simplifica la ecuación de Hamilton-Jacobi a

$$H\left(q_1, \dots, q_n, \frac{\partial S_0}{\partial q_1}, \dots, \frac{\partial S_0}{\partial q_n}\right) = E. \quad (4.7)$$

Análogamente, si H no depende de alguna coordenada (coordenada cíclica), digamos q_1 , entonces se puede plantear

$$S(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n, t) = S'(q_2, \dots, q_n, \alpha_1, \alpha_2, \dots, \alpha_n, t) + \alpha_1 q_1,$$

que lleva a

$$H\left(q_2, \dots, q_n, \alpha_1, \frac{\partial S'}{\partial q_2}, \dots, \frac{\partial S'}{\partial q_n}\right) + \frac{\partial S'}{\partial t} = 0$$

con una variable menos.

Para coordenadas no cíclicas la idea es proponer que la acción reducida es de la forma

$$S_0 = S_1(q_1, \alpha_1, \dots, \alpha_n) + \dots + S_n(q_n, \alpha_1, \dots, \alpha_n), \quad (4.8)$$

y reemplazar en (4.7). La ecuación resultante debe poder separarse en grupos que dependan de sólo una coordenada cada uno. Como la suma de estos grupos (o en general una combinación más complicada de éstos) es igual a la constante E para *cualquier* combinación de las variables q_k , esto es posible sólo si cada grupo es constante. Así, igualamos el grupo que depende de cada q_k a la constante α_k , para escribir, por ejemplo, $E = \alpha_1 + \alpha_2 + \dots + \alpha_n$ (éste es el caso más sencillo; en general se obtiene una forma más complicada). Por otro lado, cada grupo igualado al α_k correspondiente nos provee de una ecuación diferencial ordinaria de primer orden, que puede entonces resolverse por cuadraturas. Nótese que las constantes de integración (las α_k) provienen de la separación de variables y no de constantes de integración de estas cuadraturas, que no son esenciales ya que suman a una constante aditiva a S .

4.1. Variables ángulo-acción

Supongamos ahora que tenemos un sistema separable con acción del tipo (4.6), con S_0 de la forma (4.8). Cada p_k depende de sólo su variable conjugada

q_k , ya que $p_k = \partial S / \partial q_k = \partial S_k / \partial q_k$. Además, ocurre muy frecuentemente que el rango de variación de cada par (q_k, p_k) es limitado; típicamente porque q_k corresponde a un ángulo. También ocurre habitualmente por lo siguiente; al ser H cuadrático en los impulsos, la condición $H = E$ junto a la condición de separabilidad de la ecuación de Hamilton-Jacobi lleva a relaciones de la forma $p_k^2 + f(q_k) = cte$, donde f es una función de sólo q_k ; esta ecuación cuadrática generalmente tiene soluciones reales sólo si q_k varía dentro de ciertos rangos.

Por la razón que fuera, al ser el rango limitado, en su evolución temporal cada par (q_k, p_k) repite su movimiento. Esto no implica que el movimiento sea periódico ya que cuando q_k y p_k vuelven a tomar un valor dado, sus “velocidades” \dot{q}_k y \dot{p}_k son en general distintas porque éstas dependen de los valores de todas las otras variables (considere las ecuaciones de Hamilton), que no retoman sus valores en el mismo instante que el par considerado (salvo casos muy especiales).

Suponemos entonces que todos los pares de variables conjugadas tienen movimientos cíclicos, lo que nos permite calcular las integrales

$$J_k = \oint p_k dq_k, \quad (4.9)$$

para un ciclo completo de cada par de variables conjugadas. Pero como $p_k = \partial S_k(q_k, \alpha_1, \dots, \alpha_n) / \partial q_k$, resulta

$$J_k = J_k(\alpha_1, \dots, \alpha_n),$$

las cuales uno puede invertir para determinar

$$\alpha_k = \alpha_k(J_1, \dots, J_n). \quad (4.10)$$

que al reemplazarse en (4.8) nos da (usamos el mismo símbolo para la función S_0 , si bien la forma funcional ha cambiado)

$$S_0(q_1, \dots, q_n, J_1, \dots, J_n) \quad (4.11)$$

La idea es ahora tomar a la acción reducida (4.11) como una nueva función generatriz; tenemos entonces que los J_k son las coordenadas Q_k de la nueva transformación canónica, cuyas variables conjugadas P_k están dadas por

$$P_k = -\frac{\partial S_0}{\partial J_k},$$

a las que conviene denominar $-\theta_k/2\pi$, de manera que

$$\theta_k = 2\pi \frac{\partial S_0}{\partial J_k}. \quad (4.12)$$

Como $\partial S_0/\partial t = 0$, el nuevo hamiltoniano coincide con el original y, como $H = E$, al reemplazar las (4.10) en la expresión de E (ver (4.6)) tenemos

$$\tilde{H} = E(J_1, \dots, J_n), \quad (4.13)$$

de donde, consistentemente con lo hecho, las J_k resultan ser constantes,

$$\dot{J}_k = \dot{Q}_k = \frac{\partial \tilde{H}}{\partial P_k} = -2\pi \frac{\partial E}{\partial \theta_k} = 0$$

mientras que las ecuaciones de evolución de las θ_k son

$$\dot{\theta}_k = -2\pi \dot{P}_k = 2\pi \frac{\partial \tilde{H}}{\partial Q_k} = 2\pi \frac{\partial E}{\partial J_k} = cte.$$

Si llamamos ω_k a estas constantes,

$$\omega_k \equiv 2\pi \frac{\partial E}{\partial J_k}, \quad (4.14)$$

las nuevas impulsiones tienen la ecuación de movimiento

$$\theta_k(t) = \theta_k(0) + \omega_k t. \quad (4.15)$$

Para ver el significado de estas variables consideremos cuánto varía una de ellas, digamos θ_i , cuando una de las coordenadas q_k realiza un ciclo completo, manteniendo a las demás quietas,

$$\Delta\theta_i = \oint \frac{\partial \theta_i}{\partial q_k} dq_k = 2\pi \frac{\partial}{\partial J_i} \oint \frac{\partial S_0}{\partial q_k} dq_k = 2\pi \frac{\partial}{\partial J_i} \oint p_k dq_k = 2\pi \frac{\partial J_k}{\partial J_i} = 2\pi \delta_{ik}, \quad (4.16)$$

donde se usaron las definiciones (4.9) y (4.12). Tenemos entonces que cuando q_k describe un ciclo la variable θ_k varía en 2π mientras que el resto de las θ_j no son afectadas. Las θ_k son naturalmente llamadas *variables de ángulo*, y las J_k *variables de acción* (ya que tienen dimensiones de acción). Las J_k y θ_k son también conocidas como *variables de Delaunay*, por ser el astrónomo francés Delaunay (1816-1872) quien las introdujo en 1848 en su estudio del movimiento de la luna.

De su definición (4.12) es

$$\theta_k = \theta_k(q_1, \dots, q_n, J_1, \dots, J_n),$$

que podemos invertir para obtener la solución

$$q_k = q_k(\theta_1, \dots, \theta_n, J_1, \dots, J_n)$$

con $\theta_k(t)$ dado por (4.15). El hecho que cuando cada θ_k varía en 2π la variable q_k correspondiente vuelve a tener el mismo valor, indica que las q_k son funciones periódicas de las variables de ángulo y por lo tanto expresables como serie múltiple de Fourier de senos y cosenos de argumentos

$$N_1\theta_1 + N_2\theta_2 + \dots + N_n\theta_n,$$

donde los N_k son enteros. Por las expresiones (4.15) las q_k dependen entonces del tiempo como serie de Fourier de senos y cosenos de argumento

$$(N_1\omega_1 + N_2\omega_2 + \dots + N_n\omega_n)t, \quad (4.17)$$

lo que nos dice que el movimiento será periódico sólo si *todas* las frecuencias angulares ω_k dadas por (4.14) están relacionadas entre sí por factores racionales (de esa manera existen valores de t finitos en los que *todos* los argumentos de la forma (4.17) difieren de los correspondientes a un t dado en un número entero de veces 2π). Si alguna o varias frecuencias están en relaciones irracionales con las otras esto no se cumple nunca y el movimiento no es estrictamente periódico. Por otro lado, como los irracionales pueden aproximarse con tanta precisión como se quiera por números racionales, dado un t cualquiera existirán otros instantes en los que el sistema se acerque tanto como uno quiera al estado que tenía en t . De esta manera, al transcurrir el tiempo el sistema va “llenando” todo el espacio de fases a su disposición, lo que no ocurre cuando el movimiento es estrictamente periódico.

Es notable que esta información pueda ser obtenida por simples cuadraturas (las integraciones en (4.9) para el cálculo de las J_k), reemplazos algebraicos para obtener $E(J_1, \dots, J_n)$ (4.13), y derivaciones (4.14) para obtener las frecuencias angulares ω_k . Éste es el gran mérito del método de Delaunay.

Nota histórica: La posibilidad de una descripción semejante a la presentada fue demostrada por Hamilton, quien en su obra de 1834 utilizó para este fin la denominada *función principal* (por analogía con la función característica introducida por él mismo en el tratamiento de la óptica en su obra *Theory of Systems of Rays* de 1832). Sin embargo, la función principal debe satisfacer dos ecuaciones diferenciales, lo que complica su uso. Fue Jacobi en 1836 quien simplificó y a la vez extendió el método de Hamilton al desarrollar la descripción que lleva el nombre de ambos. Sin embargo, las bondades del método hamiltoniano no fueron reconocidas por la mayoría de los físicos sino hasta que Delaunay introdujo en 1848 las variables de ángulo-acción.

4.2. Invariantes adiabáticos

Una propiedad importante de las variables de acción J_k es que si algún parámetro del sistema varía muy lentamente (en tiempos largos comparados

con los períodos del movimiento a parámetro fijo), las J_k permanecen prácticamente constantes (varían mucho más lentamente que el parámetro mismo). En efecto, si el hamiltoniano depende de algún parámetro λ que varía lentamente por alguna causa externa a la dinámica considerada (por ejemplo la longitud de un péndulo, la constante elástica de un resorte, etc.), la energía ya no se conservará, sino que tendrá una variación muy lenta dada por

$$\frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \dot{\lambda}.$$

De todas maneras podemos seguir considerando a la acción reducida, $S_0(q, J) = S + Et$, como función generatriz de la transformación de las (q, p) a las variables de ángulo-acción, pero teniendo en cuenta que ahora S_0 depende (débilmente) del tiempo, con lo que

$$\tilde{H} = H + \frac{\partial S_0}{\partial t} = H + \frac{\partial S_0}{\partial \lambda} \dot{\lambda},$$

y la ecuación de evolución de las J_k ,

$$\dot{J}_k = \frac{\partial \tilde{H}}{\partial P_k} = -2\pi \frac{\partial \tilde{H}}{\partial \theta_k},$$

será, como $H = E(J_1, \dots, J_n)$,

$$\dot{J}_k = -2\pi \dot{\lambda} \frac{\partial}{\partial \theta_k} \left(\frac{\partial S_0}{\partial \lambda} \right).$$

Téngase en cuenta que sólo se considera la dependencia explícita de λ en S_0 , porque esta función tiene la misma dependencia, en las demás variables, que con λ constante, debido a que la variación de λ es una perturbación pequeña.

Por otro lado sabemos que la variación de S_0 cuando la coordenada q_k recorre un ciclo completo (y por lo tanto θ_k varía en 2π ; ver (4.16)) es

$$\Delta S_0 = \oint \frac{\partial S_0}{\partial q_k} dq_k = J_k,$$

con lo que

$$\frac{\partial \Delta S_0}{\partial \lambda} = \Delta \frac{\partial S_0}{\partial \lambda} = 0,$$

que expresa que $\partial S_0 / \partial \lambda$ es una función periódica de θ_k , por lo que también lo será su derivada respecto de θ_k . Así, si consideramos el valor de \dot{J}_k promediado sobre un ciclo completo de q_k tenemos

$$\langle \dot{J}_k \rangle = -2\pi \dot{\lambda} \left\langle \frac{\partial}{\partial \theta_k} \left(\frac{\partial S_0}{\partial \lambda} \right) \right\rangle = 0.$$

Se ha tomado que $\dot{\lambda}$ no cambia durante el promedio porque λ tiene, por hipótesis, variación despreciable en el período considerado. Así, $\langle J_k \rangle = J_k$ varía aún más lentamente que λ y puede considerarse constante durante la variación de este parámetro; se denomina a J_k invariante adiabático.

En el caso de un oscilador unidimensional de frecuencia ω y energía E la variable de acción es $J = 2\pi E/\omega$, por lo que ante cambios lentos en los parámetros la energía y la frecuencia cambiarán, pero su cociente variará mucho más lentamente, y puede entonces decirse que (con notación evidente) $\delta E = \delta\omega E_0/\omega_0$.

Capítulo 5

Mecánica relativista

5.1. Cinemática relativista

La relatividad especial de Einstein introduce un cambio conceptual en la mecánica de Newton. Por un lado mantiene el principio de relatividad introducido por Galileo en el sentido que las leyes fundamentales de la física son las mismas (invariantes en forma; o sea, covariantes) en todos los sistemas inerciales. Por otro lado, incluye explícitamente entre estas leyes al electromagnetismo de Maxwell. En particular, la velocidad de la luz en el vacío c , al ser la velocidad con la que se propagan las ondas electromagnéticas, deducida de las leyes de Maxwell, debe ser la misma en todos los sistemas inerciales. Así, dos observadores inerciales, en movimiento relativo uno al otro, deben obtener el mismo valor al medir la velocidad de un mismo frente de luz. Postulando esto último, más el principio de relatividad en sistemas inerciales puede deducirse la relatividad especial. Si un frente de luz se mueve del punto de coordenadas \mathbf{x}_1 al punto \mathbf{x}_2 entre los tiempos t_1 y t_2 , medidos en el sistema de referencia de uno de los observadores (sea este S), y en el sistema del otro observador (S') las coordenadas y tiempos correspondientes son \mathbf{x}'_1 , \mathbf{x}'_2 , t'_1 y t'_2 , debe cumplirse que

$$\frac{|\mathbf{x}_2 - \mathbf{x}_1|}{t_2 - t_1} = \frac{|\mathbf{x}'_2 - \mathbf{x}'_1|}{t'_2 - t'_1} = c. \quad (5.1)$$

Es claro que esta relación no se cumple si las coordenadas y tiempos de los dos sistemas están relacionados por una transformación de Galileo. Para determinar cuál es la transformación apropiada notemos que (5.1) puede escribirse como

$$c^2 (t_2 - t_1)^2 - |\mathbf{x}_2 - \mathbf{x}_1|^2 = c^2 (t'_2 - t'_1)^2 - |\mathbf{x}'_2 - \mathbf{x}'_1|^2 = 0,$$

o en forma diferencial, para $\mathbf{x}_2 = \mathbf{x}_1 + d\mathbf{x}$, $t_2 = t_1 + dt$, y expresiones análogas en el otro sistema,

$$c^2 dt^2 - |d\mathbf{x}|^2 = c^2 dt'^2 - |d\mathbf{x}'|^2 = 0. \quad (5.2)$$

Si se considera ahora $d\mathbf{x}$ y dt correspondientes a sucesos arbitrarios en S (no necesariamente relacionando la partida y llegada de un frente de luz), con transformados $d\mathbf{x}'$ y dt' en S' , podemos definir “intervalos” ds y ds' en cada sistema como

$$\begin{aligned} ds^2 &\equiv c^2 dt^2 - |d\mathbf{x}|^2, \\ ds'^2 &\equiv c^2 dt'^2 - |d\mathbf{x}'|^2, \end{aligned}$$

y la relación (5.2) nos dice que si uno de ellos es cero el otro también debe serlo. Como ambos son infinitesimales del mismo orden resulta de todo esto que deben ser proporcionales:

$$ds^2 = a ds'^2.$$

El coeficiente a no puede depender de la posición ni del tiempo debido a la homogeneidad espacial y temporal, y no puede depender de la dirección de la velocidad relativa por la isotropía del espacio; de manera que sólo puede depender del módulo de la velocidad relativa. Si ahora consideramos tres sistemas de referencia: S , S_1 y S_2 , tales que $S_{1,2}$ se mueven con velocidades $\mathbf{V}_{1,2}$ relativas al sistema S , se debe cumplir

$$ds^2 = a(V_1) ds_1^2 = a(V_2) ds_2^2,$$

a la vez que

$$ds_1^2 = a(V_{21}) ds_2^2,$$

donde V_{21} es el módulo de la velocidad del sistema S_2 relativa al sistema S_1 . De estas relaciones obtenemos inmediatamente que

$$\frac{a(V_2)}{a(V_1)} = a(V_{21}),$$

pero V_{21} depende no sólo de V_1 y V_2 , sino también del ángulo entre ambas velocidades. La relación anterior requiere entonces que a sea constante y, por la misma ecuación, de valor uno. Tenemos entonces la relación fundamental

$$ds^2 = ds'^2,$$

esto es, el intervalo ds es invariante ante las transformaciones correctas. Por supuesto, como vale punto a punto en forma diferencial, la igualdad vale también en el caso de intervalo finito:

$$c^2 (t_2 - t_1)^2 - |\mathbf{x}_2 - \mathbf{x}_1|^2 = c^2 (t'_2 - t'_1)^2 - |\mathbf{x}'_2 - \mathbf{x}'_1|^2. \quad (5.3)$$

Para simplificar las expresiones podemos referir los intervalos espaciales y temporales de cada sistema a sus respectivos orígenes, y considerar sistemas que en el origen de sus tiempos coinciden también en el origen de sus coordenadas espaciales; esto es, tomamos en (5.3) $t_1 = t'_1 = 0$ y $\mathbf{x}_1 = \mathbf{x}'_1 = \mathbf{0}$, para escribirla como (el subíndice 2 corresponde a un suceso genérico denotado sin subíndice)

$$c^2 t^2 - |\mathbf{x}|^2 = c^2 t'^2 - |\mathbf{x}'|^2, \quad (5.4)$$

donde las variables primadas son las transformadas de las sin primar (y viceversa). Así, la descripción de la realidad física en cada sistema inercial debe hacerse a través de una variedad de cuatro dimensiones, tres espaciales y una temporal, cuyas coordenadas espacio-temporales se transforman entre sistemas inerciales de manera tal de preservar el intervalo (5.4). La forma de (5.4) correspondería a la de una distancia en un espacio euclidiano de cuatro dimensiones si no fuera por el signo menos; podemos remediar esto, momentáneamente, pasando a una nueva variedad (compleja) en la que el tiempo t se reemplaza por la coordenada imaginaria it , de manera que (5.4) se reescribe (cancelando un signo menos global)

$$c^2 t^2 + |\mathbf{x}|^2 = c^2 t'^2 + |\mathbf{x}'|^2.$$

Podemos decir ahora que las transformaciones que preservan esta distancia euclidiana en cuatro dimensiones son las rotaciones (no consideramos reflexiones). Desde ya son rotaciones en cuatro dimensiones: tres en los planos (ict, x) , (ict, y) , (ict, z) , y tres en los planos (x, y) , (x, z) , (y, z) . Estas últimas, que dejan el tiempo t invariante corresponden a rotaciones del espacio ordinario (en un mismo sistema de referencia); las más interesantes son las tres primeras; por ejemplo, la rotación (it, x) corresponde a dejar fijos los ejes y, z . Su expresión general es, como ya sabemos,

$$\begin{aligned} x' &= x \cos(\theta) + ict \sin(\theta), \\ ict' &= -x \sin(\theta) + ict \cos(\theta), \end{aligned}$$

que podemos reescribir, multiplicando la segunda por $-i$,

$$\begin{aligned} x' &= x \cos(\theta) + ict \sin(\theta) \\ ct' &= ix \sin(\theta) + ct \cos(\theta). \end{aligned}$$

Como los tiempos y coordenadas son reales, basta definir $\theta \equiv -i\psi$, con ψ real, para tener las expresiones entre variables reales

$$\begin{aligned}x' &= x \cosh(\psi) + ct \sinh(\psi), \\ct' &= x \sinh(\psi) + ct \cosh(\psi),\end{aligned}$$

con expresiones análogas para las otras “rotaciones” (ahora de la variable real t) $(ct, y), (ct, z)$. Para determinar ψ notemos que si el sistema S' se desplaza con velocidad V respecto de S a lo largo del eje x , el origen de S' , $x' = 0$, corresponde a $x = Vt$, con lo que de la primera de las relaciones anteriores es

$$\tanh(\psi) = -\frac{V}{c},$$

de donde

$$\begin{aligned}\sinh(\psi) &= \frac{-V/c}{\sqrt{1 - (V/c)^2}}, \\ \cosh(\psi) &= \frac{1}{\sqrt{1 - (V/c)^2}},\end{aligned}$$

con lo que resulta finalmente

$$t' = \frac{t - Vx/c^2}{\sqrt{1 - (V/c)^2}} \equiv \gamma(t - Vx/c^2), \quad (5.5a)$$

$$x' = \frac{x - Vt}{\sqrt{1 - (V/c)^2}} \equiv \gamma(x - Vt), \quad (5.5b)$$

donde se introdujo el símbolo convencional $\gamma \equiv [1 - (V/c)^2]^{-1/2}$. Éstas son las transformaciones buscadas, denominadas de Lorentz, para el caso particular visto (\mathbf{V} paralela a x), que se completan con $y' = y$, $z' = z$.

Si el movimiento del sistema S' relativo al S fuese con velocidad relativa constante \mathbf{V} , pero de dirección arbitraria, basta ver que las relaciones anteriores corresponden a la transformación de la coordenada paralela a \mathbf{V} , mientras que las coordenadas perpendiculares no sufren transformación. Con esto, podemos escribir, directamente de las expresiones anteriores, proyectando en la dirección de \mathbf{V} , (ahora V simboliza el módulo de \mathbf{V})

$$\begin{aligned}t' &= \gamma(t - \mathbf{V} \cdot \mathbf{x}/c^2), \\ \frac{\mathbf{V} \cdot \mathbf{x}'}{V} &= \gamma\left(\frac{\mathbf{V} \cdot \mathbf{x}}{V} - Vt\right), \\ \mathbf{x}' - \frac{\mathbf{V} \cdot \mathbf{x}'}{V^2}\mathbf{V} &= \mathbf{x} - \frac{\mathbf{V} \cdot \mathbf{x}}{V^2}\mathbf{V}.\end{aligned}$$

Combinando las dos últimas tenemos finalmente la transformación de Lorentz más general (restringida sólo a que coincidan los orígenes de ambos sistemas de coordenadas espaciales en el origen también común de tiempos)

$$t' = \gamma (t - \mathbf{V} \cdot \mathbf{x}/c^2), \quad (5.6a)$$

$$\mathbf{x}' = \mathbf{x} - \gamma \mathbf{V} t + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{x}}{V^2} \mathbf{V}. \quad (5.6b)$$

Las transformaciones inversas pueden obtenerse sin necesidad de despejarlas de las (5.6), simplemente usando que deben tener la misma forma, sólo que la velocidad relativa tiene sentido opuesto:

$$t = \gamma (t' + \mathbf{V} \cdot \mathbf{x}'/c^2), \quad (5.7a)$$

$$\mathbf{x} = \mathbf{x}' + \gamma \mathbf{V} t' + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{x}'}{V^2} \mathbf{V}. \quad (5.7b)$$

Tenemos así que las transformaciones que dejan invariante el intervalo (5.4) son las rotaciones del espacio ordinario y las transformaciones (5.6) llamadas “impulsos” (en inglés “boosts”). Por supuesto, tenemos además las translaciones del origen del tiempo y del espacio. Con esto, transformaciones con un total de diez grados de libertad: tres de rotaciones, tres de impulsos y cuatro de translaciones.

5.1.1. Cuadrivectores

Como sucede en el caso del espacio de dimensión tres, pueden definirse magnitudes vectoriales, que son aquellas que se transforman como los vectores posición del espacio. En este caso la “posición” en el espacio-tiempo de dimensión cuatro está dada por el quadri-vector posición, que abreviamos $(ct, \mathbf{x}) \equiv (x_0, \mathbf{x})$, que se transforma según (5.6). Debe tenerse en cuenta además que la distancia al origen que es preservada por tal transformación está dada por $(ct)^2 - |\mathbf{x}|^2 \equiv x_0^2 - |\mathbf{x}|^2$, por lo que definimos el producto escalar entre quadri-vectores para que satisfaga esta expresión. Así, definimos los cuadrivectores como conjuntos de cuatro componentes, genéricamente representados como $A \equiv (A_0, \mathbf{A})$. La componente A_0 es denominada temporal, y las tres componentes \mathbf{A} son denominadas espaciales. Si λ es un escalar, definimos

$$A'_0 = \gamma (A_0 - \mathbf{V} \cdot \mathbf{A}/c), \quad (5.8a)$$

$$\mathbf{A}' = \mathbf{A} - \gamma \frac{\mathbf{V} A_0}{c} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{A}}{V^2} \mathbf{V}, \quad (5.8b)$$

$$\lambda A = (\lambda A_0, \lambda \mathbf{A}),$$

$$A \cdot B^* = A_0 B_0^* + \mathbf{A} \cdot \mathbf{B}^* = A_0 B_0 - \mathbf{A} \cdot \mathbf{B},$$

donde $B^* = (B_0^*, \mathbf{B}^*) \equiv (B_0, -\mathbf{B})$ es denominado co-cuadrivector de B . Éste se transforma según las (5.8), pero con $-\mathbf{V}$ en lugar de \mathbf{V} . El producto escalar así definido es invariante ante transformaciones de Lorentz; esto es, es un escalar de Lorentz.

Entonces, así como en la mecánica de Newton las leyes de la física deben ser invariantes de Galileo, y deben por lo tanto expresarse en términos de vectores y escalares ante transformaciones de Galileo, las leyes en la mecánica relativista deben expresarse en términos de vectores de Lorentz (cuadrivectores) y escalares de Lorentz.

Por ejemplo, nos planteamos cómo representar el movimiento de una partícula en este marco; claramente su posición cuadri-dimensional estará dada por el cuadri-vector $X = (x_0, \mathbf{x})$. Si queremos ahora representar su cuadri-velocidad notamos que si su velocidad ordinaria es \mathbf{u} , su posición habrá cambiado en un dt en $dX = (dx_0, d\mathbf{x}) = (cdt, \mathbf{u} dt)$. Si definiéramos su cuadri-velocidad como $dX/dt = (c, \mathbf{u})$ obtendríamos una expresión que no es un vector de Lorentz porque si bien dX es un cuadri-vector, dt no es un escalar de Lorentz. Lo más similar a dt que es además escalar de Lorentz es ds/c , el intervalo cuadri-dimensional recorrido por la partícula en dt , dividido por c ; en particular, para velocidades pequeñas comparadas con la de la luz ds/c es prácticamente igual a dt . Para formalizar esto consideramos la magnitud escalar de Lorentz, con dimensiones de tiempo,

$$\begin{aligned} d\tau &\equiv \frac{ds}{c} = \frac{1}{c} \sqrt{(cdt)^2 - d|\mathbf{x}|^2} \\ &= \sqrt{1 - \frac{1}{c^2} \left(\frac{d|\mathbf{x}|}{dt} \right)^2} dt = \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}} dt, \end{aligned}$$

y podemos entonces definir la cuadri-velocidad como $U \equiv dX/d\tau$. El significado físico de $d\tau$ es, de la última igualdad, que corresponde al intervalo de tiempo medido en un sistema en el que la partícula tiene, en ese instante, velocidad nula; es el denominado *intervalo de tiempo propio*. El tiempo propio es entonces el tiempo que marcaría un reloj que se mueve junto con la partícula (y es por lo tanto el tiempo que “experimenta” la partícula). La cuadri-velocidad resulta entonces

$$U = \frac{dX}{d\tau} = \frac{(c, \mathbf{u})}{\sqrt{1 - |\mathbf{u}|^2/c^2}}. \quad (5.9)$$

El tiempo propio que experimenta una partícula entre los tiempos t_1 y t_2 es

$$\tau = \int_{t_1}^{t_2} \sqrt{1 - \frac{|\mathbf{u}(t)|^2}{c^2}} dt, \quad (5.10)$$

que es menor que el intervalo de tiempo correspondiente $t_2 - t_1$, tanto menor cuanto mayor es la velocidad, la notable predicción de la relatividad de la ralentización de los relojes en movimiento (o dilatación del tiempo). Con todo esto, ¿a qué nos referimos cuando hablamos del tiempo t ? Operacionalmente podemos definir un tiempo t , común a todos los puntos espaciales de un sistema de referencia dado, como el que indican relojes iguales, sincronizados entre sí, que están en reposo permanente en esos puntos. Sabiendo que en cualquier sistema inercial la luz en el vacío se propaga con igual velocidad c , la sincronización mencionada puede hacerse con señales luminosas; basta acordar que en un instante; por ejemplo, en un tiempo que se tomará como $t = 0$, se emitirá desde el origen un pulso en forma de frente esférico de luz. Conocidas las posiciones de cada reloj puede entonces sincronizarse cada uno de estos con el del origen (y entre ellos), teniendo en cuenta el retraso en la llegada de la señal. Con lo que vimos entendemos ahora que si esta sincronización se hace en dos sistemas de referencia distintos; por ejemplo, en el momento en el que ambos orígenes coinciden se emite un único pulso de sincronización y se sincronizan los relojes dentro de cada sistema con este mismo pulso, aun así los relojes de cada sistema marcarán tiempos diferentes (justamente esto es lo que cuantifica la primera relación de las (5.6): el reloj de S' en \mathbf{x}' indica t' cuando se cruza con el de S que está en \mathbf{x} y que indica t). Aun más, lo que llamamos $t = cte$ en S no corresponde a un valor fijo de t' en S' : lo que es simultáneo en un sistema no lo es en el otro.

