Notas para un Curso de

MECÁNICA CLÁSICA

Rafael Ferraro

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Prólogo

Estas Notas reúnen el material producido por el autor para el dictado de clases a distancia durante la pandemia de los años 2020 y 2021, como parte de los cursos ofrecidos por la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires para la carrera de Licenciatura en Ciencias Físicas. Este curso de Mecánica Clásica abarca los contenidos usuales del programa de la materia, y tiene como bibliografía de base los tradicionales textos escritos por L. Landau y E. Lifshitz (Mecánica, Ed. Reverté) y H. Goldstein (Mecánica Clásica, Ed. Reverté), a los que se agrega la bibliografía específica mencionada en algunos capítulos.

Buenos Aires, Febrero de 2022.

Prólogo **i**

Introducción **vii**

1 Repaso de Mecánica Elemental **1**
1.1 Teorema de la derivada relativa 1
1.2 Cantidad de movimiento de un sistema de N partículas 4
1.2.1 Centro de masa 5
1.3 Momento angular de un sistema de N partículas 5
1.3.1 Cambio del centro de momentos y cambio del sistema de referencia 7
1.4 Transformación de la energía cinética 8

2 Trabajo-Energía. Fuerzas Conservativas. Teorema del virial **9**
2.1 Trabajo-energía 9
2.2 Fuerzas conservativas 10
2.2.1 Fuerzas conservativas internas 11
2.3 Teorema del virial 12

3 Coordenadas generalizadas. Vínculos. Trabajos virtuales. Ecuaciones de Lagrange **17**
3.1 Coordenadas generalizadas 17
3.2 Vínculos 17
3.2.1 Vínculos holónomos 18
3.2.2 Vínculos anholónomos 18
3.2.3 Vínculos esclerónomos y reónomos 20
3.3 Espacio de configuración 21
3.4 Fuerzas aplicadas y de vínculo. Trabajos virtuales 21
3.5 Ecuaciones de Lagrange 23

4 Vínculos anholónomos. Multiplicadores de Lagrange. Potenciales dependientes de la velocidad **29**
4.1 Vínculos anholónomos 29
4.1.1 Vínculos holónomos vistos como anholónomos 31
4.2 Potenciales dependientes de la velocidad 33

5 Principios variacionales **35**
5.1 Cálculo variacional 35
5.2 Principio de Hamilton 38
5.2.1 ¿Máximo o mínimo? 40

6 Simetría y Conservación. Teoremas de Noether **43**
6.1 Coordenadas cíclicas . . . 43
6.2 Simetría y conservación . . . 44
6.3 Teorema de Noether . . . 44
6.4 Conservación asociada a la homogeneidad del tiempo. . . . 47
6.4.1 ¿Cuándo coincide H con la energía mecánica E? . . . 48
6.5 Segundo teorema de Noether. Simetrías de gauge . . . 49

7 Sistema de dos cuerpos. Fuerzas centrales. Problema de Kepler **51**
7.1 Sistema de dos cuerpos . . . 51
7.2 Segunda Ley de Kepler . . . 52
7.3 Integración . . . 53
7.4 Potencial efectivo . . . 54
7.4.1 Órbita circular . . . 57
7.5 Problema de Kepler . . . 58
7.6 Ecuación de Binet . . . 59
7.7 Secciones cónicas . . . 61
7.8 Primera Ley de Kepler . . . 63
7.9 Tercera Ley de Kepler . . . 63
7.10 Regreso al problema de dos cuerpos . . . 64

8 Choque elástico. Scattering **65**
8.1 Choque elástico . . . 65
8.1.1 El caso $v_{2o} = 0$ . . . 67
8.2 Scattering . . . 69
8.2.1 Potencial coulombiano . . . 71

9 Oscilaciones **73**
9.1 Pequeñas oscilaciones . . . 73
9.1.1 El péndulo doble plano . . . 76
9.2 Ortonormalidad . . . 77
9.3 Notación matricial . . . 78
9.4 Coordenadas normales . . . 79
9.4.1 La molécula de CO_2 . . . 81
9.5 Oscilador forzado . . . 83

10 Resonancia paramétrica **87**
10.1 El oscilador armónico en el espacio de estados . . . 88
10.2 Variación de la frecuencia propia . . . 88
10.3 Ecuación de Mathieu . . . 94

11 Cinemática del cuerpo rígido **95**
11.1 Ángulos de Euler . . . 95
11.2 Grupo de rotaciones . . . 96

11.3 Relación entre $(\hat{\imath}, \hat{\jmath}, \hat{k})$ y $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ 98
11.4 Velocidad de rotación 99
11.5 Cinemática del cuerpo rígido 101

12 Tensor de inercia. Ecuaciones de Euler 103
12.1 Momento angular del cuerpo rígido 103
12.2 Tensor de inercia 104
12.3 Ejes principales de inercia 105
12.4 Teorema de Steiner 106
12.5 Descomposición del momento angular en ejes principales 106
12.6 Dinámica del cuerpo rígido. Ecuaciones de Euler 107
12.7 Energía cinética del cuerpo rígido 108
12.8 Formalismo Lagrangiano 108

13 Trompo simétrico 111
13.1 Trompo asimétrico libre de torques 111
13.2 Trompo simétrico libre de torques 113
13.3 Trompo simétrico con punto fijo bajo la acción de la gravedad 117
13.3.1 Trompo vertical 121
13.3.2 Trompo horizontal 122
13.3.3 Trompo rápido 123
13.4 Giróscopo 123
13.4.1 Girocompás de Foucault 124

14 Formalismo Hamiltoniano 127
14.1 Transformada de Legendre 127
14.2 La función Hamiltoniana 129
14.3 Evolución en el espacio de las fases 130
14.3.1 Sistemas conservativos de 1 grado de libertad 130
14.4 Conservación de H 132
14.5 Coordenadas cíclicas 133
14.6 Principio variacional de Hamilton 133
14.7 Función principal de Hamilton 134
14.8 Notación matricial 135
14.9 Corchete de Poisson 136
14.10 Transformaciones canónicas. Estructura simpléctica 137

15 Generatrices de transformaciones canónicas 141
15.1 Tipos de funciones generatrices 142
15.2 Transformaciones infinitesimales 144
15.3 La evolución del sistema como una transformación canónica 145
15.3.1 Teorema de Liouville 145
15.4 Relación entre simetría y conservación 146

16 Ecuación de Hamilton-Jacobi **147**
16.1 Integrabilidad 147
16.2 Ecuación de Hamilton-Jacobi 148
16.3 Sistema conservativo de 1 grado de libertad 150
16.4 Separación de variables 151

17 Variables ángulo-acción **153**
17.1 Variables ángulo-acción 153
17.1.1 Frecuencias del problema de Kepler 157
17.2 Sistemas periódicos y condicionalmente periódicos 158

18 Invariantes adiabáticos **161**
18.1 Invariancia adiabática 161
18.1.1 Expansión adiabática de un gas ideal 163
18.2 Formulación general 163
18.3 Oscilador armónico con frecuencia variable 166

19 Relatividad Especial **169**
19.1 El Principio de relatividad y las leyes de Maxwell 169
19.2 La relatividad de Einstein 172
19.3 Transformaciones de Lorentz 175
19.3.1 Intervalo 177
19.3.2 Cono de luz. Tipos de separación entre eventos 178
19.4 Composición relativista de movimientos 179
19.5 Tiempo propio de la partícula 179
19.5.1 Paradoja de los gemelos 181
19.6 Dinámica relativista 181
19.6.1 Acción de la partícula libre 181
19.6.2 Cantidad de movimiento y energía 182
19.6.3 Transformaciones de E y $\vec{p}$ 182
19.6.4 Fuerza de Lorentz 183
19.6.5 Fotones 184
19.7 Energía-cantidad de movimiento. Equivalencia masa-energía 185
19.7.1 Equivalencia masa-energía 185
19.7.2 Interacciones “a distancia” 187

20 Acción de Jacobi. Analogía con la óptica geométrica **189**
20.1 Analogía con la óptica geométrica 191
20.1.1 Principio de Fermat 191
20.1.2 Ecuación de la eikonal 192
20.2 Ecuación de Schrödinger 194

Introducción

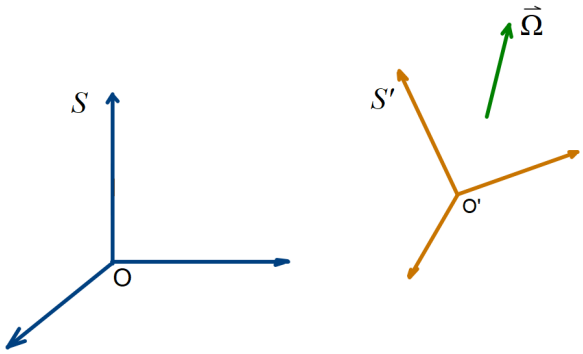
En los cien años que siguieron a la publicación de los *Principia* de Newton (Principios Matemáticos de la Filosofía Natural, 1687), la ciencia de la Mecánica se desarrolló vertiginosamente gracias a la sustitución del lenguaje matemático simple de esa obra fundamental, que se vale de relaciones geométricas, proporciones, etc., por el cálculo diferencial creado por Leibniz y el propio Newton. En ese período aparecieron las contribuciones al estudio de la dinámica del cuerpo rígido, los fluidos, la propagación de ondas en medios materiales, el cálculo variacional, etc., debidas a Euler, los Bernoulli, d'Alembert, Huygens y Maupertuis, entre otros. Estos avances en sí mismos muy importantes, no constituían todavía un único cuerpo teórico. En 1788 Lagrange publica la Mecánica Analítica, donde la ciencia de la Mecánica es reescrita desde primeros principios en el lenguaje de las ecuaciones diferenciales. En palabras de Lagrange, su objetivo era "reducir la teoría de esta Ciencia, y el arte de resolver los problemas asociados, a fórmulas generales cuyo simple desarrollo provea todas las ecuaciones necesarias para la solución de cada problema... No se encontrarán figuras en esta obra. Los métodos que aquí expongo no requieren ni construcciones, ni razonamientos geométricos o mecánicos, sino solamente las operaciones algebraicas inherentes a un proceso regular y uniforme." Medio siglo después de este magnífico legado teórico, en la primera mitad del siglo XIX Hamilton logrará una reformulación de la mecánica Lagrangiana introduciendo conceptos que iluminarán la ciencia de la Mecánica desde una perspectiva diferente, y que serán indispensables para el desarrollo de la Mecánica Cuántica en el siglo XX. La Relatividad Especial de Einstein (1905) modificará la estructura espacio-temporal sobre la que se asienta la Mecánica Clásica. No obstante, tanto la formulación Lagrangiana como la Hamiltoniana son aptas en el contexto relativista, y en ellas se basan las teorías cuánticas de campos y de cuerdas de la segunda mitad del siglo XX.

1 Repaso de Mecánica Elemental

Repasaremos algunos resultados importantes del curso de Mecánica Elemental (Física 1). No definiremos los conceptos básicos, que supondremos ya conocidos.

1.1 Teorema de la derivada relativa

Consideremos dos sistemas de referencia S y S' en movimiento relativo. Si la *orientación* de S' relativa a S cambia con el tiempo, entonces S' *rota* respecto de S. Llamaremos $\vec{\Omega}$ a la velocidad de rotación de S' respecto de S.



Queremos derivar un vector $\vec{A}(t)$ desde el sistema S. Es posible descomponer a $\vec{A}$ respecto de las bases de S y S' :

$$\begin{aligned}\vec{A}(t) &= A_x\ \hat{\imath} + A_y\ \hat{\jmath} + A_z\ \hat{k} \\ &= A_x'\ \hat{\imath}\,' + A_y'\ \hat{\jmath}\,' + A_z'\ \hat{k}\,'\end{aligned}$$

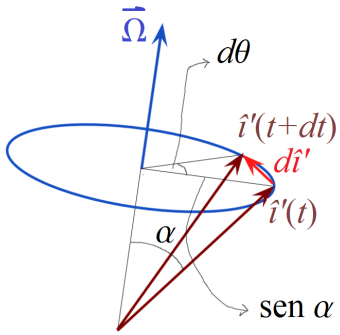
¿Cuál es la diferencia entre derivar $\vec{A}$ desde el sistema S o hacerlo desde S'? Si derivamos desde S entonces los versores $\hat{\imath},\ \hat{\jmath},\ \hat{k}$ se ven fijos (constantes en el tiempo); el resultado es

$$\frac{d\vec{A}}{dt}|_S = \frac{dA_x}{dt}\ \hat{\imath} + \frac{dA_y}{dt}\ \hat{\jmath} + \frac{dA_z}{dt}\ \hat{k}$$

Pero también podríamos derivar desde S usando la otra descomposición de $\vec{A}$:

$$\begin{aligned}\frac{d\vec{A}}{dt}|_S &= \frac{dA'_x}{dt}\,\hat{\imath}' + \frac{dA'_y}{dt}\,\hat{\jmath}' + \frac{dA'_z}{dt}\,\hat{k}' \\ &\quad + A'_x\,\frac{d\,\hat{\imath}'}{dt}|_S + A'_y\,\frac{d\,\hat{\jmath}'}{dt}|_S + A'_z\,\frac{d\,\hat{k}'}{dt}|_S\end{aligned}$$

En este caso debemos calcular las derivadas de los versores del sistema S'vistos desde S. Veámoslo en una Figura,



$$|d\hat{\imath}'| = \sin\alpha\; d\theta = \sin\alpha\;\Omega\;dt$$

Si además notamos que la dirección de $d\hat{\imath}\,'$ coincide con la de $\vec{\Omega}\times\hat{\imath}\,'$, concluimos que

$$\frac{d\hat{\imath}'}{dt}|_S = \vec{\Omega}\times\hat{\imath}'$$

y lo mismo para los otros versores. Reemplazando obtenemos

$$\begin{aligned}\frac{d\vec{A}}{dt}|_S &= \frac{dA'_x}{dt}\,\hat{\imath}' + \frac{dA'_y}{dt}\,\hat{\jmath}' + \frac{dA'_z}{dt}\,\hat{k}' \\ &\quad + A'_x\,\vec{\Omega}\times\hat{\imath}' + A'_y\,\vec{\Omega}\times\hat{\jmath}' + A'_z\,\vec{\Omega}\times\hat{k}'\end{aligned}$$

Como la primera línea no es otra cosa que derivar $\vec{A}$ desde S', concluimos que

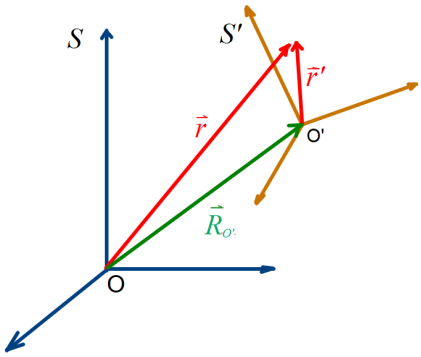
$$\boxed{\frac{d\vec{A}}{dt}|_S = \frac{d\vec{A}}{dt}|_{S'} + \vec{\Omega}\times\vec{A}}$$

que es el teorema de la derivada relativa.

Ejemplos

En el resultado anterior $\vec{A}$ es un vector arbitrario. Veremos algunos casos de interés.

Sea $\vec{A} = \vec{r}\,'$ la posición de un punto o partícula respecto del origen de S'.



Se cumple que

$$\vec{A} = \vec{r}\,' = \vec{r} - \vec{R}_{O'}$$

donde $\vec{R}_{O'}$ es la posición del origen O' de S' respecto del origen O de S. Entonces, reemplazando en el teorema:

$$\frac{d(\vec{r} - \vec{R}_{O'})}{dt}|_S = \frac{d\,\vec{r}\,'}{dt}|_{S'} \; + \; \vec{\Omega} \times \vec{r}\,'$$

Entonces

$$\vec{v} - \vec{V}_{O'} = \vec{v}\,' + \; \vec{\Omega} \times \vec{r}\,' \tag{1.1}$$

donde $\vec{v}$, $\vec{v}\,'$ son las velocidades del punto o partícula respecto de cada sistema de referencia. Veamos dos casos particulares de este resultado:

1) Si $\vec{\Omega} = 0$ (es decir, si S' se *traslada* respecto de S) entonces resulta el teorema de adición de velocidades de Galileo: $\vec{v}\,' = \vec{v} - \vec{V}$.