Además de la relatividad del tiempo, se predice también la relatividad de las longitudes. Consideremos una varilla fija en el sistema S' con un extremo en \mathbf{x}'_1 y otro en \mathbf{x}'_2 . Si en el sistema S queremos medir su longitud, determinamos las coordenadas espaciales de los extremos de la varilla en S , \mathbf{x}_1 y \mathbf{x}_2 , en un instante t_0 . De la primera de las (5.7) tenemos los tiempos en S' , correspondientes al momento de la medición en S , de las posiciones de los extremos de la varilla:

$$t'_{1,2} = \frac{t_0}{\gamma} - \frac{\mathbf{V} \cdot \mathbf{x}'_{1,2}}{c^2},$$

donde vemos claramente la no simultaneidad en S' de la medición. Al reemplazar en la segunda de las (5.7) obtenemos las posiciones de los extremos de la varilla en S

$$\mathbf{x}_{1,2} = \mathbf{x}'_{1,2} + \mathbf{V}t_0 - \gamma \frac{\mathbf{V} \cdot \mathbf{x}'_{1,2}}{c^2} \mathbf{V} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{x}'_{1,2}}{V^2} \mathbf{V},$$

de donde resulta que $\Delta \mathbf{x} \equiv \mathbf{x}_2 - \mathbf{x}_1$ está relacionado con $\Delta \mathbf{x}' \equiv \mathbf{x}'_2 - \mathbf{x}'_1$ por

$$\Delta \mathbf{x} = \Delta \mathbf{x}' + [\gamma (1 - V^2/c^2) - 1] \frac{\mathbf{V} \cdot \Delta \mathbf{x}'}{V^2} \mathbf{V}.$$

Tenemos entonces para la componente de $\Delta\mathbf{x}'$ perpendicular a \mathbf{V} que $\Delta\mathbf{x}_\perp = \Delta\mathbf{x}'_\perp$, la misma en los dos sistemas. Por otro lado, la componente de $\Delta\mathbf{x}'$ con la misma dirección que \mathbf{V} resulta

$$\Delta\mathbf{x}_\parallel = \frac{\Delta\mathbf{x}'_\parallel}{\gamma},$$

con lo que la longitud paralela a \mathbf{V} medida en S es un factor $\gamma^{-1} = \sqrt{1 - V^2/c^2}$ más pequeña que la correspondiente en el sistema donde se encuentra en reposo, S' , que denominamos naturalmente como la longitud propia, o longitud en reposo. En particular, esto nos dice que la varilla con movimiento oblicuo respecto de \mathbf{V} se observará “rotada” en S .

Esta contracción de los objetos en la dirección de su movimiento debe considerarse como real en el sentido que es el resultado de una medición física; en particular, objetos con el movimiento apropiado deben ser capaces de pasar por aberturas fijas en S que serían demasiado estrechas para permitir el paso del objeto si tuviera su longitud en reposo.

Con respecto al significado de “observar”, está usado aquí en el sentido de interpretar una medición física. Particularmente en relatividad debe tenerse muy clara la forma en que es hecha cada medición para interpretarla apropiadamente. Como regla, la medición de la propiedad de un objeto físico debe ser hecha por un observador en la proximidad inmediata del objeto (en movimiento o no relativo a éste), y a posteriori puede ser interpretada por el conjunto de observadores en reposo respecto al observador original (los que están en su sistema inercial) una vez obtenida la información. Si las características del experimento lo requieren, intervendrán conjuntos de observadores (como en el caso de la medición de la longitud de la varilla), cada uno de los cuales mide la propiedad correspondiente a su lugar. Mediciones hechas a distancia, con telescopios, detectores remotos, etc. deben ser procesadas apropiadamente para recuperar la información que hubiera obtenido el observador (en el mismo sistema) cercano al fenómeno. Sólo a partir de estos datos pueden obtenerse sin ambigüedad propiedades derivadas, como la longitud de la varilla a partir de las posiciones simultáneas de sus extremos. Un ejemplo clásico es cómo se ve un objeto dado en movimiento, lo que requiere considerar la llegada simultánea al ojo de la luz proveniente de las distintas partes del objeto, habiendo ya medido (o calculado) las trayectorias espacio-temporales de tales partes en el sistema de referencia de quien mira.

Otra consecuencia básica de la relatividad está referida a la composición de velocidades. Si la partícula se mueve con velocidad \mathbf{u} referida al sistema S su velocidad \mathbf{u}' referida al sistema S' , que se mueve a su vez con velocidad \mathbf{V} relativa a S , puede calcularse sabiendo cómo se transforma la cuadri-

velocidad. Si para abreviar notación definimos

$$\gamma_u \equiv \frac{1}{\sqrt{1 - |\mathbf{u}|^2/c^2}}, \quad \gamma'_u \equiv \frac{1}{\sqrt{1 - |\mathbf{u}'|^2/c^2}}, \quad (5.11)$$

entonces sabemos que, de (5.8) y (5.9),

$$\begin{aligned} U'_0 &= \gamma'_u c = \gamma (\gamma_u c - \gamma_u \mathbf{V} \cdot \mathbf{u}/c), \\ \gamma'_u \mathbf{u}' &= \gamma_u \left[\mathbf{u} - \gamma \mathbf{V} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{u}}{V^2} \mathbf{V} \right]. \end{aligned}$$

De la primera es entonces

$$\gamma'_u = \gamma_u \gamma (1 - \mathbf{V} \cdot \mathbf{u}/c^2), \quad (5.12)$$

que al reemplazar en la segunda nos da

$$\mathbf{u}' = \frac{\mathbf{u} - \gamma \mathbf{V} + (\gamma - 1) (\mathbf{V} \cdot \mathbf{u}) \mathbf{V}/V^2}{\gamma (1 - \mathbf{V} \cdot \mathbf{u}/c^2)}. \quad (5.13)$$

En particular, si \mathbf{u} y \mathbf{V} tienen la misma dirección (sean paralelas o antiparalelas) resulta

$$\mathbf{u}' = \frac{\mathbf{u} - \mathbf{V}}{1 - \mathbf{V} \cdot \mathbf{u}/c^2},$$

que vemos que nunca supera el valor de c cualesquiera que sean las velocidades \mathbf{u} y \mathbf{V} (siempre que éstas tampoco superen el valor de c).

5.2. Dinámica relativista

Consideremos ahora la dinámica de una partícula libre en el marco de la relatividad. En la mecánica newtoniana la acción es un escalar de Galileo; en mecánica relativista, si queremos tener leyes invariantes en este marco debemos tener una acción escalar de Lorentz. Si procedemos por analogía con lo ya visto, podemos proponer que la integral que define la acción debe hacerse en la variable τ , el tiempo propio de la partícula, en lugar de t . Por otro lado, el cuadrado de la velocidad que usamos para la energía cinética debe reemplazarse por una magnitud relacionada, que sea escalar de Lorentz. Si evaluamos $U \cdot U$ vemos que resulta igual a la constante c^2 . Proponemos entonces para la partícula libre

$$S = \alpha \int_{\tau_1}^{\tau_2} d\tau = \alpha \int_{t_1}^{t_2} \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}} dt,$$

con α una constante a determinar. Identificamos así el lagrangiano como

$$L = \alpha \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}},$$

de donde obtenemos la componente cartesiana genérica del momento lineal y la energía de la partícula

$$p_i = \frac{\partial L}{\partial u_i} = -\frac{\alpha u_i}{c^2 \sqrt{1 - |\mathbf{u}|^2/c^2}},$$

$$E = \sum_i u_i p_i - L = -\frac{\alpha}{\sqrt{1 - |\mathbf{u}|^2/c^2}}.$$

En el límite de velocidades pequeñas comparadas con la de la luz la primera igualdad tiende a $p_i \rightarrow -\alpha u_i/c^2$, con lo que debemos identificar a α como proporcional a la masa de la partícula; al menos la correspondiente a bajas velocidades, que tomamos como la masa en reposo m . Así, $\alpha = -mc^2$, y

$$\mathbf{p} = \frac{m\mathbf{u}}{\sqrt{1 - |\mathbf{u}|^2/c^2}}, \quad (5.14)$$

$$E = \frac{mc^2}{\sqrt{1 - |\mathbf{u}|^2/c^2}}. \quad (5.15)$$

Notemos que para velocidades bajas es también

$$L \simeq -mc^2 + \frac{1}{2}m|\mathbf{u}|^2,$$

$$E \simeq mc^2 + \frac{1}{2}m|\mathbf{u}|^2,$$

con lo que se recuperan todos los resultados conocidos, con el agregado notable que la energía de la partícula tiene una contribución de base de valor mc^2 .

Una relación útil es

$$E^2 - |\mathbf{p}|^2 c^2 = m^2 c^4. \quad (5.16)$$

Una observación importante es que el cuadvivector mU tiene como componentes

$$mU = \frac{m(c, \mathbf{u})}{\sqrt{1 - |\mathbf{u}|^2/c^2}} = (E/c, \mathbf{p}) \equiv P, \quad (5.17)$$

la energía y el momento lineal son así componentes del cuadrivector llamado energía-momento, y se transforman por lo tanto ante cambio de sistema de referencia como indican las (5.8):

$$E' = \gamma (E - \mathbf{V} \cdot \mathbf{p}), \quad (5.18)$$

$$\mathbf{p}' = \mathbf{p} - \gamma \frac{E}{c^2} \mathbf{V} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{p}}{V^2} \mathbf{V}. \quad (5.19)$$

Esto que hemos deducido para una partícula libre es un resultado general de la relatividad, y podemos comprenderlo en términos de la acción como función de las coordenadas espacio-temporales. En efecto, por (2.48) y (2.49) podemos escribir para la acción de un sistema genérico

$$\begin{aligned} dS &= \frac{\partial S}{\partial t} dt + \sum_i \frac{\partial S}{\partial x_i} dx_i \\ &= -E dt + \sum_i p_i dx_i = -(E/c, \mathbf{p}) \cdot dX^*. \end{aligned}$$

Como dS es un escalar de Lorentz y dX un cuadrivector, $(E/c, \mathbf{p})$ debe ser un cuadrivector.

Si queremos ahora encontrar la acción de un sistema de varias partículas nos encontramos con problemas debido a las interacciones, sea entre sí o con agentes externos. Si las partículas no interactúan la acción es simplemente la suma de las acciones de cada partícula

$$S = -c^2 \sum_n \int_{t_1}^{t_2} m^{(n)} \sqrt{1 - \frac{|\mathbf{u}^{(n)}|^2}{c^2}} dt.$$

Las interacciones requieren adicionar términos que sean invariantes de Lorentz. Para el caso de interacción con un campo electromagnético externo sabemos que el caso no relativista corresponde a un potencial dependiente de las velocidades (2.23) que contribuye a la acción de la forma

$$S^{EM} = - \sum_n \int_{t_1}^{t_2} e^{(n)} [\phi(\mathbf{x}^{(n)}, t) - \mathbf{u}^{(n)} \cdot \mathbf{A}(\mathbf{x}^{(n)}, t)] dt.$$

Esta expresión resulta ser invariante de Lorentz porque, del electromagnetismo de Maxwell, la carga $e^{(n)}$ lo es, y $A = (\phi, \mathbf{A})$ es un cuadrivector, con lo que

$$\phi - \mathbf{u} \cdot \mathbf{A} = (c, \mathbf{u}) \cdot (\phi, -\mathbf{A}) = \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}} U \cdot A^*,$$

resultando entonces

$$S^{EM} = - \sum_n \int_{t_1}^{t_2} e^{(n)} U^{(n)} \cdot A^{(n)*} \sqrt{1 - \frac{|\mathbf{u}^{(n)}|^2}{c^2}} dt.$$

Nótese que la acción del sistema puede escribirse como suma de acciones de cada partícula por separado en el campo externo.

No existen otros campos dentro de la mecánica clásica de partículas que den lugar a acciones invariantes de Lorentz. Por otro lado, las interacciones entre partículas clásicas pueden tratarse en forma correcta relativista para el caso de interacción electromagnética, en la que las mismas partículas son fuente de los potenciales ϕ , \mathbf{A} . En este último caso la acción del sistema no puede escribirse como una suma de acciones de partículas individuales porque $A^{(n)}$ depende de la historia de todas las partículas del sistema. Es claro que no puede depender de posiciones y velocidades “instantáneas”; esto es, simultáneas, lo que implicaría acción instantánea a distancia y, además, lo que es simultáneo en un sistema de referencia no lo es en otro (existiría un sistema privilegiado). Así, el lagrangiano resultante no es una función de sólo las coordenadas y velocidades de las partículas, y termina siendo él mismo una funcional de las trayectorias, lo que es en la práctica inmanejable. La solución a esto es considerar a los potenciales mismos como coordenadas generalizadas adicionales, definidas en cada punto del espacio, con lo que a la suma sobre partículas se le agrega una integral sobre las coordenadas del espacio. Esto está fuera del marco de la mecánica de partículas y no será considerado aquí.

Las únicas interacciones que pueden considerarse en una descripción de sólo partículas son a través de choques, en los que la interacción tiene lugar cuando las partículas se encuentran en el mismo punto del espacio. En estas condiciones el potencial de interacción depende de las posiciones y velocidades de las partículas en el instante del choque, que es un evento simultáneo para todas las partículas interactuantes en cualquier sistema de referencia.

Cualquiera sea la interacción, si en un dado sistema de referencia inercial observamos que una partícula cambia su momento lineal \mathbf{p} al transcurrir el tiempo, concluimos que está actuando sobre ella una fuerza, que siempre podemos definir como $\mathbf{F} = d\mathbf{p}/dt$. Para describir consistentemente la misma partícula desde otro sistema de referencia debemos decir que la fuerza en tal sistema es $\mathbf{F}' = d\mathbf{p}'/dt'$, donde \mathbf{p}' y t' son el momento de la partícula y el tiempo en este otro sistema. Puede obtenerse la relación entre \mathbf{F}' y \mathbf{F} (la transformación de la fuerza), a través de las relaciones de transformación entre \mathbf{p}' y t' , y \mathbf{p} y t . Sin embargo, la manera más simple de considerar esto es a través de cuadvectores. Si definimos un cuadvector asociado a la fuerza

sobre la partícula a partir de (5.17), usando además la expresión del tiempo propio (5.10),

$$\begin{aligned} F &\equiv \frac{dP}{d\tau} = \left(\frac{1}{c} \frac{dE}{d\tau}, \frac{d\mathbf{p}}{d\tau} \right) \\ &= \frac{1}{\sqrt{1 - |\mathbf{u}|^2/c^2}} \left(\frac{1}{c} \frac{dE}{dt}, \frac{d\mathbf{p}}{dt} \right) \\ &= \frac{1}{\sqrt{1 - |\mathbf{u}|^2/c^2}} \left(\frac{1}{c} \frac{dE}{dt}, \mathbf{F} \right). \end{aligned}$$

Si suponemos que (5.16) es válida cuando la partícula está bajo la acción de \mathbf{F} tenemos que (usando además (5.14) y (5.15))

$$\frac{dE}{dt} = \frac{c^2 \mathbf{p}}{E} \cdot \frac{d\mathbf{p}}{dt} = \mathbf{u} \cdot \frac{d\mathbf{p}}{dt} = \mathbf{u} \cdot \mathbf{F},$$

con lo que, finalmente,

$$F = \frac{1}{\sqrt{1 - |\mathbf{u}|^2/c^2}} \left(\frac{\mathbf{u} \cdot \mathbf{F}}{c}, \mathbf{F} \right).$$

Dado que F es un cuadrivector, sabemos de las transformaciones de Lorentz, usando las definiciones (5.11), que

$$\gamma_u \mathbf{F}' = \gamma_u \left[\mathbf{F} - \gamma \mathbf{V} \frac{\mathbf{u} \cdot \mathbf{F}}{c^2} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{F}}{V^2} \mathbf{V} \right],$$

y, por la (5.12), obtenemos

$$\mathbf{F}' = \frac{1}{\gamma(1 - \mathbf{V} \cdot \mathbf{u}/c^2)} \left[\mathbf{F} - \gamma \mathbf{V} \frac{\mathbf{u} \cdot \mathbf{F}}{c^2} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{F}}{V^2} \mathbf{V} \right].$$

En términos de las componentes paralela y perpendicular a \mathbf{V} es

$$\begin{aligned} F'_{\parallel} &= \frac{F_{\parallel} - \mathbf{u} \cdot \mathbf{F} \mathbf{V}/c^2}{(1 - \mathbf{V} \cdot \mathbf{u}/c^2)}, \\ \mathbf{F}'_{\perp} &= \frac{\mathbf{F}_{\perp}}{\gamma(1 - \mathbf{V} \cdot \mathbf{u}/c^2)}. \end{aligned}$$

Nótese que aun si en un sistema particular la fuerza no depende de la velocidad, en otro genérico sí dependerá.

5.2.1. Leyes de conservación en sistemas de varias partículas

Por lo dicho arriba sólo consideramos interacciones tipo choque en los sistemas relativistas de varias partículas.

Sabemos del teorema de Noether que si la acción de un sistema de partículas tiene una simetría particular; esto es, es invariante ante un cambio particular de coordenadas, se obtiene una magnitud conservada asociada a tal simetría. Este teorema es válido también en la mecánica relativista, ya que en su deducción sólo se usó la existencia de la acción y su dependencia de las coordenadas, pero con la condición que la transformación del tiempo sea la misma para todas las partículas. Como un sistema aislado debe tener una acción que, por la homogeneidad del espacio y del tiempo, no puede depender del origen usado de espacio y tiempo, las translaciones espacio-temporales deben dejar la acción invariable. La invarianza ante translaciones espaciales lleva, como sabemos, a la conservación del momento lineal del sistema, mientras que la invarianza ante translaciones temporales corresponde a la conservación de la energía. Usando la expresión (2.46) escrita ahora como (la suma sobre n corresponde a las partículas del sistema)

$$\begin{aligned} C &= \left(\sum_n \sum_{\alpha=1}^3 \frac{\partial L}{\partial u_\alpha^{(n)}} u_\alpha^{(n)} - L \right) \theta_0 - \sum_n \sum_{\alpha=1}^3 \frac{\partial L}{\partial u_\alpha^{(n)}} \theta_\alpha^{(n)} \\ &= E \theta_0 - \sum_n \sum_{\alpha=1}^3 p_\alpha^{(n)} \theta_\alpha^{(n)}. \end{aligned}$$

se obtiene lo dicho de igual manera a lo hecho en el caso no relativista, usando que para una translación espacio-temporal de magnitud ε en la dirección \mathbf{n} es $x'_i = x_i + \varepsilon n_i$ ($i = 0, 1, 2, 3$),

$$\begin{aligned} E &= cte, \\ \mathbf{p} &= \sum_n \mathbf{p}^{(n)} = cte. \end{aligned}$$

Como se mencionó, fue posible usar el resultado (2.46) porque en la translación temporal la transformación del tiempo es la misma para todas las partículas (θ_0 no depende de n).

Dado que la energía se conserva incluyendo términos de interacción de colisiones, la energía puede evaluarse como la correspondiente al momento en que ninguna partícula está colisionando con otra; esto es, cuando son todas libres; por lo que E es la suma de las energías libres de todas las partículas

$$E = \sum_n E^{(n)} = cte,$$

donde, de (5.15),

$$E^{(n)} = \frac{m^{(n)}c^2}{\sqrt{1 - |\mathbf{u}^{(n)}|^2/c^2}}.$$

La isotropía del espacio-tiempo nos dice además que la acción de un sistema cerrado debe ser invariante ante rotaciones espaciales e impulsos (ambas constituyen el grupo de “rotaciones” en el espacio-tiempo tetra-dimensional, denominado grupo de Lorentz no reflexivo). Ante una rotación infinitesimal generalizada (de parámetro ε) el cambio de coordenadas espacio-temporales es de la forma general

$$x'_i = x_i + \varepsilon \sum_{k=0}^3 \Omega_{ik} x_k, \quad (5.20)$$

donde la expresión precisa de la matriz tetradimensional Ω_{ik} no es importante, salvo el hecho de que es antisimétrica ($\Omega_{ik} = -\Omega_{ki}$ para las rotaciones ordinarias, como sabemos; para los impulsos basta considerar las (5.6) con parámetro infinitesimal $\mathbf{V}/c = \varepsilon \mathbf{n}$). Esta expresión puede invertirse trivialmente al mismo orden de aproximación para dar

$$x_i = x'_i - \varepsilon \sum_{k=0}^3 \Omega_{ik} x'_k.$$

Ahora parece no ser posible usar el resultado (2.46) para los impulsos porque el tiempo se transforma de manera distinta para cada partícula. En efecto, tenemos ($\alpha = 1, 2, 3$)

$$\begin{aligned} \theta_0^{(n)} &= \left. \frac{1}{c} \frac{\partial x_0^{(n)}}{\partial \varepsilon} \right|_{\varepsilon=0} = -\frac{1}{c} \sum_{k=0}^3 \Omega_{0k} x_k^{(n)}, \\ \theta_\alpha^{(n)} &= \left. \frac{\partial x_\alpha^{(n)}}{\partial \varepsilon} \right|_{\varepsilon=0} = -\sum_{k=0}^3 \Omega_{\alpha k} x_k^{(n)}. \end{aligned}$$

Sin embargo, como hemos considerado interacciones tipo choque, las partículas que interactúan lo hacen en el mismo punto del espacio-tiempo, con lo que para todas ellas el tiempo se transforma de igual manera ante (5.20). Así, la acción puede considerarse como suma de términos correspondientes a distintos grupos de partículas (cada grupo con igual θ_0 para sus partículas), que no interactúan con las demás, con lo que cada término es invariante por separado. Conceptualmente, podemos entonces considerar que aplicamos (2.46) para cada término de la acción y luego sumamos todos ellos, con lo

que obtenemos

$$\sum_n \left[-\frac{E^{(n)}}{c} \sum_{k=0}^3 \Omega_{0k} x_k^{(n)} + \sum_{\alpha=1}^3 p_\alpha^{(n)} \sum_{k=0}^3 \Omega_{\alpha k} x_k^{(n)} \right] = - \sum_{j,k=0}^3 \Omega_{jk} \sum_n p_j^{*(n)} x_k^{(n)},$$

donde usamos, como antes, la energía libre de la partícula n porque la energía se conserva para cada grupo de partículas interactuantes, por lo que debe ser igual a la suma de las energías de las mismas cuando no interactúan.

Como Ω_{jk} es arbitraria (pero antisimétrica) podemos siempre elegir los ejes de rotación espacial o la dirección del impulso para que sean distintos de cero sólo un par de componentes, por ejemplo, $\Omega_{02} = -\Omega_{20}$, con lo que se conserva en tal caso $\sum_n (p_2^{*(n)} x_0^{(n)} - p_0^{*(n)} x_2^{(n)})$. Como lo mismo es válido para todo par de componentes distintas entre sí, debe conservarse en general la matriz antisimétrica (el cuatri-tensor momento angular, ya que está expresado en términos del producto (externo) de cuatri-vectores)

$$M_{jk} = -M_{kj} = \sum_n \left(p_j^{(n)*} x_k^{(n)} - p_k^{(n)*} x_j^{(n)} \right). \quad (5.21)$$

Las componentes puramente espaciales son

$$M_{\alpha\beta} = \sum_n \left(x_\alpha^{(n)} p_\beta^{(n)} - x_\beta^{(n)} p_\alpha^{(n)} \right),$$

que corresponden a las tres componentes del momento angular ordinario.

Las componentes cruzadas

$$\begin{aligned} M_{0\alpha} &= \sum_n \left(x_0^{(n)} p_\alpha^{(n)} - x_\alpha^{(n)} p_0^{(n)} \right) \\ &= \sum_n \left(ct p_\alpha^{(n)} - E^{(n)} x_\alpha^{(n)} / c \right), \end{aligned}$$

son las componentes del vector espacial $ct \sum_n \mathbf{p}^{(n)} - \sum_n E^{(n)} \mathbf{x}^{(n)} / c$. La constancia de este vector puede interpretarse multiplicándolo por $c / \sum_n E^{(n)}$ para tener

$$\frac{\sum_n E^{(n)} \mathbf{x}^{(n)}}{\sum_n E^{(n)}} - c^2 \frac{\sum_n \mathbf{p}^{(n)}}{\sum_n E^{(n)}} t = \mathbf{cte}.$$

De donde interpretamos, sabiendo además que $\sum_n E^{(n)}$ y $\sum_n \mathbf{p}^{(n)}$ son constantes, que el vector generalización del centro de masas

$$\mathbf{R} \equiv \frac{\sum_n E^{(n)} \mathbf{x}^{(n)}}{\sum_n E^{(n)}},$$

se mueve con velocidad constante

$$\mathbf{V}_R = c^2 \frac{\sum_n \mathbf{P}^{(n)}}{\sum_n E^{(n)}}. \quad (5.22)$$

5.2.2. Desintegración de partículas

Consideremos la desintegración espontánea de una partícula de masa m en dos partículas de masas m_1 y m_2 . Supongamos que la partícula inicial está en reposo en el sistema considerado. La conservación de la energía en este proceso (que puede ser considerado el proceso inverso del choque perfectamente plástico de dos partículas en el sistema de referencia en que la cantidad de movimiento total es nula, sistema del centro de inercia) se escribe como (el subíndice 0 se refiere al sistema centro de inercia)

$$mc^2 = E_{01} + E_{02}, \quad (5.23)$$

mientras que la conservación de la cantidad de movimiento conduce a

$$\mathbf{p}_{01} = -\mathbf{p}_{02}. \quad (5.24)$$

De esta última y la relación (5.16) tenemos que

$$E_{01}^2 - m_1^2 c^4 = E_{02}^2 - m_2^2 c^4,$$

que, escrita como

$$E_{01}^2 - E_{02}^2 = (E_{01} - E_{02})(E_{01} + E_{02}) = m_1^2 c^4 - m_2^2 c^4,$$

junto con (5.23) nos permite obtener inmediatamente

$$E_{01} - E_{02} = \frac{m_1^2 - m_2^2}{m} c^2,$$

y finalmente, entre ésta y (5.23),

$$\begin{aligned} E_{01} &= \frac{m^2 + m_1^2 - m_2^2}{2m} c^2, \\ E_{02} &= \frac{m^2 - m_1^2 + m_2^2}{2m} c^2. \end{aligned}$$

De (5.16) y (5.15) determinamos fácilmente la magnitud de los momentos y velocidades, respectivamente:

$$\begin{aligned} |\mathbf{p}_{01,2}|^2 &= E_{01,2}^2/c^2 - m_{1,2}^2 c^2, \\ |\mathbf{u}_{01,2}|^2 &= c^2 - \frac{m_{1,2}^2 c^4}{E_{1,2}}, \end{aligned}$$

que sabemos tienen la misma dirección y sentidos opuestos por (5.24).

Si la partícula original se mueve en el laboratorio con velocidad \mathbf{V} , lo calculado hasta aquí vale en un sistema de referencia que se mueve con tal velocidad, el sistema centro de inercia, de manera que las magnitudes luego del choque en el sistema de laboratorio pueden ser calculadas haciendo una transformación de Lorentz con velocidad $-\mathbf{V}$ (ecuaciones (5.18) y (5.19):

$$\begin{aligned} E_{1,2} &= \gamma (E_{01,2} + \mathbf{V} \cdot \mathbf{p}_{01,2}), \\ \mathbf{p}_{1,2} &= \mathbf{p}_{01,2} + \gamma \frac{E_{01,2}}{c^2} \mathbf{V} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{p}_{01,2}}{V^2} \mathbf{V}. \end{aligned}$$

Se requiere entonces conocer el ángulo entre \mathbf{V} y \mathbf{p}_{01} en el sistema centro de inercia (por supuesto $\mathbf{V} \cdot \mathbf{p}_{02} = -\mathbf{V} \cdot \mathbf{p}_{01}$), que es una variable que no puede ser predicha sin conocer la dinámica interna de la partícula y se debe tomar como un dato de la observación; llamémoslo χ , de manera que $\mathbf{V} \cdot \mathbf{p}_{01} = |\mathbf{p}_{01}| V \cos \chi$. Tenemos así (paralelo y perpendicular se refieren a la dirección de \mathbf{V})

$$\begin{aligned} E_1 &= \gamma (E_{01} + |\mathbf{p}_{01}| V \cos \chi), \\ E_2 &= \gamma (E_{02} - |\mathbf{p}_{01}| V \cos \chi), \\ p_{1\parallel} &= \gamma \left(|\mathbf{p}_{01}| \cos \chi + \frac{E_{01} V}{c^2} \right), \\ p_{2\parallel} &= \gamma \left(-|\mathbf{p}_{01}| \cos \chi + \frac{E_{02} V}{c^2} \right), \\ p_{1,2\perp} &= p_{01,2\perp} = |\mathbf{p}_{01}| \sin \chi. \end{aligned}$$

Desde ya, $\mathbf{p}_{2\perp} = -\mathbf{p}_{1\perp}$ al no haber componente del momento transversal a \mathbf{V} .

El valor de χ puede ser entonces obtenido de la medición de la energía de alguna de las partículas resultantes, lo que permite evaluar todo lo demás. La medición de tal energía puede hacerse a través del ángulo en que es eyectada la partícula en el sistema de laboratorio respecto de \mathbf{V} . Sea ésta, por ejemplo, la partícula 1, de manera que $\mathbf{V} \cdot \mathbf{p}_1 = |\mathbf{p}_1| V \cos \theta_1$. Usando la transformación de Lorentz inversa a la que usamos en los cálculos anteriores es

$$\begin{aligned} E_{01} &= \gamma (E_1 - \mathbf{V} \cdot \mathbf{p}_1) \\ &= \gamma (E_1 - |\mathbf{p}_1| V \cos \theta_1) \\ &= \gamma \left(E_1 - \sqrt{E_1^2/c^2 - m_1^2 c^2} V \cos \theta_1 \right), \end{aligned}$$

de donde puede obtenerse una expresión para E_1 en función de θ_1 , y completar así la evaluación de todas las magnitudes.

5.2.3. Choque de partículas

Como vimos en el caso de la desintegración de partículas es conveniente resolver el problema en el sistema centro de inercia en el que la cantidad total de movimiento es nula. Consideremos en tal sistema dos partículas con valores iniciales de masa y momento m_1, m_2 , y $\mathbf{p}_{01} = -\mathbf{p}_{02}$, que por lo tanto tendrán energías

$$\begin{aligned} E_{01} &= \sqrt{m_1^2 c^4 + |\mathbf{p}_{01}|^2 c^2}, \\ E_{02} &= \sqrt{m_2^2 c^4 + |\mathbf{p}_{01}|^2 c^2}, \end{aligned}$$

que interactúan en un choque para dar dos partículas de masas y momentos m'_1, m'_2 , y $\mathbf{p}'_{01} = -\mathbf{p}'_{02}$. Si las masas iniciales son iguales a las finales, la estructura interna de las partículas no se habrá afectado y por lo tanto el choque es elástico. La conservación del impulso está expresada por ser $\mathbf{p}'_{01} = -\mathbf{p}'_{02}$, pero en general la dirección final estará rotada en un ángulo χ respecto de la inicial.

La conservación de la energía impone que

$$\begin{aligned} E_{01} + E_{02} &= E'_{01} + E'_{02} \\ &= \sqrt{m_1^2 c^2 + |\mathbf{p}'_{01}|^2} + \sqrt{m_2^2 c^2 + |\mathbf{p}'_{01}|^2}, \end{aligned}$$

que puede resolverse para obtener $|\mathbf{p}'_{01}|$ dadas las otras variables, que deben ser conocidas por las condiciones iniciales y por la observación del tipo de partículas resultantes (en el caso de choque elástico es simplemente $|\mathbf{p}'_{01}| = |\mathbf{p}_{01}|$). Se tienen así los momentos finales en términos de su módulo y dirección observada a través del ángulo χ .

En el sistema de laboratorio supondremos que la partícula 2 se encuentra inicialmente en reposo ($\mathbf{p}_2 = \mathbf{0}$), y que la partícula 1 tiene momento inicial \mathbf{p}_1 , con lo que será

$$E_1 = \sqrt{m_1^2 c^4 + |\mathbf{p}_1|^2 c^2}; \quad E_2 = m_2 c^2.$$

Para encontrar cuál es el sistema centro de inercia a partir de estos datos basta considerar que si éste se mueve con velocidad \mathbf{V} respecto del sistema de laboratorio debe ser

$$\begin{aligned} \mathbf{p}_{01} &= \mathbf{p}_1 - \gamma \frac{E_1}{c^2} \mathbf{V} + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{p}_1}{V^2} \mathbf{V} \\ &= -\mathbf{p}_{02} = \gamma \frac{E_2}{c^2} \mathbf{V}. \end{aligned}$$

Vemos inmediatamente que \mathbf{V} es paralelo a \mathbf{p}_1 (y a \mathbf{p}_{01} y \mathbf{p}_{02}) con lo que obtenemos fácilmente

$$\begin{aligned}\mathbf{V} &= \frac{c^2}{E_1 + E_2} \mathbf{p}_1, \\ \mathbf{p}_{01} &= -\mathbf{p}_{02} = \gamma \frac{E_2}{E_1 + E_2} \mathbf{p}_1.\end{aligned}$$

Para obtener las magnitudes finales (luego del choque) en el sistema de laboratorio hacemos la transformación a éste desde el sistema centro de inercia (transformación de Lorentz con velocidad $-\mathbf{V}$), exactamente como hicimos en el caso de desintegración de una partícula, y obtenemos expresiones del todo análogas, (paralelo y perpendicular se refieren a la dirección de \mathbf{V} o \mathbf{p}_1)

$$\begin{aligned}E'_1 &= \gamma (E'_{01} + |\mathbf{p}'_{01}| V \cos \chi), \\ E'_2 &= \gamma (E'_{02} - |\mathbf{p}'_{01}| V \cos \chi), \\ p'_{1\parallel} &= \gamma \left(|\mathbf{p}'_{01}| \cos \chi + \frac{E'_{01} V}{c^2} \right), \\ p'_{2\parallel} &= \gamma \left(-|\mathbf{p}'_{01}| \cos \chi + \frac{E'_{02} V}{c^2} \right), \\ p'_{1,2\perp} &= p_{1,2\perp} = |\mathbf{p}'_{01}| \sin \chi.\end{aligned}$$

Análogamente al caso de la desintegración, el ángulo χ puede obtenerse midiendo el ángulo θ en que es eyectada una de las partículas, por ejemplo la 1, en el sistema de laboratorio respecto de \mathbf{p}_1 para tener la relación que permite calcular E'_1 conocido θ_1 :

$$E'_{01} = \gamma \left(E'_1 - \sqrt{E_1'^2/c^2 - m_1'^2 c^2} V \cos \theta_1 \right).$$

Capítulo 6

Fuerzas centrales

6.1. Problema de Kepler

Son las derivadas de un potencial de la forma $V(|\mathbf{x}_2 - \mathbf{x}_1|) \equiv V(r)$, con \mathbf{x}_1 y \mathbf{x}_2 las posiciones de las partículas interactuantes. Así, la fuerza \mathbf{f}_{12} sobre la partícula 1 debida a la 2 vale

$$\mathbf{f}_{12} = -\frac{\partial V}{\partial \mathbf{x}_1} = -\frac{\partial V}{\partial r} \frac{\partial r}{\partial \mathbf{x}_1} = -\frac{\partial V}{\partial r} \frac{(\mathbf{x}_1 - \mathbf{x}_2)}{r},$$

con dirección a lo largo de la recta que las une.

Cuando el sistema tiene sólo dos partículas es conveniente, en lugar de \mathbf{x}_1 y \mathbf{x}_2 , usar las coordenadas

$$\begin{aligned}\mathbf{X}_{CM} &= \frac{m_1 \mathbf{x}_1 + m_2 \mathbf{x}_2}{m_1 + m_2}, \\ \mathbf{r} &= \mathbf{x}_2 - \mathbf{x}_1,\end{aligned}$$

en términos de las cuales las coordenadas originales se escriben

$$\begin{aligned}\mathbf{x}_1 &= \mathbf{X}_{CM} - \frac{m_2 \mathbf{r}}{m_1 + m_2}, \\ \mathbf{x}_2 &= \mathbf{X}_{CM} + \frac{m_1 \mathbf{r}}{m_1 + m_2}.\end{aligned}$$

Si además no existe fuerza externa, el centro de masas se moverá con velocidad constante y es entonces posible describir el movimiento respecto de un sistema inercial en el que la velocidad del centro de masas es nula. Considerando de ahora en más que estamos en tal sistema de referencia, si el sistema de coordenadas se elige con origen en el centro de masas es $\mathbf{X}_{CM} = 0$,

con lo que escribimos ($M \equiv m_1 + m_2$)

$$\mathbf{x}_1 = -\frac{m_2}{M} \mathbf{r}, \quad (6.1a)$$

$$\mathbf{x}_2 = \frac{m_1}{M} \mathbf{r}. \quad (6.1b)$$

El momento angular es constante y vale, respecto del origen (el CM),

$$\begin{aligned} \mathbf{L} &= m_1 \mathbf{x}_1 \times \dot{\mathbf{x}}_1 + m_2 \mathbf{x}_2 \times \dot{\mathbf{x}}_2 \\ &= \left[m_1 \left(\frac{m_2}{M} \right)^2 + m_2 \left(\frac{m_1}{M} \right)^2 \right] \mathbf{r} \times \dot{\mathbf{r}} \\ &= \frac{m_1 m_2}{M} \mathbf{r} \times \dot{\mathbf{r}} \equiv \mu \mathbf{r} \times \dot{\mathbf{r}}, \end{aligned}$$

donde se ha definido la *masa reducida* μ

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2}. \quad (6.2)$$

Si en un instante dado t se tienen los vectores \mathbf{r} y $\dot{\mathbf{r}}$, como el incremento de \mathbf{r} en un dt es justamente $\dot{\mathbf{r}}dt$, el \mathbf{r} en $t+dt$ sigue permaneciendo en el plano definido por los originales \mathbf{r} y $\dot{\mathbf{r}}$; por otro lado, como el vector constante \mathbf{L} es siempre perpendicular a \mathbf{r} y $\dot{\mathbf{r}}$, el nuevo $\dot{\mathbf{r}}$ debe seguir estando en el plano original, y todo el movimiento está entonces contenido en tal plano.

Otra forma de ver esto es escribiendo el Lagrangiano. Como

$$\begin{aligned} T &= \frac{1}{2} m_1 |\dot{\mathbf{x}}_1|^2 + \frac{1}{2} m_2 |\dot{\mathbf{x}}_2|^2 \\ &= \frac{1}{2} \left[m_1 \left(\frac{m_2}{M} \right)^2 + m_2 \left(\frac{m_1}{M} \right)^2 \right] |\dot{\mathbf{r}}|^2 \\ &= \frac{1}{2} \mu |\dot{\mathbf{r}}|^2, \end{aligned} \quad (6.3)$$

se tiene

$$L = \frac{1}{2} \mu |\dot{\mathbf{r}}|^2 - V(|\mathbf{r}|).$$

Usando coordenadas esféricas es

$$\begin{aligned} \mathbf{r} &= r \mathbf{e}_r, \\ \dot{\mathbf{r}} &= \dot{r} \mathbf{e}_r + r \dot{\theta} \mathbf{e}_\theta + r \sin \theta \dot{\varphi} \mathbf{e}_\varphi, \end{aligned}$$

con lo que

$$L = \frac{1}{2} \mu \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \right) - V(r),$$

y las ecuaciones de movimiento son

$$\begin{aligned}\mu\ddot{r} - \mu r \left(\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2 \right) + \frac{dV}{dr} &= 0, \\ \frac{d}{dt} \left(\mu r^2 \dot{\theta} \right) - \mu r^2 \sin \theta \cos \theta \dot{\varphi}^2 &= 0, \\ \mu r^2 \sin^2 \theta \dot{\varphi} &= cte.\end{aligned}$$

De la segunda se ve fácilmente que si $\dot{\theta} = 0$ en $\theta = 0$ ó $\pi/2$ es $\ddot{\theta} = 0$ y θ debe entonces permanecer constante. Si elegimos $\theta = \pi/2$ donde $\dot{\theta} = 0$ tenemos entonces un movimiento plano y las ecuaciones se reducen a

$$\mu\ddot{r} - \mu r \dot{\varphi}^2 + \frac{dV}{dr} = 0, \quad (6.4a)$$

$$\mu r^2 \dot{\varphi} = cte. \quad (6.4b)$$

Las coordenadas r y φ pueden interpretarse como coordenadas polares en el plano (x,y) en las que

$$\mathbf{r} = r \mathbf{e}_r, \quad (6.5a)$$

$$\dot{\mathbf{r}} = \dot{r} \mathbf{e}_r + r \dot{\varphi} \mathbf{e}_\varphi, \quad (6.5b)$$

con lo que

$$\mathbf{L} = \mu \mathbf{r} \times \dot{\mathbf{r}} = \mu r^2 \dot{\varphi} \mathbf{e}_z,$$

de manera que la segunda de las (6.4) indica precisamente la constancia de \mathbf{L} y se escribe

$$\mu r^2 \dot{\varphi} = L_z. \quad (6.6)$$

Dado que en un dt el vector \mathbf{r} barre un área $r^2 \dot{\varphi} dt / 2$, podemos decir de (6.6) que la *velocidad areolar* $\dot{A} = r^2 \dot{\varphi} / 2$ es constante, lo que constituye la segunda ley de Kepler para el movimiento de los planetas alrededor del sol. Nótese que dicha ley vale para cualquier tipo de potencial central, no sólo para el caso gravitatorio. Si usamos (6.6) para despejar $\dot{\varphi}$ y reemplazamos este valor en la primera de las (6.4) obtenemos la ecuación de movimiento de la coordenada r

$$\mu\ddot{r} - \frac{L_z^2}{\mu r^3} + \frac{dV}{dr} = 0,$$

que podemos reescribir

$$\mu\ddot{r} + \frac{dV_{ef}}{dr} = 0, \quad (6.7)$$

donde el *potencial efectivo* V_{ef} está definido por

$$V_{ef}(r) \equiv V(r) + \frac{L_z^2}{2\mu r^2}, \quad (6.8)$$

el segundo término se denomina *potencial centrífugo*. La ecuación (6.7) representa un movimiento unidimensional equivalente en la variable unidimensional r relacionando la “aceleración” de ésta, \ddot{r} , con su valor instantáneo r . Para resolver (6.7) se puede usar entonces la técnica de relacionar la aceleración con la “velocidad” \dot{r} como

$$\ddot{r} = \frac{d\dot{r}}{dt} = \frac{d\dot{r}}{dr} \frac{dr}{dt} = \frac{d}{dr} \left(\frac{1}{2} \dot{r}^2 \right),$$

con lo que la (6.7) se reescribe

$$\frac{d}{dr} \left(\frac{1}{2} \mu \dot{r}^2 + V_{ef} \right) = 0,$$

o sea

$$\frac{1}{2} \mu \dot{r}^2 + V_{ef} = cte. \quad (6.9)$$

Usando las (6.3) y las (6.5) la energía del sistema (que constante por ser L independiente del tiempo, y que corresponde a la energía mecánica, por ser T bicuadrática en las velocidades generalizadas) se escribe

$$E = T + V = \frac{1}{2} \mu (\dot{r}^2 + r^2 \dot{\varphi}^2) + V,$$

que, con (6.6) y (6.8), se escribe como

$$E = \frac{1}{2} \mu \dot{r}^2 + V_{ef}. \quad (6.10)$$

Esta es entonces la primera integral de (6.7) y nos permite escribir

$$\frac{dr}{dt} = \pm \sqrt{\frac{2}{\mu} [E - V_{ef}(r)]}, \quad (6.11)$$

para tener

$$dt = \pm \frac{dr}{\sqrt{\frac{2}{\mu} [E - V_{ef}(r)]}},$$

que puede resolverse por una cuadratura. Téngase en cuenta que si al avanzar el tiempo r crece se debe usar el signo $+$, mientras que si r decrece corresponde el signo $-$ en la integración. Así, partiendo de un t_0 con un r_0 , para cada par de valores constantes E y L_z , se obtiene la función $t(r)$ que al invertirse da $r(t)$. Con ésta se puede integrar (6.6) escrita como

$$\frac{d\varphi}{dt} = \frac{L_z}{\mu [r(t)]^2}, \quad (6.12)$$

para obtener $\varphi(t)$, lo que resuelve el problema.

Una alternativa es obtener directamente la relación $r(\varphi)$ que nos da la forma geométrica de la trayectoria, eliminando el tiempo del problema. Para esto basta escribir

$$\frac{dr}{dt} = \frac{dr}{d\varphi} \frac{d\varphi}{dt},$$

y usar (6.11) y (6.12) para escribir

$$\begin{aligned} \frac{dr}{d\varphi} &= \pm \frac{\mu r^2}{L_z} \sqrt{\frac{2}{\mu} [E - V_{ef}(r)]} \\ &= \pm r^2 \sqrt{\frac{2\mu}{L_z^2} [E - V(r)] - \frac{1}{r^2}}, \end{aligned}$$

con lo que se obtiene

$$\frac{d\varphi}{dr} = \pm \frac{1}{r^2 \sqrt{\frac{2\mu}{L_z^2} [E - V(r)] - \frac{1}{r^2}}}, \quad (6.13)$$

que puede integrarse para obtener $\varphi(r)$ y, por inversión, $r(\varphi)$. En particular, vemos de (6.13) que la trayectoria será simétrica respecto del punto donde hay cambio de signo; esto es, en los puntos donde r alcanza un valor extremo, máximo o mínimo. El punto de máximo acercamiento se denomina periapsis (perihelio en el caso del sol) y el de máximo alejamiento, cuando es finito, apoapsis (aphelio o afelio para el sol).

Veamos como ejemplo importante el caso del potencial gravitatorio. En este caso es

$$V(r) = -\frac{G m_1 m_2}{r},$$

donde $G = 6,672 \times 10^{-11} N m^2 kg^{-2}$ es la constante de la gravitación de Newton. Llamando $\alpha \equiv G m_1 m_2$ podemos escribir (6.11) para este caso como

$$\frac{dr}{dt} = \pm \sqrt{\frac{2}{\mu} \left[E + \frac{\alpha}{r} - \frac{L_z^2}{2\mu r^2} \right]},$$

del que podemos decir que el rango de valores posibles de r estará limitado a aquellos para los cuales el argumento de la raíz es no negativo:

$$E \geq -\frac{\alpha}{r} + \frac{L_z^2}{2\mu r^2}. \quad (6.14)$$

La expresión de la derecha tiene un valor mínimo en

$$r_0 = \frac{L_z^2}{\mu\alpha}, \quad (6.15)$$

que vale

$$-\frac{\mu\alpha^2}{2L_z^2},$$

por lo que (6.14) implica que

$$E \geq -\frac{\mu\alpha^2}{2L_z^2}. \quad (6.16)$$

En estas condiciones el rango de r está determinado por los ceros de la expresión en (6.14). Si $E \geq 0$ es claro de (6.14) que r puede llegar a ∞ y existe entonces sólo un valor mínimo de r , $r_{\text{mín}}$ dado por

$$r_{\text{mín}} = \frac{\alpha}{2E} \left[\sqrt{1 + \frac{2L_z^2 E}{\mu\alpha^2}} - 1 \right]; \quad E > 0 \quad (6.17)$$

$$r_{\text{mín}} = \frac{L_z^2}{2\mu\alpha}; \quad E = 0.$$

Cuando $E < 0$ (estado ligado) existe un rango superior también, con lo que el movimiento está restringido por los valores (periapsis y apoapsis, respectivamente)

$$r_{\text{mín}} = \frac{\alpha}{2|E|} \left[1 - \sqrt{1 - \frac{2L_z^2 |E|}{\mu\alpha^2}} \right], \quad (6.18a)$$

$$r_{\text{máx}} = \frac{\alpha}{2|E|} \left[1 + \sqrt{1 - \frac{2L_z^2 |E|}{\mu\alpha^2}} \right]. \quad (6.18b)$$

Obtengamos ahora la órbita correspondiente a $E < 0$. Escribimos (6.13) para este caso como

$$\frac{d\varphi}{dr} = \pm \frac{1}{r^2 \sqrt{-\frac{2\mu|E|}{L_z^2} + \frac{2\mu\alpha}{L_z^2 r} - \frac{1}{r^2}}}$$

Tomemos el valor $\varphi = 0$ correspondiendo a $r = r_{\text{mín}}$ y consideremos que φ crece al variar r con lo que tendremos

$$\varphi = \int_{r_{\text{mín}}}^r \frac{dr}{r^2 \sqrt{-\frac{2\mu|E|}{L_z^2} + \frac{2\mu\alpha}{L_z^2 r} - \frac{1}{r^2}}}, \quad (6.19)$$

para $r \leq r_{\text{máx}}$. Alcanzado $r_{\text{máx}}$ debe seguir integrándose entre $r_{\text{máx}}$ y r , usando ahora el signo menos hasta alcanzar otra vez $r_{\text{mín}}$, etc.

La integral puede realizarse explícitamente haciendo en (6.19) el cambio de variable $u = r^{-1}$:

$$\varphi = \int_u^{u_{\min}} \frac{du}{\sqrt{-\frac{2\mu|E|}{L_z^2} + \frac{2\mu\alpha}{L_z^2}u - u^2}},$$

con $u_{\min} = r_{\min}^{-1}$. Usando la integral de tablas

$$\int \frac{dx}{\sqrt{ax^2 + bx + c}} = -\frac{1}{\sqrt{-a}} \arccos \left[-\frac{2ax + b}{\sqrt{b^2 - 4ac}} \right], \quad (6.20)$$

el resultado es

$$\varphi = \arccos \left[\left(\frac{L_z^2}{\mu\alpha} u - 1 \right) \left(\sqrt{1 - \frac{2L_z^2|E|}{\mu\alpha^2}} \right)^{-1} \right],$$

que se invierte inmediatamente para dar (volviendo a la variable r y usando la definición (6.15) de r_0)

$$\frac{1}{r} = \frac{1}{r_0} (1 + e \cos \varphi), \quad (6.21)$$

donde se ha definido la *excentricidad* $0 < e < 1$

$$e \equiv \sqrt{1 + \frac{2L_z^2 E}{\mu\alpha^2}}. \quad (6.22)$$

La ecuación (6.21) describe una elipse de excentricidad e y foco en el origen de coordenadas, de semiejes mayor y menor dados respectivamente por

$$\begin{aligned} a &= \frac{r_0}{(1 - e^2)} = \frac{\alpha}{2|E|}, \\ b &= \frac{r_0}{\sqrt{1 - e^2}} = \frac{L_z}{\sqrt{2\mu|E|}}. \end{aligned}$$

Si se aplica esto a la órbita de los planetas alrededor del sol tomando en las fórmulas anteriores m_1 la masa del sol y m_2 la del planeta considerado, vemos de (6.1) que cada cuerpo describe también una elipse, ambas con foco en el centro de masas. Si se considera además que la masa del sol es mucho mayor que la de los planetas (para la tierra el cociente de masas es $\simeq 3,5 \times 10^5$) se ve entonces que la elipse descrita por el sol es de tamaño despreciable y el

sol se encuentra entonces prácticamente en el foco de la elipse descripta por cada planeta. Ésta es la primera ley de Kepler.

La tercera ley de Kepler resulta de la ley de las áreas (6.6) escrita como

$$\frac{dA}{dt} = \frac{1}{2}r^2\dot{\varphi} = \frac{L_z}{2\mu},$$

que integrada en un período da

$$\begin{aligned} T &= \frac{2\mu A}{L_z} = \frac{2\mu\pi ab}{L_z} \\ &= \frac{\pi\alpha}{2} \frac{\sqrt{2\mu}}{|E|^{3/2}} = 2\pi a^{3/2} \sqrt{\frac{\mu}{\alpha}}. \end{aligned}$$

La última igualdad dice que el período de revolución de cada cuerpo alrededor del centro de masas es proporcional a la potencia $3/2$ del tamaño de la órbita. Aplicada al sistema solar ésta es la tercera ley de Kepler que dice que el cuadrado del período de revolución de cada planeta es proporcional al cubo de su distancia al sol. Nótese que en el caso de ser $m_1 \gg m_2$ es $\mu \simeq m_2$ y, por lo tanto, $\alpha/\mu \simeq Gm_1$, por lo que la constante de proporcionalidad es la misma para todos los planetas.

Si $E \geq 0$ las (6.21) y (6.22) siguen valiendo, sólo que para $E = 0$ es $e = 1$, con lo que (6.21) representa una parábola con distancia mínima al origen $r_0/2$ y en la que el infinito se alcanza con velocidad nula. Cuando $E > 0$ es $e > 1$ que corresponde a una hipérbola en la que la partícula se acerca al origen hasta $r_0/(1+e)$ y alcanza el infinito con velocidad no nula.

6.2. Choque elástico

Consideramos ahora el choque elástico ($E = cte$) entre dos partículas con masas m_1 y m_2 que interactúan a través de un potencial central. La idea es considerar que en el sistema centro de masas la energía mecánica $E_{CM} > 0$ (desde ya, esto valdrá entonces para cualquier otro sistema) de manera que, aun para un potencial atractivo, las partículas no están ligadas. De esta manera, consideraremos sus estados antes y después de la interacción; esto es, cuando estén muy separadas de manera que la energía potencial entre ellas pueda considerarse despreciable frente a las correspondientes energías cinéticas (se toma el cero de la energía potencial correspondiendo a distancia infinita entre partículas).

Como la energía mecánica se conserva, y antes y después de la interacción se puede considerar que ésta es sólo cinética, escribimos

$$\frac{1}{2}m_1 |\mathbf{v}_1|^2 + \frac{1}{2}m_2 |\mathbf{v}_2|^2 = \frac{1}{2}m_1 |\mathbf{v}'_1|^2 + \frac{1}{2}m_2 |\mathbf{v}'_2|^2, \quad (6.23)$$

donde las velocidades sin primar corresponden a antes del choque y las primadas a luego del choque. Por otro lado, como no existen fuerzas externas se conserva también la cantidad de movimiento total

$$m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 = m_1 \mathbf{v}'_1 + m_2 \mathbf{v}'_2. \quad (6.24)$$

La conservación del momento angular total indica, como se vio más arriba, que el movimiento de ambas partículas puede siempre considerarse contenido en un plano (si las velocidades iniciales no son coplanares puede siempre pasarse a un sistema, a infinitos en realidad, en el que sean coplanares, y así permanecerán).

Es conveniente describir el movimiento desde el sistema centro de masas; distinguiremos entonces el llamado sistema de laboratorio como aquél en el que las velocidades iniciales son coplanares y valen \mathbf{v}_1 y \mathbf{v}_2 , y el sistema centro de masas que distinguiremos con el subíndice 0 (en lugar de CM). Así, las velocidades iniciales en el sistema CM valen

$$\mathbf{v}_{01} = \mathbf{v}_1 - \mathbf{V}_{CM} = \mathbf{v}_1 - \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2} = \frac{m_2}{m_1 + m_2} \mathbf{v}, \quad (6.25a)$$

$$\mathbf{v}_{02} = \mathbf{v}_2 - \mathbf{V}_{CM} = \mathbf{v}_2 - \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2} = -\frac{m_1}{m_1 + m_2} \mathbf{v}, \quad (6.25b)$$

donde hemos definido la velocidad de la partícula 1 relativa a la de la partícula 2: $\mathbf{v} \equiv \mathbf{v}_1 - \mathbf{v}_2$.

La conservación de la cantidad de movimiento en el CM es claramente

$$m_1 \mathbf{v}_{01} + m_2 \mathbf{v}_{02} = m_1 \mathbf{v}'_{01} + m_2 \mathbf{v}'_{02} = \mathbf{0},$$

de donde

$$\mathbf{v}_{02} = -\frac{m_1}{m_2} \mathbf{v}_{01}, \quad (6.26a)$$

$$\mathbf{v}'_{02} = -\frac{m_1}{m_2} \mathbf{v}'_{01}, \quad (6.26b)$$

que al usarse en la ecuación de conservación de la energía en el CM da

$$\begin{aligned} \frac{1}{2} m_1 |\mathbf{v}_{01}|^2 + \frac{1}{2} m_2 |\mathbf{v}_{02}|^2 &= \frac{1}{2} m_1 |\mathbf{v}'_{01}|^2 + \frac{1}{2} m_2 |\mathbf{v}'_{02}|^2 = \\ \frac{1}{2} m_1 |\mathbf{v}_{01}|^2 \left(1 + \frac{m_1}{m_2}\right) &= \frac{1}{2} m_1 |\mathbf{v}'_{01}|^2 \left(1 + \frac{m_1}{m_2}\right), \end{aligned}$$

de donde se deduce que

$$|\mathbf{v}_{01}| = |\mathbf{v}'_{01}|, \quad (6.27)$$

esto es, en el sistema CM la interacción sólo cambia la dirección de cada velocidad, pero no su módulo.

Llamemos entonces \mathbf{n}_0 al versor que indica la dirección y sentido de la velocidad de la partícula 1 después del choque en el CM :

$$\mathbf{n}_0 \equiv \frac{\mathbf{v}'_{01}}{|\mathbf{v}'_{01}|},$$

que nos permite escribir, usando las (6.25) aplicadas a las velocidades luego del choque, y la segunda de las (6.26),

$$\begin{aligned}\mathbf{v}'_{01} &= \frac{m_2}{m_1 + m_2} |\mathbf{v}| \mathbf{n}_0, \\ \mathbf{v}'_{02} &= -\frac{m_1}{m_1 + m_2} |\mathbf{v}| \mathbf{n}_0,\end{aligned}$$

que nos permite finalmente escribir las velocidades luego del choque en el sistema de laboratorio (SL)

$$\mathbf{v}'_1 = \frac{m_2}{m_1 + m_2} |\mathbf{v}| \mathbf{n}_0 + \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2}, \quad (6.28a)$$

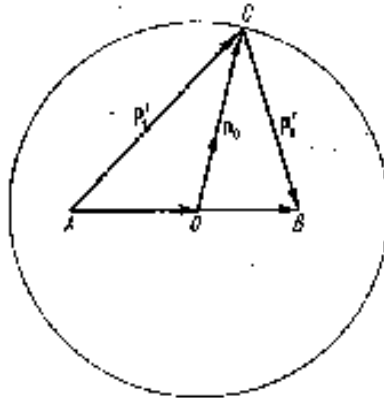
$$\mathbf{v}'_2 = -\frac{m_1}{m_1 + m_2} |\mathbf{v}| \mathbf{n}_0 + \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2}. \quad (6.28b)$$

Nótese que todas las magnitudes son conocidas de antes del choque, salvo \mathbf{n}_0 que es determinado por la interacción en sí.

Las expresiones (6.28) son más sencillas escritas en términos de las cantidades de movimiento; multiplicando la primera por m_1 y la segunda por m_2 , se obtiene inmediatamente (μ es la masa reducida, ver (6.2))

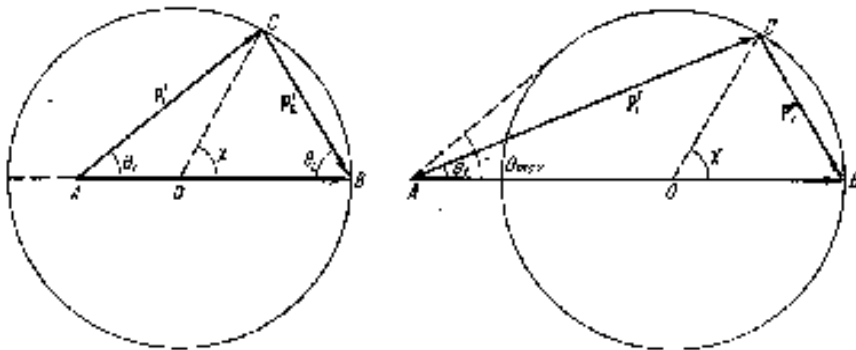
$$\begin{aligned}\mathbf{p}'_1 &= \mu |\mathbf{v}| \mathbf{n}_0 + \frac{\mu}{m_2} (\mathbf{p}_1 + \mathbf{p}_2), \\ \mathbf{p}'_2 &= -\mu |\mathbf{v}| \mathbf{n}_0 + \frac{\mu}{m_1} (\mathbf{p}_1 + \mathbf{p}_2).\end{aligned}$$

Es muy conveniente representar estas expresiones gráficamente.



Donde $\vec{AO} = \mu(\mathbf{p}_1 + \mathbf{p}_2)/m_2$, y $\vec{OB} = \mu(\mathbf{p}_1 + \mathbf{p}_2)/m_1$. La circunferencia tiene radio $\mu|\mathbf{v}|$ por lo que $\vec{OC} = \mu|\mathbf{v}| \mathbf{n}_0$. Nótese que el vector $\vec{AB} = \vec{AO} + \vec{OB} = \mathbf{p}_1 + \mathbf{p}_2$, que es la cantidad de movimiento total en el SL .

De ahora en más consideraremos sólo el caso en que la partícula 2 se encuentra en reposo antes del choque en el SL ; esto es $\mathbf{p}_2 = \mathbf{0}$. En tal caso es $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{v}_1$ y, por lo tanto, $\vec{OB} = \mu\mathbf{v}_1 = \mu\mathbf{v}$; esto es, el punto B se encuentra sobre la circunferencia. Por otro lado, $\vec{AO} = (m_1/m_2)\mu\mathbf{v}$, por lo que A estará dentro de la circunferencia si $m_1 < m_2$, y fuera si $m_1 > m_2$, como se representa en la figura.