2) Pensemos a S' como un cuerpo rígido cuyo movimiento es observado desde S, y sea $\vec{v}$ la velocidad de un punto P perteneciente al cuerpo rígido (entonces $\vec{v}\,' = 0$). Ubiquemos el origen O' en un punto Q cualquiera del cuerpo rígido ($\vec{r}_P' = \vec{r}_P - \vec{r}_Q$). Entonces obtenemos

$$\vec{v}_P - \vec{v}_Q = \; \vec{\Omega} \times (\vec{r}_P - \vec{r}_Q)$$

que no es otra cosa que la cinemática del cuerpo rígido.

Volviendo a derivar el resultado obtenido más arriba,

$$\frac{d}{dt}(\vec{v} - \vec{V}_{O'})|_S = \frac{d\vec{v}\,'}{dt}|_S + \; \dot{\vec{\Omega}} \times \vec{r}\,' + \vec{\Omega} \times \frac{d\vec{r}\,'}{dt}|_S$$

En el miembro de la derecha aplicamos el teorema de la derivada relativa para obtener

$$\vec{a} - \vec{A}_{O'} = \vec{a}\,' + \vec{\Omega} \times \vec{v}\,' + \dot{\vec{\Omega}} \times \vec{r}\,' + \vec{\Omega} \times (\vec{v}\,' + \vec{\Omega} \times \vec{r}\,')$$

Es decir,

$$\vec{a} = \vec{a}\,' + \vec{A}_{O'} + 2\,\vec{\Omega} \times \vec{v}\,' + \vec{\Omega} \times (\vec{\Omega} \times \vec{r}\,') + \dot{\vec{\Omega}} \times \vec{r}\,' \tag{1.2}$$

donde reconocemos los términos que van a dar lugar a las fuerzas de inercia en la Dinámica.

1.2 Cantidad de movimiento de un sistema de N partículas

$$\vec{P} \doteq \sum_{i=1}^{N} \vec{p}_i = \sum_{i=1}^{N} m_i\,\vec{v}_i$$

$$\frac{d\vec{P}}{dt} = \sum_{i=1}^{N} \frac{d\vec{p}_i}{dt} = \sum_{i=1}^{N} \vec{F}_i$$

donde $\vec{F}_i$ es la fuerza *resultante* sobre la i-ésima partícula. Podemos descomponerla en la suma de la resultante de las fuerzas externas, $\vec{F}_i^{ext}$, y la resultante de las fuerzas internas

$$\vec{F}_i = \vec{F}_i^{ext} + \sum_{j \neq i} \vec{f}_{ij}$$

donde $\vec{f}_{ij}$ es la fuerza (interna) sobre i debida a j. Ahora bien, en la suma $\sum_{i=1}^{N} \vec{F}_i$ de las fuerzas que actúan sobre todas las partículas del sistema aparecerá

$$(\,\vec{f}_{12} + \vec{f}_{13} + \ldots) + (\vec{f}_{21} + \vec{f}_{23} + \ldots) + \ldots$$

Pero $\vec{f}_{ij} = -\,\vec{f}_{ji}$ por la 3ra. ley de Newton (principio de acción y reacción). Entonces la suma de fuerzas internas del sistema es nula:

$$\boxed{\frac{d\vec{P}}{dt} = \sum_{i=1}^{N} \vec{F}_i^{ext}}$$

$\vec{P}$ se conserva si y sólo si la resultante de fuerzas externas es cero (vale también "componente a componente").

1.2.1 Centro de masa

Notemos que

$$\vec{P} = \sum_{i=1}^{N} m_i \, \vec{v}_i = \frac{d}{dt} \sum_{i=1}^{N} m_i \, \vec{r}_i$$

Entonces resulta útil definir una posición $\vec{R}_{CM}$ característica del sistema, tal que

$$M \, \vec{R}_{CM} \equiv \sum_{i=1}^{N} m_i \, \vec{r}_i$$

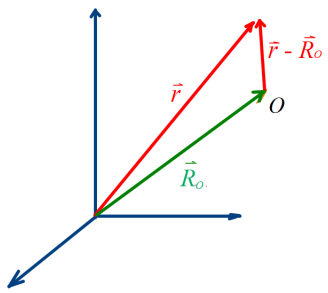
donde M es la masa total del sistema ($M = \sum m_i$). Como vemos, $\vec{R}_{CM}$ es un promedio pesado de las posiciones de las partículas que forman el sistema. Llamaremos a $\vec{R}_{CM}$ posición del *centro de masa* del sistema de partículas. La definición hace posible ver a $\vec{P}$ como la cantidad de movimiento de una única partícula de masa M que se mueve con velocidad $\vec{V}_{CM} = \dot{\vec{R}}_{CM}$:

$$\vec{P} = M \, \vec{V}_{CM}$$

La conservación de $\vec{P}$ implica que la velocidad del centro de masa es constante.

1.3 Momento angular de un sistema de N partículas

$$\vec{L}_O \doteq \sum_{i=1}^{N} \vec{l}_{O\,i} = \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times \vec{p}_i$$



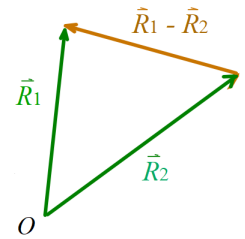
$$\begin{aligned}
\frac{d\vec{L}_O}{dt} &= \sum_{i=1}^{N} \left[\frac{d}{dt}(\vec{r}_i - \vec{R}_O) \times \vec{p}_i + (\vec{r}_i - \vec{R}_O) \times \frac{d\vec{p}_i}{dt} \right] \\
&= -\frac{d\vec{R}_O}{dt} \times \sum_{i=1}^{N} \vec{p}_i + \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times \vec{F}_i \\
&= -\vec{V}_O \times \vec{P} + \sum_{i=1}^{N} \vec{N}_{O\,i}
\end{aligned}$$

(en la primera línea hemos usado que $\dot{\vec{r}}_i \times \vec{p}_i = 0$) donde $\vec{N}_{O\,i}$ es la resultante de los momentos de las fuerzas (torques) sobre la partícula i. Podemos preguntarnos si los momentos de todas las fuerzas internas se cancelan de a pares debido al principio de acción y reacción. Veamos esta cuestión:

$$\sum_{i=1}^{N} \vec{N}^{int}_{O\,i} = \sum_{i=1}^{N} \sum_{j \neq i} \vec{R}_i \times \vec{f}_{ij}$$

donde $\vec{R}_i \equiv \vec{r}_i - \vec{R}_O$. En la doble sumatoria aparecen los términos

$$\vec{R}_1 \times \vec{f}_{12} + \vec{R}_2 \times \vec{f}_{21} + \ldots = (\vec{R}_1 - \vec{R}_2) \times \vec{f}_{12} + \ldots$$



Vemos que la cancelación de a pares se produce sólo si

$$\vec{r}_i - \vec{r}_j = (\vec{R}_i - \vec{R}_j) \;\; || \;\; \vec{f}_{ij} \qquad \forall\, i, j \;,$$

es decir, cuando las fuerzas de interacción entre las partículas del sistema están dirigidas a lo largo de la recta que une las partículas. En ese caso el resultado anterior queda

$$\boxed{\frac{d\vec{L}_O}{dt} = -\vec{V}_O \times \vec{P} + \sum_{i=1}^{N} \vec{N}^{ext}_{O\,i}}$$

Notemos que el término $\vec{V}_O \times \vec{P}$ se anula en dos casos sencillos: i) si O es fijo, ii) si O es el CM. En cualquiera de esos dos casos vale que $\vec{L}_O$ se conserva si y sólo si se anula la suma de los torques externos respecto de O (en la hipótesis que $(\vec{r}_i - \vec{r}_j) \;||\; \vec{f}_{ij} \quad \forall\, i, j$).

La magnitud $\vec{L}_O$ depende de la elección del centro de momentos O, por un lado, y de la elección del sistema de referencia por el otro (porque la elección del sistema de referencia determina los valores de las velocidades $\vec{v}_i$). Veamos cómo cambia $\vec{L}_O$ ante el cambio de alguno de estos dos factores.

1.3.1 Cambio del centro de momentos y cambio del sistema de referencia

Cambiaremos el centro de momento sin cambiar de sistema de referencia:

$$\vec{L}_{O'} - \vec{L}_O = \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_{O'}) \times \vec{p}_i - \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times \vec{p}_i = -(\vec{R}_{O'} - \vec{R}_O) \times \sum_{i=1}^{N} \vec{p}_i$$

Entonces

$$\vec{L}_{O'} = \vec{L}_O - (\vec{R}_{O'} - \vec{R}_O) \times \vec{P}$$

Cuando $\vec{P} = 0$ (es decir, en el sistema CM) el momento angular del sistema no depende del centro de momentos (resultado típico de todo momento).

Ahora veamos qué sucede si cambiamos el sistema de referencia manteniendo el mismo centro de momentos O. Supondremos dos sistemas de referencia S y S' que **no están en rotación relativa**, de manera que vale la adición de velocidades galileana: $\vec{v}_i{}' = \vec{v}_i - \vec{V}$, donde $\vec{V}$ es la velocidad de S' relativa a S.

$$\begin{aligned}
\vec{L}_O' - \vec{L}_O &= \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times \vec{p}_i{}' - \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times \vec{p}_i = \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times (\vec{p}_i{}' - \vec{p}_i) \\
&= -\sum_{i=1}^{N} (\vec{r}_i - \vec{R}_O) \times m_i \vec{V} = -\sum_{i=1}^{N} m_i \, \vec{r}_i \times \vec{V} + \vec{R}_O \times \vec{V} \sum_{i=1}^{N} m_i \\
&= -\vec{R}_{CM} \times M\vec{V} + \vec{R}_O \times M\vec{V} = (\vec{R}_O - \vec{R}_{CM}) \times M\vec{V}
\end{aligned}$$

Tenemos aquí algunos casos particulares interesantes:

1) Eligiendo $O \equiv CM$ obtenemos

$$\vec{L}_{CM}' = \vec{L}_{CM}$$

El momento angular respecto del CM tiene el mismo valor en distintos sistemas de referencia en traslación relativa.

2) Eligiendo S' como el sistema CM (donde, como dijimos, el valor del momento angular no depende del centro de momentos), tendremos $\vec{V} = \vec{V}_{CM}$; entonces obtenemos

$$\vec{L}_O = \vec{L}^{CM} + (\vec{R}_{CM} - \vec{R}_O) \times M\vec{V}_{CM} = \vec{L}^{CM} + \vec{R}_{CM-O} \times \vec{P}$$

Este resultado dice que el momento angular de un sistema de partículas puede descomponerse en el momento angular de una única partícula de cantidad de movimiento $\vec{P}$ ubicada en el CM (**momento angular orbital**) más el momento angular medido en un sistema CM que se traslada respecto de S (**momento angular intrínseco o spin**).

1.4 Transformación de la energía cinética

Consideremos nuevamente dos sistemas de referencia en traslación relativa, de manera que vale la adición galileana de velocidades $\vec{v}_i{}' = \vec{v}_i - \vec{V}$. La energía cinética en el sistema S' es

$$\begin{aligned} T\,' &= \sum_{i=1}^{N} \frac{1}{2}\, m_i\, v_i'^2 = \sum_{i=1}^{N} \frac{1}{2}\, m_i\, (\vec{v}_i - \vec{V}) \cdot (\vec{v}_i - \vec{V}) \\ &= \sum_{i=1}^{N} \frac{1}{2}\, m_i\, {v_i}^2 - \sum_{i=1}^{N} m_i\, \vec{v}_i \cdot \vec{V} + \frac{1}{2} \sum_{i=1}^{N} m_i\, V^2 \\ &= T \;-\; M\, \vec{V}_{CM} \cdot \vec{V} \;+\; \frac{1}{2}\, M\, V^2 \end{aligned}$$

Si el sistema S' es el sistema CM, entonces $\vec{V} = \vec{V}_{CM}$ y obtenemos

$$T = T^{CM} \;+\; \frac{1}{2}\, M\, V_{CM}^2$$

La energía cinética de un sistema de partículas puede descomponerse en la energía cinética de una única partícula de masa M moviéndose con velocidad V_{CM} más la energía cinética medida en un sistema CM que se traslada respecto de S.

2 Trabajo-Energía. Fuerzas Conservativas. Teorema del virial

2.1 Trabajo-energía

Derivemos la energía cinética sobre las trayectorias dinámicas de un sistema de partículas:

$$\frac{dT}{dt} = \frac{d}{dt}\sum_{i=1}^{N} \frac{1}{2}\, m_i\, {v_i}^2 = \sum_{i=1}^{N} m_i\, \vec{v}_i \cdot \frac{d\,\vec{v}_i}{dt} = \sum_{i=1}^{N} \vec{F}_i \cdot \vec{v}_i$$

Integrando en el tiempo entre dos configuraciones del sistema (inicial y final, o "A" y "B"):

$$\boxed{T(B) - T(A) \;=\; \int_A^B \sum_{i=1}^{N} \vec{F}_i \cdot d\,\vec{r}_i \;\equiv\; W_{A\to B}}$$

En el lado derecho hemos definido el **trabajo** que realizan *todas* las fuerzas sobre el sistema, a lo largo de la evolución desde A hasta B.

En general las fuerzas internas realizan trabajo. Por ejemplo, consideremos la interacción entre las partículas 1 y 2. La suma de los trabajos del par de interacción es

$$\vec{f}_{12} \cdot d\,\vec{r}_1 \;+\; \vec{f}_{21} \cdot d\,\vec{r}_2 = \;\; \vec{f}_{12} \cdot (d\,\vec{r}_1 - d\,\vec{r}_2) \;=\; \vec{f}_{12} \cdot d(\,\vec{r}_1 - \,\vec{r}_2)$$

El resultado es distinto de cero salvo en casos donde la variación de la posición relativa $\vec{r}_1 - \,\vec{r}_2$ sea perpendicular a la fuerza de interacción.

Cuerpo rígido. Si el sistema de partículas es un cuerpo rígido entonces sabemos que las distancias entre los puntos del cuerpo no cambian. Por lo tanto los módulos de los vectores $\vec{r}_i - \,\vec{r}_j$ son fijos. Un vector $\vec{A}$ cuyo módulo es fijo sólo puede experimentar cambios de su dirección; en ese caso es $d\vec{A} \perp \vec{A}$ (si $d\vec{A}$ tuviese una componente a lo largo de $\vec{A}$ significaría que el módulo de $\vec{A}$ cambia). Entonces diremos que un cuerpo rígido se caracteriza porque las posiciones relativas entre sus partículas satisfacen

$$d(\,\vec{r}_i - \,\vec{r}_j) \perp (\,\vec{r}_i - \,\vec{r}_j) \qquad \forall\, i, j$$

Aceptando que las fuerzas internas en un rígido son paralelas a las direcciones entre las partículas ($\vec{f}_{ij} \parallel (\vec{r}_i - \vec{r}_j)$), de la relación anterior surge que no hay trabajo de fuerzas internas en un cuerpo rígido. (Nota: las fuerzas de interacción *no* son paralelas a la dirección entre partículas en la interacción magnética, en el rozamiento entre superficies en movimiento relativo, etc.)

2.2 Fuerzas conservativas

En general el trabajo de una fuerza depende del camino. Dicho de otro modo, $\vec{F} \cdot d\vec{r}$ no es, en general, un diferencial exacto. En cambio, cuando $\vec{F} \cdot d\vec{r}$ es un diferencial exacto, digamos

$$\vec{F} \cdot d\vec{r} = -dV$$

entonces el trabajo de $\vec{F}$ será

$$\int_A^B \vec{F} \cdot d\vec{r} = -\int_A^B dV = V(A) - V(B)$$

cuyo resultado sólo depende de los extremos A y B, sin que importe el camino seguido para unir esos extremos.