En estas figuras θ_1 y θ_2 representan los ángulos que forman las velocidades (en el SL) de las respectivas partículas respecto de la dirección horizontal correspondiente a la del vector \vec{AB} , cantidad de movimiento total en el SL .

El ángulo χ es el ángulo que forma \mathbf{n}_0 con \overrightarrow{AB} ; esto es, el ángulo que forma la velocidad de la partícula 1 en el sistema CM con \overrightarrow{AB} .

Vemos en particular que cuando $m_1 > m_2$ existe un ángulo de desviación máxima $\theta_{\text{máx}}$ de la partícula 1, que corresponde al formado por la tangente a la circunferencia que pasa por A , que es fácilmente calculable como

$$\sin \theta_{\text{máx}} = \frac{\mu |\mathbf{v}|}{\frac{m_1}{m_2} \mu |\mathbf{v}|} = \frac{m_2}{m_1}.$$

Para relacionar los ángulos $\theta_{1,2}$ y χ basta ver que, por estar B y C sobre la circunferencia, el triángulo OCB nos dice que $2\theta_2 + \chi = \pi$; o sea,

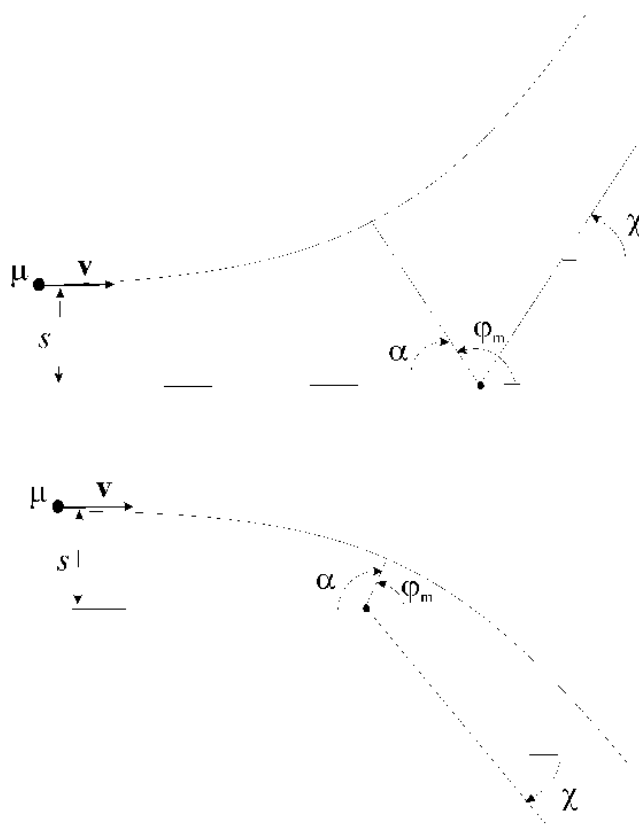
$$\theta_2 = \frac{\pi - \chi}{2}. \quad (6.29)$$

Por otro lado, podemos ver que

$$\tan \theta_1 = \frac{\mu |\mathbf{v}| \sin \chi}{AO + \mu |\mathbf{v}| \cos \chi} = \frac{\mu |\mathbf{v}| \sin \chi}{\frac{m_1}{m_2} \mu |\mathbf{v}| + \mu |\mathbf{v}| \cos \chi} = \frac{\sin \chi}{\frac{m_1}{m_2} + \cos \chi}. \quad (6.30)$$

6.3. Dispersión (Scattering)

Vimos en la sección anterior que para definir completamente el choque debe determinarse el versor \mathbf{n}_0 o, equivalentemente, el ángulo χ formado por \mathbf{v}'_{01} con la dirección de la cantidad de movimiento en el SL . Seguimos considerando que $\mathbf{v}_2 = \mathbf{0}$, de manera que la dirección de referencia es la dada por la velocidad \mathbf{v}_1 . Asimismo, el vector \mathbf{v}'_{01} es (anti)paralelo al \mathbf{v}'_{02} (por la segunda de las (6.26)) de manera que la velocidad relativa después del choque $\mathbf{v}' = \mathbf{v}'_{01} - \mathbf{v}'_{02}$ es paralela a \mathbf{v}'_{01} y tiene por lo tanto el mismo χ . De esta manera, para calcular χ basta resolver el problema equivalente, visto en la primera sección, de una partícula de masa μ (la masa reducida) sometida al potencial de interacción, pero respecto a un centro fijo (el CM). La velocidad de esta partícula corresponde a la relativa entre las partículas originales, y vale inicialmente $\mathbf{v} = \mathbf{v}_1$ (siempre consideramos $\mathbf{v}_2 = \mathbf{0}$). Tomamos como dirección de referencia (horizontal) la de esta \mathbf{v} que corresponde a distancia muy grande al origen antes de la interacción. Para fijar el origen decimos que si la \mathbf{v} se mantuviese constante la partícula μ pasaría a distancia s de éste; s es denominado parámetro de impacto.



En la figura se representa la situación para un potencial repulsivo y uno atractivo. El φ_m es el ángulo correspondiente a la mínima distancia al centro (atractor o repulsor). Como la trayectoria es simétrica respecto del punto de máximo acercamiento, el ángulo α es $\alpha = \varphi_m - \chi$ para el potencial repulsivo, y $\alpha = \varphi_m + \chi$ para el atractivo. Como en ambos casos es $\alpha + \varphi_m = \pi$, resulta $\chi = 2\varphi_m - \pi$ en el caso repulsivo y $\chi = \pi - 2\varphi_m$ en el atractivo. Tenemos entonces en general

$$\chi = 2\varphi_m - \pi, \quad (6.31)$$

que resultará positivo para repulsión y negativo para atracción en las condiciones de la figura.

El valor de φ_m se obtiene fácilmente de (6.13), teniendo en cuenta que cuando $r \rightarrow \infty$ antes del choque es $\varphi = \pi$ (ver figura) y que entre este valor y φ_m tanto r como φ disminuyen; esto es, el signo apropiado en (6.13) es el positivo. Con esto,

$$\varphi_m = \pi + \int_{\infty}^{r_{\text{mín}}} \frac{dr}{r^2 \sqrt{\frac{2\mu}{L_z^2} [E - V(r)] - \frac{1}{r^2}}}, \quad (6.32)$$

con $r_{\text{mín}}$ dado por la condición $\dot{r} = 0$ en (6.10): $E = V_{ef}(r_{\text{mín}})$.

El valor de φ_m depende de la energía mecánica en el sistema CM y del valor absoluto de L_z en el mismo sistema. De (6.6) y (6.10) y la figura podemos entonces escribir (recordemos que $\mathbf{v} = \mathbf{v}_1$ por ser $\mathbf{v}_2 = 0$)

$$|L_z| = \mu |\mathbf{v}| s, \quad E = \frac{1}{2} \mu |\mathbf{v}|^2. \quad (6.33)$$

Así, con (6.31), (6.32) y (6.33) el choque queda completamente determinado.

Consideremos ahora lo que sucede si un haz uniforme de partículas de masa m_1 incide sobre una distribución también uniforme de partículas de masa m_2 en reposo. Si el haz incide desde la izquierda, suficientemente lejos a la derecha se tendrán partículas de los tipos 1 y 2 que han sido dispersadas por las interacciones que tuvieron lugar. El objetivo que nos planteamos es determinar la distribución de estas partículas; esto es, la proporción de partículas en cada intervalo de velocidades (módulo y dirección).

La hipótesis que hacemos es que el haz de partículas 1 y el conjunto de partículas 2 son suficientemente diluidos para que puedan despreciarse las interacciones entre partículas del mismo tipo y, a la vez, considerar que la interacción entre partículas de distinto tipo ocurre sólo una vez para cada par de ellas. Así, en forma general podemos decir que el haz tiene una densidad uniforme en volumen de partículas 1, n_1 , que se mueven originalmente todas con velocidad \mathbf{v}_1 ; luego de interactuar con las partículas 2 el haz tendrá toda una distribución de velocidades \mathbf{v}'_1 ; sabemos que esta velocidad está determinada unívocamente por el ángulo χ que forma en el sistema CM respecto de la dirección original. De esta manera, caracterizamos la distribución de partículas dispersadas tipo 1 por la densidad de ellas, dn_χ , que es dispersada por unidad de tiempo con ángulos entre χ y $\chi + d\chi$ en el CM . Esta densidad (por segundo) es proporcional al intervalo angular considerado, $d\chi$, y a la intensidad del haz incidente, definida como el número de partículas del haz que atraviesan la unidad de área en una unidad de tiempo: $I_1 = n_1 |\mathbf{v}_1| = n_1 |\mathbf{v}|$, de manera que es útil definir la llamada *sección eficaz de dispersión* $\sigma(\chi)$ que relaciona dn_χ con $d\chi$ y I_1 como

$$dn_\chi = \sigma(\chi) I_1 d\chi. \quad (6.34)$$

El objetivo entonces es calcular $\sigma(\chi)$ que puede hacerse si calculamos dn_χ . Para esto basta tener en cuenta que las partículas equivalentes de masa μ son dispersadas con un dado ángulo χ si pasan a la distancia correcta del centro dispersor; esto es, si tienen un valor del parámetro de impacto s apropiado. Así, todas aquellas partículas cuyo valor de s está entre s y $s + ds$ se desviarán entre χ y $\chi + d\chi$. dn_χ corresponderá entonces al número de

partículas de haz incidente que por unidad de tiempo atraviesan un anillo de radio interior s y radio exterior $s + ds$; o sea, de área $2\pi s ds$ (no nos interesa en qué plano se desarrolla el movimiento; estamos considerando todos los posibles):

$$dn_\chi = I_1 2\pi s ds; \quad (6.35)$$

comparando con (6.34) deducimos entonces

$$\sigma(\chi) = 2\pi s \left| \frac{ds}{d\chi} \right|. \quad (6.36)$$

La precaución de tomar valor absoluto de la derivada viene del hecho que $\sigma(\chi)$ es una magnitud positiva, mientras que, en general, al aumentar s χ disminuye. Como hemos considerado la desviación en todos los planos posibles de movimiento es más correcto hablar de dispersión dentro de un dado ángulo sólido $d\Omega_\chi = 2\pi \sin \chi d\chi$, que nos define una sección eficaz diferente

$$dn_\chi = \sigma(\Omega_\chi) I_1 d\Omega_\chi = 2\pi \sigma(\Omega_\chi) I_1 \sin \chi d\chi,$$

que por comparación con (6.35) da

$$\sigma(\Omega_\chi) = \frac{s}{\sin \chi} \left| \frac{ds}{d\chi} \right|. \quad (6.37)$$

Vemos entonces que pueden calcularse en forma sistemática las secciones eficaces de dispersión en términos de la dispersión angular en el sistema CM . En las aplicaciones es más útil conocer la distribución de partículas dispersadas en el SL ; por ejemplo, el número de partículas tipo 1 dispersadas en ángulos entre θ_1 y $\theta_1 + d\theta_1$, independientemente del plano donde ocurre la dispersión; esto es, en el ángulo sólido $d\Omega_1 = 2\pi \sin \theta_1 d\theta_1$:

$$dn_{\theta_1} = 2\pi \sigma(\Omega_1) I_1 \sin \theta_1 d\theta_1.$$

Como existe una relación unívoca entre θ_1 y χ , la (6.30), basta escribir que dn_{θ_1} es igual al correspondiente dn_χ :

$$2\pi \sigma(\Omega_1) I_1 \sin \theta_1 d\theta_1 = 2\pi \sigma(\Omega_\chi) I_1 \sin \chi d\chi,$$

de donde se tiene

$$\sigma(\Omega_1) = \sigma(\Omega_\chi) \frac{\sin \chi}{\sin \theta_1} \left| \frac{d\chi}{d\theta_1} \right|,$$

en donde el χ debe evaluarse en el valor dado por (6.30).

Análogamente, para el número de partículas tipo 2 dispersadas en $d\Omega_2 = 2\pi \sin \theta_2 d\theta_2$ es, por (6.29),

$$\begin{aligned}\sigma(\Omega_2) &= \sigma(\Omega_\chi) \frac{\sin \chi}{\sin \theta_2} \left| \frac{d\chi}{d\theta_2} \right| \\ &= \frac{2}{\sin \theta_2} [\sigma(\Omega_\chi) \sin \chi]_{\chi=\pi-2\theta_2}.\end{aligned}$$

Finalmente, calculemos la sección eficaz de dispersión en el caso de un potencial electrostático entre partículas de cargas eléctricas Z_1e y Z_2e (e es la carga del protón). El problema es totalmente análogo al problema de Kepler donde ahora es (en unidades MKS, con ε_0 la permeabilidad eléctrica del vacío)

$$\alpha = -\frac{Z_1 Z_2 e^2}{4\pi \varepsilon_0}.$$

De (6.31) y (6.32), se debe calcular

$$\chi = \pi + 2 \int_{\infty}^{r_{\min}} \frac{dr}{r^2 \sqrt{\frac{2\mu}{L_z^2} \left[E + \frac{\alpha}{r} \right] - \frac{1}{r^2}}},$$

con r_{\min} dado por el único cero positivo de (6.14) (téngase en cuenta que $\alpha < 0$),

$$r_{\min} = \frac{|\alpha|}{2E} \left[\sqrt{1 + \frac{2L_z^2 E}{\mu \alpha^2}} + 1 \right].$$

Haciendo como antes, $u = r^{-1}$, es

$$\chi = \pi + 2 \int_{r_{\min}^{-1}}^0 \frac{du}{\sqrt{\frac{2\mu}{L_z^2} [E + \alpha u] - u^2}},$$

que, usando (6.20), resulta en

$$\chi = \pi - 2 \arccos \left[- \left(1 + \frac{2L_z^2 E}{\mu \alpha^2} \right)^{-1/2} \right],$$

o sea,

$$- \left(1 + \frac{2L_z^2 E}{\mu \alpha^2} \right)^{-1/2} = \cos \left(\frac{\pi}{2} - \frac{\chi}{2} \right) = \sin \left(\frac{\chi}{2} \right),$$

de donde es

$$|L_z| = \sqrt{\frac{\mu \alpha^2}{2E} \left[\frac{1}{\sin^2 \left(\frac{\chi}{2} \right)} - 1 \right]^{1/2}},$$

y, usando que $|L_z| = \mu |\mathbf{v}| s = \sqrt{2\mu E} s$,

$$s = \frac{|\alpha|}{2E} \cot\left(\frac{\chi}{2}\right).$$

Con esto calculamos inmediatamente

$$\sigma(\Omega_\chi) = \frac{s}{\sin \chi} \left| \frac{ds}{d\chi} \right| = \frac{\alpha^2}{16E^2} \csc^4\left(\frac{\chi}{2}\right), \quad (6.38)$$

que es la sección eficaz para la dispersión de Rutherford. En el caso estudiado por él las partículas tipo 1 fueron partículas alfa (masa atómica 4) y las tipo 2 núcleos pesados (por ejemplo, oro de masa atómica 200), de manera que $m_1/m_2 \ll 1$, por lo que (6.30) indica que $\chi \simeq \theta_1$ y la (6.38) es prácticamente igual a $\sigma(\Omega_1)$, que es lo medido en el SL .

Capítulo 7

Pequeñas oscilaciones

Son las que ocurren, con pequeña amplitud, alrededor de los estados de equilibrio estable del sistema. Supongamos el sistema de n grados de libertad con vínculos independientes del tiempo, descrito por un lagrangiano genérico de la forma

$$L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = T(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) - V(q_1, \dots, q_n). \quad (7.1)$$

Llamamos puntos de equilibrio a aquellos en los que la fuerza generalizada es nula: $Q_k = -\partial V / \partial q_k = 0$. Supongamos que existe equilibrio para los n grados de libertad; esto es, las n fuerzas generalizadas se anulan simultáneamente cuando las variables toman los valores $q_{01}, q_{02}, \dots, q_{0n}$. Esto significa que el potencial $V(q_1, \dots, q_n)$ tiene un extremo en ese punto del espacio de configuración.

La independencia del tiempo de los vínculos lleva a que la energía cinética es de la forma genérica

$$T(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = \frac{1}{2} \sum_{k,l=1}^n f_{kl}(q_1, \dots, q_n) \dot{q}_k \dot{q}_l. \quad (7.2)$$

Por la misma razón la energía se conserva y es de la forma $E = T + V$. Si el sistema se encuentra entonces en la posición de equilibrio con, además, velocidades nulas, debe permanecer indefinidamente allí (las \dot{q}_k y \ddot{q}_k son ambas nulas), mientras que si se encuentra en un punto ligeramente apartado del de equilibrio, también con velocidades inicialmente nulas, la no anulación de las \ddot{q}_k implica que el sistema ganará velocidad y se apartará entonces del punto inicial. Sin embargo, si el punto de equilibrio corresponde a un mínimo del potencial, el estado inicial apartado de éste tiene energía potencial mayor, por lo que la conservación de la energía implica que al aumentar la energía cinética el sistema se mueve hacia zonas de menor energía potencial; esto es,

hacia el punto de equilibrio. En este caso se dice que el punto de equilibrio es estable. Nótese que si el potencial tuviese un máximo en el punto de equilibrio, el movimiento del sistema sería apartándose de éste; claramente el equilibrio es inestable en este caso.

Así, pequeños apartamientos del estado de equilibrio (con velocidades iniciales nulas o muy pequeñas) producirán movimientos limitados alrededor del punto de equilibrio. Estudiemos entonces la teoría general para este tipo de movimientos desarrollando (7.1) alrededor del punto de equilibrio $(q_{01}, q_{02}, \dots, q_{0n})$. Llamemos $\eta_k \equiv q_k - q_{0k}$ a los apartamientos del equilibrio, con lo que

$$q_k = q_{0k} + \eta_k, \quad \dot{q}_k = \dot{\eta}_k. \quad (7.3)$$

La pequeñez del apartamiento implica pequeñez de la fuerza (las derivadas de orden uno del potencial son finitas alrededor del equilibrio) con la consecuente pequeñez de las aceleraciones y velocidades. Con esto en mente, desarrollamos (7.1) en serie de Taylor hasta el orden dos (que es el más bajo no trivial) alrededor de $(q_{01}, q_{02}, \dots, q_{0n})$, teniendo en cuenta la forma de la energía cinética (7.2) y las (7.3), para escribir

$$L = \frac{1}{2} \sum_{k,l=1}^n f_{kl}(q_{01}, \dots, q_{0n}) \dot{\eta}_k \dot{\eta}_l - \frac{1}{2} \sum_{k,l=1}^n \left. \frac{\partial^2 V}{\partial q_k \partial q_l} \right|_{q_{01}, \dots, q_{0n}} \eta_k \eta_l - V(q_{01}, \dots, q_{0n}),$$

donde se tuvo en cuenta además que las derivadas de orden uno de V en $(q_{01}, q_{02}, \dots, q_{0n})$ son nulas, y que las \dot{q}_k son en sí mismas de orden uno. Si dejamos de lado la constante $V(q_{01}, \dots, q_{0n})$ (o redefinimos el cero del potencial en coincidencia con este mínimo) y definimos las constantes

$$\begin{aligned} T_{kl} &\equiv f_{kl}(q_{01}, \dots, q_{0n}), \\ V_{kl} &\equiv \left. \frac{\partial^2 V}{\partial q_k \partial q_l} \right|_{q_{01}, \dots, q_{0n}}, \end{aligned}$$

podemos escribir que, al orden de aproximación considerado, es

$$L = \frac{1}{2} \sum_{k,l=1}^n T_{kl} \dot{\eta}_k \dot{\eta}_l - \frac{1}{2} \sum_{k,l=1}^n V_{kl} \eta_k \eta_l. \quad (7.4)$$

Este lagrangiano determina las trayectorias de las nuevas coordenadas generalizadas η_k a partir de las ecuaciones de Lagrange correspondientes

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_s} \right) - \frac{\partial L}{\partial \eta_s} = 0.$$

Notemos que por la forma (7.2) sólo la parte simétrica de las f_{kl} contribuye a la energía cinética, de manera que siempre podemos tomar simétricas a las f_{kl} ($f_{kl} = f_{lk}$) y, consecuentemente, a las T_{kl} . Por otro lado, de su propia definición, las V_{kl} son también simétricas. Así, cuando derivamos L respecto de una particular $\dot{\eta}_s$, la misma aparecerá ya sea como una de las $\dot{\eta}_k$ o una de las $\dot{\eta}_l$ en (7.4); explícitamente,

$$\frac{\partial L}{\partial \dot{\eta}_s} = \frac{1}{2} \sum_{k=1}^n T_{ks} \dot{\eta}_k + \frac{1}{2} \sum_{l=1}^n T_{sl} \dot{\eta}_l,$$

que, aprovechando la simetría de las T_{kl} , podemos escribir (renombrando l al índice k de la primera sumatoria)

$$\frac{\partial L}{\partial \dot{\eta}_s} = \sum_{l=1}^n T_{sl} \dot{\eta}_l.$$

Análogamente,

$$\frac{\partial L}{\partial \eta_s} = - \sum_{l=1}^n V_{sl} \eta_l.$$

Con lo que las ecuaciones de movimiento son

$$\sum_{l=1}^n (T_{sl} \ddot{\eta}_l + V_{sl} \eta_l) = 0. \quad (7.5)$$

Éste es un sistema de n ecuaciones diferenciales ordinarias, lineales y acopladas. La solución general puede obtenerse proponiendo soluciones de la forma

$$\eta_l(t) = C a_l \exp(-i\omega t), \quad (7.6)$$

donde i es la unidad imaginaria, C una constante independiente del índice l , y a_l una constante. Por ser el problema original real debe tomarse la parte real de (7.6). Dado que las (7.5) son lineales y sus coeficientes reales, las partes real e imaginaria de las η_l no se mezclan entre sí, por lo que puede usarse la forma compleja de η_l propuesta y tomarse la parte real al final del cálculo. Introduciendo entonces (7.6) en (7.5) se tiene (dejando de lado el factor global $C \exp(-i\omega t)$)

$$\sum_{l=1}^n (-\omega^2 T_{sl} + V_{sl}) a_l = 0, \quad (7.7)$$

que es un sistema algebraico, lineal y homogéneo para las incógnitas a_l . Este sistema tendrá solución no trivial sólo si el determinante de la matriz de coeficientes es nulo

$$\det(-\omega^2 T_{sl} + V_{sl}) = 0, \quad (7.8)$$

que determina que la incógnita ω^2 corresponde a las raíces de un polinomio de grado n . Éstas tienen en general n valores distintos. Veamos antes que nada que todas estas raíces son positivas, lo que implica que las ω 's son reales. Para esto multiplicamos cada una de las (7.7) por a_s^* (el asterisco simboliza conjugación compleja) y sumemos todas ellas para obtener

$$\sum_{l,s=1}^n (-\omega^2 T_{sl} + V_{sl}) a_l a_s^* = 0.$$

Si intercambiamos los índices l y s entre sí, y usamos la simetría de T_{sl} y V_{sl} obtenemos

$$\sum_{l,s=1}^n (-\omega^2 T_{sl} + V_{sl}) a_s a_l^* = 0,$$

que sumada a la anterior y pasando a la derecha los términos con ω^2 resulta en

$$\sum_{l,s=1}^n V_{sl} (a_l a_s^* + a_s a_l^*) = \omega^2 \sum_{l,s=1}^n T_{sl} (a_l a_s^* + a_s a_l^*),$$

de donde

$$\omega^2 = \frac{\sum_{l,s=1}^n V_{sl} (a_l a_s^* + a_s a_l^*)}{\sum_{l,s=1}^n T_{sl} (a_l a_s^* + a_s a_l^*)}.$$

Cada uno de los factores que multiplica a cada V_{sl} y T_{sl} en estas sumatorias es real, con lo que tenemos ante todo que ω^2 es real. Además, como la energía cinética es positiva definida, la suma del denominador debe ser positiva. De igual manera, como el punto de equilibrio es un mínimo del potencial, la suma del numerador también debe ser positiva definida (indica la variación del potencial ante un apartamiento genérico desde el mínimo); por lo tanto, ω^2 debe ser positivo. Así, como los coeficientes en el sistema (7.7) son reales, las a_l pueden elegirse siempre reales (ésta es una de las razones por las que se incluyó una constante global C en (7.6), para incluir posibles fases del argumento de la exponencial en ella).

Tenemos entonces que existen en general n valores distintos de ω , y que para cada uno de ellos se obtiene un conjunto de n valores a_l que pueden además considerarse reales. Si denominamos ω_j a las n posibles soluciones de (7.8), y con a_{lj} a los n valores obtenidos para cada frecuencia ω_j ($1 \leq l \leq$

n), la solución general de (7.5) se obtiene sumando para todos las posibles soluciones de la forma (7.6) (y tomando la parte real):

$$\eta_l(t) = \operatorname{Re} \left[\sum_{j=1}^n C_j a_{lj} \exp(-i\omega_j t) \right]. \quad (7.9)$$

Nótese que siempre podemos elegir a los $\omega_j > 0$, porque el correspondiente valor $-\omega_j$ equivale a un cambio en la fase de C_j (ténganse en cuenta que $\operatorname{Re}[C_j \exp(-i\omega_j t)] = \operatorname{Re}[C_j^* \exp(i\omega_j t)]$).

En este punto conviene trabajar con notación matricial. Llamemos \mathbf{V} a la matriz de componentes V_{sl} , \mathbf{T} a la de componentes T_{sl} , y \mathbf{a}_j al vector columna de componentes a_{lj} asociados a la frecuencia ω_j . Agrupamos además estos n vectores columna en una matriz de $n \times n$ componentes (cada columna está asociada a un ω_j distinto)

$$\mathbf{A} \equiv \left(\left(\begin{array}{c} a_{11} \\ a_{21} \\ \dots \\ a_{n1} \end{array} \right) \left(\begin{array}{c} a_{12} \\ a_{22} \\ \dots \\ a_{n2} \end{array} \right) \dots \left(\begin{array}{c} a_{1n} \\ a_{2n} \\ \dots \\ a_{nn} \end{array} \right) \right),$$

y definimos una matriz diagonal con los valores de ω_j

$$\mathbf{\Omega} \equiv \left(\begin{array}{cccc} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \omega_n \end{array} \right).$$

Con estas definiciones las (7.7) para cada ω_j se pueden escribir

$$\mathbf{V}\mathbf{a}_j = \omega_j^2 \mathbf{T}\mathbf{a}_j, \quad 1 \leq j \leq n, \quad (7.10)$$

y todas éstas, a su vez, como una única ecuación matricial (se entiende el cuadrado como el producto matricial de la matriz consigo misma)

$$\mathbf{V}\mathbf{A} = \mathbf{T}\mathbf{A}\mathbf{\Omega}^2. \quad (7.11)$$

Multipliquemos ahora (7.10) a izquierda por el vector fila \mathbf{a}_l^T (el supraíndice T indica transposición matricial, y llamamos $\lambda_j \equiv \omega_j^2$)

$$\mathbf{a}_l^T \mathbf{V}\mathbf{a}_j = \lambda_j \mathbf{a}_l^T \mathbf{T}\mathbf{a}_j, \quad (7.12)$$

y escribamos la misma ecuación con los índices l y j intercambiados,

$$\mathbf{a}_j^T \mathbf{V}\mathbf{a}_l = \lambda_l \mathbf{a}_j^T \mathbf{T}\mathbf{a}_l. \quad (7.13)$$

Si tenemos en cuenta que

$$\begin{aligned}(\mathbf{a}_j^T \mathbf{V} \mathbf{a}_l)^T &= \mathbf{a}_l^T \mathbf{V}^T \mathbf{a}_j = \mathbf{a}_l^T \mathbf{V} \mathbf{a}_j, \\ (\mathbf{a}_j^T \mathbf{T} \mathbf{a}_l)^T &= \mathbf{a}_l^T \mathbf{T}^T \mathbf{a}_j = \mathbf{a}_l^T \mathbf{T} \mathbf{a}_j\end{aligned}$$

donde las últimas igualdades en cada renglón resultan de la simetría de las matrices \mathbf{V} y \mathbf{T} , al transponer la (7.13) y restarla a la (7.12) tenemos

$$(\lambda_j - \lambda_l) \mathbf{a}_l^T \mathbf{T} \mathbf{a}_j = 0,$$

que indica que si $j \neq l$ y, además, $\lambda_j \neq \lambda_l$, entonces debe ser

$$\mathbf{a}_l^T \mathbf{T} \mathbf{a}_j = 0; \quad (7.14)$$

esto es, los vectores correspondientes a valores distintos de λ_j son ortogonales a través de la matriz \mathbf{T} .

Dado que siempre podemos incluir un factor global en cada vector \mathbf{a}_j (las (7.7) que los determinan son homogéneas) es posible elegir los \mathbf{a}_j de manera que estén normalizados de la siguiente manera

$$\mathbf{a}_j^T \mathbf{T} \mathbf{a}_j = 1,$$

que, junto con las (7.14), nos permiten decir que

$$\mathbf{A}^T \mathbf{T} \mathbf{A} = \mathbf{1}, \quad (7.15)$$

donde $\mathbf{1}$ representa la matriz identidad. Notablemente, si premultiplicamos (7.11) por \mathbf{A}^T tenemos, usando (7.15),

$$\mathbf{A}^T \mathbf{V} \mathbf{A} = \mathbf{\Omega}^2, \quad (7.16)$$

que es una matriz diagonal de elementos ω_j^2 . Así, la matriz \mathbf{A} diagonaliza simultáneamente a \mathbf{V} y \mathbf{T} .

Hasta aquí hemos supuesto que para $j \neq l$ es $\lambda_j \neq \lambda_l$. En general, si para $j \neq l$ resulta que $\lambda_j = \lambda_l$ (autovalor degenerado) esto indica que en las ecuaciones para los \mathbf{a}_j y \mathbf{a}_l asociados, más de una de las componentes de cada uno de éstos puede elegirse arbitrariamente (es lo mismo que sucede en el problema usual de autovalores), lo que da la libertad de poder siempre elegir que se cumpla la condición de ortogonalidad (7.14), aun para autovalores $\lambda_j = \lambda_l$. Consideramos entonces que siempre vale esto de ahora en más.

Resumiendo, una vez determinadas las matrices \mathbf{V} y \mathbf{T} se resuelve el problema de autovalores (7.8), que reescribimos como

$$\det(\mathbf{V} - \omega^2 \mathbf{T}) = 0,$$

para calcular los n autovalores ω_j^2 . Con éstos resolvemos las (7.10) para obtener los \mathbf{a}_j que normalizamos para que cumplan (7.15). Las n trayectorias $\eta_l(t)$ resultan entonces de (7.9). Para determinar las constantes C_j deben usarse las condiciones iniciales, que podemos escribir en forma matricial si definimos vectores columna $\boldsymbol{\eta}_0$, $\dot{\boldsymbol{\eta}}_0$ y \mathbf{C}

$$\boldsymbol{\eta}_0 \equiv \begin{pmatrix} \eta_1(t=0) \\ \eta_2(t=0) \\ \dots \\ \eta_n(t=0) \end{pmatrix}, \quad \dot{\boldsymbol{\eta}}_0 \equiv \begin{pmatrix} \dot{\eta}_1(t=0) \\ \dot{\eta}_2(t=0) \\ \dots \\ \dot{\eta}_n(t=0) \end{pmatrix}, \quad \mathbf{C} \equiv \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_n \end{pmatrix}.$$

En efecto, evaluando (7.9) y su derivada en $t=0$,

$$\begin{aligned} \eta_l(t=0) &= \operatorname{Re} \left[\sum_{j=1}^n C_j a_{lj} \right], \\ \dot{\eta}_l(t=0) &= \operatorname{Re} \left[\sum_{j=1}^n -i\omega_j C_j a_{lj} \right] = \operatorname{Im} \left[\sum_{j=1}^n \omega_j C_j a_{lj} \right], \end{aligned}$$

que, por ser los únicos factores complejos los C_j , podemos escribir

$$\begin{aligned} \eta_l(t=0) &= \sum_{j=1}^n \operatorname{Re}[C_j] a_{lj}, \\ \dot{\eta}_l(t=0) &= \sum_{j=1}^n \omega_j \operatorname{Im}[C_j] a_{lj}, \end{aligned}$$

que en forma matricial se expresa

$$\begin{aligned} \boldsymbol{\eta}_0 &= \mathbf{A} \operatorname{Re}[\mathbf{C}], \\ \dot{\boldsymbol{\eta}}_0 &= \boldsymbol{\Omega} \mathbf{A} \operatorname{Im}[\mathbf{C}]. \end{aligned}$$

Usando la condición (7.15), multiplicamos a izquierda cada una de estas ecuaciones por $\mathbf{A}^T \mathbf{T}$, y usamos además que $\boldsymbol{\Omega}^{-1}$, la inversa de $\boldsymbol{\Omega}$, es simplemente

$$\boldsymbol{\Omega}^{-1} = \begin{pmatrix} \omega_1^{-1} & 0 & \dots & 0 \\ 0 & \omega_2^{-1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \omega_n^{-1} \end{pmatrix},$$

para obtener inmediatamente

$$\begin{aligned} \operatorname{Re}[\mathbf{C}] &= \mathbf{A}^T \mathbf{T} \boldsymbol{\eta}_0, \\ \operatorname{Im}[\mathbf{C}] &= \mathbf{A}^T \mathbf{T} \boldsymbol{\Omega}^{-1} \dot{\boldsymbol{\eta}}_0. \end{aligned}$$

La solución se obtiene entonces de forma muy sistemática operando matricialmente.

7.1. Modos normales

El formalismo de la sección anterior nos permite determinar de manera también sistemática los denominados modos normales del sistema oscilante considerado. Los modos normales son en primer lugar nuevas coordenadas generalizadas que se obtienen como combinaciones lineales de las η_l consideradas arriba. Escribamos las ecuaciones de movimiento (7.5) en forma matricial para el vector columna

$$\boldsymbol{\eta}(t) \equiv \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \\ \dots \\ \eta_n(t) \end{pmatrix}$$

y sus derivadas,

$$\mathbf{T}\ddot{\boldsymbol{\eta}} + \mathbf{V}\boldsymbol{\eta} = \mathbf{0}. \quad (7.17)$$

Esta expresión, junto a la condición de ortogonalidad (7.15), nos sugiere entonces definir una relación lineal entre el vector $\boldsymbol{\eta}$ y el vector de coordenadas normales $\boldsymbol{\varsigma}$,

$$\boldsymbol{\varsigma}(t) \equiv \begin{pmatrix} \varsigma_1(t) \\ \varsigma_2(t) \\ \dots \\ \varsigma_n(t) \end{pmatrix},$$

de la forma

$$\boldsymbol{\eta} = \mathbf{A}\boldsymbol{\varsigma}, \quad (7.18)$$

ya que al reemplazar esta expresión en (7.17) y premultiplicar por \mathbf{A}^T obtenemos (usando (7.15) y (7.16))

$$\ddot{\boldsymbol{\varsigma}} + \boldsymbol{\Omega}^2\boldsymbol{\varsigma} = \mathbf{0},$$

que en componentes se escribe,

$$\ddot{\varsigma}_j + \omega_j^2\varsigma_j = 0;$$

esto es, las coordenadas normales están desacopladas entre sí; cada una vibra con su frecuencia ω_j .

La (7.18) se invierte fácilmente premultiplicándola por $\mathbf{A}^T\mathbf{T}$ y usando (7.15) para escribir

$$\boldsymbol{\varsigma} = \mathbf{A}^T\mathbf{T}\boldsymbol{\eta}, \quad (7.19)$$

que nos da la forma sistemática de obtener los modos normales en términos de las coordenadas η_j .

Vemos de (7.19) que cada modo normal es una superposición lineal de las n coordenadas η_j , con coeficientes dados por los elementos de la matriz $\mathbf{A}^T \mathbf{T}$ que es independiente de las particulares condiciones iniciales; sólo depende del sistema en sí. Lo notable es que esta precisa superposición representa una oscilación particular del sistema en la que *todos* los componentes de éste vibran con la misma frecuencia; propiedad que tienen todos los sistemas que ejecutan pequeñas oscilaciones como las estudiadas, por complejos que sean. Cada ω_j representa entonces una frecuencia a la que vibra el sistema en conjunto (cada componente con su propia fase y amplitud) y las ω_j son entonces llamadas frecuencias propias del sistema. La propiedad notable de oscilar en conjunto con igual frecuencia permite identificar los modos normales en sistemas sencillos si se puede determinar cómo darle condiciones iniciales apropiadas para que esto ocurra. Desde ya, la forma más general de oscilación es una combinación lineal de los modos normales, que es justamente lo que (7.9) representa.

7.2. Oscilaciones de sistemas aislados (moléculas)

Cuando el sistema que puede sostener oscilaciones está aislado, algunos de los modos posibles de movimiento corresponden a translación y rotación uniformes. Estas posibilidades de movimiento no oscilatorio se reflejan en la aparición de frecuencias nulas al aplicar el formalismo general desarrollado en la primera sección. La forma de evitar estas frecuencias nulas es descartar de entrada la posibilidad de movimientos translatorios y rotatorios globales, imponiendo que la cantidad de movimiento y el momento angular totales son nulos, tanto en el equilibrio como en los movimientos alrededor de éste.

Si las N partículas del sistema están caracterizadas por posiciones \mathbf{x}_i , las condiciones de cantidad de movimiento y el momento angular totales nulos se escriben, respectivamente,

$$\sum_{i=1}^N m_i \dot{\mathbf{x}}_i = \sum_{i=1}^N m_i \mathbf{x}_i \times \dot{\mathbf{x}}_i = \mathbf{0}. \quad (7.20)$$

Para apartamientos pequeños de las posiciones de equilibrio $\mathbf{x}_{0i} = \mathbf{x}_i(q_{01}, q_{02}, \dots, q_{0n})$,

podemos escribir

$$\begin{aligned}\mathbf{x}_i &\simeq \mathbf{x}_{0i} + \sum_{k=1}^n \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 (q_k - q_{0k}) = \mathbf{x}_{0i} + \sum_{k=1}^n \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \eta_k, \\ \dot{\mathbf{x}}_i &\simeq \sum_{k=1}^n \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \dot{\eta}_k,\end{aligned}$$

con lo que las condiciones (7.20) se escriben, a orden uno en los apartamientos η_k (o sus derivadas $\dot{\eta}_k$),

$$\begin{aligned}\sum_{i=1}^N \sum_{k=1}^n m_i \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \dot{\eta}_k &= \frac{d}{dt} \left[\sum_{i=1}^N \sum_{k=1}^n m_i \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \eta_k \right] = \mathbf{0}, \\ \sum_{i=1}^N \sum_{k=1}^n m_i \mathbf{x}_{0i} \times \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \dot{\eta}_k &= \frac{d}{dt} \left[\sum_{i=1}^N \sum_{k=1}^n m_i \mathbf{x}_{0i} \times \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \eta_k \right] = \mathbf{0},\end{aligned}$$

que indican que los términos entre corchetes son constantes que como en el equilibrio, $\eta_k = 0$, tienen valor cero, escribimos (intercambiamos de paso el orden de las sumatorias)

$$\begin{aligned}\sum_{k=1}^n \left(\sum_{i=1}^N m_i \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \right) \eta_k &= \mathbf{0}, \\ \sum_{k=1}^n \left(\sum_{i=1}^N m_i \mathbf{x}_{0i} \times \left. \frac{\partial \mathbf{x}_i}{\partial q_k} \right|_0 \right) \eta_k &= \mathbf{0}.\end{aligned}$$

Notemos que estas dos ecuaciones vectoriales representan en general seis ecuaciones de vínculo (holónomas) entre las coordenadas η_k (si el sistema es libre de moverse sólo en un plano serán tres ecuaciones de vínculo, y sólo una si el sistema puede moverse libremente en una dimensión). Usando estas ecuaciones lineales podemos reducir fácilmente el número de coordenadas η_k para quedarnos con un sistema que sólo puede sostener oscilaciones; esto es, en el que todas las frecuencias propias son no nulas. Esto es especialmente útil cuando se estudian los modos propios de moléculas.

7.3. Oscilaciones forzadas y amortiguadas

Si el sistema que puede sostener oscilaciones está además sometido a fuerzas externas, podemos escribir las ecuaciones de Lagrange incluyendo las fuerzas como

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = Q_k^{ext} (q_1, q_2, \dots, q_n, t).$$

Para que los apartamientos puedan considerarse pequeños la fuerza aplicada debe serlo, de manera que los pequeños apartamientos pueden estudiarse como en la primera sección evaluando la expresión de la fuerza generalizada, consistentemente con la aproximación original, en las posiciones de equilibrio para tener el análogo de (7.5)

$$\sum_{l=1}^n (T_{sl}\ddot{\eta}_l + V_{sl}\eta_l) = Q_{0s}(t),$$

donde

$$Q_{0s}(t) \equiv Q_s^{ext}(q_{01}, q_{02}, \dots, q_{0n}, t).$$

Escribiendo las ecuaciones en forma matricial, introduciendo el vector columna de fuerzas

$$\mathbf{Q}^{ext} \equiv \begin{pmatrix} Q_{01} \\ Q_{02} \\ \dots \\ Q_{0n} \end{pmatrix},$$

tenemos

$$\mathbf{T}\ddot{\boldsymbol{\eta}} + \mathbf{V}\boldsymbol{\eta} = \mathbf{Q}^{ext}.$$

Del que podemos obtener las ecuaciones de los modos normales como se hizo arriba, reemplazando $\boldsymbol{\eta}$ por $\mathbf{A}\boldsymbol{\zeta}$ y premultiplicando por \mathbf{A}^T para obtener (usando las (7.15) y (7.16))

$$\ddot{\boldsymbol{\zeta}} + \boldsymbol{\Omega}^2\boldsymbol{\zeta} = \mathbf{A}^T\mathbf{Q}^{ext} \equiv \mathbf{Q}',$$

que nos indica que todavía puede resolverse el sistema de ecuaciones lineales desacopladas para los modos normales, ahora forzados por las fuerzas Q'_k :

$$\ddot{\zeta}_k + \omega_k^2\zeta_k = Q'_k.$$

Veamos qué sucede si agregamos fuerzas amortiguadoras. El tipo que consideraremos es el de fuerzas proporcionales a la velocidad y que se oponen a ésta. Éste es el tipo de fuerzas esperables cuando el movimiento es en el seno de un fluido o gas y las velocidades son pequeñas, como son las que estudiamos. Si la velocidad de la partícula i es $\dot{\mathbf{x}}_i$ (hay N partículas en el sistema), la fuerza amortiguadora sobre ella que estudiamos es de la forma

$$\mathbf{F}_i = -\nu_i\dot{\mathbf{x}}_i,$$

donde ν_i es un coeficiente positivo que depende de las características de la partícula misma y del medio en el que se mueve (por ejemplo, $\nu \simeq 6\pi\mu R$

para una partícula esférica de radio R que se mueve “lentamente” en un fluido newtoniano de viscosidad dinámica μ , y la densidad de la partícula es grande comparada con la del fluido). En las condiciones que estudiamos (vínculos independientes del tiempo) es

$$\dot{\mathbf{x}}_i = \sum_{l=1}^n \frac{\partial \mathbf{x}_i}{\partial q_l} \dot{q}_l,$$

y la fuerza generalizada asociada a esta fuerza amortiguadora es

$$\begin{aligned} Q_k^{dis} &= - \sum_{i=1}^N \nu_i \dot{\mathbf{x}}_i \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \\ &= - \sum_{l=1}^n \sum_{i=1}^N \nu_i \frac{\partial \mathbf{x}_i}{\partial q_l} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \dot{q}_l. \end{aligned}$$

Definiendo entonces los coeficientes simétricos en los índices k y l

$$\alpha_{kl} \equiv \sum_{i=1}^N \nu_i \left(\frac{\partial \mathbf{x}_i}{\partial q_l} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right)_0, \quad (7.21)$$

donde las derivadas están evaluadas en las posiciones de equilibrio, tenemos

$$Q_k^{dis} = - \sum_{l=1}^n \alpha_{kl} \dot{q}_l. \quad (7.22)$$

Definiendo entonces la matriz simétrica $\boldsymbol{\alpha}$ de coeficientes α_{kl} podemos escribir el análogo de (7.5) como

$$\mathbf{T}\ddot{\boldsymbol{\eta}} + \mathbf{V}\dot{\boldsymbol{\eta}} = -\boldsymbol{\alpha}\dot{\boldsymbol{\eta}}. \quad (7.23)$$

La solución general de este sistema de ecuaciones acopladas se obtiene proponiendo que cada η_k es de la forma (análoga a la (7.6))

$$\eta_k = C a_k \exp(rt), \quad (7.24)$$

donde r es en general complejo. Reemplazando estas expresiones en (7.23) se obtiene el sistema lineal para los a_k (agrupados en el vector columna \mathbf{a})

$$(r^2\mathbf{T} + r\boldsymbol{\alpha} + \mathbf{V}) \mathbf{a} = \mathbf{0},$$

que tendrá solución no trivial si el determinante de la matriz de coeficientes es nulo,

$$\det(r^2\mathbf{T} + r\boldsymbol{\alpha} + \mathbf{V}) = 0,$$

que determina los posibles valores de r . Esta ecuación es un polinomio de grado $2n$ en r , con coeficientes reales, de manera que sus raíces son reales o pares de complejos conjugados. En todo caso las partes reales son siempre negativas ya que de otro modo, por la forma de (7.24), la coordenada η_k asociada crecería indefinidamente, lo que no es posible en el sistema considerado.

Calculemos cómo varía la energía de un sistema con fuerzas disipativas dadas por las (7.22). Sabemos que

$$E = \sum_{k=1}^n \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L,$$

por lo que

$$\begin{aligned} \frac{dE}{dt} &= \sum_{k=1}^n \ddot{q}_k \frac{\partial L}{\partial \dot{q}_k} + \sum_{k=1}^n \dot{q}_k \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) \\ &\quad - \sum_{k=1}^n \frac{\partial L}{\partial q_k} \dot{q}_k - \sum_{k=1}^n \frac{\partial L}{\partial \dot{q}_k} \ddot{q}_k \\ &= \sum_{k=1}^n \dot{q}_k \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} \right] \\ &= - \sum_{k=1}^n \sum_{l=1}^n \alpha_{kl} \dot{q}_k \dot{q}_l. \end{aligned}$$

La expresión del último renglón es claramente definida negativa ya que la doble suma es definida positiva, lo que es evidente si se usa la (7.21) para reescribirla como

$$\begin{aligned} &\sum_{i=1}^N \nu_i \sum_{k=1}^n \sum_{l=1}^n \left(\frac{\partial \mathbf{x}_i}{\partial q_l} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right)_0 \dot{q}_k \dot{q}_l \\ &= \sum_{i=1}^N \nu_i \left| \sum_{k=1}^n \frac{\partial \mathbf{x}_i}{\partial q_k} \Big|_0 \dot{q}_k \right|^2. \end{aligned}$$

La función definida positiva

$$\mathcal{F} \equiv \frac{1}{2} \sum_{k=1}^n \sum_{l=1}^n \alpha_{kl} \dot{q}_k \dot{q}_l,$$

es denominada función de disipación de Rayleigh, y tenemos finalmente

$$\frac{dE}{dt} = -2\mathcal{F} < 0.$$

Notemos que la fuerza generalizada (7.22) puede escribirse también como

$$Q_k^{dis} = -\frac{\partial \mathcal{F}}{\partial \dot{q}_k}.$$

Finalmente, si combinamos forzado externo y amortiguamiento,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = Q_k^{ext} - \frac{\partial \mathcal{F}}{\partial \dot{q}_k},$$

tenemos el sistema de pequeñas oscilaciones (en forma matricial)

$$\mathbf{T}\ddot{\boldsymbol{\eta}} + \boldsymbol{\alpha}\dot{\boldsymbol{\eta}} + \mathbf{V}\boldsymbol{\eta} = \mathbf{Q}^{ext}. \quad (7.25)$$

Sólo consideraremos el caso particular en el que las matrices \mathbf{T} y $\boldsymbol{\alpha}$ sean proporcionales; esto no es tan extraño teniendo en cuenta que

$$T_{kl} = \sum_{i=1}^N m_i \left(\frac{\partial \mathbf{x}_i}{\partial q_l} \cdot \frac{\partial \mathbf{x}_i}{\partial q_k} \right)_0,$$

tiene una forma del todo análoga a la expresión (7.21) de los α_{kl} donde aparece ν_i en lugar de m_i . Así, si $\nu_i = \kappa m_i$, con κ una constante positiva independiente de i , tenemos $\boldsymbol{\alpha} = \kappa \mathbf{T}$, y podemos escribir (7.25) como

$$\mathbf{T}\ddot{\boldsymbol{\eta}} + \kappa \mathbf{T}\dot{\boldsymbol{\eta}} + \mathbf{V}\boldsymbol{\eta} = \mathbf{Q}^{ext},$$

con lo que, procediendo como al principio de esta sección, obtenemos los modos normales desacoplados

$$\ddot{\zeta}_k + \kappa \dot{\zeta}_k + \omega_k^2 \zeta_k = Q'_k,$$

donde las ω_k son las frecuencias propias del sistema no amortiguado. Una condición menos restrictiva es que la matriz \mathbf{A} diagonalice también a $\boldsymbol{\alpha}$ con lo que se obtienen ecuaciones iguales a las anteriores, pero con un κ_k en general distinto para cada modo. Las soluciones de estas ecuaciones pueden entonces estudiarse considerando el sistema de una dimensión con coordenada x

$$\ddot{x} + \kappa \dot{x} + \omega_0^2 x = A \cos(\omega t), \quad (7.26)$$

donde consideramos que el forzante es un modo puro de Fourier, con frecuencia angular ω , en términos del cual puede desarrollarse un forzante de dependencia más genérica.

Como todas las constantes que aparecen en esta ecuación son reales puede estudiarse la versión compleja

$$\ddot{z} + \kappa \dot{z} + \omega_0^2 z = A \exp(i\omega t), \quad (7.27)$$

de la cual (7.26) es la parte real, siendo $x = \text{Re}(z)$. La solución general de (7.27) consiste de la solución homogénea, de la forma $z_h \sim \exp(ct)$, más la particular $z_p = B \exp(i\omega t)$. Reemplazando z_p en (7.27) tenemos inmediatamente que

$$B = \frac{A}{\omega_0^2 - \omega^2 + i\kappa\omega},$$

que podemos escribir en forma exponencial $B = b \exp(i\delta)$, donde

$$b = \frac{A}{\sqrt{(\omega_0^2 - \omega^2)^2 + \kappa^2\omega^2}}, \quad \tan \delta = \frac{\kappa\omega}{\omega^2 - \omega_0^2}. \quad (7.28)$$

Por otro lado, la solución homogénea reemplazada en (7.27) resulta en

$$c^2 + \kappa c + \omega_0^2 = 0,$$

cuya solución es $c = -\frac{\kappa}{2} \pm \sqrt{\frac{\kappa^2}{4} - \omega_0^2}$, que en el caso $\kappa < 2\omega_0$ (subamortiguado) podemos escribir $c = -\frac{\kappa}{2} \pm i\gamma$, donde $\gamma \equiv \sqrt{\omega_0^2 - \frac{\kappa^2}{4}}$. La solución general se escribe entonces

$$z = C \exp\left(-\frac{\kappa}{2}t \pm i\gamma t\right) + B \exp(i\omega t),$$

cuya parte real escribimos como

$$x = a \exp\left(-\frac{\kappa}{2}t\right) \cos(\gamma t + \alpha) + b \cos(\omega t + \delta),$$

donde a y α se determinan por las condiciones iniciales, y b y δ están dados por (7.28). Tenemos entonces que la solución corresponde a una oscilación amortiguada de frecuencia $\gamma = \sqrt{\omega_0^2 - \frac{\kappa^2}{4}}$, cercana a la frecuencia propia si el amortiguamiento es muy pequeño, más una oscilación a la frecuencia del forzante ω . En tiempos suficientemente largos, $t \gg 2\kappa^{-1}$, sólo sobrevive la oscilación forzada

$$x = b \cos(\omega t + \delta).$$

Concentrémonos en cómo se comporta esta solución forzada en función de la frecuencia del forzante. Para comenzar, la amplitud b tiene un máximo, de valor $A \left(\kappa \sqrt{\omega_0^2 - \frac{\kappa^2}{4}}\right)^{-1}$, cuando la frecuencia del forzante ω coincide con la frecuencia $\sqrt{\omega_0^2 - \frac{\kappa^2}{4}}$. En cuanto al desfase δ entre la oscilación del sistema y el forzante, éste es claramente negativo cuando $\omega < \omega_0$ (va a cero para $\omega \rightarrow 0$) y tiende a $-\pi/2$ para $\omega = \omega_0$ (condición de resonancia); cuando $\omega > \omega_0$ varía desde el valor de resonancia, $-\pi/2$, hasta $-\pi$ cuando $\omega \rightarrow \infty$. Vemos entonces que δ es siempre negativo, lo que indica que la oscilación “atrassa” respecto al forzante.

7.4. Oscilaciones no lineales en una dimensión

En una dimensión cartesiana el término de energía cinética es siempre de la forma $T = m\dot{x}^2/2$ sea pequeño o no el apartamiento de la posición de equilibrio, mientras que la forma del potencial $V = m\omega_0^2 x^2/2$ sólo es en general válida para apartamientos pequeños del equilibrio (salvo, por supuesto, en el caso elástico puro). En general, alrededor del equilibrio el potencial tendrá un desarrollo genérico (las constantes se eligen así por conveniencia de escritura)

$$V = \frac{1}{2}m\omega_0^2 x^2 + \frac{1}{3}m\alpha x^3 + \frac{1}{4}m\beta x^4 + \dots \quad (7.29)$$

Consideremos entonces una oscilación en la que los términos cúbico y cuártico del potencial tengan importancia, si bien su influencia pueda considerarse pequeña. La ecuación de movimiento para (7.29) es

$$\ddot{x} + \omega_0^2 x = -\alpha x^2 - \beta x^3, \quad (7.30)$$

ecuación conocida como del oscilador de Duffing. Resolvámosla suponiendo que los términos de la derecha son pequeños, proponiendo que la solución dominante es la homogénea, la que anula el lado izquierdo de (7.30) (sin pérdida de generalidad elegimos el $t = 0$ para anular la fase inicial)

$$x_1 = a \cos(\omega t),$$

donde ω no es necesariamente igual a ω_0 , aunque su parte dominante sí lo es. Proponemos entonces

$$\begin{aligned} x &= x_1 + x_2 + x_3 + \dots, \\ \omega &= \omega_0 + \omega_1 + \omega_2 + \dots, \end{aligned}$$

donde cada término es de orden (indicado por el subíndice) superior al anterior. El orden uno está dado por la amplitud a que se supone “pequeña”.

El lado derecho de (7.30) es entonces, desarrollado hasta el orden tres,

$$-\alpha x_1^2 - 2\alpha x_1 x_2 - \beta x_1^3,$$

que, teniendo en cuenta que

$$\begin{aligned} \cos^2(\omega t) &= \left(\frac{e^{i\omega t} + e^{-i\omega t}}{2} \right)^2 = \frac{e^{2i\omega t} + 2 + e^{-2i\omega t}}{4} \\ &= \frac{1}{2} + \frac{1}{2} \cos(2\omega t), \end{aligned}$$

y que

$$\begin{aligned}\cos^3(\omega t) &= \left(\frac{e^{i\omega t} + e^{-i\omega t}}{2}\right)^3 = \frac{e^{3i\omega t} + 3e^{i\omega t} + 3e^{-i\omega t} + e^{-3i\omega t}}{8} \\ &= \frac{3}{4}\cos(\omega t) + \frac{1}{4}\cos(3\omega t),\end{aligned}$$

se escribe

$$-\alpha\frac{a^2}{2}[1 + \cos(2\omega t)] - 2\alpha a \cos(\omega t)x_2 - \beta\frac{a^3}{4}[3\cos(\omega t) + \cos(3\omega t)]. \quad (7.31)$$

Por otro lado, el lado izquierdo se escribe, también al tercer orden de aproximación,

$$-\omega^2 a \cos(\omega t) + \ddot{x}_2 + \ddot{x}_3 + \omega_0^2 [a \cos(\omega t) + x_2 + x_3],$$

que al desarrollar ω^2 hasta términos de orden dos (no hace falta más porque ω^2 multiplica a x_1 que es ya de orden uno): $\omega^2 = \omega_0^2 + 2\omega_0\omega_1 + \omega_1^2 + 2\omega_0\omega_2$, nos permite escribir el lado izquierdo desarrollado explícitamente hasta el orden tres (vemos que el término de orden uno se cancela, como debe ser)

$$-(2\omega_0\omega_1 + \omega_1^2 + 2\omega_0\omega_2)a \cos(\omega t) + \ddot{x}_2 + \ddot{x}_3 + \omega_0^2(x_2 + x_3). \quad (7.32)$$

Su igualamos ahora los términos de orden dos en (7.31) con los correspondientes de (7.32) tenemos

$$\ddot{x}_2 + \omega_0^2 x_2 - 2\omega_0\omega_1 a \cos(\omega t) = -\alpha\frac{a^2}{2}[1 + \cos(2\omega t)],$$

que reescribimos

$$\ddot{x}_2 + \omega_0^2 x_2 = 2\omega_0\omega_1 a \cos(\omega t) - \alpha\frac{a^2}{2}[1 + \cos(2\omega t)]. \quad (7.33)$$

Lo primero que notamos es que la ecuación para x_2 corresponde a la de un oscilador de frecuencia propia ω_0 , forzado por una fuerza constante y por dos oscilantes de frecuencias ω y 2ω . Como $\omega \simeq \omega_0$ el forzante de frecuencia ω estará casi en resonancia y generará una oscilación de amplitud muy grande (considérese (7.28) con $\kappa = 0$); esto no es físicamente correcto y la manera de evitar esta inconsistencia es que la amplitud de este forzante sea nula; esto es, $\omega_1 = 0$, lo que nos dice que la no linealidad no induce correcciones de orden uno a la frecuencia fundamental de oscilación. La solución de (7.33) con $\omega_1 = 0$ es directa (basta proponer una particular ya que la solución

homogénea puede incorporarse en x_1 , y ω puede considerarse igual a ω_0 ya que las diferencias entre ambas son de orden dos)

$$x_2 = -\frac{\alpha a^2}{2\omega_0^2} + \frac{\alpha a^2}{6\omega_0^2} \cos(2\omega t). \quad (7.34)$$

Igualando ahora los términos de orden tres en (7.31) y (7.32) tenemos (usamos que $\omega_1 = 0$)

$$-2\omega_0\omega_2 a \cos(\omega t) + \ddot{x}_3 + \omega_0^2 x_3 = -2\alpha a \cos(\omega t) x_2 - \beta \frac{a^3}{4} [3 \cos(\omega t) + \cos(3\omega t)].$$

Si reemplazamos aquí x_2 por su expresión (7.34) y usamos que

$$\begin{aligned} \cos(\omega t) \cos(2\omega t) &= \frac{e^{i\omega t} + e^{-i\omega t}}{2} \frac{e^{2i\omega t} + e^{-2i\omega t}}{2} \\ &= \frac{1}{4} (e^{3i\omega t} + e^{i\omega t} + e^{-i\omega t} + e^{-3i\omega t}) \\ &= \frac{1}{2} [\cos(\omega t) + \cos(3\omega t)], \end{aligned}$$

tenemos para x_3 la ecuación

$$\begin{aligned} \ddot{x}_3 + \omega_0^2 x_3 &= \left(\frac{5\alpha^2 a^3}{6\omega_0^2} + 2a\omega_0\omega_2 - \frac{3}{4}\beta a^3 \right) \cos(\omega t) \\ &\quad - \left(\frac{\alpha^2 a^3}{6\omega_0^2} + \frac{\beta a^3}{4} \right) \cos(3\omega t). \end{aligned}$$

La condición de que no haya forzante en resonancia lleva a que

$$\omega_2 = \frac{3\beta a^2}{8\omega_0} - \frac{5\alpha^2 a^2}{12\omega_0^3}, \quad (7.35)$$

y la ecuación resultante es fácilmente resuelta para la solución particular

$$x_3 = \frac{a^3}{16\omega_0^2} \left(\frac{\beta}{2} + \frac{\alpha^2}{3\omega_0^2} \right) \cos(3\omega t). \quad (7.36)$$

Recapitulando, vemos que el efecto de los términos no lineales de la ecuación (7.30), los llamados términos anarmónicos, es la aparición de oscilaciones con frecuencias múltiplo de la frecuencia fundamental, las llamadas armónicas, junto a que la frecuencia fundamental no es exactamente la frecuencia propia ω_0 , sino que difiere de ésta en términos cuadráticos en la amplitud de la oscilación fundamental, dados por (7.35).