Como para toda función V definida en el espacio e independiente del tiempo, se tiene que

$$dV = \frac{\partial V}{\partial x}\, dx + \frac{\partial V}{\partial y}\, dy + \frac{\partial V}{\partial z}\, dz = \vec{\nabla} V \cdot d\vec{r}\ ,$$

entonces diremos que una fuerza es **conservativa** si puede expresarse como

$$\vec{F} = -\vec{\nabla} V$$

para alguna **energía potencial** V. El trabajo de las fuerzas conservativas no depende del camino; equivalentemente,

$$\vec{\nabla} \times \vec{F} = 0\ , \qquad \oint \vec{F} \cdot d\vec{r} = 0$$

(Nota: a modo de contraejemplo consideremos la fuerza de rozamiento dinámico, la cual se opone siempre al desplazamiento. Entonces siempre resulta que $\vec{F}_{roz} \cdot d\vec{r} < 0 \Rightarrow \oint \vec{F}_{roz} \cdot d\vec{r} < 0$).

Si en un sistema de partículas existen fuerzas conservativas (ya sean internas o externas) entonces habrá una energía potencial dependiente de las posiciones de las N partículas, de forma tal que la fuerza conservativa neta sobre la partícula i se calcula como

$$\vec{F}_i^{cons} = -\vec{\nabla}_i V(\vec{r}_1, \vec{r}_2, ...\vec{r}_N)$$

donde $\vec{\nabla}_i$ deriva respecto de las componentes x, y, z de la posición $\vec{r}_i$. El trabajo de todas las fuerzas que actúan sobre el sistema puede descomponerse en la suma del trabajo de las fuerzas conservativas más el trabajo del resto de las fuerzas:

$$W_{A\to B} = W^{cons}_{A\to B} \ + \ W^{no\ cons}_{A\to B}$$

Como $W_{A\to B} = T(B) - T(A)$ y $W^{cons}_{A\to B} = V(A) - V(B)$, obtenemos

$$[T(B) + V(B)] - [T(A) + V(A)] = W^{no\ cons}_{A\to B}$$

donde A y B significan dos configuraciones del sistema ("inicial " y "final") dentro de una evolución dinámica del mismo. Definimos la **energía mecánica** del sistema como la suma de las energías cinética y potencial, $E \equiv T + V$. Entonces

$$E(B) - E(A) = W^{no\ cons}_{A\to B}$$

Si el trabajo de las fuerzas no conservativas es cero, entonces se conserva la energía mecánica del sistema.

2.2.1 Fuerzas conservativas internas

Si existen fuerzas internas conservativas habrá una energía potencial V_{ij} asociada a cada par de interacción. Así la fuerza sobre la partícula i debida a la partícula j será

$$\vec{f}_{ij} = -\vec{\nabla}_i V_{ij}$$

En caso que V_{ij} dependa sólo de la distancia entre partículas $r_{ij} \equiv |\vec{r}_{ij}| \equiv |\vec{r}_i - \vec{r}_j|$,

$$V_{ij} = V_{ij}(r_{ij}) \ ,$$

entonces $\vec{f}_{ij}$ tendrá la dirección de la recta que une las partículas ($\vec{f}_{ij} \parallel \vec{r}_{ij}$). En efecto

$$\vec{f}_{ij} = -\vec{\nabla}_i V_{ij} = -\frac{\partial V_{ij}}{\partial r_{ij}} \ \vec{\nabla}_i \, |\vec{r}_i - \vec{r}_j| = -\frac{\partial V_{ij}}{\partial r_{ij}} \ \frac{\vec{r}_{ij}}{r_{ij}}$$

Como vemos, $\vec{f}_{ij} \parallel \vec{r}_{ij}$. Para que se cumpla el principio de acción y reacción debe ser

$$V_{ij} = V_{ji} \qquad \Rightarrow \qquad \vec{f}_{ij} = -\ \vec{f}_{ji}$$

Trabajo de las fuerzas conservativas internas. Consideremos la interacción entre las partículas 1 y 2, y usemos que $V_{12} = V_{21}$:

$$\int_A^B \vec{f}_{12}\cdot d\vec{r}_1 + \int_A^B \vec{f}_{21}\cdot d\vec{r}_2 = -\int_A^B \vec{\nabla}_1 V_{12}\cdot d\vec{r}_1 - \int_A^B \vec{\nabla}_2 V_{21}\cdot d\vec{r}_2 = -\int_A^B dV_{12} = V_{12}(A) - V_{12}(B)$$

Esto significa que el trabajo total de las fuerzas internas conservativas es

$$W_{A\to B}^{cons\ interno} = \sum_{i<j} (V_{ij}(A) - V_{ij}(B))$$

($\sum_{i<j}$ es una suma doble). Es decir que la energía potencial interna es

$$V^{interna} = \sum_{i<j} V_{ij} = \frac{1}{2} \sum_{i\neq j} V_{ij}$$

La suma doble restringida a $i < j$ significa que por cada interacción entra un solo potencial. Por ejemplo, si dos partículas interactúan mediante un resorte, la energía potencial de la interacción es $V = 1/2\ k\ \eta^2$ (η es el estiramiento del resorte).

Si además existen fuerzas externas conservativas, la energía potencial total será la suma de $V^{interna}$ más $V^{externa}$.

2.3 Teorema del virial

Los sistemas formados por muchas partículas en interacción son sumamente complicados. Ya el problema de un sistema aislado de 3 cuerpos tiene solamente algunas soluciones exactas particulares. El teorema del virial nos da algunas características de la evolución de sistemas de N partículas. Definimos la **función virial**:

$$G \equiv \sum_{i=1}^{N} m_i\ \vec{v}_i \cdot \vec{r}_i$$

cuya derivada temporal es

$$\frac{dG}{dt} \equiv \sum_{i=1}^{N} m_i\ v_i^2 + \sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i = 2T + \sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i$$

Promediemos este resultado en un intervalo de tiempo τ:

$$\frac{1}{\tau} \int_0^\tau \frac{dG}{dt}\ dt = \frac{G(\tau) - G(0)}{\tau} = \ 2 < T > + < \sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i >$$

El primer miembro se anula en dos casos:

i) Si el sistema es periódico con período τ (entonces $G(\tau) = G(0)$) ,

ii) Si las posiciones y velocidades permanecen acotadas (G permanece acotado), y $\tau \to \infty$.

En cualquiera de los dos casos resulta el **teorema del virial**:

$$\boxed{< T > = \; -\frac{1}{2} < \sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i >}$$

Cuando se cumple esta relación hablamos de un sistema "virializado".
Nota histórica: en 1870 Clausius enunció el teorema diciendo que "la *vis viva* promedio es igual a su *virial*", por lo que el segundo miembro se llama *virial de Clausius*.

Sistema aislado. Si el sistema está aislado entonces

$$\vec{F}_i = \sum_{j \neq i} \vec{f}_{ij}$$

Nótese que

$$\vec{f}_{12} \cdot \vec{r}_1 + \; \vec{f}_{21} \cdot \vec{r}_2 = \; \vec{f}_{12} \cdot (\vec{r}_1 - \vec{r}_2) = \; \vec{f}_{12} \cdot \vec{r}_{12}$$

En este caso es

$$\sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i = \sum_{i<j} \vec{f}_{ij} \cdot \vec{r}_{ij}$$

($\sum_{i<j}$ es una suma doble).

Sistema aislado conservativo. En este caso las fuerzas derivan de los potenciales de interacción $V_{ij}(r_{ij})$:

$$\vec{f}_{ij} = -\frac{\partial V_{ij}}{\partial r_{ij}} \frac{\vec{r}_{ij}}{r_{ij}}$$

Si los potenciales tienen la forma $V_{ij} \propto \; r_{ij}^{\alpha}$ ($\alpha = -1$ para la interacción gravitatoria; $\alpha = 2$ para la interacción elástica), entonces

$$\vec{f}_{ij} \cdot \vec{r}_{ij} = -\alpha \; V_{ij}$$

Por lo tanto

$$\sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i = -\alpha \sum_{i<j} V_{ij} = -\alpha \; V$$

En este caso el teorema del virial dice que

$$< T > = \; \frac{\alpha}{2} \; < V > \tag{2.1}$$

Además podemos valernos de la conservación de la energía mecánica, $< T > + < V >= E$ para obtener

$$< T > = \frac{\alpha \; E}{\alpha + 2} \; , \qquad\qquad < V > = \frac{2 \; E}{\alpha + 2}$$

En el caso gravitatorio ($\alpha = -1$) el sistema virializado tiene necesariamente $E < 0$. En los sistemas auto-gravitantes con $E > 0$ al menos una partícula del sistema se alejará indefinidamente a medida que el sistema evoluciona (como situación particular, también podría suceder que el sistema colapse en un tiempo finito).[1]

Aplicaciones

1) En un sistema auto-gravitante (por ejemplo, una galaxia) hagamos una estimación de $< T >$ y $< V >$ suponiendo que las partículas son de la misma masa m :

$$< T > \approx \frac{N\ m}{2}\ < v^2 > , \qquad\qquad < V > \approx - \frac{N(N-1)}{2}\ \frac{G\ m^2}{< r >}$$

donde la energía potencial tiene en cuenta el número de interacciones (el número de pares de partículas entre N), y el promedio de V se toma como el promedio de distancias entre pares de partículas. Reemplazamos en la ecuación (2.1); si $N >> 1$ resulta

$$2 < v^2 > < r > \approx\ G\ N\ m\ =\ G\ M$$

lo cual da una forma de estimar la masa M de una galaxia a partir de la velocidad de sus estrellas y la distancia entre las mismas. De esta manera Zwicky (1933, 1937) estimó masas de cúmulos de galaxias que resultaron unas 800 veces más grandes que las estimaciones que se obtenían de la luminosidad de los mismos. Esta fue la primera evidencia de la existencia de grandes cantidades de materia oscura (no luminosa) en galaxias y cúmulos de galaxias.

1. Nótese que G es la derivada del **momento de inercia escalar** I:

$$G \equiv \frac{1}{2}\frac{d}{dt}\sum_{i=1}^{N} m_i\ r_i^2 \equiv \frac{1}{2}\frac{dI}{dt}$$

Entonces nuestra primera ecuación dice que

$$\frac{dG}{dt} \equiv \frac{1}{2}\frac{d^2 I}{dt^2} = 2T + \sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i = 2T - \alpha V = (2+\alpha)T - \alpha E$$

En el caso gravitatorio queda

$$\frac{1}{2}\frac{d^2 I}{dt^2} = T + E$$

Por lo tanto, $E > 0 \Rightarrow d^2 I/dt^2 > 0$. $I(t)$ es una función positiva cóncava hacia arriba. Viene entonces de infinito y va a infinito, lo que significa que al menos una partícula alcanza posiciones infinitamente alejadas. Como caso particular, también podría suceder que todas las partículas llegaran a una misma posición en un tiempo finito y el sistema auto-gravitante colapsara.

2) Veamos un *gas ideal* confinado en un recipiente. En un gas ideal despreciamos las interacciones a distancia entre las moléculas; suponemos que las moléculas (vistas como partículas) sólo interactúan cuando chocan entre sí. En ese caso la interacción no contribuye al virial de Clausius:

$$\vec{f}_{12} \cdot \vec{r}_1 + \vec{f}_{21} \cdot \vec{r}_2 = \vec{f}_{12} \cdot \vec{r}_{12} = 0 \ , \qquad \textit{porque} \quad \vec{r}_{12} = 0$$

Además debemos considerar las fuerzas externas, que se producen cuando las moléculas chocan con la pared del recipiente. Si pensamos el gas como un fluido, entonces la fuerza externa es $-p\, d\vec{S}$, donde p es la presión y $d\vec{S}$ es la superficie de la pared (con orientación exterior). Entonces calcularemos el virial de Clausius como

$$\sum_{i=1}^{N} \vec{F}_i \cdot \vec{r}_i \quad \longrightarrow \quad -\oint p\, d\vec{S} \cdot \vec{r} = -p \int \vec{\nabla} \cdot \vec{r}\, dv = -3\, p \times \textit{volumen}$$

y reemplazando en el teorema del virial,

$$< T > = \frac{3}{2}\, p \ \times \textit{volumen}$$

Pero la ley de los gases ideales dice que $p \times \textit{volumen} = N\ k_B \times \textit{temperatura}$ (k_B es la constante de Boltzmann). Por lo tanto concluímos que la temperatura de un gas es una medida de la energía cinética media de las partículas que lo componen.

3 Coordenadas generalizadas. Vínculos. Trabajos virtuales. Ecuaciones de Lagrange

3.1 Coordenadas generalizadas

La posición de una partícula se determina mediante 3 coordenadas, que pueden ser de distinto tipo (distancias, ángulos, etc.). Las llamaremos **coordenadas generalizadas** q_k. Un sistema de N partículas se describe con $3N$ coordenadas generalizadas.

3.2 Vínculos

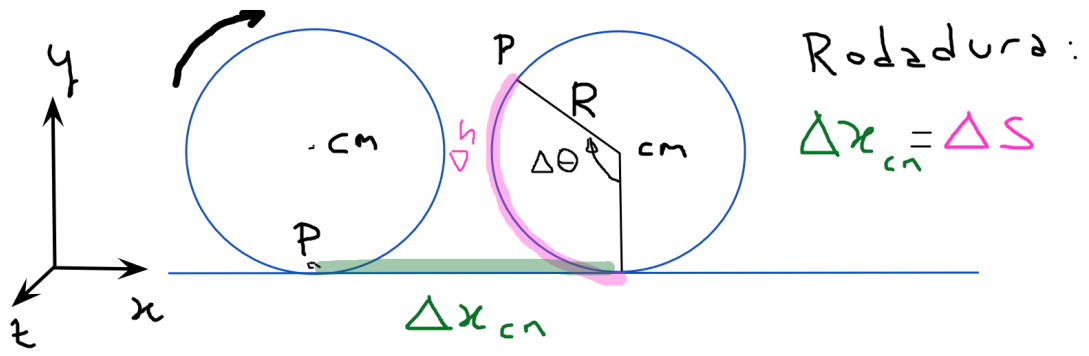
A menudo las $3N$ coordenadas generalizadas no son todas necesarias porque existen **condiciones de vínculo** entre las mismas, que hacen que las $3N$ coordenadas q_k no sean independientes.

Por ejemplo, en un cuerpo rígido alcanzan 6 coordenadas para establecer la posición de las partículas que lo componen. Conociendo la posición de un punto cualquiera del mismo (3 coordenadas) y la orientación del cuerpo en el espacio (3 ángulos que corresponden a tres rotaciones independientes) podemos saber la posición de cualquiera de sus puntos. Decimos que el cuerpo rígido tiene 6 **grados de libertad**.

El cuerpo rígido puede estar sometido a vínculos adicionales. Por ejemplo, un cilindro que rueda sin deslizar y sin modificar la orientación de su eje de rotación (ver Figura) cumple con las ecuaciones

$$x_{CM} = R\,\theta\ , \qquad y_{CM} = R\ , \qquad z_{CM} = 0$$

A estas tres ecuaciones de vínculo debemos agregar dos más que corresponden a congelar las otras dos posibles rotaciones que el cuerpo puede tener. En total son 5 vínculos sobre seis posibles grados de libertad. Esto reduce los grados de libertad a sólo uno, en este caso. Está claro que ese único grado de libertad puede describirse con el ángulo θ, la coordenada x_{CM}, o cualquier otra coordenada generalizada equivalente.



La existencia de m ecuaciones de vínculo (o *ligadura*) independientes reduce el número de grados de libertad a

$$n = 3N - m$$

3.2.1 Vínculos holónomos

Son aquellos que se expresan como relaciones entre las coordenadas (como en el ejemplo anterior). Eso permite describir la dinámica del sistema en términos de n coordenadas generalizadas independientes, que llamaremos q_μ ($\mu = 1, ..., n$).