Capítulo 8

Cuerpo rígido

8.1. Cinemática

Llamamos cuerpo rígido al sólido continuo indeformable, en el sentido que la distancia entre cada par de partículas que lo componen es constante, independiente de las fuerzas a las que esté sometido el cuerpo. De esta manera, una forma natural de caracterizar al sólido es dar las posiciones de sus (infinitas) partículas referidas a un sistema fijo al cuerpo; estas posiciones serán entonces siempre las mismas; llamemos S' a tal sistema. La posición de cualquiera de las partículas del sólido respecto de un sistema externo S se podrá entonces determinar conociendo la posición y orientación del sistema fijo al cuerpo, S' . La posición estará dada por la del origen de S' y la orientación por la de los ejes de S' respecto de los de S .

Tomemos un punto cualquiera del cuerpo, caracterizado por su posición \mathbf{x} en el sistema S . Tomemos el sistema de referencia fijo al cuerpo S' con origen en el punto \mathbf{x}_0 (no necesariamente dentro del cuerpo); como dijimos, cada uno de los puntos del cuerpo está caracterizado por un vector de componentes constantes en este sistema; en particular, llamemos \mathbf{x}' al vector posición del punto \mathbf{x} en S' . Tenemos entonces que

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}' ,$$

y la velocidad de \mathbf{x} en el sistema S será

$$\mathbf{v} = \left. \frac{d\mathbf{x}}{dt} \right)_S = \left. \frac{d\mathbf{x}_0}{dt} \right)_S + \left. \frac{d\mathbf{x}'}{dt} \right)_S ,$$

donde se ha puesto de manifiesto respecto de qué sistema se considera la variación en el tiempo del vector en cuestión. Sabemos que, en general,

$$\left. \frac{d\mathbf{x}'}{dt} \right)_S = \left. \frac{d\mathbf{x}'}{dt} \right)_{S'} + \boldsymbol{\Omega}_0 \times \mathbf{x}' ,$$

donde en el lado derecho tenemos en cuenta posibles variaciones del vector \mathbf{x}' en S' (el primer término), y las variaciones debidas a la rotación del sistema S' alrededor de su origen O . Como el vector \mathbf{x}' no varía en S' , tenemos finalmente que

$$\mathbf{v} = \left. \frac{d\mathbf{x}_0}{dt} \right)_S + \boldsymbol{\Omega}_0 \times \mathbf{x}',$$

o sea,

$$\mathbf{v} = \mathbf{v}_0 + \boldsymbol{\Omega}_0 \times (\mathbf{x} - \mathbf{x}_0). \quad (8.1)$$

Si en lugar de tomar el origen de S' en el punto O lo tomáramos en, digamos, el punto O' , la velocidad de \mathbf{x} se escribiría

$$\mathbf{v} = \mathbf{v}_{0'} + \boldsymbol{\Omega}_{0'} \times (\mathbf{x} - \mathbf{x}_{0'}). \quad (8.2)$$

Por otro lado, podemos escribir que

$$\mathbf{v}_{0'} = \mathbf{v}_0 + \boldsymbol{\Omega}_0 \times (\mathbf{x}_{0'} - \mathbf{x}_0),$$

que podemos reescribir, sumando y restando \mathbf{x} dentro del paréntesis,

$$\mathbf{v}_{0'} = \mathbf{v}_0 + \boldsymbol{\Omega}_0 \times (\mathbf{x} - \mathbf{x}_0) + \boldsymbol{\Omega}_0 \times (\mathbf{x}_{0'} - \mathbf{x}),$$

y que al reemplazar en (8.2) nos da

$$\mathbf{v} = \mathbf{v}_0 + \boldsymbol{\Omega}_0 \times (\mathbf{x} - \mathbf{x}_0) + (\boldsymbol{\Omega}_{0'} - \boldsymbol{\Omega}_0) \times (\mathbf{x} - \mathbf{x}_{0'}),$$

que al comparar con (8.1) indica que

$$(\boldsymbol{\Omega}_{0'} - \boldsymbol{\Omega}_0) \times (\mathbf{x} - \mathbf{x}_{0'}) = 0$$

para cualesquiera puntos O , O' y \mathbf{x} , que sólo es posible si

$$\boldsymbol{\Omega}_{0'} = \boldsymbol{\Omega}_0.$$

De esta manera, cualquiera sea el origen que elijamos para el sistema S' , la velocidad de rotación de los ejes alrededor de este origen (y con ello la orientación instante a instante de los ejes de S' respecto de los de S) será la misma. Podemos entonces identificar la rotación del sólido con la de cualquier sistema de referencia unido a él, y denotaremos simplemente como $\boldsymbol{\Omega}$ a esta velocidad para escribir la (8.1) en general como

$$\mathbf{v} = \mathbf{v}_0 + \boldsymbol{\Omega} \times (\mathbf{x} - \mathbf{x}_0). \quad (8.3)$$

Nótese que ni siquiera es necesario que el origen O de S' esté contenido dentro del cuerpo; lo único que se usó es que S' está fijo al sólido.

Si multiplicamos escalarmente (8.3) por $\boldsymbol{\Omega}$ tenemos

$$\boldsymbol{\Omega} \cdot \mathbf{v} = \boldsymbol{\Omega} \cdot \mathbf{v}_0 \quad (8.4)$$

para cualquier O , por lo que, si en un instante dado la velocidad de una de las partículas del sólido satisface $\mathbf{v} \perp \boldsymbol{\Omega}$, entonces, por ser O arbitrario, $\boldsymbol{\Omega}$ es perpendicular a la velocidad de todos los puntos del sistema fijo al cuerpo. Puede entonces elegirse un punto O' de este sistema tal que su velocidad sea nula (basta elegir el $\mathbf{x}_{O'}$ que cumple $\mathbf{v}_0 + \boldsymbol{\Omega} \times (\mathbf{x}_{O'} - \mathbf{x}_0) = 0$, lo que es siempre posible por la condición de velocidades todas perpendiculares a $\boldsymbol{\Omega}$). Así, el cuerpo en el instante considerado está rotando con $\boldsymbol{\Omega}$ alrededor de O' sin translación de este punto. El eje paralelo a $\boldsymbol{\Omega}$ que pasa por O' es denominado eje instantáneo de rotación.

Si la velocidad de los puntos del cuerpo no fuese perpendicular a $\boldsymbol{\Omega}$ la (8.4) nos dice que todos los puntos del cuerpo tienen igual componente de velocidad en la dirección de $\boldsymbol{\Omega}$, por lo que, en el instante considerado, en un sistema que se moviera con la velocidad apropiada en la dirección de $\boldsymbol{\Omega}$ tendríamos las condiciones del punto anterior; esto es, velocidades todas perpendiculares a $\boldsymbol{\Omega}$. Así, siempre puede descomponerse (instantáneamente) el movimiento del rígido ya sea en una rotación pura sola, o en una rotación más una translación a lo largo del eje de rotación (movimiento de tirabuzón).

8.1.1. Matrices de rotación

Recordemos brevemente algunas características de estas matrices. Para fijar ideas consideraremos matrices asociadas a espacios vectoriales de dimensión tres. El punto es que si tenemos dos sistemas de coordenadas, S y S' , cuyos orígenes coinciden, pero la orientación de sus ejes es distinta, un mismo vector \mathbf{x} tendrá distintas componentes en cada uno de ellos. Llamemos (x, y, z) a las componentes de \mathbf{x} en S , y (x', y', z') a sus componentes en S' . Si definimos vectores columna

$$\mathbf{x} \equiv \begin{pmatrix} x \\ y \\ z \end{pmatrix}; \quad \mathbf{x}' \equiv \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix},$$

lo que llamamos matriz de rotación \mathbf{R} es la que relaciona ambos conjuntos de componentes como

$$\mathbf{x}' = \mathbf{R} \mathbf{x}.$$

Como el módulo del vector es el mismo en ambos sistemas tenemos que (el supraíndice T indica transposición matricial)

$$\mathbf{x}^T \mathbf{x} = \mathbf{x}'^T \mathbf{x}' = (\mathbf{R} \mathbf{x})^T \mathbf{R} \mathbf{x} = \mathbf{x}^T \mathbf{R}^T \mathbf{R} \mathbf{x},$$

lo que nos dice que ($\mathbf{1}$ es la matriz identidad)

$$\mathbf{R}^T \mathbf{R} = \mathbf{1},$$

por lo que la transpuesta de \mathbf{R} es su propia inversa; esto es, \mathbf{R} es ortogonal. Una propiedad inmediata es que

$$\det(\mathbf{R}^T \mathbf{R}) = \det(\mathbf{R}^T) \det(\mathbf{R}) = [\det(\mathbf{R})]^2 = \det(\mathbf{1}) = 1,$$

por lo que el determinante de \mathbf{R} vale ± 1 . El signo $+$ corresponde a rotaciones puras y el $-$ a rotaciones más inversión de ejes (transformación esta última que también preserva el módulo de los vectores, que fue lo único usado para deducir la ortogonalidad de \mathbf{R}).

Escribamos la ecuación para calcular los autovectores \mathbf{u} de \mathbf{R}

$$\mathbf{R} \mathbf{u} = \lambda \mathbf{u},$$

y su transpuesta conjugada (el supraíndice \dagger indica transposición matricial seguida de conjugación compleja; el supraíndice $*$ indica conjugación compleja)

$$\mathbf{u}^\dagger \mathbf{R}^\dagger = \lambda^* \mathbf{u}^\dagger.$$

Como \mathbf{R} es una matriz real, es $\mathbf{R}^\dagger = \mathbf{R}^T$, por lo que, multiplicando ambas ecuaciones miembro a miembro y usando la condición de ortogonalidad de \mathbf{R} , tenemos

$$\mathbf{u}^\dagger \mathbf{R}^\dagger \mathbf{R} \mathbf{u} = \mathbf{u}^\dagger \mathbf{u} = \lambda^* \lambda \mathbf{u}^\dagger \mathbf{u} = |\lambda|^2 \mathbf{u}^\dagger \mathbf{u},$$

lo que nos dice que $|\lambda|^2 = 1$; todos los autovalores de una matriz ortogonal tienen módulo uno.

La ecuación de los autovalores

$$\det(\mathbf{R} - \lambda \mathbf{1}) = 0,$$

es una ecuación con coeficientes reales (por estar determinados por \mathbf{R} que es real) de manera que sus raíces λ son reales o complejas, y en este caso aparecen de a pares conjugados. En el caso que nos interesa, de dimensión tres, la ecuación es cúbica y debe haber entonces tres raíces (no necesariamente todas distintas). Así, al menos una de las raíces será real y por lo tanto debe valer ± 1 . El signo $-$ indicaría que al aplicar \mathbf{R} al autovector correspondiente se obtiene a éste con sentido invertido, lo cual no corresponde a una rotación, sino a una reflexión; como consideramos sólo matrices de rotaciones puras el signo debe ser $+$. La consecuencia importante es entonces que para una matriz de rotación existe al menos un vector cuyas componentes no cambian al

serle aplicada la matriz; la dirección de este vector invariante corresponde claramente al eje de rotación.

Finalmente, si tenemos una matriz \mathbf{A} generada como el producto matricial de dos matrices de rotación \mathbf{R}_1 y \mathbf{R}_2 , $\mathbf{A} = \mathbf{R}_1\mathbf{R}_2$, tenemos

$$\begin{aligned}\mathbf{A}^T\mathbf{A} &= (\mathbf{R}_1\mathbf{R}_2)^T\mathbf{R}_1\mathbf{R}_2 \\ &= \mathbf{R}_2^T\mathbf{R}_1^T\mathbf{R}_1\mathbf{R}_2 = \mathbf{R}_2^T\mathbf{R}_2 = \mathbf{1},\end{aligned}$$

o sea que el producto de dos matrices ortogonales es también una matriz ortogonal; en particular, si ambas son de rotación pura (determinante +1), la matriz producto también lo será (el determinante de un producto es el producto de los determinantes). Recuérdese que el producto de matrices es no conmutativo ($\mathbf{R}_1\mathbf{R}_2 \neq \mathbf{R}_2\mathbf{R}_1$), salvo que ambas matrices correspondan a rotaciones alrededor del mismo eje.

Más generalmente, sabemos que las matrices ortogonales de dimensión $d \times d$, con la operación producto de matrices, forman un grupo (no conmutativo o no abeliano), el grupo ortogonal $O(d)$, que por depender de parámetros continuos, los ángulos de rotación, es también un grupo de Lie. La restricción a sólo rotaciones puras es un subgrupo de $O(d)$, denominado grupo especial y denotado $SO(d)$. El grupo que nos interesará es el de rotaciones puras en el espacio de dimensión tres, el $SO(3)$.

8.1.2. Ángulos de Euler

Tenemos entonces que la orientación de un cuerpo rígido está dada por la de los ejes del sistema S' solidario al cuerpo, y su posición por la del origen \mathbf{x}_0 de tal sistema. Consideremos que en $t = 0$ el sistema S' coincide con uno fijo al espacio, el S . Al transcurrir el tiempo la translación se describe por la posición del origen de S' , $\mathbf{x}_0(t)$, lo que no presenta dificultades especiales. Concentrémonos en las rotaciones; sabemos que si $\mathbf{x}(t)$ es el vector posición de un punto del cuerpo en el instante t , el vector $\mathbf{x}(t) - \mathbf{x}_0(t)$ tiene siempre las mismas componentes en S' ; la rotación se describirá por cómo cambian estas componentes en el sistema S (nótese que la translación no afecta a estas últimas). Si el cuerpo tiene en t una velocidad angular $\boldsymbol{\Omega}(t)$, el cambio del vector $\mathbf{x}(t) - \mathbf{x}_0(t)$ en un dt es

$$d[\mathbf{x}(t) - \mathbf{x}_0(t)] = \boldsymbol{\Omega}(t) \times [\mathbf{x}(t) - \mathbf{x}_0(t)] dt;$$

por lo que, integrando desde $t = 0$,

$$\mathbf{x}(t) - \mathbf{x}_0(t) = \mathbf{x}(0) - \mathbf{x}_0(0) + \int_0^t \boldsymbol{\Omega}(t') \times [\mathbf{x}(t') - \mathbf{x}_0(t')] dt'.$$

Cada $\boldsymbol{\Omega}(t') \times [\mathbf{x}(t') - \mathbf{x}_0(t')] dt'$ corresponde a una rotación infinitesimal (de eje definido por la dirección de $\boldsymbol{\Omega}(t')$ y ángulo $|\boldsymbol{\Omega}(t')| dt'$), y el valor de $\mathbf{x}(t) - \mathbf{x}_0(t)$ es la composición de todas estas rotaciones, que por la propiedad de grupo es también una rotación. Esto significa que, por complicada que haya sido la sucesión de rotaciones del sólido, el estado resultante en un tiempo cualquiera es el de una rotación equivalente, con su matriz correspondiente, con un dado eje de rotación (dado por el autovector de autovalor 1 de la matriz), tal que un único giro de ángulo apropiado alrededor de este eje lleva la orientación del sólido del estado inicial al actual. Esto se conoce como teorema de Euler. Dado que, obviamente, el origen del sistema S' puede llevarse de su estado inicial al actual a través de una única translación, podemos decir que el estado genérico de un sólido puede ser descrito en cada instante por una dada translación y una dada rotación desde su estado inicial, lo que se denomina teorema de Chasles.

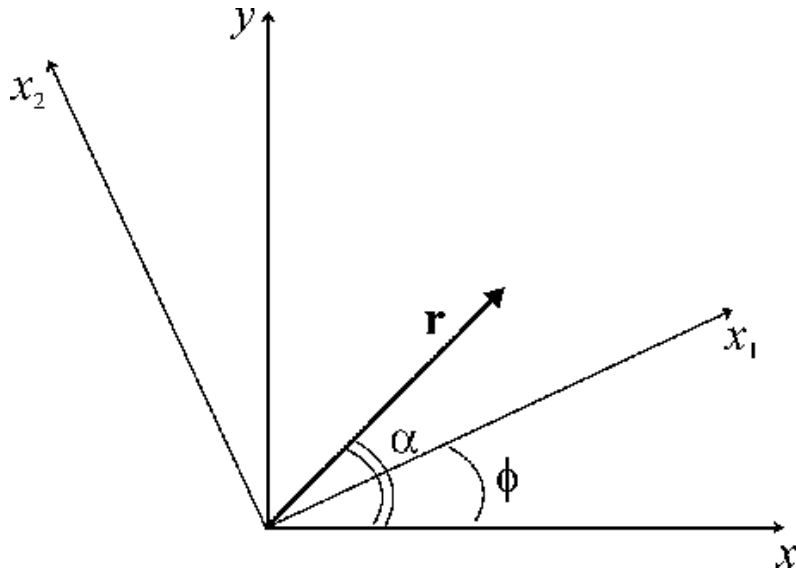
Así, una posible descripción del movimiento del sólido podría consistir en la posición instantánea del origen de S' , el $\mathbf{x}_0(t)$, junto con el eje de la rotación equivalente, dado por el versor $\mathbf{n}(t)$, y el ángulo equivalente $\Phi(t)$ girado alrededor de éste; un total de seis parámetros.

La determinación de $\mathbf{n}(t)$ y $\Phi(t)$ es un problema difícil, y resulta en la práctica más conveniente usar otra parametrización de la rotación equivalente, que es la dada por los ángulos de Euler.

Para definir éstos llamemos (x, y, z) a los ejes del sistema S y (x_1, x_2, x_3) a los de S' . Si un vector \mathbf{r} tiene componentes (x, y, z) en S , y componentes (x_1, x_2, x_3) en S' , sabemos que en cada instante una única rotación, de matriz \mathbf{A} , relaciona a ambas:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (8.5)$$

Veamos cómo sería \mathbf{A} si los ejes z y x_3 coincidieran y S' estuviese girado un ángulo ϕ alrededor de este eje común (en la figura se representa la proyección de \mathbf{r} sobre el plano x, y)



Vemos en la figura que es

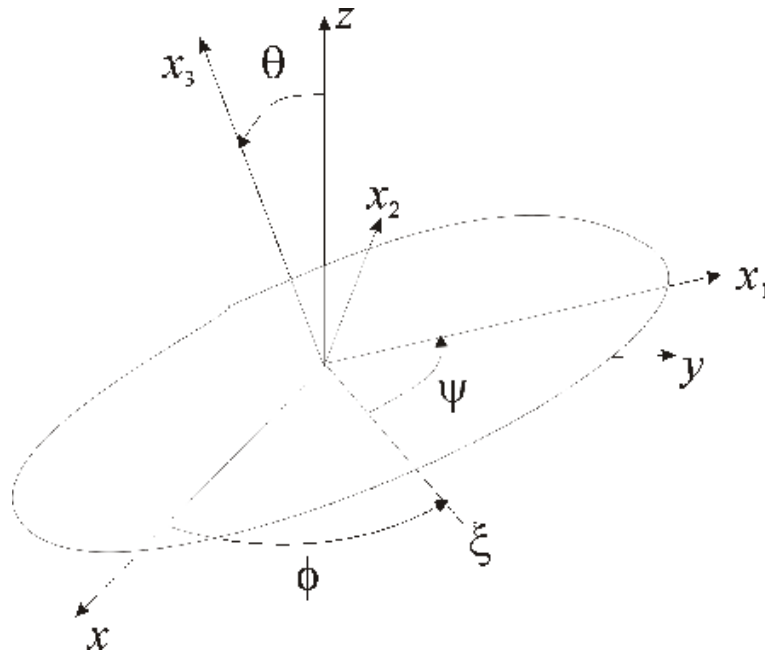
$$\begin{aligned}
 x_1 &= r \cos(\alpha - \phi) = r \cos \alpha \cos \phi + r \sin \alpha \sin \phi \\
 &= x \cos \phi + y \sin \phi, \\
 x_2 &= r \sin(\alpha - \phi) = r \sin \alpha \cos \phi - r \cos \alpha \sin \phi \\
 &= y \cos \phi - x \sin \phi,
 \end{aligned}$$

con lo que podemos escribir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

La parametrización en ángulos de Euler corresponde a escribir la matriz \mathbf{A} como el producto de tres rotaciones de S' similares a la vista (una de ellas es exactamente ésta), parametrizada cada una por un único ángulo:

$$\mathbf{A} = \mathbf{B}(\psi) \mathbf{C}(\theta) \mathbf{D}(\phi). \quad (8.6)$$



$\mathbf{D}(\phi)$ corresponde al giro ya visto de S' con ángulo ϕ alrededor del eje z de S , y tiene entonces componentes

$$\mathbf{D}(\phi) = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad (8.7)$$

el sistema S' así rotado se hace girar ahora un ángulo θ alrededor de su eje x_1 , que se designa convencionalmente ξ en este estado intermedio para no confundirlo con el x_1 definitivo. La matriz correspondiente es entonces

$$\mathbf{C}(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}. \quad (8.8)$$

Finalmente, se gira S' alrededor de su eje x_3 un ángulo ψ , con lo que se tiene

$$\mathbf{B}(\psi) = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (8.9)$$

Nótese que el eje denominado ξ es simultáneamente perpendicular a z y a x_3 , y determina la llamada línea de nodos. El giro arbitrario de S' se parametriza entonces como un giro alrededor de z , seguido de uno alrededor de la línea de nodos ξ , seguido a su vez por otro alrededor de x_3 .

Esta parametrización es conceptualmente sencilla y, además, la velocidad instantánea de rotación $\mathbf{\Omega}$ es muy sencilla de escribir en términos de las velocidades de variación de ϕ , θ y ψ . En efecto, si la orientación de S' pasa de estar descripta en t por ϕ , θ y ψ , a serlo en $t + dt$ por $\phi + d\phi$, $\theta + d\theta$ y $\psi + d\psi$, para calcular $\mathbf{\Omega}$ cada una de estas variaciones puede considerarse que ocurre alrededor del eje correspondiente tal como se encuentra en t , y la velocidad es simplemente la suma vectorial de cada una de estas velocidades.

Así, $\dot{\psi}$ tiene la dirección del eje x_3 ; $\dot{\theta}$ tiene la dirección de la línea de nodos, o sea, $\dot{\theta} \cos \psi$ en x_1 y $-\dot{\theta} \sin \psi$ en x_2 ; finalmente, $\dot{\phi}$ tiene la dirección de z , con lo que es $\dot{\phi} \cos \theta$ en x_3 , $\dot{\phi} \sin \theta \sin \psi$ en x_1 , y $\dot{\phi} \sin \theta \cos \psi$ en x_2 . Reuniendo todo esto, las componentes en S' de la velocidad angular $\mathbf{\Omega}$ son

$$\begin{aligned}\Omega_1 &= \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi, \\ \Omega_2 &= \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi, \\ \Omega_3 &= \dot{\psi} + \dot{\phi} \cos \theta.\end{aligned}\tag{8.10}$$

De igual manera, las componentes en S se obtienen fácilmente. $\dot{\phi}$ tiene la dirección de z ; $\dot{\theta}$ tiene la dirección de la línea de nodos, o sea, $\dot{\theta} \cos \phi$ en x y $\dot{\theta} \sin \phi$ en y ; finalmente, $\dot{\psi}$ tiene la dirección del eje x_3 , con lo que es $\dot{\psi} \cos \theta$ en z , $\dot{\psi} \sin \theta \sin \phi$ en x , y $-\dot{\psi} \sin \theta \cos \phi$ en y . Reuniendo esto es

$$\begin{aligned}\Omega_x &= \dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi, \\ \Omega_y &= \dot{\theta} \sin \phi - \dot{\psi} \sin \theta \cos \phi, \\ \Omega_z &= \dot{\phi} + \dot{\psi} \cos \theta.\end{aligned}\tag{8.11}$$

Desde ya, (8.10) y (8.11) están relacionadas por la forma general (8.5) con la \mathbf{A} dada por (8.6), (8.7), (8.8) y (8.9), como es fácil comprobar.

8.2. Dinámica

8.2.1. Energía cinética

Calculemos la energía cinética del sólido. Para esto denominamos con un subíndice i cada una de las partículas que lo componen. Desde ya, consideramos el sólido como un continuo de materia, pero la notación es menos pesada y más clara si tratamos al cuerpo como un conjunto discreto de partículas; el paso al continuo es evidente. Escribimos entonces la velocidad de cada partícula del cuerpo usando (8.3) como

$$\mathbf{v}_i = \mathbf{v}_0 + \mathbf{\Omega} \times \mathbf{r}_i,\tag{8.12}$$

donde hemos escrito

$$\mathbf{r}_i \equiv \mathbf{x}_i - \mathbf{x}_0,$$

con lo que

$$\begin{aligned} |\mathbf{v}_i|^2 &= (\mathbf{v}_0 + \boldsymbol{\Omega} \times \mathbf{r}_i) \cdot (\mathbf{v}_0 + \boldsymbol{\Omega} \times \mathbf{r}_i) \\ &= |\mathbf{v}_0|^2 + 2\mathbf{v}_0 \cdot (\boldsymbol{\Omega} \times \mathbf{r}_i) + |\boldsymbol{\Omega} \times \mathbf{r}_i|^2. \end{aligned}$$

Usamos ahora la propiedad de conmutatividad cíclica del producto mixto para escribir los últimos dos términos como

$$2\mathbf{v}_0 \cdot (\boldsymbol{\Omega} \times \mathbf{r}_i) = 2\mathbf{r}_i \cdot (\mathbf{v}_0 \times \boldsymbol{\Omega}),$$

y

$$\begin{aligned} |\boldsymbol{\Omega} \times \mathbf{r}_i|^2 &= (\boldsymbol{\Omega} \times \mathbf{r}_i) \cdot (\boldsymbol{\Omega} \times \mathbf{r}_i) \\ &= \mathbf{r}_i \cdot [(\boldsymbol{\Omega} \times \mathbf{r}_i) \times \boldsymbol{\Omega}], \end{aligned} \quad (8.13)$$

respectivamente, y usamos la identidad vectorial

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \quad (8.14)$$

para escribir, de la (8.13),

$$\begin{aligned} (\boldsymbol{\Omega} \times \mathbf{r}_i) \times \boldsymbol{\Omega} &= -\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}_i) \\ &= \mathbf{r}_i |\boldsymbol{\Omega}|^2 - \boldsymbol{\Omega}(\boldsymbol{\Omega} \cdot \mathbf{r}_i), \end{aligned}$$

con lo que la (8.13) resulta

$$|\boldsymbol{\Omega} \times \mathbf{r}_i|^2 = |\mathbf{r}_i|^2 |\boldsymbol{\Omega}|^2 - (\boldsymbol{\Omega} \cdot \mathbf{r}_i)^2.$$

Podemos entonces escribir la energía cinética del sólido

$$\begin{aligned} T &= \frac{1}{2} \sum_i m_i |\mathbf{v}_i|^2 = \frac{1}{2} \sum_i m_i |\mathbf{v}_0|^2 \\ &\quad + \sum_i m_i \mathbf{r}_i \cdot (\mathbf{v}_0 \times \boldsymbol{\Omega}) \\ &\quad + \frac{1}{2} \sum_i m_i [|\mathbf{r}_i|^2 |\boldsymbol{\Omega}|^2 - (\boldsymbol{\Omega} \cdot \mathbf{r}_i)^2]. \end{aligned}$$

Llamando $M = \sum_i m_i$ a la masa total del cuerpo, y teniendo en cuenta que la posición del centro de masas \mathbf{X}_{CM} cumple con $\sum_i m_i \mathbf{x}_i = M\mathbf{X}_{CM}$, reescribimos a T (recordemos que $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_0$)

$$\begin{aligned} T &= \frac{1}{2} M |\mathbf{v}_0|^2 + M (\mathbf{X}_{CM} - \mathbf{x}_0) \cdot (\mathbf{v}_0 \times \boldsymbol{\Omega}) \\ &\quad + \frac{1}{2} \sum_i m_i [|\mathbf{r}_i|^2 |\boldsymbol{\Omega}|^2 - (\boldsymbol{\Omega} \cdot \mathbf{r}_i)^2]. \end{aligned}$$

Se identifican claramente la contribución del movimiento de translación puro (primer término), un término mixto, y una contribución de la rotación pura (último término). Para reescribir este último de manera más útil, considerémoslo en términos de las componentes cartesianas de los vectores intervinientes

$$\begin{aligned}\boldsymbol{\Omega} &= (\Omega_1, \Omega_2, \Omega_3), \\ \mathbf{r}_i &= (r_1^{(i)}, r_2^{(i)}, r_3^{(i)}),\end{aligned}$$

que designamos con un subíndice alfabético (que toma los valores 1, 2 ó 3), para escribir

$$\begin{aligned}|\mathbf{r}_i|^2 |\boldsymbol{\Omega}|^2 &= |\mathbf{r}_i|^2 \Omega_k \Omega_l \delta_{lk}, \\ (\boldsymbol{\Omega} \cdot \mathbf{r}_i)^2 &= \Omega_k \Omega_l r_k^{(i)} r_l^{(i)},\end{aligned}$$

donde se ha usado la delta de Kronecker δ_{lk} , y la convención de Einstein de sumar sobre los índices vectoriales repetidos. Así, podemos escribir

$$\sum_i m_i [|\mathbf{r}_i|^2 |\boldsymbol{\Omega}|^2 - (\boldsymbol{\Omega} \cdot \mathbf{r}_i)^2] = \Omega_k \Omega_l \sum_i m_i [|\mathbf{r}_i|^2 \delta_{lk} - r_k^{(i)} r_l^{(i)}],$$

y definimos el tensor de inercia, referido al origen \mathbf{x}_0 (por ser $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_0$)

$$I_{kl}^0 \equiv \sum_i m_i [|\mathbf{r}_i|^2 \delta_{lk} - r_k^{(i)} r_l^{(i)}], \quad (8.15)$$

que en el caso de distribución continua de materia escribimos como

$$I_{kl}^0 \equiv \int \rho(\mathbf{r}) [|\mathbf{r}|^2 \delta_{lk} - r_k r_l] d^3 r, \quad (8.16)$$

donde $\rho(\mathbf{r})$ es la densidad de masa local del cuerpo, y la integral está extendida a todo su volumen. Notemos que este tensor depende el origen \mathbf{x}_0 respecto del cual se miden los \mathbf{r} de cada punto del cuerpo, y que es una propiedad del cuerpo mismo (no de su estado de movimiento) que se requiere calcular para cada sólido estudiado.

La energía cinética se escribe entonces de manera compacta como

$$T = \frac{1}{2} M |\mathbf{v}_0|^2 + M (\mathbf{X}_{CM} - \mathbf{x}_0) \cdot (\mathbf{v}_0 \times \boldsymbol{\Omega}) + \frac{1}{2} I_{kl}^0 \Omega_k \Omega_l. \quad (8.17)$$

Un punto útil es que si se elige el origen \mathbf{x}_0 coincidiendo con el centro de masas el término mixto se anula y la energía cinética se descompone en dos

términos, uno de translación y otro de rotación (también sucede esto, para cualquier \mathbf{x}_0 , si $\mathbf{v}_0 \parallel \boldsymbol{\Omega}$). Es conveniente entonces elegir $\mathbf{x}_0 = \mathbf{X}_{CM}$, con lo que (no usamos supraíndice alguno cuando el tensor de inercia se refiere al CM)

$$T = \frac{1}{2}M |\mathbf{v}_{CM}|^2 + \frac{1}{2}I_{kl}\Omega_k\Omega_l. \quad (8.18)$$

Si representamos con la matriz \mathbf{I} al tensor de inercia

$$\mathbf{I} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix},$$

y con el vector columna $\boldsymbol{\Omega}$ la velocidad angular

$$\boldsymbol{\Omega} = \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix},$$

podemos escribir en representación matricial la energía cinética de rotación como

$$T_{rot} = \frac{1}{2}\boldsymbol{\Omega}^T\mathbf{I}\boldsymbol{\Omega}.$$

8.2.2. Teorema de Steiner

Útil si conocido el tensor de inercia respecto del CM quisiera calcularse respecto de un origen \mathbf{x}_0 genérico. Supongamos que $\mathbf{x}_0 = \mathbf{X}_{CM} + \mathbf{a}$, con lo que si \mathbf{r} es la posición de un punto genérico del cuerpo respecto del CM , o sea, $\mathbf{x} = \mathbf{X}_{CM} + \mathbf{r}$, tenemos que su posición respecto de \mathbf{x}_0 es

$$\mathbf{r}' = \mathbf{x} - \mathbf{x}_0 = \mathbf{r} - \mathbf{a};$$

de la definición (8.16), debemos calcular

$$\begin{aligned} |\mathbf{r}'|^2 &= |\mathbf{r}|^2 - 2\mathbf{r} \cdot \mathbf{a} + |\mathbf{a}|^2, \\ r'_k r'_l &= r_k r_l - r_k a_l - a_k r_l + a_k a_l. \end{aligned}$$

De los términos a la derecha de estas igualdades los que son lineales en \mathbf{r} dan contribución nula a la integral en (8.16) por ser posiciones referidas al CM :

$$\int \rho(\mathbf{r}) \mathbf{r} d^3r = 0,$$

con lo que

$$\begin{aligned}
 I_{kl}^0 &= \int \rho(\mathbf{r}') \left[|\mathbf{r}'|^2 \delta_{lk} - r'_k r'_l \right] d^3 r' \\
 &= \int \rho(\mathbf{r}) \left[|\mathbf{r}|^2 \delta_{lk} - r_k r_l \right] d^3 r + [|\mathbf{a}|^2 \delta_{lk} - a_k a_l] \int \rho(\mathbf{r}) d^3 r \\
 &= I_{kl} + M [|\mathbf{a}|^2 \delta_{lk} - a_k a_l].
 \end{aligned}$$

Igualdad útil conocida como teorema de Steiner.

8.2.3. Momento angular

Hemos visto al principio del curso que el momento angular \mathbf{L}'_0 respecto de un sistema cuyo origen \mathbf{x}_0 se traslada respecto de un sistema fijo, está relacionado con el referido a este sistema fijo, \mathbf{L} , por (todo lo no primado se refiere a este sistema; $\mathbf{P} = M \dot{\mathbf{X}}_{CM}$)

$$\mathbf{L} = \mathbf{L}'_0 + \mathbf{x}_0 \times \mathbf{P} + M \mathbf{X}_{CM} \times \dot{\mathbf{x}}_0 - M \mathbf{x}_0 \times \dot{\mathbf{x}}_0.$$

En particular, si $\mathbf{x}_0 = \mathbf{X}_{CM}$, esta relación se simplifica a

$$\mathbf{L} = \mathbf{L}'_{CM} + \mathbf{X}_{CM} \times \mathbf{P}. \quad (8.19)$$

Vemos entonces que es muy conveniente calcular el momento angular respecto al centro de masas

$$\mathbf{L}'_{CM} = \sum_i m_i \mathbf{r}_i \times \mathbf{v}'_i,$$

en la que \mathbf{v}'_i es la velocidad de la partícula del cuerpo respecto del centro de masas. Tengamos en cuenta que consideramos un sistema que se traslada con el centro de masas, pero que no rota, de manera que \mathbf{v}'_i contiene sólo la contribución de la rotación del cuerpo:

$$\mathbf{v}'_i = \boldsymbol{\Omega} \times \mathbf{r}_i,$$

con lo que tenemos (sobrentendemos el subíndice CM de \mathbf{L}'_{CM})

$$\mathbf{L}' = \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\Omega} \times \mathbf{r}_i).$$

Aplicando la identidad (8.14) escribimos

$$\mathbf{r}_i \times (\boldsymbol{\Omega} \times \mathbf{r}_i) = \boldsymbol{\Omega} |\mathbf{r}_i|^2 - \mathbf{r}_i (\boldsymbol{\Omega} \cdot \mathbf{r}_i),$$

que en componentes cartesianas es

$$\begin{aligned} [\mathbf{r}_i \times (\boldsymbol{\Omega} \times \mathbf{r}_i)]_k &= \Omega_k |\mathbf{r}_i|^2 - r_k^{(i)} \Omega_l r_l^{(i)} \\ &= \Omega_l \left[|\mathbf{r}_i|^2 \delta_{kl} - r_k^{(i)} r_l^{(i)} \right], \end{aligned}$$

con lo que, finalmente,

$$L'_k = \sum_i m_i \left[|\mathbf{r}_i|^2 \delta_{kl} - r_k^{(i)} r_l^{(i)} \right] \Omega_l = I_{kl} \Omega_l. \quad (8.20)$$

En representación matricial es

$$\mathbf{L}' = \mathbf{I}\boldsymbol{\Omega}.$$

8.2.4. Ejes principales del tensor de inercia

El tensor de inercia es entonces fundamental para determinar la energía cinética y el momento angular del sólido. De su definición (8.15) o (8.16) vemos que es un tensor simétrico; como también es real sabemos entonces que puede ser siempre diagonalizable; esto es, existen tres ejes ortogonales, tal que si es calculado respecto de ellos su expresión es una matriz diagonal (los elementos de la diagonal son en general distintos); estos ejes se denominan *ejes principales*. Tomemos los ejes principales como los del sistema fijo a cuerpo, S' , respecto de los cuales podemos escribir (en notación matricial)

$$\mathbf{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}.$$

Para un cuerpo genérico los ejes principales se obtienen calculando el tensor de inercia en un sistema ortogonal dado y luego buscando sus autovectores. Éstos definen entonces la matriz ortogonal que representa la transformación desde los ejes originales usados a los ejes principales.

Si el cuerpo tiene una distribución de masa que es simétrica respecto de un plano, es fácil ver que dos de los ejes principales están contenidos en este plano y el tercero es perpendicular a él. En efecto, si elegimos el eje x_3 perpendicular al plano de simetría de la distribución de masa del cuerpo, y su origen en él, tenemos que, para $l \neq 3$,

$$I_{3l} = - \int \rho(\mathbf{r}) r_3 r_l d^3r = 0,$$

ya que para cada valor de r_l existen valores simétricos de $\rho(\mathbf{r})$ por encima y por debajo del plano de simetría, en los cuales r_3 toma valores de igual módulo y signo distinto. La forma del tensor en este sistema de ejes es entonces

$$\mathbf{I} = \begin{pmatrix} I_{11} & I_{12} & 0 \\ I_{21} & I_{22} & 0 \\ 0 & 0 & I_3 \end{pmatrix},$$

por lo que basta diagonalizar la submatriz correspondiente al plano de simetría.

Un subcaso particular importante corresponde a distribución de masa con simetría de revolución. En tal caso cualquier plano que contiene al eje de revolución es un plano de simetría, lo que nos dice que este eje es un eje principal y, además, que cualesquiera dos ejes perpendiculares a éste (y por lo tanto contenidos cada uno de ellos en alguno de los infinitos planos de simetría) son también ejes principales.

Expresados el tensor de inercia y la velocidad angular en componentes sobre los ejes principales, las expresiones de la energía cinética y del momento angular son muy sencillas:

$$T = \frac{1}{2}M |\mathbf{v}_{CM}|^2 + \frac{1}{2} (I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2), \quad (8.21)$$

y

$$\mathbf{L}' = I_1\Omega_1\mathbf{e}_1 + I_2\Omega_2\mathbf{e}_2 + I_3\Omega_3\mathbf{e}_3. \quad (8.22)$$

Damos a continuación una lista de tensores de inercia correspondientes a sólidos con densidad de masa uniforme y diferentes formas, referidos a sus ejes principales respectivos; el origen está siempre en el centro de masas (la masa del cuerpo se denomina M en todos los casos):

1) Varilla fina de longitud l , eje x_3 a lo largo de ella:

$$\begin{aligned} I_1 &= I_2 = \frac{Ml^2}{12}, \\ I_3 &= 0. \end{aligned}$$

2) Disco delgado de radio R , eje x_3 perpendicular al disco:

$$\begin{aligned} I_1 &= I_2 = \frac{MR^2}{4}, \\ I_3 &= \frac{MR^2}{2}. \end{aligned}$$

3) Cilindro de radio R y longitud l , eje x_3 a lo largo de él:

$$\begin{aligned} I_1 &= I_2 = \frac{M}{4} \left(R^2 + \frac{l^2}{3} \right), \\ I_3 &= \frac{MR^2}{2}. \end{aligned}$$

4) Esfera de radio R :

$$I_1 = I_2 = I_3 = \frac{2}{5}MR^2.$$

5) Paralelepípedo de lado a en la dirección de x_1 , lado b en la de x_2 , y lado c en la de x_3 :

$$\begin{aligned} I_1 &= \frac{M}{12}(b^2 + c^2), \\ I_2 &= \frac{M}{12}(a^2 + c^2), \\ I_3 &= \frac{M}{12}(a^2 + b^2). \end{aligned}$$

6) Elipsoide de semieje a en la dirección de x_1 , semieje b en la de x_2 , y semieje c en la de x_3 :

$$\begin{aligned} I_1 &= \frac{M}{5}(b^2 + c^2), \\ I_2 &= \frac{M}{5}(a^2 + c^2), \\ I_3 &= \frac{M}{5}(a^2 + b^2). \end{aligned}$$

7) Cono circular de altura h y radio de base R , eje x_3 a lo largo de su eje. El centro de masas está sobre el eje a $h/4$ de la base.

$$\begin{aligned} I_1 &= I_2 = \frac{3}{20}M \left(R^2 + \frac{h^2}{4} \right), \\ I_3 &= \frac{3}{10}MR^2. \end{aligned}$$

8.3. Ecuaciones de Euler

Las ecuaciones de Euler son simplemente la ley de variación del momento angular. Si \mathbf{K} es la cupla actuante sobre el sólido, referida al origen de S , sabemos que

$$\frac{d\mathbf{L}}{dt} = \mathbf{K}, \quad (8.23)$$

que usando (8.19) escribimos como (siempre sobreentendemos el subíndice CM de \mathbf{L}')

$$\frac{d\mathbf{L}'}{dt} + \dot{\mathbf{X}}_{CM} \times \mathbf{P} + \mathbf{X}_{CM} \times \dot{\mathbf{P}} = \mathbf{K}.$$

Por supuesto, el segundo término del lado izquierdo es nulo y $\dot{\mathbf{P}}$ es igual a la fuerza \mathbf{F} sobre el sólido, por lo que tenemos

$$\frac{d\mathbf{L}'}{dt} = \mathbf{K} - \mathbf{X}_{CM} \times \mathbf{F} \equiv \mathbf{K}'.$$

La derivada temporal corresponde a la variación, referida al sistema S , de \mathbf{L}' en el tiempo, que podemos escribir en términos de la variación referida a S'

$$\left. \frac{d\mathbf{L}'}{dt} \right)_S = \left. \frac{d\mathbf{L}'}{dt} \right)_{S'} + \boldsymbol{\Omega} \times \mathbf{L}',$$

que nos permite escribir

$$\left. \frac{d\mathbf{L}'}{dt} \right)_{S'} + \boldsymbol{\Omega} \times \mathbf{L}' = \mathbf{K}'.$$

Si el sistema S' es el de los ejes principales de inercia estas ecuaciones tienen una representación muy sencilla. En estos ejes es, de (8.22),

$$\begin{aligned} L'_1 &= I_1 \Omega_1, \\ L'_2 &= I_2 \Omega_2, \\ L'_3 &= I_3 \Omega_3, \end{aligned}$$

y es justamente la variación de estas componentes la que corresponde a la variación de \mathbf{L}' respecto de S' ; o sea,

$$\left. \frac{d\mathbf{L}'}{dt} \right)_{S'} = I_1 \frac{d\Omega_1}{dt} \mathbf{e}_1 + I_2 \frac{d\Omega_2}{dt} \mathbf{e}_2 + I_3 \frac{d\Omega_3}{dt} \mathbf{e}_3,$$

con lo que (haciendo el producto vectorial $\boldsymbol{\Omega} \times \mathbf{L}'$) podemos escribir en componentes sobre los ejes principales

$$\begin{aligned} I_1 \frac{d\Omega_1}{dt} + \Omega_2 \Omega_3 (I_3 - I_2) &= K'_1, \\ I_2 \frac{d\Omega_2}{dt} + \Omega_1 \Omega_3 (I_1 - I_3) &= K'_2, \\ I_3 \frac{d\Omega_3}{dt} + \Omega_1 \Omega_2 (I_2 - I_1) &= K'_3, \end{aligned} \tag{8.24}$$

que son las ecuaciones de Euler.

Notemos que estamos usando ejes móviles (fijos al cuerpo) para escribir sobre ellos las componentes de una ecuación vectorial válida en el sistema de referencia de ejes fijos al espacio, S .

Si en estas ecuaciones introducimos las expresiones (8.10) y damos la forma explícita de las cuplas K'_i tendremos ecuaciones para los ángulos de Euler, la solución de las cuales nos da la orientación del cuerpo al transcurrir el tiempo. Por supuesto, siempre podemos escribir las ecuaciones de Lagrange usando la expresión (8.21) para la energía cinética junto con las (8.10) y obtendríamos ecuaciones equivalentes. Sin embargo, las ecuaciones de Euler son en sí mismas de primer orden de derivación y pueden muchas veces integrarse para obtener las componentes de $\mathbf{\Omega}$, lo que lleva a que las (8.10) resulten ecuaciones de primer orden para los ángulos de Euler. Además, el conocimiento de sólo $\mathbf{\Omega}$ proporciona información útil por sí misma.

8.4. Movimiento del cuerpo sólido libre

8.4.1. Construcción de Poinsot

Consideremos el movimiento general de un sólido arbitrario no sometido a fuerzas ni torques externos. En particular, aprovecharemos que el centro de masas se mueve con velocidad constante para describir el movimiento desde un sistema en el que el centro de masas está quieto. La energía cinética, que es constante, tiene en este sistema la contribución rotacional solamente, que usando ejes principales podemos escribir

$$T = \frac{1}{2} (I_1 \Omega_1^2 + I_2 \Omega_2^2 + I_3 \Omega_3^2). \quad (8.25)$$

Asimismo, el momento angular, también constante, corresponde al \mathbf{L}'_{CM} :

$$\mathbf{L} = I_1 \Omega_1 \mathbf{e}_1 + I_2 \Omega_2 \mathbf{e}_2 + I_3 \Omega_3 \mathbf{e}_3. \quad (8.26)$$

Notemos que podemos escribir la relación útil

$$T = \frac{1}{2} \mathbf{L} \cdot \mathbf{\Omega}. \quad (8.27)$$

La constancia de (8.25) y (8.26) permite hacer una descripción geométrica del movimiento a partir de la llamada construcción de Poinsot. Para esto escribamos (8.25) como

$$\frac{\Omega_1^2}{\frac{2T}{I_1}} + \frac{\Omega_2^2}{\frac{2T}{I_2}} + \frac{\Omega_3^2}{\frac{2T}{I_3}} = 1,$$

que corresponde, en un espacio de ejes $(\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3)$, a un elipsoide de semiejes $\sqrt{2T/I_1}$, $\sqrt{2T/I_2}$ y $\sqrt{2T/I_3}$, denominado *elipsoide de inercia*.

Si consideramos la función

$$F(\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3) \equiv \frac{\tilde{\Omega}_1^2}{I_1} + \frac{\tilde{\Omega}_2^2}{I_2} + \frac{\tilde{\Omega}_3^2}{I_3},$$

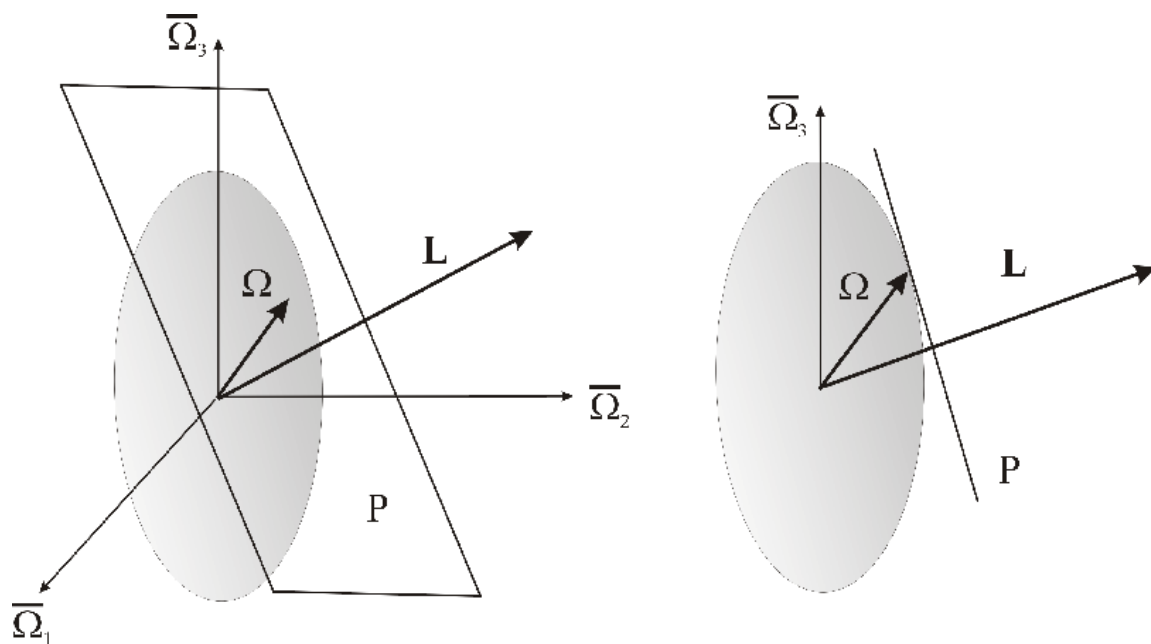
tenemos que $F(\Omega_1, \Omega_2, \Omega_3) = 1$ y que, por ejemplo,

$$\frac{\partial F}{\partial \tilde{\Omega}_1} = \frac{I_1 \tilde{\Omega}_1}{T};$$

con lo que, como $I_1 \Omega_1 = L_1$, el gradiente de F evaluado en $(\Omega_1, \Omega_2, \Omega_3)$ es paralelo a \mathbf{L} . Así, el plano tangente al elipsoide de inercia en el punto $(\Omega_1, \Omega_2, \Omega_3)$ del espacio $(\tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3)$ es perpendicular a \mathbf{L} . Como \mathbf{L} es constante, este plano es fijo y el elipsoide se mueve teniendo un punto de contacto sobre él. Es más, como por el punto de contacto pasa el eje de giro instantáneo ($\boldsymbol{\Omega}$ une el centro del elipsoide con el punto de contacto) este punto está instantáneamente en reposo, por lo que el elipsoide de inercia rueda sin deslizar sobre el plano invariante. Finalmente, la “distancia” del centro del elipsoide al plano está dada por

$$\boldsymbol{\Omega} \cdot \frac{\mathbf{L}}{|\mathbf{L}|} = \frac{2T}{|\mathbf{L}|},$$

(por (8.27)) que es claramente constante.



Esta construcción geométrica de Poincaré nos dice entonces que el elipsoide de inercia se mueve sobre el plano invariante perpendicular a \mathbf{L} , rodando por él sin deslizar y manteniendo fija la distancia de su centro al plano. Como el elipsoide de inercia está asociado unívocamente al cuerpo (el eje $\tilde{\Omega}_1$ es paralelo al x_1 , etc.), el movimiento de aquél nos determina el de éste.

La curva que “pinta” el punto de contacto sobre el elipsoide es denominada *polodia*, la que pinta sobre el plano *herpolodia*.

8.4.2. Otra representación geométrica

Otra representación geométrica útil puede hacerse en el espacio de ejes de momento angular $(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3)$. En efecto, escribiendo que

$$|\mathbf{L}|^2 = L_1^2 + L_2^2 + L_3^2,$$

y

$$2T = \frac{L_1^2}{I_1} + \frac{L_2^2}{I_2} + \frac{L_3^2}{I_3},$$

es claro que el movimiento debe efectuarse en el espacio $(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3)$ en la intersección de la esfera de radio $|\mathbf{L}|$ y el elipsoide de semiejes $\sqrt{2TI_1}$, $\sqrt{2TI_2}$ y $\sqrt{2TI_3}$. Que estos cuerpos se intersectan se ve al considerar que, tomando $I_1 > I_2 > I_3$,

$$\begin{aligned} 2TI_1 &= I_1^2\Omega_1^2 + I_1I_2\Omega_2^2 + I_1I_3\Omega_3^2 \\ &\geq I_1^2\Omega_1^2 + I_2^2\Omega_2^2 + I_3^2\Omega_3^2 = |\mathbf{L}|^2, \end{aligned}$$

y, análogamente,

$$2TI_3 \leq |\mathbf{L}|^2.$$

La curva que resulta de la intersección nos indica los puntos por los que transita el vector \mathbf{L} en el espacio $(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3)$ al moverse el sólido libre. Por supuesto, \mathbf{L} es un vector constante en el sistema fijo al espacio, pero sus componentes sobre el sistema fijo al cuerpo cambian siguiendo la curva mencionada. El movimiento del cuerpo es entonces interpretable teniendo en cuenta que el eje \tilde{L}_1 es paralelo al x_1 , etc..

8.4.3. Estabilidad de la rotación alrededor de los ejes principales

De las construcciones anteriores es claro que si \mathbf{L} es paralelo a alguno de los ejes principales el sólido permanece siempre en esa posición. Estudiemos entonces la estabilidad de este tipo particular de movimiento libre, proponiendo

que la velocidad angular tiene su componente más importante a lo largo de un eje principal, con componentes infinitesimales sobre los demás. Tenemos entonces, por ejemplo, para rotación a lo largo del eje x_1

$$\boldsymbol{\Omega} = (\Omega_0 + \varepsilon_1) \mathbf{e}_1 + \varepsilon_2 \mathbf{e}_2 + \varepsilon_3 \mathbf{e}_3,$$

con los ε 's velocidades infinitesimales. Elegimos además los ejes principales para que se satisfaga

$$I_1 > I_2 > I_3. \quad (8.28)$$

Escribimos ahora las ecuaciones de Euler (8.24) reteniendo sólo términos de primer orden en los ε 's

$$\begin{aligned} I_1 \frac{d\varepsilon_1}{dt} &= 0, \\ I_2 \frac{d\varepsilon_2}{dt} + (I_1 - I_3) \Omega_0 \varepsilon_3 &= 0, \\ I_3 \frac{d\varepsilon_3}{dt} + (I_2 - I_1) \Omega_0 \varepsilon_2 &= 0. \end{aligned}$$

De la primera es sencillamente $\varepsilon_1 = cte$. Mientras que si derivamos respecto del tiempo una de las otras y usamos la restante para eliminar la derivada primera obtenemos inmediatamente

$$\begin{aligned} \frac{d^2\varepsilon_2}{dt^2} + \frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3} \Omega_0^2 \varepsilon_2 &= 0, \\ \frac{d^2\varepsilon_3}{dt^2} + \frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3} \Omega_0^2 \varepsilon_3 &= 0. \end{aligned}$$

Por las condiciones (8.28) vemos que el factor que multiplica a ε_2 y ε_3 en las ecuaciones anteriores es positivo, por lo que ε_2 y ε_3 oscilan armónicamente alrededor del valor cero con frecuencia angular

$$\omega_1 = \Omega_0 \sqrt{\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3}}.$$

Podemos decir entonces que pequeños apartamientos iniciales del giro puro alrededor del eje x_1 siguen permaneciendo pequeños al transcurrir el tiempo.

Si repetimos esta deducción para giros alrededor del eje x_3 ($\boldsymbol{\Omega} = \varepsilon_1 \mathbf{e}_1 + \varepsilon_2 \mathbf{e}_2 + (\Omega_0 + \varepsilon_3) \mathbf{e}_3$), obtenemos que ε_1 y ε_2 oscilan armónicamente alrededor del valor cero con frecuencia angular

$$\omega_3 = \Omega_0 \sqrt{\frac{(I_3 - I_1)(I_3 - I_2)}{I_1 I_2}}.$$

Sin embargo, para giros alrededor del eje con valor intermedio del momento de inercia, el x_2 , al proponer que $\boldsymbol{\Omega} = \varepsilon_1 \mathbf{e}_1 + (\Omega_0 + \varepsilon_2) \mathbf{e}_2 + \varepsilon_3 \mathbf{e}_3$ obtenemos para ε_1 y ε_3

$$\frac{d^2 \varepsilon_{1,3}}{dt^2} + \frac{(I_2 - I_3)(I_2 - I_1)}{I_1 I_3} \Omega_0^2 \varepsilon_{1,3} = 0,$$

en la que el factor que multiplica a $\varepsilon_{1,3}$ es ahora negativo, con lo que la solución de esta ecuación corresponde a exponenciales, una creciente y otra decreciente, con tiempo característico τ_2 dado por

$$\tau_2^{-1} = \Omega_0 \sqrt{\frac{(I_2 - I_3)(I_1 - I_2)}{I_1 I_3}}.$$

La rotación alrededor del eje x_2 es entonces inestable; un apartamiento inicial pequeño no permanece así.

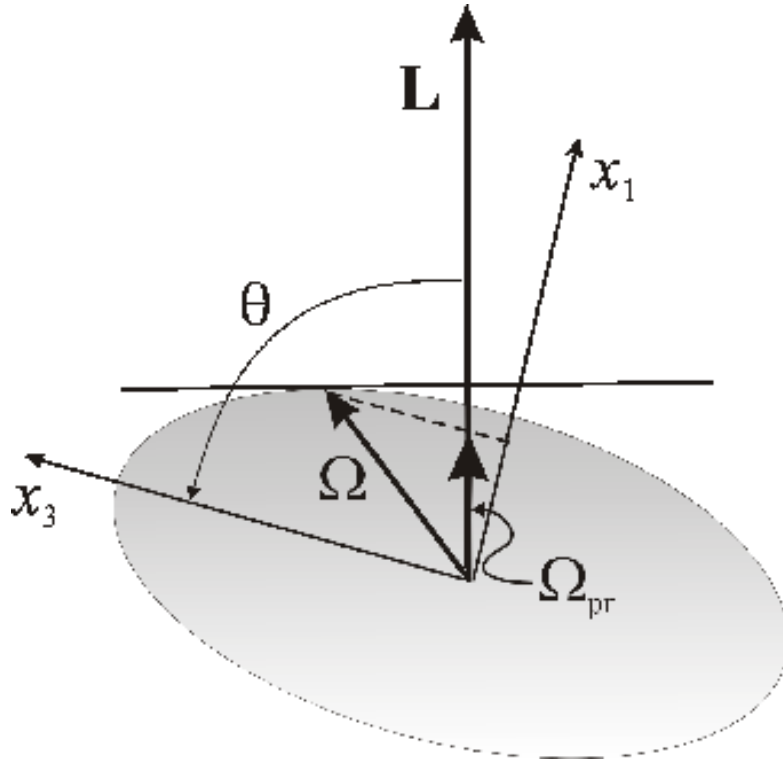
De la construcción geométrica en el espacio $(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3)$ puede verse lo anterior. Cuando $\boldsymbol{\Omega}$ tiene componente sólo sobre uno de los ejes principales, la intersección entre la esfera y el elipsoide es un punto, por lo que el cuerpo permanece rotando alrededor del eje considerado. Cuando existen pequeñas componentes sobre los otros ejes, sin embargo, la intersección es una curva cerrada muy similar a una elipse cuando la componente principal de $\boldsymbol{\Omega}$ es sobre los ejes x_1 o x_3 , con lo que el movimiento será acotado alrededor del eje dado. Si la componente más importante de $\boldsymbol{\Omega}$ es sobre el eje x_2 , la curva intersección de esfera y elipsoide es similar a una hipérbola, con lo que el vector \mathbf{L} se aleja de x_2 .

8.4.4. Elipsoide con simetría de revolución

Un caso relativamente sencillo de estudiar es el del sólido con dos momentos de inercia iguales, digamos I_1 e I_2 , con lo que el elipsoide de inercia asociado es de revolución alrededor del eje $\tilde{\Omega}_3$. En este caso, es inmediato ver que la rodadura del elipsoide en las condiciones de Poinot describe una herpolodia circular alrededor de \mathbf{L} , con lo que $\boldsymbol{\Omega}$ precede alrededor del vector fijo \mathbf{L} manteniendo un ángulo constante con éste y con el eje $\tilde{\Omega}_3$ (o el x_3), por lo que Ω_3 es también constante. Llamando θ al ángulo fijo entre x_3 y \mathbf{L} tenemos que

$$\Omega_3 = \frac{L_3}{I_3} = \frac{|\mathbf{L}| \cos \theta}{I_3}.$$

Nótese que Ω_3 **no** es la velocidad de rotación del cuerpo sobre su eje x_3 . En realidad, Ω_3 es esta velocidad más la componente sobre x_3 de la velocidad angular de precesión Ω_{pr} alrededor de \mathbf{L} .