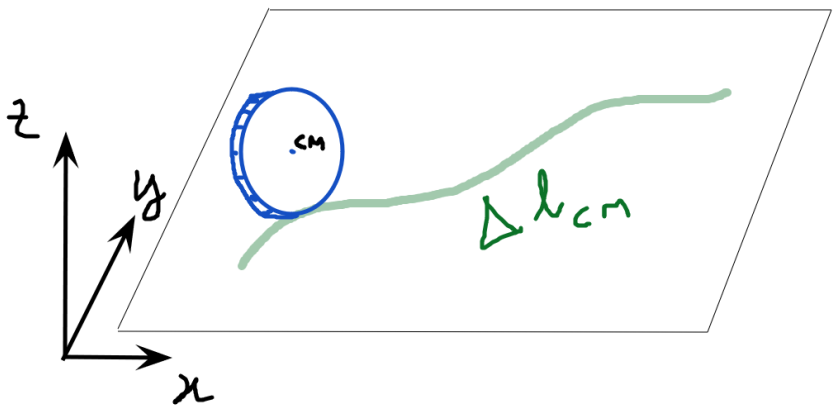
3.2.2 Vínculos anholónomos

En otros casos, las ligaduras se expresan como relaciones entre desplazamientos infinitesimales que no pueden integrarse. Por ejemplo, en la Figura anterior, el primer vínculo podría expresarse en la forma

$$dx_{CM} = R\ d\theta\ ,$$

pero esta relación es integrable y da $x_{CM} = R\ \theta + const$. En cambio, si la dirección del eje del cilindro variara a lo largo de la evolución entonces la situación cambiaría sustancialmente. Consideremos el caso de una moneda que rueda sin deslizar con su eje condicionado a mantenerse horizontal (es decir, la moneda no puede inclinarse); veamos la Figura:

.



Ahora el CM no se desplaza sobre una línea determinada, como lo hacía en el ejemplo anterior. El problema no es que la trayectoria sea curva sino que no está prefijada; la moneda es libre de torcer su rumbo, sin inclinarse, dependiendo de las fuerzas que actúen sobre ella. Todavía podemos decir que

$$d\ell_{CM} = R\ d\theta$$

donde $\Delta\ell_{CM}$ se mide sobre una curva en el plano $x - y$:

$$\Delta\ell = \int \sqrt{dx^2 + dy^2} = \int \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\ dx$$

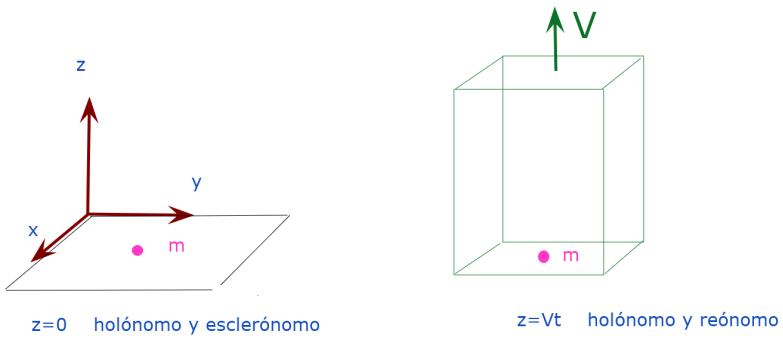
Esta expresión sólo podría ser integrada si conociéramos la curva $y(x)$ seguida por el CM. Pero, a diferencia del ejemplo anterior, no hemos establecido una ligadura para esta curva. En otras palabras, $d\ell = \sqrt{dx^2 + dy^2}$ no es un diferencial exacto. Así el vínculo permanece como **anholónomo** y tiene la forma

$$\sqrt{dx_{CM}^2 + dy_{CM}^2} = R\ d\theta$$

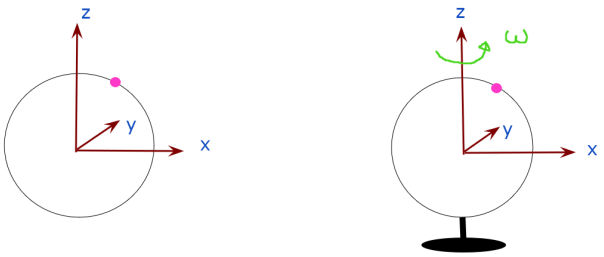
3.2.3 Vínculos esclerónomos y reónomos

Una segunda clasificación de los vínculos se refiere a la presencia o no del tiempo t en la ecuación de vínculo. Los vínculos **esclerónomos** no dependen del tiempo (del griego: $\sigma\kappa\lambda\eta\rho\acute{o} \rightarrow$ *duro*). Los vínculos **reónomos** dependen del tiempo (del griego: $\rho\acute{\varepsilon}\omega \rightarrow$ *fluir*).

Veamos algunos ejemplos. En este primer ejemplo, la partícula tiene dos grados de libertad en ambos casos.



En el segundo ejemplo, el aro de la primera figura está fijo.



Hay dos vínculos holónomos y esclerónomos:

$$x^2 + z^2 = R^2 \ , \qquad y = 0$$

En la segunda figura el aro gira con una velocidad ω preestablecida. Hay dos vínculos holónomos y reónomos:[1]

$$x = \sqrt{R^2 - z^2}\ \cos\omega t\ , \qquad y = \sqrt{R^2 - z^2}\ \sin\omega t$$

La partícula enhebrada en el aro tiene un único grado de libertad en ambos casos.

3.3 Espacio de configuración

Es el espacio de las coordenadas generalizadas utilizadas para describir el sistema. Si hay m vínculos holónomos el espacio de configuración tiene dimensión $n = 3N - m$.

3.4 Fuerzas aplicadas y de vínculo. Trabajos virtuales

Las fuerzas *aplicadas* son fuerzas que están determinadas en forma independiente del resto. Dependen de las posiciones de las partículas del sistema, y a veces de sus velocidades.

Las fuerzas de vínculo son las fuerzas que se encargan de mantener las ligaduras. No se las conoce a priori sino que dependen del resto las fuerzas que intervienen en el problema.

La fuerza de rozamiento estático es de vínculo porque garantiza que no haya deslizamiento entre dos superficies que están en contacto. La fuerza de rozamiento dinámico entre superficies secas escapa a esta clasificación; no es de vínculo pero tampoco es conocida de antemano (depende de la fuerza normal). En cambio, el rozamiento entre un cuerpo y un fluido puede ser descripto por una ley dependiente de la velocidad relativa entre la superficie del cuerpo y el fluido.

Cuando trabajamos con la leyes de Newton escribimos 3 ecuaciones diferenciales por partícula (las tres componentes de la ecuación vectorial $\vec{F} = m\ \ddot{\vec{r}}$); las componentes de los vectores posición $\vec{r_i}$ son las incógnitas del problema. Pero hay incógnitas adicionales: las fuerzas de vínculo. Para resolver el conjunto de incógnitas debemos entonces agregar más ecuaciones: las m ecuaciones de vínculo. Así tendremos un total de $3N + m$ ecuaciones para igual número de incógnitas. Ahora bien, en caso que no nos interese resolver las fuerzas de vínculo, ¿será posible lograr una reformulación de las ecuaciones dinámicas donde las únicas incógnitas sean los grados de libertad? Si esto se pudiese conseguir, el número de ecuaciones a resolver se reduciría a $n = 3N - m$. Vamos a ver que ese objetivo se puede alcanzar cuando los vínculos son holónomos, introduciendo el

1. En coordenadas cilíndricas ρ, φ, z los dos vínculos son $\rho^2 + z^2 = R^2$, $\varphi = \omega\ t$. Por otro lado las coordenadas cartesianas son $x = \rho\ \cos\varphi$, $y = \rho\ \sin\varphi$.

Principio de los trabajos virtuales. La idea básica consiste en notar que el trabajo total de las fuerzas de vínculo puede ser declarado nulo cuando los desplazamientos son compatibles con los vínculos, con el agregado de una sutileza que pasaremos a analizar.

1) En el caso de la partícula que se mueve libremente sobre una mesa (el vínculo es $z = 0$), el desplazamiento compatible con el vínculo es $\delta\vec{r} = (\delta x, \delta y, \delta z = 0) = (\delta x, \delta y, 0)$. Como la fuerza de vínculo es normal a la mesa, entonces no hace trabajo.

2) Consideremos dos bloques que se desplazan a lo largo de una recta horizontal (eje x), unidos por una soga tensa de masa despreciable y longitud L. La soga establece un vínculo holónomo y esclerónomo entre los bloques (fija la distancia entre los mismos: $x_2 - x_1 = L$), de manera que los desplazamientos de los bloques compatibles con los vínculos son iguales $\delta\vec{r}_1 = \delta\vec{r}_2 = (\delta x, 0, 0)$. Por otro lado, las fuerzas de las sogas sobre los bloques son iguales y opuestas. Cada una por separado realiza trabajo; pero, como los desplazamientos son iguales, la suma de los trabajos de las fuerzas de la soga sobre los bloques se anula. Además se anulan los trabajos de las fuerzas normales, como en el ejemplo anterior.

3) Analicemos ahora nuestro primer ejemplo de vínculo reónomo, donde una partícula se mueve libremente sobre el piso de una cabina que asciende con velocidad V. El vínculo es $z = Vt$, y el desplazamiento compatible con el vínculo es $\delta\vec{r} = (\delta x, \delta y\,, \delta z = V\ \delta t) = (\delta x, \delta y\,, V\ \delta t)$. La fuerza de vínculo $\vec{R}$ es normal al piso: $\vec{R} = R\ \hat{k} = (0, 0, R)$; su trabajo ante un desplazamiento compatible con el vínculo es $\delta W = \vec{R} \cdot \delta\vec{r} = R\ V\ \delta t$, que es distinto de cero.

Podremos incluir los vínculos reónomos en el Principio de los trabajos virtuales si introducimos los **desplazamientos virtuales** $\delta\vec{r}^{\,V}$, que son desplazamientos compatibles con los vínculos "a tiempo fijo" (es decir, con $\delta t = 0$). En el caso anterior, el desplazamiento virtual será $\delta\vec{r}^{\,V} = (\delta x, \delta y\,, V\ \delta t)|_{\delta t=0} = (\delta x, \delta y\,, 0)$. Así, el **trabajo virtual** de $\vec{R}$ se anulará como en el primer ejemplo.

La reformulación de las leyes de la dinámica partirá del **Principio de los trabajos virtuales**, que afirma que la suma de los trabajos virtuales de las fuerzas de vínculo es cero. En caso que los vínculos sean holónomos, llegaremos a $n = 3N - m$ ecuaciones dinámicas que reflejarán la evolución de los n grados de libertad.

En la segunda ley de Newton reescribiremos la resultante de fuerzas $\vec{F}_i$ como $\vec{F}_i + \vec{R}_i$, para reservar el nombre de $\vec{F}_i$ a la resultante de las fuerzas *aplicadas* sobre la masa m_i, mientras que $\vec{R}_i$ será la resultante de las fuerzas de vínculo sobre dicha masa. De esa forma escribiremos $\vec{F}_i + \vec{R}_i = m_i\ \vec{a}_i$. Si multiplicamos por los desplazamientos virtuales $\delta\vec{r}_i^{\,V}$ que cada partícula puede realizar, y sumamos las N ecuaciones obtenemos

$$\sum_{i=1}^{N} (\vec{F}_i + \vec{R}_i) \cdot \delta\vec{r}_i^{\,V} = \sum_{i=1}^{N} m_i\ \vec{a}_i \cdot \delta\vec{r}_i^{\,V}$$

Valiéndonos del principio de trabajos virtuales, que dice que $\sum_{i=1}^{N} \vec{R}_i \cdot \delta\vec{r}_i{}^V = 0$, resulta

$$\sum_{i=1}^{N} \vec{F}_i \cdot \delta\vec{r}_i{}^V = \sum_{i=1}^{N} m_i \; \vec{a}_i \cdot \delta\vec{r}_i{}^V$$

Esta expresión se conoce como **Principio de d'Alembert**.

3.5 Ecuaciones de Lagrange

En el **Principio de d'Alembert** las fuerzas de vínculo $\vec{R}_i$ han sido excluidas por aplicación del Principio de los trabajos virtuales. Las fuerzas que permanecen en la expresión son las fuerzas *aplicadas* (aquellas que se expresan en función de las posiciones y velocidades de las partículas). Por otro lado, los desplazamientos $\delta\vec{r}_i{}^V$ son desplazamientos *virtuales* (a $\delta t = 0$) compatibles con los vínculos.

Para entender el significado de esta expresión es esencial comprender que los desplazamientos virtuales no son independientes, sino que están sujetos a cumplir las ecuaciones de vínculo a tiempo fijo. Porque si la igualdad fuera válida para desplazamientos arbitrarios entonces se concluiría que los coeficientes de cada desplazamiento de un lado y otro de la igualdad deben ser iguales. Es decir, se concluiría que $\vec{F}_i = m_i \; \vec{a}_i$, lo cual es falso porque lo verdadero es que $\vec{F}_i + \vec{R}_i = m_i \; \vec{a}_i$.

Para dar el siguiente paso, debemos darle forma a los desplazamientos virtuales compatibles con los vínculos. Si los vínculos son holónomos, entonces tendremos m relaciones entre las coordenadas (y el tiempo) que permitirán pasar de las $3N$ coordenadas originales a $n = 3N - m$ coordenadas que podremos variar independientemente sin afectar los vínculos. Equivalentemente, podemos introducir n coordenadas generalizadas q_μ ($\mu = 1, ..., n$) que describan los grados de libertad a los que queda reducido el sistema una vez que los vínculos son tenidos en cuenta. Es decir que las coordenadas q_μ pueden ser variadas libre e independientemente sin afectar los vínculos.

Una vez elegido un conjunto de n coordenadas generalizadas q_μ para describir los grados de libertad del sistema, escribiremos los N vectores posición en función de las q_μ's (y de t, en caso que haya vínculos reónomos),

$$\vec{r}_i = \vec{r}_i(q_1, ..., q_n, t)$$

Los desplazamientos

$$\delta\vec{r}_i = \sum_{\mu=1}^{n} \frac{\partial \; \vec{r}_i}{\partial q_\mu} \; \delta q_\mu + \frac{\partial \; \vec{r}_i}{\partial t} \; \delta t$$

son automáticamente compatibles con los vínculos, pues sólo entran en juego las variaciones de las coordenadas asociadas a los grados de libertad. Para que sean virtuales eliminamos el último término:

$$\delta\vec{r}_i{}^V = \sum_{\mu=1}^{n} \frac{\partial \; \vec{r}_i}{\partial q_\mu} \; \delta q_\mu$$

Reemplazando en el Principio de d 'Alembert,

$$\sum_{\mu=1}^{n}\sum_{i=1}^{N}\left(\vec{F}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right)\,\delta q_\mu=\sum_{\mu=1}^{n}\sum_{i=1}^{N}\left(m_i\,\vec{a}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right)\,\delta q_\mu \tag{3.1}$$

Vamos a trabajar con el miembro de la derecha para darle una forma más interesante. Comencemos escribiendo

$$\vec{a}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}=\frac{d}{dt}\left(\vec{v}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right)-\vec{v}_i\cdot\frac{d}{dt}\left(\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right) \tag{3.2}$$

Veamos que

$$\vec{v}_i=\frac{d}{dt}\,\vec{r}_i(q_\mu(t),\ t)=\frac{\partial\,\vec{r}_i}{\partial t}+\sum_{\nu=1}^{n}\frac{\partial\,\vec{r}_i}{\partial q_\nu}\,\dot{q}_\nu$$

Entonces $\vec{v}_i=\vec{v}_i(q_\mu,\dot{q}_\mu,t)$, donde la única dependencia en las **velocidades generalizadas** $\dot{q}_\mu$ es la que se observa en el último término. Entonces obtenemos que

$$\frac{\partial\,\vec{v}_i}{\partial\dot{q}_\mu}=\frac{\partial\,\vec{r}_i}{\partial q_\mu}$$

Por otro lado, también podemos ver que

$$\frac{d}{dt}\left(\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right)=\frac{\partial^2\,\vec{r}_i}{\partial t\,\partial q_\mu}+\sum_{\nu=1}^{n}\frac{\partial^2\,\vec{r}_i}{\partial q_\nu\,\partial q_\mu}\,\dot{q}_\nu$$

Viendo a las coordenadas y velocidades generalizadas como variables independientes, entonces

$$\frac{d}{dt}\left(\frac{\partial\,\vec{r}_i}{\partial q_\mu}\right)=\frac{\partial}{\partial q_\mu}\left(\frac{\partial\,\vec{r}_i}{\partial t}+\sum_{\nu=1}^{n}\frac{\partial\,\vec{r}_i}{\partial q_\nu}\,\dot{q}_\nu\right)=\frac{\partial\,\vec{v}_i}{\partial q_\mu}$$

reemplazando estos resultados en la ecuación (3.2) llegamos a

$$\vec{a}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}=\frac{d}{dt}\left(\vec{v}_i\cdot\frac{\partial\,\vec{v}_i}{\partial\dot{q}_\mu}\right)-\vec{v}_i\cdot\frac{\partial\,\vec{v}_i}{\partial q_\mu}=\frac{1}{2}\frac{d}{dt}\left(\frac{\partial\,v_i^2}{\partial\dot{q}_\mu}\right)-\frac{1}{2}\,\frac{\partial\,v_i^2}{\partial q_\mu}$$

Entonces

$$\sum_{i=1}^{N}m_i\,\vec{a}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}=\frac{d}{dt}\left(\frac{\partial\,T}{\partial\dot{q}_\mu}\right)-\frac{\partial\,T}{\partial q_\mu}$$

Por otro lado, llamaremos **fuerza generalizada** Q_μ a

$$Q_\mu\equiv\sum_{i=1}^{N}\vec{F}_i\cdot\frac{\partial\,\vec{r}_i}{\partial q_\mu}$$

(nótese que el índice no alude a partículas sino a grados de libertad). Reemplazando en la ecuación (3.1) llegamos a

$$\sum_{\mu=1}^{n} \left(\frac{d}{dt} \left(\frac{\partial\, T}{\partial \dot{q}_\mu} \right) - \frac{\partial\, T}{\partial q_\mu} - Q_\mu \right) \delta q_\mu = 0 \tag{3.3}$$

Pero como los δq_μ son arbitrarios e independientes, entonces concluimos que la igualdad vale porque se anulan cada uno de los coeficientes de los δq_μ. Entonces obtenemos n ecuaciones dinámicas que se escriben

$$\boxed{\frac{d}{dt} \left(\frac{\partial\, T}{\partial \dot{q}_\mu} \right) - \frac{\partial\, T}{\partial q_\mu} = Q_\mu \qquad \mu = 1, ..., n}$$

Son tantas ecuaciones como grados de libertad, y no aparecen las fuerzas de vínculo. Para usar estas ecuaciones es suficiente conocer la energía cinética del sistema como una función de las coordenadas y velocidades generalizadas, $T = T(q_\mu, \dot{q}_\mu, t)$, calculada en un sistema de referencia inercial. Estas son las **ecuaciones de Lagrange** (o de Euler-Lagrange), aunque ese nombre suele reservarse al caso conservativo que veremos enseguida. Nótese que la *forma* de las ecuaciones es independiente del carácter de las coordenadas generalizadas utilizadas (*covariancia*). Claro que para cada problema habrá coordenadas más aptas, que otorguen una forma más sencilla a la función T, y de ese modo faciliten la resolución de las ecuaciones.

Las ecuaciones de Lagrange son ecuaciones diferenciales de segundo orden para las funciones $q_\mu(t)$. Su resolución permite obtener la evolución del sistema físico en el espacio de configuración. Cada solución particular corresponde a un determinado conjunto de valores iniciales para q_μ y $\dot{q}_\mu$.

Si las fuerzas aplicadas son conservativas, $\vec{F}_i = -\vec{\nabla}_i V$, las fuerzas generalizadas resultan

$$Q_\mu = -\sum_{i=1}^{N} \vec{\nabla}_i V \cdot \frac{\partial\, \vec{r}_i}{\partial q_\mu} = -\frac{\partial\, V}{\partial q_\mu}$$

Entonces las fuerzas generalizadas pueden pasarse al primer miembro de las ecuaciones de Lagrange y combinarse con el segundo término de las mismas para dar $\partial(T - V)/\partial q_\mu$. Pero como $\partial V/\partial \dot{q}_\mu = 0$, no cometemos error si V es incluido también en el primer término; así obtendremos las ecuaciones que más usualmente se denominan **ecuaciones de Lagrange** (o de Euler-Lagrange):

$$\boxed{\frac{d}{dt} \left(\frac{\partial\, L}{\partial \dot{q}_\mu} \right) - \frac{\partial\, L}{\partial q_\mu} = 0 \qquad \mu = 1, ..., n}$$

donde

$$\boxed{L(q_\mu, \dot{q}_\mu, t) \;\equiv\; T - V}$$

es la función Lagrangiana del sistema físico.

Ejemplos

1) La oscilación de un **péndulo simple** en un plano implica un único grado de libertad. La coordenada más apta para describir ese grado de libertad es el ángulo θ que va desde la vertical hasta el hilo que sujeta la partícula de masa m (pero también podríamos usar la altura de la partícula u otra coordenada equivalente).

El vínculo es holónomo y esclerónomo: la distancia de la partícula al punto de suspensión es igual a la longitud del hilo ℓ (involucra coordenadas en forma independiente del tiempo), y, además, el movimiento está restringido a ocurrir en un plano.

El vector velocidad es $\vec{v} = \ell\ \dot{\theta}\ \hat{\theta}$, entonces la energía cinética es $T = (1/2)\ m\ \ell^2\ \dot{\theta}^2$.

La energía potencial gravitatoria es $V = -mg\ell\ \cos\theta$. Entonces el Lagrangiano es

$$L(\theta, \dot{\theta}) = \frac{1}{2}\ m\ \ell^2\ \dot{\theta}^2 + mg\ell\ \cos\theta$$

La (única) ecuación de Lagrange queda así:

$$\frac{d}{dt}\left(m\ \ell^2\ \dot{\theta}\right) + mg\ell\ \sin\theta = 0$$

es decir

$$\ddot{\theta} + \frac{g}{\ell}\ \sin\theta = 0$$

que es la ecuación de movimiento del péndulo.

Para finalizar, utilicemos este mismo ejemplo para mencionar un aspecto de la fuerza generalizada. Si bien en este ejemplo no fue necesario calcularla porque la sustituimos por el conocimiento del potencial, veamos cuánto vale. La única fuerza *aplicada* es el peso (además hay una fuerza de vínculo: la tensión del hilo). Entonces la fuerza generalizada es

$$Q = \vec{F} \cdot \frac{\partial \vec{r}}{\partial \theta} = m\vec{g} \cdot \frac{\partial \vec{r}}{\partial \theta} = -mg\ \hat{\jmath} \cdot \frac{\partial \vec{r}}{\partial \theta}$$

donde

$$\frac{\partial \vec{r}}{\partial \theta} = \frac{\partial}{\partial \theta}\ (\ell\ \sin\theta\ \hat{\imath} - \ell\ \cos\theta\ \hat{\jmath}) = -\ell\ \cos\theta\ \hat{\imath} + \ell\ \sin\theta\ \hat{\jmath}$$

Entonces

$$Q = -mg\ell\ \sin\theta$$

Podemos verificar que $Q = -\partial V/\partial\theta$.

En este ejemplo la fuerza generalizada no es una fuerza sino un torque; es el torque del peso respecto del punto de suspensión del péndulo. Como vemos, el carácter de la fuerza generalizada depende del carácter de la coordenada generalizada q que utilicemos.

2) En el ejemplo de la partícula enhebrada en el aro rotante (ver §3.2.3), los vínculos son

$$x = \sqrt{R^2 - z^2}\ \cos\omega t\ , \qquad y = \sqrt{R^2 - z^2}\ \sin\omega t$$

La partícula tiene entonces un grado de libertad. Una coordenada generalizada adecuada para describir ese grado de libertad es el ángulo θ entre el eje de rotación del aro y el radio que une el centro del aro con la partícula. Como $z = R\cos\theta$, la posición de la partícula es

$$\vec{r} = x\ \hat{\imath} + y\ \hat{\jmath} + z\ \hat{k} = R\ \sin\theta\ \cos\omega t\ \hat{\imath} + R\ \sin\theta\ \sin\omega t\ \hat{\jmath} + R\ \cos\theta\ \hat{k}$$

Así calculamos la velocidad

$$\vec{v}\ =\ \frac{d\vec{r}}{dt}\ =\ \frac{\partial\vec{r}}{\partial t} + \frac{\partial\vec{r}}{\partial\theta}\ \dot{\theta}\ =\ ...$$

cuyo módulo al cuadrado resulta

$$v^2 = R^2\omega^2\sin^2\theta + R^2\dot{\theta}^2$$

La energía potencial es

$$V = mgz = mg\ R\ \cos\theta$$

y el Lagrangiano resulta

$$L = \frac{1}{2}\ m\ (R^2\omega^2\sin^2\theta + R^2\dot{\theta}^2) - mg\ R\ \cos\theta$$

Reemplazando en la ecuación de Lagrange se obtiene la ecuación dinámica para $\theta(t)$:

$$R\,\ddot{\theta} = \omega^2 R\ \sin\theta\ \cos\theta + g\ \sin\theta$$

Debe notarse que hemos trabajado desde un sistema inercial. En tal sistema se ve rotar el aro, y los vínculos quedan como los hemos escrito. Así se llegó a la forma de la energía cinética, donde vemos un término debido a la contribución de la rotación ω del aro. Es igualmente posible trabajar en el sistema no inercial fijo al aro. En ese caso los vínculos son distintos (corresponden a la situación donde $\omega = 0$). La energía cinética carecerá del término asociado a ω; pero en su lugar habrá que agregar las fuerzas de inercia, que entrarán en la fuerza generalizada. De una forma o de la otra se obtendrá la misma ecuación dinámica (la evolución de θ es independiente del carácter del sistema de referencia). Notoriamente, la ecuación dinámica iguala la aceleración tangencial $R\,\ddot{\theta}$ a la suma de las componentes tangenciales del peso y la fuerza centrífuga $-m\ \vec{\Omega}\times(\vec{\Omega}\times\vec{r}\,')$. La fuerza de vínculo es perpendicular al aro (la aceleración tangencial no depende de ella). Como en el sistema no inercial aparece también la fuerza de Coriolis $-2m\ \vec{\Omega}\times\vec{v}\,'$, que es perpendicular al plano del aro, concluimos que la fuerza de vínculo tiene una componente perpendicular al plano del aro para cancelar la fuerza de Coriolis (no hay aceleración perpendicular al plano del aro en el sistema no inercial). Esa componente de la fuerza de vínculo realiza trabajo; pero el trabajo *v*irtual es nulo pues se refiere a desplazamientos compatibles con los vínculos *a tiempo fijo*. Por otro lado, es fácil verificar que la fuerza de Coriolis no contribuye a la fuerza generalizada en el sistema no inercial.

Material histórico

J.L. Lagrange, *Méchanique Analitique* (1788).
https://gallica.bnf.fr/ark:/12148/bpt6k862625/

4 Vínculos anholónomos. Multiplicadores de Lagrange. Potenciales dependientes de la velocidad

4.1 Vínculos anholónomos

Recordemos el **Principio de d 'Alembert**

$$\sum_{i=1}^{N} \vec{F}_i \cdot \delta \vec{r}_i^{\,V} = \sum_{i=1}^{N} m_i \ \vec{a}_i \cdot \delta \vec{r}_i^{\,V}$$

Si algunos de los vínculos son anholónomos entonces no resulta posible reducir las coordenadas generalizadas a un número igual al de grados de libertad n. En ese caso, las posiciones de las partículas se expresarán en función de $M = n + \bar{m}$ coordenadas q_k, donde $\bar{m}$ es el número de vínculos anholónomos,[1]

$$\vec{r}_i = \vec{r}_i(q_k, t)$$

Como las coordenadas q_k ya no están asociadas a los grados de libertad, sus variaciones compatibles con los vínculos no son independientes. Por lo tanto, los desplazamientos virtuales

$$\delta \vec{r}_i^{\,V} = \sum_{k=1}^{M} \frac{\partial \ \vec{r}_i}{\partial q_k} \ \delta q_k^V$$

no resultan automáticamente compatibles con los vínculos, porque las coordenadas q_k todavía están restringidas por los vínculos anholónomos. Esto significa que cuando reemplacemos estos desplazamientos en el Principio de d 'Alembert,

$$\sum_{k=1}^{M} \sum_{i=1}^{N} \left(\vec{F}_i \cdot \frac{\partial \ \vec{r}_i}{\partial q_k} \right) \ \delta q_k^V = \sum_{k=1}^{M} \sum_{i=1}^{N} \left(m_i \ \vec{a}_i \cdot \frac{\partial \ \vec{r}_i}{\partial q_k} \right) \ \delta q_k^V \ ,$$

1. Si $\bar{m} = 0$ volvemos al caso anterior, donde el número de coordenadas en uso es igual al número de grados de libertad. Si todos los vínculos son anholónomos entonces es $\bar{m} = m = 3N - n$, y resulta $M = 3N$ (es decir que no podemos reducir el número de coordenadas).

aunque podamos trabajar cada miembro como lo hicimos en §3.5, y llegar al resultado de la ecuación (3.3),

$$\sum_{k=1}^{M} \left[\frac{d}{dt} \left(\frac{\partial\, T}{\partial \dot{q}_k} \right) - \frac{\partial\, T}{\partial q_k} - Q_k \right] \; \delta q_k^V = 0 \; , \tag{4.1}$$

ya no obtendremos las ecuaciones de Lagrange porque los δq_k^V no son independientes.

Consideraremos vínculos anholónomos que consistan en relaciones lineales entre diferenciales de las coordenadas,

$$\sum_{k=1}^{M} A_{ak}(q,t)\; dq_k + A_a(q,t)\; dt = 0 \; , \qquad a = 1, ..., \bar{m} \tag{4.2}$$

De modo que para desplazamientos virtuales tendremos:

$$\sum_{k=1}^{M} A_{ak}\; \delta q_k^V = 0 \; , \qquad a = 1, ..., \bar{m} \tag{4.3}$$

La ecuación (4.1) debe ser vista a la luz de los desplazamientos virtuales, que cumplen las ecuaciones (4.3). Una forma de hacerlo es utilizar el método de los **multiplicadores de Lagrange**. La idea es introducir $\bar{m}$ nuevas variables $\lambda_a(t)$ y combinarlas con las ecuaciones (4.3) para formar una única ecuación,

$$\sum_{a=1}^{\bar{m}} \lambda_a \; \left(\sum_{k=1}^{M} A_{ak}\; \delta q_k^V \right) = 0 \; ,$$

que sumaremos a (4.1):

$$\sum_{k=1}^{M} \left[\frac{d}{dt} \left(\frac{\partial\, T}{\partial \dot{q}_k} \right) - \frac{\partial\, T}{\partial q_k} - Q_k - \sum_{a=1}^{\bar{m}} \lambda_a\; A_{ak} \right] \; \delta q_k^V = 0$$

Ahora podríamos valernos de los $\bar{m}$ multiplicadores de Lagrange $\lambda_a(t)$, y elegirlos de tal forma de anular $\bar{m}$ términos de la suma sobre k; así quedarían n términos no nulos en la suma sobre k. Si bien los $M = n + \bar{m}$ desplazamientos virtuales δq_k^V no son independientes, n de ellos sí pueden ser considerados independientes, pues n es el número de grados de libertad. De esa forma sería lícito afirmar que los n términos que quedan en la suma deben anularse por separado, como hicimos en §3.5. En síntesis, por una razón o por otra todos los términos de la suma se anulan por separado:

$$\frac{d}{dt} \left(\frac{\partial\, T}{\partial \dot{q}_k} \right) - \frac{\partial\, T}{\partial q_k} = Q_k + \sum_{a=1}^{\bar{m}} \lambda_a\; A_{ak} \; , \qquad k = 1, ..., M$$

(en un caso conservativo usaremos el Lagrangiano L). Pero aquí tenemos M ecuaciones para $M + \bar{m}$ incógnitas (M q_k's junto con $\bar{m}$ λ_a's). Estas ecuaciones deben resolverse junto con las ecuaciones de vínculo (4.2) escritas en la forma

$$\sum_{k=1}^{M} A_{ak}(q,t)\ \dot{q}_k + A_a(q,t) = 0\ , \qquad a = 1, ..., \bar{m}$$

La conclusión es que en el caso de vínculos anholónomos no nos podemos librar de resolver las fuerzas de vínculo asociadas. En efecto, debimos agregar nuevas incógnitas, las $\bar{m}$ λ_a's, que aparecen en las ecuaciones dinámicas jugando el papel de fuerzas generalizadas de vínculo

$$P_k \equiv \sum_{a=1}^{\bar{m}} \lambda_a\ A_{ak}\ ,$$

junto con las propias ecuaciones de vínculo para poder resolver la dinámica del sistema.[2]