Si elegimos en el instante considerado al eje x_1 en el plano de \mathbf{L} y x_3 , lo que es siempre posible por ser el elipsoide de revolución, vemos en la figura que

$$\Omega_{pr} \sin \theta = \Omega_1 = \frac{L_1}{I_1} = \frac{|\mathbf{L}| \sin \theta}{I_1},$$

con lo cual,

$$\Omega_{pr} = \frac{|\mathbf{L}|}{I_1}. \quad (8.29)$$

La velocidad de rotación del cuerpo alrededor de su eje x_3 , que en términos de los ángulos de Euler es $\dot{\psi}$, será entonces

$$\dot{\psi} = \Omega_3 - \Omega_{pr} \cos \theta = \left(\frac{1}{I_3} - \frac{1}{I_1} \right) |\mathbf{L}| \cos \theta. \quad (8.30)$$

Veamos esto mismo usando las ecuaciones de Euler. Como $I_1 = I_2$ las ecuaciones de Euler (8.24) se escriben en este caso

$$\begin{aligned} I_1 \frac{d\Omega_1}{dt} + \Omega_2 \Omega_3 (I_3 - I_1) &= 0, \\ I_1 \frac{d\Omega_2}{dt} - \Omega_1 \Omega_3 (I_3 - I_1) &= 0, \\ I_3 \frac{d\Omega_3}{dt} &= 0, \end{aligned}$$

de la última de las cuales es inmediato que $\Omega_3 = cte$. Definiendo

$$\omega_0 \equiv \Omega_3 \frac{I_3 - I_1}{I_1},$$

escribimos las dos primeras como

$$\begin{aligned} \frac{d\Omega_1}{dt} + \omega_0 \Omega_2 &= 0, \\ \frac{d\Omega_2}{dt} - \omega_0 \Omega_1 &= 0. \end{aligned}$$

Si definimos la función compleja $W \equiv \Omega_1 + i\Omega_2$, multiplicando la segunda de las ecuaciones anteriores por la unidad imaginaria i , y sumando ambas resulta

$$\frac{dW}{dt} - i\omega_0 W = 0,$$

cuya solución inmediata es

$$W = W_0 \exp(i\omega_0 t),$$

con $W_0 = \Omega_{01} + i\Omega_{02}$. Elijamos los ejes para que en $t = 0$ sea $\Omega_{02} = 0$, con lo que, tomando partes real e imaginaria de W tenemos en este caso

$$\begin{aligned} \Omega_1 &= \Omega_{01} \cos(\omega_0 t), \\ \Omega_2 &= \Omega_{01} \sin(\omega_0 t). \end{aligned}$$

De esta manera conocemos $\boldsymbol{\Omega}(t)$ en componentes sobre el sistema móvil. Debemos todavía determinar cómo es el movimiento del sistema móvil.

Usemos que \mathbf{L} es constante y elijamos el eje fijo z con la dirección y sentido de \mathbf{L} para escribir $\mathbf{L} = |\mathbf{L}| \mathbf{e}_z$. Las componentes de \mathbf{L} en el sistema móvil estarán dadas entonces (ver (8.5)) por

$$\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \mathbf{A} \begin{pmatrix} 0 \\ 0 \\ |\mathbf{L}| \end{pmatrix},$$

que, usando las expresiones (8.6), (8.7), (8.8) y (8.9), resulta fácilmente en

$$\begin{aligned} L_1 &= |\mathbf{L}| \sin \theta \sin \psi, \\ L_2 &= |\mathbf{L}| \sin \theta \cos \psi, \\ L_3 &= |\mathbf{L}| \cos \theta. \end{aligned} \tag{8.31}$$

Como $L_3 = I_3\Omega_3$ tenemos de la última ecuación que

$$\cos \theta = \frac{I_3\Omega_3}{|\mathbf{L}|} \quad (8.32)$$

es constante; o sea $\dot{\theta} = 0$, con lo que las (8.10) se escriben en este caso

$$\begin{aligned} \dot{\phi} \sin \theta \sin \psi &= \Omega_1, \\ \dot{\phi} \sin \theta \cos \psi &= \Omega_2, \\ \dot{\psi} + \dot{\phi} \cos \theta &= \Omega_3. \end{aligned} \quad (8.33)$$

Como $L_1 = I_1\Omega_1$ y $L_2 = I_1\Omega_2$, cualquiera de las dos primeras de las (8.33) y la correspondiente de las (8.31) dan

$$\dot{\phi} = \frac{|\mathbf{L}|}{I_1}, \quad (8.34)$$

que es justamente la velocidad de giro del sistema móvil alrededor del eje z (o \mathbf{L}); esto es, la velocidad angular de precesión Ω_{pr} , la (8.29) obtenida antes.

Usando la última de las (8.33), la (8.32) y la (8.34) tenemos entonces la velocidad de rotación del cuerpo alrededor de su eje x_3 ,

$$\dot{\psi} = \Omega_3 - \dot{\phi} \cos \theta = \left(\frac{1}{I_3} - \frac{1}{I_1} \right) |\mathbf{L}| \cos \theta,$$

que es justamente la (8.30).

8.5. Movimiento de trompos y giróscopos

8.5.1. Trompo

Consideremos el movimiento del trompo en presencia de gravedad, con el punto de apoyo fijo. Aprovechando esto último, es conveniente elegir el sistema fijo al cuerpo con origen en este punto al que designaremos O . Como $\mathbf{v}_0 = 0$ la (8.17) se escribe sencillamente como

$$T = \frac{1}{2} I_{kl}^0 \Omega_k \Omega_l.$$

Como el sólido es de revolución un eje principal es el de revolución mismo, que llamamos x_3 , al que corresponde momento de inercia I_3 respecto al CM . Por supuesto es $I_1 = I_2$, también respecto al CM . Si el sólido tiene masa

M y el CM se encuentra a una distancia l (a lo largo del eje) del punto de apoyo, el teorema de Steiner nos permite calcular inmediatamente que

$$\begin{aligned} I_1^0 &= I_2^0 = I_1 + Ml^2, \\ I_3^0 &= I_3, \end{aligned}$$

con todo lo cual podemos escribir que

$$T = \frac{1}{2} (I_1 + Ml^2) (\Omega_1^2 + \Omega_2^2) + \frac{1}{2} I_3 \Omega_3^2.$$

Usando las (8.10) se obtiene fácilmente

$$\Omega_1^2 + \Omega_2^2 = \dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2.$$

Como además la energía potencial se escribe

$$V = Mgl \cos \theta,$$

el lagrangiano es entonces

$$L = \frac{1}{2} (I_1 + Ml^2) (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2} I_3 (\dot{\psi} + \dot{\phi} \cos \theta)^2 - Mgl \cos \theta.$$

En lugar de escribir las ecuaciones de Lagrange basta ver que ϕ y ψ son variables cíclicas, y que L no depende explícitamente del tiempo. Tenemos entonces tres constantes de movimiento

$$\frac{\partial L}{\partial \dot{\phi}} = (I_1 + Ml^2) \dot{\phi} \sin^2 \theta + I_3 (\dot{\psi} + \dot{\phi} \cos \theta) \cos \theta = L_z, \quad (8.35)$$

$$\frac{\partial L}{\partial \dot{\psi}} = I_3 (\dot{\psi} + \dot{\phi} \cos \theta) = L_3, \quad (8.36)$$

y

$$\dot{\phi} \frac{\partial L}{\partial \dot{\phi}} + \dot{\theta} \frac{\partial L}{\partial \dot{\theta}} + \dot{\psi} \frac{\partial L}{\partial \dot{\psi}} - L = E,$$

que, como T es homogéneo de orden dos en las velocidades, resulta simplemente igual a $T + V$:

$$\frac{1}{2} (I_1 + Ml^2) (\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2} I_3 (\dot{\psi} + \dot{\phi} \cos \theta)^2 + Mgl \cos \theta = E. \quad (8.37)$$

De las (8.35) y (8.36) tenemos inmediatamente que

$$\dot{\psi} + \dot{\phi} \cos \theta = \frac{L_3}{I_3}, \quad (8.38)$$

$$\dot{\phi} = \frac{L_z - L_3 \cos \theta}{(I_1 + Ml^2) \sin^2 \theta}, \quad (8.39)$$

que al reemplazar en (8.37) nos da

$$\frac{1}{2} (I_1 + Ml^2) \left[\dot{\theta}^2 + \frac{(L_z - L_3 \cos \theta)^2}{(I_1 + Ml^2)^2 \sin^2 \theta} \right] + Mgl \cos \theta = E - \frac{L_3^2}{2I_3} \equiv E'.$$

Definiendo las constantes

$$\begin{aligned} a &\equiv \frac{L_3}{I_1 + Ml^2}, \\ b &\equiv \frac{L_z}{I_1 + Ml^2}, \\ \alpha &\equiv \frac{2E'}{I_1 + Ml^2}, \\ \beta &\equiv \frac{2Mgl}{I_1 + Ml^2}, \end{aligned}$$

esta ecuación se reescribe (después de multiplicar por $\sin^2 \theta$ y reordenar)

$$\dot{\theta}^2 \sin^2 \theta = \alpha \sin^2 \theta - \beta \cos \theta \sin^2 \theta - (b - a \cos \theta)^2.$$

Conviene escribir esta ecuación diferencial para θ en términos de la variable transformada

$$\mu = \cos \theta,$$

para la cual es

$$\dot{\mu} = -\dot{\theta} \sin \theta,$$

con lo que resulta inmediatamente

$$\begin{aligned} \dot{\mu}^2 &= \alpha (1 - \mu^2) - \beta \mu (1 - \mu^2) - (b - a\mu)^2 \\ &= (\alpha - \beta\mu) (1 - \mu^2) - (b - a\mu)^2 \equiv F(\mu). \end{aligned} \quad (8.40)$$

El movimiento en μ corresponde entonces al movimiento unidimensional de una partícula con “masa” de valor 2 en un potencial efectivo $V_{ef}(\mu) = -F(\mu)$ y energía mecánica cero:

$$\dot{\mu}^2 + V_{ef}(\mu) = 0. \quad (8.41)$$

Por supuesto, los valores físicos de $\mu = \cos \theta$, están limitados al intervalo $[-1, 1]$ y a valores de $V_{ef} < 0$, o $F > 0$. Es importante entonces analizar los ceros (o raíces) de F , que serán puntos de retorno del movimiento, entre los que se desarrollará éste.

La función F es polinómica de grado tres en μ , y cuando $|\mu| \rightarrow \infty$, $F \rightarrow \beta\mu^3$; como $\beta > 0$, F es positiva para μ positivo y suficientemente grande y

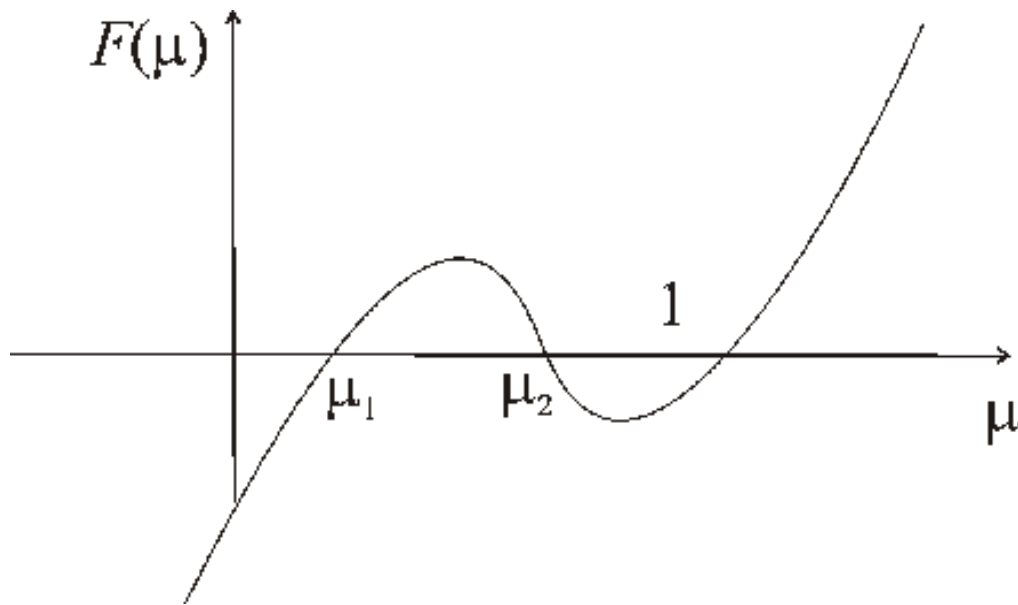
negativa para valores de μ suficientemente negativos. Además, $F(\mu = \pm 1) = -(b \mp a)^2 < 0$ (salvo el caso $b = a$; o sea, $L_3 = L_z$, que corresponde a un trompo vertical, que no consideramos ahora), por todo lo cual la función tiene al menos una raíz en $\mu > 1$. Las distintas posibilidades son entonces:

i) Las otras dos raíces son complejas conjugadas. Significa entonces que $F < 0$ en $[-1, 1]$ y no existe entonces solución física.

ii) Las otras dos raíces son reales y están ambas fuera del intervalo $[-1, 1]$, con lo que $F < 0$ en $[-1, 1]$ como en i).

iii) Ambas raíces están dentro del intervalo $[-1, 1]$, llamémoslas $\mu_{1,2}$, por lo que entre ellas es $F > 0$ y el movimiento real es posible.

Tenemos entonces que sólo es posible para un trompo real el caso iii), con movimiento limitado entre $\mu_1 = \cos \theta_1$ y $\mu_2 = \cos \theta_2$. Este movimiento periódico de “cabeceo” del trompo entre estos ángulos es denominado de nutación.



Como (8.39), reescrita en términos de las constantes definidas es

$$\dot{\phi} = \frac{b - a\mu}{1 - \mu^2}, \quad (8.42)$$

vemos, ante todo, que la variación periódica de μ inducirá una variación, también periódica de la velocidad de precesión $\dot{\phi}$; además, ésta cambia de signo si $b - a\mu$ lo hace en el intervalo $[\mu_1, \mu_2]$, lo que da tres distintos regímenes de precesión:

i) $b - a\mu$ no se anula ni en el intervalo $[\mu_1, \mu_2]$ ni en ninguno de sus extremos, lo que da precesión siempre en el mismo sentido y con velocidad nunca nula.

ii) $b - a\mu$ se anula dentro del intervalo $[\mu_1, \mu_2]$ con lo que la precesión se revierte periódicamente y el trompo se mueve “hacia atrás” durante esos intervalos.

iii) $b - a\mu$ se anula en uno de los extremos, llamemoslo μ_2 , con lo que la precesión se frena en estos puntos.

El regimen iii) es muy común en la práctica porque las condiciones iniciales que le dan lugar son las más naturales: girando el trompo sobre su eje x_3 , se lo deja libre con una inclinación inicial θ_0 de x_3 respecto del eje vertical z , sin velocidad $\dot{\theta}$. En este caso es, por supuesto, $\mu_2 = \cos \theta_0$; las condiciones iniciales son $\dot{\phi} = \dot{\theta} = 0$, $\dot{\psi} = \dot{\psi}_0$, con lo que, en $t = 0$,

$$\begin{aligned}\Omega_1 &= \Omega_2 = 0, \\ \Omega_3 &= \dot{\psi}_0,\end{aligned}$$

y tenemos entonces

$$\begin{aligned}L_3 &= I_3 \dot{\psi}_0, \\ L_z &= I_3 \dot{\psi}_0 \cos \theta_0, \\ E &= \frac{1}{2} \frac{L_3^2}{I_3} + Mgl \cos \theta_0.\end{aligned}$$

De la última es $E' = Mgl \cos \theta_0$ y tenemos entonces

$$\begin{aligned}b &= a \cos \theta_0 = a\mu_2, \\ \alpha &= \beta \cos \theta_0 = \beta\mu_2,\end{aligned}$$

con lo que escribimos

$$\begin{aligned}F(\mu) &= (1 - \mu^2) \beta (\mu_2 - \mu) - a^2 (\mu_2 - \mu)^2 \\ &= [\beta (1 - \mu^2) - a^2 (\mu_2 - \mu)] (\mu_2 - \mu),\end{aligned}$$

de donde se obtiene el otro extremo del movimiento, μ_1 , como la raíz del corchete

$$\mu_1 = \frac{a^2}{2\beta} \pm \sqrt{\left(\frac{a^2}{2\beta}\right)^2 - \mu_2 \frac{a^2}{\beta} + 1}.$$

Como

$$\frac{a^2}{\beta} = \frac{L_3^2}{(I_1 + Ml^2) 2Mgl},$$

nos da una idea del cociente entre la energía cinética de rotación y la energía potencial gravitatoria, llamamos trompo rápido al que satisface

$$\frac{a^2}{\beta} \gg 1,$$

que nos permite escribir

$$\begin{aligned} \mu_1 &= \frac{a^2}{2\beta} \pm \frac{a^2}{2\beta} \sqrt{1 - \mu_2 \frac{4\beta}{a^2} + \left(\frac{2\beta}{a^2}\right)^2} \\ &\simeq \frac{a^2}{2\beta} \pm \frac{a^2}{2\beta} \left(1 - \mu_2 \frac{2\beta}{a^2} + O\left[\left(\frac{\beta}{a^2}\right)^2\right]\right), \end{aligned}$$

de donde claramente la raíz físicamente aceptable es la correspondiente al signo menos que da

$$\mu_1 = \mu_2 + O\left[\frac{\beta}{a^2}\right],$$

que nos dice que la nutación es muy pequeña. Si escribimos entonces que

$$\mu = \mu_2 - \varepsilon,$$

con ε pequeño, tenemos que

$$\begin{aligned} F &= \beta(1 - \mu^2)(\mu_2 - \mu) - a^2(\mu_2 - \mu)^2 \\ &\simeq \beta(1 - \mu_2^2)\varepsilon - a^2\varepsilon^2, \end{aligned}$$

y como $\dot{\mu} = -\dot{\varepsilon}$, escribimos (8.40) como

$$\dot{\varepsilon}^2 = \beta(1 - \mu_2^2)\varepsilon - a^2\varepsilon^2,$$

que es integrable por una simple cuadratura y da (con la condición inicial $\varepsilon = 0$)

$$\varepsilon = \frac{\beta \sin^2 \theta_0}{2a^2} [1 - \cos(at)].$$

La frecuencia angular de la nutación es entonces

$$a = \frac{L_3}{I_1 + Ml^2} = \frac{I_3}{I_1 + Ml^2} \dot{\psi}_0.$$

Para la precesión es, de (8.42),

$$\begin{aligned} \dot{\phi} &= \frac{b - a\mu}{1 - \mu^2} = \frac{a(\mu_2 - \mu)}{1 - \mu^2} \\ &\simeq \frac{a\varepsilon}{1 - \mu_2^2} = \frac{\beta}{2a} [1 - \cos(at)]. \end{aligned} \tag{8.43}$$

Promediando sobre la rápida oscilación de frecuencia a obtenemos una precesión lenta de valor

$$\dot{\phi} = \frac{\beta}{2a} = \frac{Mgl}{L_3} = \frac{Mgl}{I_3 \dot{\psi}_0}.$$

En la práctica, los movimientos rápidos de nutación y de oscilación de la precesión son rápidamente amortiguados por fricción en el punto de apoyo del trompo y son prácticamente inobservables, lo que da lugar a la descripción usual del movimiento del trompo como de precesión pura. En realidad (8.43) nos dice que el movimiento de precesión se genera, sin discontinuidades, a partir del aumento de θ (disminución de μ) desde su valor θ_0 debido a la caída del trompo por acción de la gravedad hasta el valor máximo θ_1 ($\mu_1 = \cos \theta_1$).

8.5.2. Estabilidad del trompo vertical

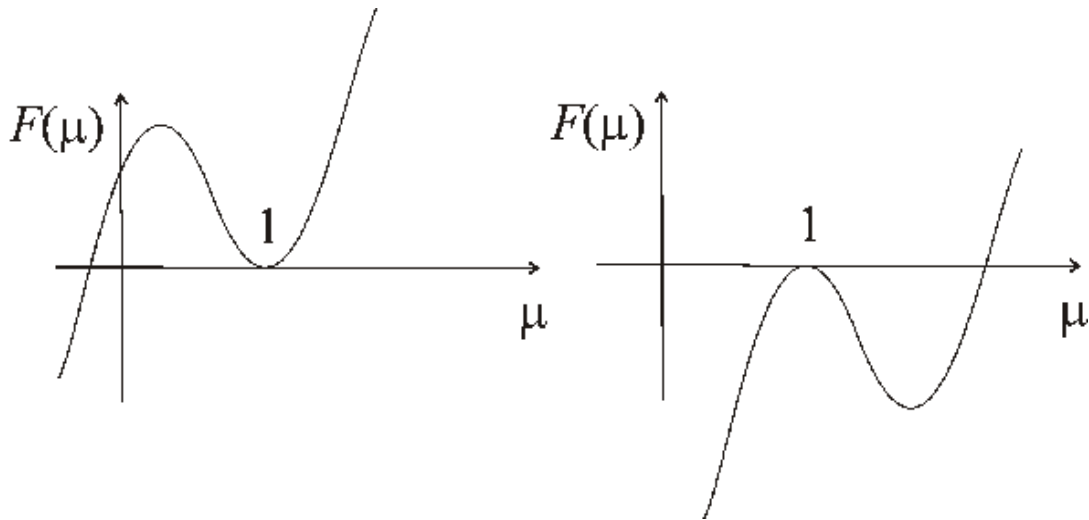
Consideremos ahora el caso excluido del estudio anterior, el del trompo vertical ($\theta_0 = 0$, $\dot{\theta}_0 = 0$) al que le corresponde $a = b$ y $\alpha = \beta$. Tenemos entonces que

$$\begin{aligned} F &= \beta(1 - \mu^2)(1 - \mu) - a^2(1 - \mu)^2 \\ &= \beta(1 + \mu)(1 - \mu)^2 - a^2(1 - \mu)^2 \\ &= [\beta(1 + \mu) - a^2](1 - \mu)^2, \end{aligned}$$

con lo que es inmediato que, además de la raíz doble $\mu_{1,2} = 1$, tenemos la raíz

$$\mu_3 = \frac{a^2}{\beta} - 1 = \frac{I_3^2 \dot{\psi}_0^2}{2(I_1 + Ml^2)Mgl} - 1.$$

El trompo se mueve entonces simplemente rotando con velocidad constante $\dot{\psi}_0$ alrededor de su eje vertical, siempre con $\mu = 1$. Nos interesa estudiar la estabilidad de este movimiento particular. Si lo consideramos en términos de la forma (8.41), podemos decir que el movimiento será estable si V_{ef} tiene un mínimo en $\mu = 1$, ya que entonces pequeños apartamientos (que no cambien cualitativamente la forma de F) generarán movimientos acotados en el entorno de $\mu = 1$.



Que V_{ef} tenga un mínimo equivale a que F tenga un máximo en su propia raíz $\mu = 1$; esto requiere que tenga valores negativos en la zona cercana con $\mu > 1$, que significa a su vez que la raíz μ_3 (donde al decrecer μ la función F pasa de ser positiva a negativa) debe satisfacer $\mu_3 > 1$; esto es, la condición de estabilidad es

$$\frac{I_3^2 \dot{\psi}_0^2}{2(I_1 + Ml^2)Mgl} > 2.$$

Tenemos entonces una velocidad crítica por debajo de la cual el movimiento es inestable:

$$\dot{\psi}_{cr} = \frac{2\sqrt{(I_1 + Ml^2)Mgl}}{I_3}.$$

Cuando un trompo vertical disminuye el valor de su velocidad por debajo este valor, debido al rozamiento, su movimiento se torna inestable y tiende a caer, generando una precesión como la estudiada en el punto anterior.

8.5.3. Giróscopo

Un giróscopo es esencialmente un trompo cuyo punto de suspensión es el mismo centro de masas, lo que se logra en la práctica con el llamado acoplamiento cardánico. El movimiento del giróscopo ante torques externos es bien conocido, deducible de consideraciones simples a través de la ecuación (8.23).

Aquí estudiaremos solamente la tendencia a alinear su eje de rotación propia (sentido incluido) con el de la rotación impuesta, cuando es forzado a girar.

Consideremos entonces el giróscopo en un sistema rotante con velocidad angular constante $\boldsymbol{\Omega}_T$ dirigida a lo largo del eje z de tal sistema. Tomemos al eje principal del giróscopo x_3 en el plano (x, z) , formando un ángulo θ con el eje z , con $\dot{\theta} = 0$ en el instante considerado. De esta manera, el giróscopo gira sólo alrededor de su eje x_3 con velocidad angular $\dot{\psi}$, lo que significa que $\Omega_1 = \Omega_2 = 0$, $\Omega_3 = \dot{\psi}$. Tomemos además el eje x_1 en el instante considerado contenido en el mismo plano (x, z) , tal que en el sistema de ejes propios es $\Omega_{T1} = -|\boldsymbol{\Omega}_T| \sin \theta$, $\Omega_{T2} = 0$, $\Omega_{T3} = |\boldsymbol{\Omega}_T| \cos \theta$ en el dado instante. En este sistema rotante las fuerzas externas que producen torque alrededor del centro de masas son las fuerzas de Coriolis y centrífuga; para un elemento de masa del giróscopo dm , con velocidad \mathbf{v} y posición $\mathbf{x} = \mathbf{X}_{CM} + \mathbf{r}$, tenemos que el elemento de fuerza de inercia sobre él es

$$d\mathbf{F} = -2dm \boldsymbol{\Omega}_T \times \mathbf{v} - dm \boldsymbol{\Omega}_T \times (\boldsymbol{\Omega}_T \times \mathbf{x}),$$

y el correspondiente elemento de torque respecto al centro de masas es

$$d\mathbf{K} = \mathbf{r} \times d\mathbf{F}.$$

Al escribir el término correspondiente a la fuerza centrífuga, reemplazando \mathbf{x} por $\mathbf{X}_{CM} + \mathbf{r}$,

$$d\mathbf{K}^{cen} = -dm \mathbf{r} \times \{ \boldsymbol{\Omega}_T \times [\boldsymbol{\Omega}_T \times (\mathbf{X}_{CM} + \mathbf{r})] \},$$

e integrar sobre todo el cuerpo, la contribución del término con \mathbf{X}_{CM} será nula porque $\int \mathbf{r} dm = 0$, de manera que podemos dejarlo de lado. Además, como la velocidad del CM es nula, tenemos que $\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r}$, con todo lo cual

$$\begin{aligned} d\mathbf{K} &= -2dm \mathbf{r} \times [\boldsymbol{\Omega}_T \times (\boldsymbol{\Omega} \times \mathbf{r})] - dm \mathbf{r} \times [\boldsymbol{\Omega}_T \times (\boldsymbol{\Omega}_T \times \mathbf{r})] \\ &= -2dm \mathbf{r} \times [\boldsymbol{\Omega}_T \times (\boldsymbol{\Omega}' \times \mathbf{r})], \end{aligned}$$

donde hemos definido a

$$\boldsymbol{\Omega}' \equiv \boldsymbol{\Omega} + \frac{1}{2}\boldsymbol{\Omega}_T.$$

En las aplicaciones la velocidad de rotación del giróscopo es mucho mayor que la del sistema rotante (sea el vehículo que lo transporta, o la tierra misma), por lo que podemos tomar $\boldsymbol{\Omega}' \simeq \boldsymbol{\Omega}$ (que significa que el torque debido a la fuerza centrífuga es despreciable comparado con el debido a la fuerza de Coriolis). Con esto, como (de (8.14))

$$\boldsymbol{\Omega}_T \times (\boldsymbol{\Omega} \times \mathbf{r}) = \boldsymbol{\Omega} (\boldsymbol{\Omega}_T \cdot \mathbf{r}) - \mathbf{r} (\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}_T),$$

tenemos finalmente

$$d\mathbf{K} = 2dm (\boldsymbol{\Omega}_T \cdot \mathbf{r}) \boldsymbol{\Omega} \times \mathbf{r}.$$

En componentes sobre los ejes propios es, como se discutió arriba,

$$\begin{aligned}\boldsymbol{\Omega} &= (0, 0, \dot{\psi}), \\ \boldsymbol{\Omega}_T &= (-|\boldsymbol{\Omega}_T| \sin \theta, 0, |\boldsymbol{\Omega}_T| \cos \theta),\end{aligned}$$

con lo que se obtiene muy sencillamente

$$\begin{aligned}dK_1 &= 2dm |\boldsymbol{\Omega}_T| \dot{\psi} (r_1 r_2 \sin \theta - r_2 r_3 \cos \theta), \\ dK_2 &= 2dm |\boldsymbol{\Omega}_T| \dot{\psi} (-r_1^2 \sin \theta + r_1 r_3 \cos \theta), \\ dK_3 &= 0.\end{aligned}$$

Debemos ahora integrar para todos los elementos de masa del giróscopo, usando que éstos están distribuidos con simetría de revolución alrededor del eje x_3 , por lo que

$$\begin{aligned}\int r_1 r_2 dm &= \int r_1 r_3 dm = \int r_2 r_3 dm = 0, \\ \int r_1^2 dm &= \int r_2^2 dm = \frac{I_3}{2},\end{aligned}$$

de donde resulta inmediatamente

$$\begin{aligned}K_1 &= K_3 = 0, \\ K_2 &= -I_3 |\boldsymbol{\Omega}_T| \dot{\psi} \sin \theta,\end{aligned}$$

que en componentes sobre los ejes fijos se escribe

$$\begin{aligned}K_x &= K_1 \cos \theta = 0, \\ K_y &= K_2 = -I_3 |\boldsymbol{\Omega}_T| \dot{\psi} \sin \theta, \\ K_z &= K_1 \sin \theta = 0.\end{aligned}$$

Vemos entonces que la componente K_y genera una precesión anticiclónica (en sentido opuesto a la rotación del sistema). La inclusión de una fricción débil ante esta precesión genera entonces torques en las distintas direcciones, que cambian la orientación del giróscopo hasta que no se genere precesión; esto es, hasta que se encuentren alineados el eje de rotación del giróscopo con el del sistema rotante. El equilibrio con ambos vectores antiparalelos, $\theta = \pi$, es inestable.

La sensibilidad del giróscopo (que aumenta con su velocidad de rotación $\dot{\psi}$) a las rotaciones del sistema que lo transporta lo hace muy adecuado como base de sistemas de guía de navegación, muy empleados en naves espaciales, barcos, submarinos, etc.. (la insensibilidad del giróscopo ante rotaciones paralelas a su propio eje se remedia usando varios giróscopos no alineados)

Por otro lado, si fuera posible aislar el giróscopo del efecto de las rotaciones del vehículo aéreo, marítimo o terrestre que lo transporta, todavía tendería a alinearse con el eje de rotación de la tierra, indicando así el norte geográfico; ésta es la base del compás giroscópico; el aislamiento se logra montando el giróscopo en una esfera que flota libremente dentro de otra esfera contenedora; la viscosidad del fluido provee la fricción necesaria para la alineación.