4.1.1 Vínculos holónomos vistos como anholónomos

Los vínculos holónomos pueden llevarse a la forma (4.2) por mera diferenciación. Esto puede ser útil si se desea conocer las fuerzas de vínculo. Los vínculos holónomos son relaciones entre las coordenadas generalizadas y el tiempo, que podemos escribir genéricamente así:

$$G_a(q_k, t) = 0\ , \qquad a = 1, ..., m$$

Si los diferenciamos resulta

$$\sum_{k=1}^{3N} \frac{\partial G_a}{\partial q_k}\ dq_k + \frac{\partial G_a}{\partial t}\ dt = 0\ ,$$

donde identificamos las funciones A_{ak} y A_a ,

$$A_{ak} = \frac{\partial G_a}{\partial q_k}\ , \qquad A_a = \frac{\partial G_a}{\partial t}\ ,$$

que deberemos utilizar en las fuerzas de vínculo generalizadas P_k . Si usamos como coordenadas generalizadas las componentes cartesianas de las posiciones $\vec{r}_i$ de las N partículas, entonces

$$A_{ak} = \frac{\partial G_a}{\partial q_k}\ \rightarrow \frac{\partial G_a}{\partial x_1}\ , \frac{\partial G_a}{\partial y_1}\ , \frac{\partial G_a}{\partial z_1}\ , \frac{\partial G_a}{\partial x_2}, \frac{\partial G_a}{\partial y_2}, ...,$$

2. El método de los multiplicadores de Lagrange se aplica a la búsqueda de extremos de una función $f(x_1, ..., x_n)$ en el subespacio definido por los m vínculos $g_a(x_1, ..., x_n) = 0$, $m < n$. La solución se obtiene extremando la función de $n + m$ variables $F(x_1, ..., x_n, \lambda_1, ..., \lambda_m) \equiv f(x_1, ..., x_n) + \sum_{a=1}^{m} \lambda_a\ g_a(x_1, ..., x_n)$.

que podemos agrupar vectorialmente como

$$\vec{A}_{ai} = \vec{\nabla}_i \, G_a$$

Ahora bien, cuando las coordenadas generalizadas son las componentes cartesianas de las posiciones de las partículas, las fuerzas generalizadas son las componentes cartesianas de la fuerza sobre cada partícula. En el caso de las fuerzas de vínculo generalizadas obtendremos que

$$\vec{R}_i = \sum_{a=1}^{m} \lambda_a \; \vec{A}_{ai} = \sum_{a=1}^{m} \lambda_a \; \vec{\nabla}_i \, G_a$$

lo que muestra que cada vínculo G_a genera sobre cada partícula una fuerza *normal* al mismo.

Ejemplos

1) Consideremos un movimiento plano de rodadura de un cilindro sobre una superficie horizontal. Aunque en este caso el vínculo es integrable, lo expresaremos como

$$dX_{CM} - R \; d\theta = 0 \qquad \Rightarrow \qquad A_k = (1, -R)$$

El Lagrangiano es

$$L = T - V = \frac{1}{2} \; M \; \dot{X}_{CM}^2 + \frac{1}{2} \; I_{CM}^{eje} \; \dot{\theta}^2 - cte$$

donde M y I_{CM}^{eje} son la masa del cilindro y su momento de inercia respecto del centro de masa y la dirección del eje de rotación. Las ecuaciones dinámicas son

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \; L}{\partial \dot{X}_{CM}}\right) - \frac{\partial \; L}{\partial X_{CM}} &= \lambda \; A_{X_{CM}} \\ \frac{d}{dt}\left(\frac{\partial \; L}{\partial \dot{\theta}}\right) - \frac{\partial \; L}{\partial \theta} &= \lambda \; A_\theta \\ \dot{X}_{CM} &= R \; \dot{\theta} \end{aligned}$$

es decir,

$$M \; \ddot{X}_{CM} = \lambda \; , \qquad I_{CM}^{eje} \; \ddot{\theta} = -\lambda \; R \; , \qquad \dot{X}_{CM} = R \; \dot{\theta}$$

Como vemos, λ es la componente x de la resultante de fuerzas (es la fuerza de rozamiento estático). La solución es $X_{CM}(t) = V_{CM} \; t + X_{CMo}$, $\theta(t) = (V_{CM}/R) \; t + \theta_o$, $\lambda = 0$. De modo que la fuerza de rozamiento estático se anula (en el caso horizontal el rozamiento estático sólo se necesita para establecer la rodadura).

2) En el problema de la partícula enhebrada en el aro rotante (ver §3.2.3), los vínculos son

$$x = \sqrt{R^2 - z^2}\ \cos\omega t\ , \qquad y = \sqrt{R^2 - z^2}\ \sin\omega t$$

que pueden combinarse para reescribirse en la siguiente forma:

$$G_1 = x^2 + y^2 + z^2 - R^2 = 0$$

$$G_2 = x\ \sin\omega t - y\ \cos\omega t = 0$$

Podemos calcular el gradiente de los vínculos para conocer la dirección de las fuerzas de vínculo respectivas (los módulos y sentidos se conocerán resolviendo las ecuaciones dinámicas y obteniendo los valores de $\lambda_1(t)$ y $\lambda_2(t)$):

$$\vec{\nabla} G_1 = 2x\ \hat{\imath} + 2y\ \hat{\jmath} + 2z\ \hat{k} = 2\ \vec{r}$$

$$\vec{\nabla} G_2 = \sin\omega t\ \hat{\imath}\ - \cos\omega t\ \hat{\jmath}$$

G_1 implica una fuerza de vínculo radial, y G_2 implica una fuerza de vínculo perpendicular al plano del aro.

4.2 Potenciales dependientes de la velocidad

El caso de fuerzas conservativas no es el único donde el término de fuerzas generalizadas puede incorporarse al Lagrangiano a través de un potencial. Podría ocurrir también que exista un **potencial dependiente de la velocidad** $U(q_\mu, \dot{q}_\mu, t)$ tal que

$$Q_\mu = \frac{d}{dt}\left(\frac{\partial\ U}{\partial \dot{q}_\mu}\right) - \frac{\partial\ U}{\partial q_\mu} \tag{4.4}$$

En ese caso el Lagrangiano sería

$$L = T - U$$

Vamos a ver que el caso de una carga eléctrica e moviéndose en un campo electromagnético externo es de este tipo. Recordemos que los campos $\vec{E}$, $\vec{B}$ se escriben en función de los potenciales escalar y vectorial como

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi\ , \qquad \vec{B} = \vec{\nabla} \times \vec{A}$$

Veremos que la carga está sujeta al potencial

$$U = e\ (\phi - \vec{v} \cdot \vec{A})$$

Usando coordenadas cartesianas para describir la dinámica de la carga nos aseguramos que las fuerzas generalizadas sean las componentes cartesianas de la fuerza sobre la carga. Comencemos por la componente x :

$$\frac{d}{dt}\left(\frac{\partial\ U}{\partial \dot{x}}\right) - \frac{\partial\ U}{\partial x} = -e\ \frac{dA_x}{dt} - e\frac{\partial\ \phi}{\partial x} + e\ \vec{v}\cdot\frac{\partial\ \vec{A}}{\partial x}$$

donde

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + \vec{v}\cdot\vec{\nabla}A_x$$

El primer término de dA_x/dt tiene en cuenta la variación temporal del campo externo $A_x(\vec{r},t)$ en cada posición. El segundo término da cuenta de la variación del campo porque la partícula se desplaza mientras transcurre el tiempo: $\vec{v}\cdot\vec{\nabla} = \dot{x}\ \partial/\partial x + \dot{y}\ \partial/\partial y + \dot{z}\ \partial/\partial z$. Debemos tener siempre presente que la derivada d/dt es una derivada a lo largo de la evolución del sistema. Entonces

$$\frac{d}{dt}\left(\frac{\partial\ U}{\partial \dot{x}}\right) - \frac{\partial\ U}{\partial x} = e\ E_x + e\ \vec{v}\cdot\left(-\vec{\nabla}A_x + \frac{\partial\ \vec{A}}{\partial x}\right)$$

donde aparece la componente x de la fuerza eléctrica. Veamos si el segundo término es la componente x de la fuerza magnética:

$$-\vec{\nabla}A_x + \frac{\partial\ \vec{A}}{\partial x} = -\frac{\partial A_x}{\partial y}\ \hat{\jmath} - \frac{\partial A_x}{\partial z}\ \hat{k} + \frac{\partial A_y}{\partial x}\ \hat{\jmath} + \frac{\partial A_z}{\partial x}\ \hat{k} = B_z\ \hat{\jmath} - B_y\ \hat{k}$$

Entonces

$$\vec{v}\cdot\left(-\vec{\nabla}A_x + \frac{\partial\ \vec{A}}{\partial x}\right) = v_y\ B_z - v_z\ B_y = (\vec{v}\times\vec{B})_x$$

Lo que obtuvimos para la componente x puede replicarse para las otras componentes. Finalmente, hemos demostrado que la fuerza de Lorentz se escribe como

$$e\ (\vec{E} + \vec{v}\times\vec{B}) = \frac{d}{dt}(\vec{\nabla}_{\vec{v}}\ U) - \vec{\nabla}_{\vec{r}}\ U$$

donde $\vec{\nabla}_{\vec{v}}$ es una notación para el operador vectorial cuyas componentes cartesianas son las derivadas parciales respecto de las componentes cartesianas de la velocidad: $\vec{\nabla}_{\vec{v}} = (\frac{\partial}{\partial v_x}, \frac{\partial}{\partial v_y}, \frac{\partial}{\partial v_z})$.

Siempre que usemos las componentes cartesianas de la posición de la partícula, las fuerzas generalizadas son iguales a las componentes cartesianas de la resultante de las fuerzas aplicadas sobre la partícula. El resultado obtenido muestra que en el caso de la fuerza de Lorentz las fuerzas generalizadas tienen la forma de la ecuación (4.4).

Otro ejemplo importante es el de las fuerzas de inercia en los sistemas no inerciales, según resultan de la ecuación (1.2). Las mismas se obtienen del potencial

$$U(\vec{r}\,',\vec{v}\,',t) = m\ \vec{A}_{O'}\cdot\vec{r}\,' - m\ \vec{v}\,'\cdot(\vec{\Omega}\times\vec{r}\,') - \frac{1}{2}\ m\ |\ \vec{\Omega}\times\vec{r}\,'|^2$$

que puede verse como la consecuencia directa de aplicar la relación (1.1) al Lagrangiano.

5 Principios variacionales

5.1 Cálculo variacional

El cálculo diferencial es la herramienta para encontrar máximos y mínimos de funciones. Pero hay otro tipo de problemas donde se buscan extremos o puntos estacionarios no de funciones sino de integrales de funciones. Por ejemplo:

1) El problema de encontrar el camino de menor longitud sobre un plano para unir dos puntos dados. En este caso habrá que minimizar la integral

$$\ell[y(x)] = \int_A^B \sqrt{dx^2 + dy^2} = \int_{x_A}^{x_B} \sqrt{1 + y'(x)^2}\, dx$$

En este caso $\ell[y(x)]$ depende de la función $y(x)$ a través de una integral que involucra la derivada de $y(x)$. Debe tenerse en cuenta que la expresión para la longitud de una curva depende de la geometría del espacio donde se traza la curva. Por ejemplo, la longitud sobre una superficie esférica, escrita en términos de las coordenadas angulares usuales en la esfera, tendría otra forma; la **funcional** ℓ (función de función) sería diferente.

2) El problema de encontrar el camino del descenso más rápido entre dos puntos A y B a distintas alturas y, que no estén sobre la misma vertical ($x_A \neq x_B$). Este es el problema de la *braquistócrona*, y fue estudiado por los hermanos Bernoulli en 1696. En este caso queremos minimizar

$$\Delta t = \int_A^B \frac{d\ell}{v} = \int_{x_A}^{x_B} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g\ y(x)}}\, dx$$

donde el denominador es $v(y)$, tal como resulta de la conservación de la energía mecánica $E = 0$. El tiempo que demanda recorrer un camino $y(x)$ es una funcional de $y(x)$: $\Delta t = \Delta t[y(x)]$. Como vemos, una funcional se construye con una función y sus derivadas.

3) El Principio de Fermat de la óptica geométrica (1679), que dice que el recorrido del rayo extremiza el camino óptico definido como $\int_A^B n\ d\ell$. El índice de refracción, que

se relaciona con la velocidad de fase ($n = c/v_f$), depende del camino del rayo a través de las propiedades del medio donde la luz se propaga.

La rama de las matemáticas que se ocupa del problema de extremar **funcionales** es el **cálculo variacional**, y fue fundada por Euler en 1733.

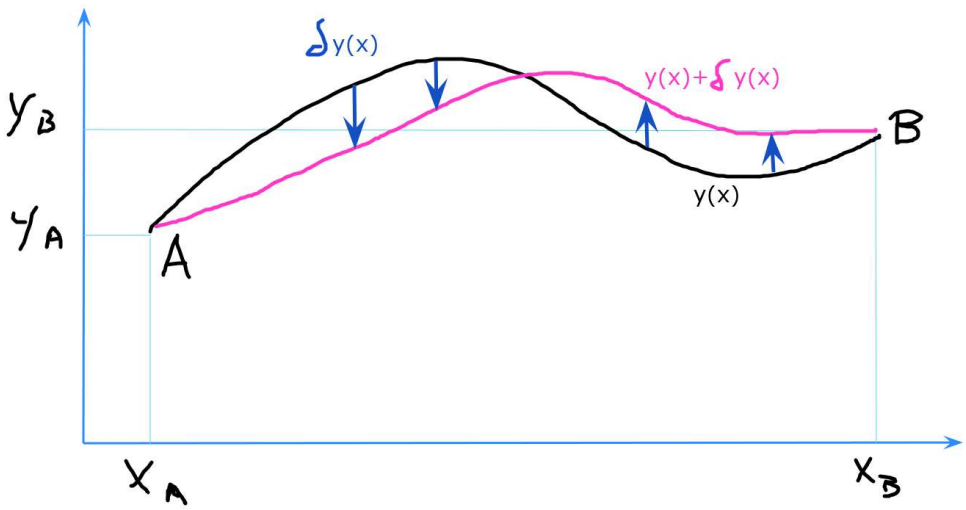
Vamos a ver el método general para extremar (o hacer estacionaria) una funcional de la forma

$$I[y(x)] = \int_{x_A}^{x_B} F(y(x),\ y'(x),\ x)\ dx$$

Mientras que en el cálculo diferencial extremamos una función $f(x)$ buscando el punto x donde una variación $x \to x+\delta x$ no afecte la función a primer orden ($\delta f = 0$), en el cálculo variacional debemos variar la función y, $y(x) \to y(x) + \delta y(x)$, en busca de la función $y(x)$ cuya variación no afecte la integral $I[y(x)]$ al primer orden en δy. Normalmente la variación $\delta y(x)$ está sujeta a condiciones de contorno; en nuestro caso, trabajaremos con variaciones que se anulan en los extremos de integración:

$$\delta y(x_A) = 0 = \delta y(x_B)$$

Entonces, las variaciones serán del tipo que se muestra en la Figura:[1]



1. La variación $\delta y(x)$ de una función $y(x)$ no debe confundirse con su diferencial: $dy = y'(x)\ dx$. La variación δy toma valores infinitesimales arbitrarios en cada x (salvo en los extremos, en nuestro caso). En cambio el diferencial dy está determinado por la derivada de la propia función y.

Nótese en la Figura que la variación de y también produce una variación de su derivada en cada valor de x:

$$\delta y' = \frac{d}{dx}(y(x) + \delta y(x)) - \frac{dy(x)}{dx} = \frac{d}{dx}\delta y(x)$$

Así, el cambio del valor de la integral $I[y(x)]$ debido a la variación de la función $y(x)$ es

$$\begin{aligned}
\delta I &= I[y + \delta y] - I[y] = \int_{x_A}^{x_B} [F(y + \delta y,\ y' + \delta y',\ x) - F(y,\ y',\ x)]\ dx \\
&= \int_{x_A}^{x_B} \left[\frac{\partial F}{\partial y}\ \delta y\ +\ \frac{\partial F}{\partial y'}\ \delta y'\right]\ dx = \int_{x_A}^{x_B} \left[\frac{\partial F}{\partial y}\ \delta y\ +\frac{\partial F}{\partial y'}\ \frac{d}{dx}\delta y\right]\ dx\ ,
\end{aligned}$$

donde reemplazamos el valor de $\delta y'$ en la última integral. Precisamente en ese término podemos hacer una integración por partes,

$$\delta I = \int_{x_A}^{x_B} \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\ \delta y(x)\ dx + \left[\frac{\partial F}{\partial y'}\ \delta y\right]_{x_A}^{x_B}$$

Pero el término de borde se anula por la condición de contorno $\delta y(x_A) = 0 = \delta y(x_B)$. Entonces queda

$$\delta I = \int_{x_A}^{x_B} \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\ \delta y(x)\ dx$$

Para que la integral I sea estacionaria ($\delta I = 0$) ante variaciones infinitesimales arbitrarias $\delta y(x)$, la función $y(x)$ debe ser entonces aquella que anula el integrando:

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

Esta es una ecuación diferencial de segundo orden que implica dos condiciones de contorno a ser fijadas para obtener una solución particular. En nuestro caso las condiciones de contorno fueron fijadas al momento de hacer la variación como $y(x_A) = y_A$, $y(x_B) = y_B$.

Este método del cálculo variacional se replica para funcionales que dependen de n funciones y sus primeras derivadas.[2] Como vemos, la ecuación obtenida es una ecuación

2. El método se extiende fácilmente para funcionales que dependan de derivadas de orden superior, con el agregado de las condiciones de contorno necesarias.

de Euler-Lagrange. Esto significa que las leyes de la Mecánica Analítica de Lagrange pueden verse como el resultado de un cálculo de variaciones. Entonces podríamos decir que las leyes de la Dinámica provienen de un **principio variacional**.

5.2 Principio de Hamilton

Los sistemas holónomos conservativos (o con potenciales dependientes de la velocidad como el que ya hemos visto) evolucionan de tal manera de hacer estacionaria la funcional **acción**

$$\boxed{S[q_\mu(t)] \equiv \int_{t_1}^{t_2} L(q_\mu \ , \dot{q}_\mu \ , t) \ dt}$$

entre configuraciones inicial y final fijas. Este es el **Principio de Hamilton** (a veces llamado Principio de *mínima acción*). En efecto, el enunciado implica que la evolución del sistema satisfará las ecuaciones de Euler-Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\mu}\right) - \frac{\partial L}{\partial q_\mu} = 0 \ , \qquad \mu = 1, ..., n$$

Alternativamente, podríamos trabajar con $3N$ coordenadas q_k, junto con los m vínculos holónomos $G_a(q_k, t)$, y los respectivos multiplicadores de Lagrange λ_a. Entonces, partiendo de la acción

$$S[q_k(t), \lambda_a(t)] \equiv \int_{t_1}^{t_2} \left(L(q_k \ , \dot{q}_k \ , t) + \sum_{a=1}^{m} \lambda_a \ G_a(q_k, t) \right) \ dt \ ,$$

la variación de las funciones $q_k(t)$, $\lambda_a(t)$ conduciría a las $3N + m$ ecuaciones

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} - \sum_{a=1}^{m} \lambda_a \ \frac{\partial G_a}{\partial q_k} &= 0 \ , \qquad k = 1, ..., 3N \\ G_a(q_k, t) &= 0 \ , \qquad a = 1, ..., m \end{aligned}$$

que son las ecuaciones que se obtuvieron cuando reintrodujimos las fuerzas de vínculo mediante el método de los multiplicadores de Lagrange.

La covariancia de las ecuaciones de Euler-Lagrange (el hecho que poseen una forma independiente de las coordenadas que se utilicen) no hace sino reflejar la **invariancia** de

la acción ante cambios de coordenadas generalizadas, $q_\mu \to q'_\mu = q'_\mu(q_1, ..., q_n, t)$ (también llamadas transformaciones de punto o de contacto). En efecto, las ecuaciones de Euler-Lagrange provienen de hacer estacionaria la integral del Lagrangiano con respecto al tiempo. Pero el valor de esa integral es independiente de las coordenadas generalizadas que se utilicen para describir la configuración del sistema (es decir, el valor del integrando $L = T - V$ no cambia frente a cambios de coordenadas generalizadas).

Nótese que el resultado del Principio variacional de Hamilton no se altera si agregamos al Lagrangiano L la derivada temporal de una función arbitraria $M(q_\mu, t)$:

$$L \to L + \frac{dM}{dt}$$

En tal caso la acción cambia por términos de borde:

$$S \to S + M \mid_{t=t_2} - M \mid_{t=t_1}$$

Los términos de borde no molestan porque el Principio involucra variaciones a extremos fijos,

$$\delta M \mid_{t=t_2} \; = \; 0 \; = \; \delta M \mid_{t=t_1}$$

Por lo tanto esos términos no contribuyen a la variación, ni tampoco aparecen en las ecuaciones de movimiento.

Veamos un ejemplo de esta situación. En §4.2 vimos que el Lagrangiano de una carga e en un campo electromagnético externo es $L = T - U$ donde

$$U = e\ (\phi - \vec{v} \cdot \vec{A})$$

y ϕ, $\vec{A}$ son los potenciales escalar y vectorial:

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi\ , \qquad\qquad \vec{B} = \vec{\nabla} \times \vec{A}$$

Como sabemos, los potenciales están definidos a menos de una *transformación de gauge*:

$$\vec{A} \to \vec{A} + \vec{\nabla} f\ , \qquad\qquad \phi \to \phi - \frac{\partial f}{\partial t}$$

Ante una transformación de gauge, el potencial U en el Lagrangiano cambia según

$$U \to U - e\ \left(\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f\right) = U - e\ \frac{df}{dt}$$

Por lo tanto, el Lagrangiano cambia por una derivada temporal total, que no afecta el resultado de las ecuaciones dinámicas. Esto es lo esperado, pues la dinámica de la carga depende de $\vec{E}$ y $\vec{B}$, que son invariantes de gauge.

Para finalizar, consideremos el cambio de la energía cinética de un sistema de partículas ante una transformación de Galileo. En §1.4 vimos que

$$T' = T - M\ \vec{V}_{CM} \cdot \vec{V} + \frac{1}{2} M\ V^2\ ,$$

que podemos escribir como

$$T' = T + \frac{d}{dt}(-M\ \vec{R}_{CM} \cdot \vec{V} + \frac{1}{2} M\ V^2 t)$$

Por otra parte, las distancias no cambian ante transformaciones de Galileo; de modo que los potenciales que dependan de distancias no serán afectados. Esto significa que los Lagrangianos de la Mecánica Clásica evaluados en dos sistemas de referencia inerciales distintos (en un caso será $L = T - V$, y en el otro será $L' = T' - V$) difieren en una derivada total. Entonces, si la acción que corresponde al primer sistema inercial es estacionaria, también lo será la acción que corresponde al otro sistema inercial. La misma evolución del sistema que hace estacionaria la acción en un sistema inercial, lo hará también en cualquier otro sistema inercial. Esta conclusión no es otra cosa que el **Principio de relatividad** ante transformaciones de Galileo.

5.2.1 ¿Máximo o mínimo?

La anulación de la variación de la acción a primer orden sólo dice que la acción es estacionaria sobre la evolución real del sistema. Para saber si estamos ante un máximo o un mínimo deberíamos analizar la variación al siguiente orden. Analicemos dos casos simples:

Movimiento unidimensional de la partícula libre

Sea $\bar{x}(t)$ la evolución real de la partícula libre, es decir el movimiento rectilíneo uniforme. Si variamos este movimiento, $\bar{x}(t) \to \bar{x}(t) + \delta x(t)$, el Lagrangiano evaluado sobre la evolución variada resulta

$$L = \frac{1}{2}\ m\ \dot{x}^2 = \frac{1}{2}\ m\ (\bar{v} + \delta\dot{x})^2 = \frac{1}{2}\ m\ \bar{v}^2 +\ m\ \bar{v}\ \delta\dot{x} + \frac{1}{2}\ m\ \delta\dot{x}^2$$

Por tratarse de una variación a partir de la evolución real, sabemos que la acción no sufrirá cambios a primer orden (siempre que la variación se anule en los bordes: $\delta x(t_1) = 0 = \delta x(t_2)$). A segundo orden el cambio es

$$\delta^2 S = \frac{m}{2} \int_{t_1}^{t_2} \delta\dot{x}^2\ dt > 0$$

Por lo tanto, la acción evaluada sobre la evolución real es un mínimo. El mismo resultado se obtiene para potenciales lineales, como el potencial gravitatorio en la vecindad de la superficie terrestre.

Movimiento oscilatorio armónico unidimensional

En este caso, el Lagrangiano evaluado sobre una variación de la evolución real queda

$$L = \frac{1}{2}\ m\ \dot{x}^2 - \frac{1}{2}\ k\ x^2 = \frac{1}{2}\ m\ (\bar{v} + \delta\dot{x})^2 - \frac{1}{2}\ k\ (\bar{x} + \delta x)^2$$

y la variación de la acción al segundo orden es

$$\delta^2 S = \int_{t_1}^{t_2} \left(\frac{1}{2}\ m\ \delta\dot{x}^2 - \frac{1}{2}\ k\ \delta x^2 \right)\ dt = \frac{m}{2} \int_{t_1}^{t_2} \left(\delta\dot{x}^2 - \omega_o^2\ \delta x^2 \right)\ dt$$

Para juzgar el signo de $\delta^2 S$ escojamos una variación $\delta x(t)$ sencilla,

$$\delta x(t) = \epsilon\ \sin\left[\omega(t - t_1)\right]$$

donde ϵ es un parámetro infinitesimal, y ω es la frecuencia de la variación (distinta de la frecuencia propia $\omega_o = \sqrt{k/m}$ del oscilador). Luego es

$$\delta\dot{x}(t) = \frac{d}{dt}\delta x = \epsilon\ \omega\ \cos\left[\omega(t - t_1)\right]$$

Para que se cumpla la condición de contorno

$$\delta x(t_1) = 0 = \delta x(t_2)$$

tomaremos $\omega = 2\pi/(t_2 - t_1)$. Por lo tanto

$$\delta^2 S = \frac{m\ \epsilon^2}{2} \int_{t_1}^{t_2} \left(\omega^2\ \cos^2\left[\omega(t - t_1)\right] - \omega_o^2\ \sin^2\left[\omega(t - t_1)\right] \right)\ dt = \frac{\pi\ m\ \epsilon^2}{2\ \omega}\ (\omega^2 - \omega_o^2)$$

donde vemos que el signo de $\delta^2 S$ puede ser tanto positivo como negativo, quedando en este caso supeditado a la relación entre el intervalo de integración $t_2 - t_1 = 2\pi/\omega$ y el período del oscilador.

Como ejercicio, considérese el signo de $\delta^2 S$ para la variación $\delta x(t) = \epsilon(t - t_2)(t - t_1)$, donde ϵ es un parámetro infinitesimal.

Bibliografía adicional

C. Lanczos, *The Variational Principles of Mechanics*, Dover.
https://archive.org/details/VariationalPrinciplesOfMechanicsLanczos/mode/2up

W. Yourgrau y S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory*, Dover.
https://archive.org/details/variationalprinc0000your

6 Simetría y Conservación. Teoremas de Noether

6.1 Coordenadas cíclicas

Supongamos que el Lagrangiano de un sistema físico, $L(q_\mu, \dot{q}_\mu, t)$ no depende de una determinada coordenada generalizada (pero depende de la respectiva velocidad generalizada). Diremos que esa coordenada es **cíclica**. Para fijar ideas, sea q_1 la coordenada cíclica. ¿Qué implicaría para la evolución del sistema? Si tomamos la ecuación de Lagrange correspondiente a $\mu = 1$ tendremos:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) = 0$$

pues $\partial L/\partial q_1 = 0$. Entonces obtenemos una ley de conservación:

$$\frac{\partial L}{\partial \dot{q}_1} = \text{constante}$$

Será útil definir el **momento canónicamente conjugado** a q_μ :

$$p_\mu \equiv \frac{\partial L}{\partial \dot{q}_\mu} \ .$$

Si una coordenada es cíclica en el Lagrangiano, entonces se conserva su momento conjugado.

Las cantidades conservadas son primeras integrales de las ecuaciones de movimiento (combinaciones de las coordenadas y las velocidades). Su conocimiento previo ayuda a la integración de las ecuaciones.

Ejemplo

El Lagrangiano de una partícula libre es $L = (m/2)(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$; por lo tanto x, y, z son cíclicas. Entonces la evolución de una partícula libre conserva los valores de $p_x = m\dot{x}$, $p_y = m\dot{y}$, $p_z = m\dot{z}$, que resultan ser las componentes de la cantidad de movimiento de la partícula.

Nota. No siempre la relación entre momentos conjugados y las cantidades de movimiento resulta tan directa como en el ejemplo anterior; eso depende de las coordenadas generalizadas que se utilicen. Más aun, si las velocidades participan en los potenciales entonces el momento conjugado recibirá contribuciones no sólo de la energía cinética sino también del potencial. Así, por ejemplo, para una carga e en un campo electromagnético externo, cuyo Lagrangiano es $L = (m/2)\ \vec{v}\cdot\vec{v} - e(\phi - \vec{v}\cdot\vec{A})$ los momentos conjugados a x, y, z son

$$p_x = m\ \dot{x} + e\ A_x\ , \qquad p_y = m\ \dot{y} + e\ A_y\ , \qquad p_z = m\ \dot{z} + e\ A_z$$

En consecuencia, los momentos conjugados resultan depender del gauge que se elija para el campo externo.

Por ejemplo, si el campo externo es un campo eléctrico uniforme $\vec{E} = E\ \hat{\imath}$ podríamos usar el gauge $\phi = -E\ x$, $\vec{A} = 0$, tanto como $\phi = 0$, $\vec{A} = -E\ t\ \hat{\imath}$ (recordemos que $\vec{E} = -\vec{\nabla}\phi - \partial\vec{A}/\partial t$). En el primer caso la coordenada x no es cíclica, pues aparece en el potencial; entonces concluimos que no se conserva $p_x = m\ \dot{x}$ (pues hay aceleración en la dirección del campo). En el segundo caso x es cíclica; entonces se conserva $p_x = m\ \dot{x} - eE\ t$, que es una primera integral de movimiento. Integramos una vez, y obtenemos $x(t) = x_o + (p_x/m)t + (eE/m)t^2/2$.

6.2 Simetría y conservación

El caso de la coordenada cíclica es un ejemplo trivial de la relación entre *simetría y conservación*. La existencia de una coordenada cíclica q implica una **simetría** del Lagrangiano: L no cambia ante la transformación

$$q(t) \to q(t) + \epsilon$$

donde ϵ es una constante. En efecto, la transformación no altera el valor de L pues, por un lado, q no participa en L, y por el otro lado $\dot{q}(t)$ no cambia porque ϵ es constante. Esta simetría de L vale para cualquier evolución $q(t)$, cumpla o no con las ecuaciones de Lagrange. En cambio, la **conservación** de p, que se deriva de la simetría en cuestión, significa que p se mantiene constante a lo largo de la evolución real del sistema físico (aquella evolución $q(t)$ que satisface las ecuaciones de Lagrange).

6.3 Teorema de Noether

En 1918 Emmy Noether publicó dos teoremas acerca de la relación entre simetría y conservación en el contexto de las teorías de campos. Estos teoremas constituyen una piedra angular de la Física del siglo XX. Nosotros veremos aquí una versión del primer teorema adaptada al contexto de la Mecánica, y haremos alguna referencia al segundo teorema.

La relación entre simetría y conservación que surge de la existencia de coordenadas cíclicas es un aspecto muy parcial de esa relación, que mostrada de esa forma resulta fuertemente dependiente de las coordenadas generalizadas elegidas. Por ejemplo, el Lagrangiano de partícula libre en coordenadas cilíndricas, $L = (m/2)(\dot{\rho}^2 + \rho^2\dot{\theta}^2 + \dot{z}^2)$, exhibe sólo dos coordenadas cíclicas mientras que las tres coordenadas cartesianas de una partícula libre son cíclicas. Entonces, trataremos de enfocar la cuestión de fondo, en una forma que no dependa de las coordenadas generalizadas que se elijan.

Como sabemos, la acción $S[q_\mu(t)]$ es estacionaria sobre la evolución *específica* que satisface las ecuaciones de Euler-Lagrange, sujeta a valores dados de las coordenadas en t_1 y t_2. Esto significa que una variación *arbitraria* $\delta q(t)$ de esta solución, sin variación en los extremos, no produce cambios de la acción a primer orden ($\delta S = 0$). Por otro lado, la acción posee una simetría si su valor sobre evoluciones *arbitrarias* es invariante (no cambia) ante ciertas variaciones *específicas* $\delta_s q(t)$.

Para que $\delta_s q_\mu(t)$ sea una simetría del Lagrangiano (y, por lo tanto, de la acción) hace falta que $\delta_s L = 0$ sobre evoluciones arbitrarias. Podemos ampliar esta definición de simetría permitiendo que $\delta_s L$ sea igual a una derivada temporal total. En ese caso la acción cambia por términos de borde que no afectan las ecuaciones dinámicas (tendremos una simetría de las ecuaciones). Entonces la situación de simetría más general es que exista una transformación específica $\delta_s q_\mu(t)$ tal que

$$\delta_s L = \frac{d\varepsilon}{dt} \tag{6.1}$$

para alguna función $\varepsilon(q,t)$. Para simetrías continuas, que se pueden descomponer en una sucesión de transformaciones infinitesimales como ocurre con las traslaciones y las rotaciones,[1] $\delta_s L$ se escribe como

$$\delta_s L = \sum_{\mu=1}^{n} \left(\frac{\partial L}{\partial q_\mu}\, \delta_s q_\mu + \frac{\partial L}{\partial \dot{q}_\mu}\, \delta_s \dot{q}_\mu \right) = \sum_{\mu=1}^{n} \left(\frac{\partial L}{\partial q_\mu}\, \delta_s q_\mu + \frac{\partial L}{\partial \dot{q}_\mu}\, \frac{d}{dt}\delta_s q_\mu \right)$$

Entonces $\delta_s q_\mu(t)$ es una simetría si

$$\sum_{\mu=1}^{n} \left(\frac{\partial L}{\partial q_\mu}\, \delta_s q_\mu + \frac{\partial L}{\partial \dot{q}_\mu}\, \frac{d}{dt}\delta_s q_\mu \right) = \frac{d\varepsilon}{dt}$$

Veamos qué sucede cuando este resultado se aplica sobre las ecuaciones de movimiento, donde vale que $\partial L/\partial q_\mu = d(\partial L/\partial \dot{q}_\mu)/dt$. Reemplazando se obtiene que

$$\frac{d}{dt}\left(\sum_{\mu=1}^{n} \frac{\partial L}{\partial \dot{q}_\mu}\, \delta_s q_\mu - \varepsilon \right) = 0$$

Es decir que a lo largo de la evolución real del sistema físico se conserva la cantidad

$$Q(q,\dot{q},t) \equiv \sum_{\mu=1}^{n} p_\mu\, \delta_s q_\mu - \varepsilon$$

1. A diferencia de las transformaciones discretas como, por ejemplo, $q \to -q$.

Veremos dos ejemplos importantes que involucran un sistema aislado de partículas que interactúan a través de potenciales que dependen de las distancias entre las mismas. En ese caso el Lagrangiano del sistema será invariante ante traslaciones y rotaciones:

1) Traslación. En una traslación los vectores posición $\vec{r}_i$ sufren un mismo desplazamiento $\vec{\xi}$ fijo (independiente del tiempo): $\vec{r}_i \to \vec{r}_i + \vec{\xi}$. Por lo tanto las velocidades $\vec{v}_i$ no cambian:

$$\delta_s \vec{r}_i = \vec{\xi}\ , \qquad \Rightarrow \qquad \delta_s \vec{v}_i = 0$$

Para su uso en las ecuaciones previas, $\vec{\xi}$ debe verse como un desplazamiento infinitesimal. La traslación es una simetría del Lagrangiano del sistema aislado porque no afecta la energía cinética (las velocidades no cambian) ni tampoco la energía potencial (las distancias entre partículas no cambian).

2) Rotación. En una rotación de ángulo fijo todas las posiciones y velocidades cambian de dirección en la misma forma. Pero esto no afecta ni los módulos de las velocidades, que entran en la energía cinética, ni las distancias entre partículas. Por lo tanto la rotación es una simetría del Lagrangiano de un sistema aislado. Una rotación de ángulo infinitesimal $\delta\alpha$ fijo (independiente de t) se escribe

$$\delta_s \vec{r}_i = \delta\vec{\alpha} \times \vec{r}_i\ , \qquad \Rightarrow \qquad \delta_s \vec{v}_i = \delta\vec{\alpha} \times \vec{v}_i$$

donde $\delta\vec{\alpha}$ tiene la dirección del eje de rotación (véase la rotación infinitesimal de un versor en §1.1).

Está claro que un campo externo rompería total o parcialmente estas simetrías, porque establecería direcciones privilegiadas en el espacio, y los distintos puntos del espacio se distinguirían por el valor del potencial externo. Pero las simetrías ante traslaciones y rotaciones que caracterizan los sistemas aislados no son tan sólo una consecuencia de la ausencia de campos externos, sino de la **homogeneidad e isotropía** atribuidas al espacio físico en la Mecánica Clásica (aceptamos que el espacio físico cumple los axiomas de la geometría plana de Euclides). Por lo tanto, las leyes de conservación que vamos a obtener pueden verse como una consecuencia de la homogeneidad e isotropía del espacio. Todo sistema físico aislado en un espacio isótropo y homogéneo poseerá seis primeras integrales de movimiento provenientes de las simetrías del espacio.[2]

Veamos cuáles son las respectivas magnitudes conservadas. Destaquemos que estas transformaciones de simetría son invariancias del Lagrangiano ($\delta_s L = 0$). Por lo tanto es $\varepsilon = 0$ en ambos casos, y las magnitudes conservadas tienen la forma $Q = \sum_{\mu=1}^{n} p_\mu\ \delta_s q_\mu$. Si usamos como coordenadas generalizadas las componentes cartesianas de las posiciones de las partículas tendremos

$$Q = \sum_{i=1}^{N} m_i \vec{v}_i \cdot\ \delta_s \vec{r}_i$$

2. Un sistema de cargas eléctricas no puede considerarse *aislado* porque emite radiación electromagnética que transporta energía y cantidad de movimiento.

1) Traslación.

$$Q = \sum_{i=1}^{N} m_i \vec{v}_i \cdot \ \vec{\xi} = \vec{P} \cdot \vec{\xi}$$

Como $\vec{\xi}$ es arbitrario y constante en el tiempo, concluimos que las magnitudes conservadas son las componentes de la cantidad de movimiento total; es decir, se conserva el vector $\vec{P}$.

2) Rotación.

$$Q = \sum_{i=1}^{N} m_i \vec{v}_i \cdot \ (\delta\vec{\alpha} \times \vec{r}_i) = \delta\vec{\alpha} \cdot \sum_{i=1}^{N} (\vec{r}_i \times m_i \vec{v}_i) = \delta\vec{\alpha} \cdot \vec{L}_o$$

donde usamos la propiedad cíclica del producto mixto. Como $\delta\vec{\alpha}$ es arbitrario y constante en el tiempo, concluimos que las magnitudes conservadas son las tres componentes del momento angular total; es decir, se conserva el vector $\vec{L}_o$ (en este caso O es el origen de coordenadas, pero podría ser cualquier punto fijo del sistema de referencia).

Así, las conservaciones de la cantidad de movimiento total y el momento angular total de sistemas aislados pueden verse como una consecuencia de la homogeneidad e isotropía del espacio.

6.4 Conservación asociada a la homogeneidad del tiempo.

Los sistemas cuyos Lagrangianos no dependen explícitamente del tiempo dan lugar a un teorema de conservación. Calculemos la derivada temporal de L:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \sum_{\mu=1}^{n} \left(\frac{\partial L}{\partial q_\mu} \ \dot{q}_\mu + \frac{\partial L}{\partial \dot{q}_\mu} \ \ddot{q}_\mu \right)$$

Sobre la evolución real del sistema vale que $\partial L/\partial q_\mu = d(\partial L/\partial \dot{q}_\mu)/dt$. Entonces

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \sum_{\mu=1}^{n} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\mu} \ \dot{q}_\mu \right)$$

es decir,

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}$$

donde hemos definido

$$H \equiv \sum_{\mu=1}^{n} p_\mu \ \dot{q}_\mu - L$$

Si L no depende explícitamente del tiempo es $\partial L/\partial t = 0$, y se obtiene que H se conserva.

La simetría involucrada en la conservación de H es la traslación temporal. Es una simetría propia de cualquier sistema donde las interacciones (internas y externas) no dependan del tiempo. En ese caso L no dependerá explícitamente de t, y tendremos que si $q_\mu(t)$ resuelve las ecuaciones de Euler-Lagrange entonces $q_\mu(t+\epsilon)$ también será una solución (ϵ es una constante). En otras palabras, si las interacciones no dependen del tiempo entonces el resultado de un experimento realizado hoy no cambiaría si el mismo experimento se realizara dentro de cientos de millones de años. Pero esto es también consecuencia de la **homogeneidad** que atribuimos al **tiempo** de la Mecánica Clásica. La naturaleza atribuida al espacio y el tiempo en la Mecánica Clásica son decisivos en los teoremas de conservación de $\vec{P}$, $\vec{L}$ y H.

6.4.1 ¿Cuándo coincide H con la energía mecánica E?

En varios casos de interés resulta $H = E$. Sea un sistema con vínculos holónomos y esclerónomos. Entonces

$$\vec{r}_i = \vec{r}_i(q_\mu) \qquad \Rightarrow \qquad \vec{v}_i = \frac{d\vec{r}_i}{dt} = \sum_{\mu=1}^{n} \frac{\partial \vec{r}_i}{\partial q_\mu}\, \dot{q}_\mu$$

(pues $\partial \vec{r}_i/\partial t = 0$). Entonces $v_i^2 = \vec{v}_i \cdot \vec{v}_i$ será una función cuadrática y homogénea de las velocidades generalizadas, y otro tanto sucederá con la energía cinética T.[3] En ese caso vale que

$$\sum_{\mu=1}^{n} \frac{\partial T}{\partial \dot{q}_\mu}\, \dot{q}_\mu = 2T$$

Si además V no depende de las velocidades generalizadas, entonces $p_\mu = \partial L/\partial \dot{q}_\mu = \partial T/\partial \dot{q}_\mu$. Entonces H resulta ser

$$H = \sum_{\mu=1}^{n} p_\mu\, \dot{q}_\mu - L = \sum_{\mu=1}^{n} \frac{\partial T}{\partial \dot{q}_\mu}\, \dot{q}_\mu - L = 2T - L = T + V = E$$

Ejemplo

Veamos un ejemplo donde H no coincide con E. En el problema del elevador que asciende con velocidad constante V, el vínculo es $z = Vt$ (ver §3.2.3). El Lagrangiano es

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + V^2) - mgVt$$

3. $T = \sum_\nu \sum_\mu m_{\mu\nu}(q)\, \dot{q}_\mu\, \dot{q}_\nu$ donde $m_{\mu\nu}(q) = \sum_{i=1}^{N} m_i\, \frac{\partial \vec{r}_i}{\partial q_\mu} \cdot \frac{\partial \vec{r}_i}{\partial q_\nu}$.

Nótese que la energía cinética no es cuadrática y homogénea en las velocidades generalizadas $\dot{x}$, $\dot{y}$ debido a que el vínculo es reónomo. H es igual a

$$\begin{aligned} H &= \sum_{\mu=1}^{n} p_\mu \, \dot{q}_\mu - L = m\dot{x} \;\; \dot{x} + m\dot{y} \;\; \dot{y} - \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + V^2) + mgVt \\ &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2 - V^2) + mgVt \neq E \end{aligned}$$

Como L depende explícitamente de t, entonces H no se conserva. En este caso tampoco se conserva E, porque la fuerza de vínculo realiza trabajo.

6.5 Segundo teorema de Noether. Simetrías de gauge

Cuando la simetría del Lagrangiano no es "rígida" sino que contiene funciones arbitrarias del tiempo, resultarán otras propiedades. Por ejemplo, consideremos un sistema descripto por coordenadas q_1, q_2 cuyo Lagrangiano es

$$L = \frac{1}{2} \, (\dot{q}_1 - q_2)^2$$

que posee la simetría

$$\delta_s q_1 = \epsilon(t) \; , \qquad \delta_s q_2 = \dot{\epsilon}(t)$$

donde ahora $\epsilon(t)$ no es una constante sino que es una *función* arbitraria. Esto es un aspecto típico de las **simetrías de gauge**.

Las ecuaciones de Lagrange de este sistema son

$$\begin{aligned} \frac{d}{dt}(\dot{q}_1 - q_2) &= 0 \\ (\dot{q}_1 - q_2) &= 0 \end{aligned}$$

Estas ecuaciones de movimiento no son independientes. Por lo tanto no alcanzan para determinar la evolución de las dos variables dinámicas; la solución general contiene una función arbitraria:

$$q_1(t) = f(t) \; , \qquad q_2(t) = \dot{f}(t)$$

que puede ser fijada mediante una "**elección de gauge**".[4]

Otro aspecto típico de este tipo de sistema es que las coordenadas y momentos no pueden ser elegidos libremente para caracterizar el estado inicial del sistema. En nuestro

4. Como la dinámica de q_1 está contenida en la ecuación para q_2 (la ecuación dinámica para q_1 no agrega nueva información), lo correcto es fijar el gauge eligiendo q_1. Esta fijación de gauge puede hacerse a nivel del Lagrangiano.

ejemplo resulta que p_2 es idénticamente nulo. Esto es un vínculo sobre los momentos, que evidencia otra vez que los dos grados de libertad q_1 y q_2 están ligados entre sí. Además, las ecuaciones del ejemplo dicen que $p_1 = \dot{q}_1 - q_2$ no puede valer otra cosa que cero. En síntesis, las simetrías de gauge conducen a relaciones entre las ecuaciones de movimiento (también puede haberlas con las derivadas de las ecuaciones de movimiento), que implican **libertades de gauge** para algunos grados de libertad. El segundo teorema de Noether se ocupa de las teorías de campo que poseen este tipo de comportamiento.[5]

Bibliografía adicional

E. Noether, *Invariante Variationsprobleme*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse, 235–257 (1918).
Traducido al inglés en arxiv.org/pdf/physics/0503066v1.pdf

M. Bañados e I. Reyes, *A short review on Noether's theorems, gauge symmetries and boundary terms*, IJMP D 25 (10) 1630021 (2016). arxiv.org/pdf/1601.03616.pdf

5. Las ecuaciones de Maxwell para los cuatro potenciales ϕ, $\vec{A}$ pueden obtenerse de una acción $S[\phi, \vec{A}]$ que posee simetría de gauge. Así resultan cuatro ecuaciones que dicen que los campos $\vec{C}_1 \doteq \vec{\nabla} \cdot \vec{E}$ y $\vec{C}_2 \doteq \vec{\nabla} \times \vec{B} - \mu_o \epsilon_o \, \partial \vec{E}/\partial t$ se anulan en ausencia de fuentes. Estos campos están ligados por la identidad $\mu_o \epsilon_o \, \partial \vec{C}_1/\partial t + \vec{\nabla} \cdot \vec{C}_2 \equiv 0$. En consecuencia las ecuaciones dinámicas no determinan completamente los potenciales (los potenciales poseen "libertades de gauge"). Cuando existen fuentes, esta identidad impone la ecuación de continuidad para las fuentes.

7 Sistema de dos cuerpos. Fuerzas centrales. Problema de Kepler

7.1 Sistema de dos cuerpos

Consideremos un sistema aislado de dos cuerpos que interactúan a través de un potencial dependiente de la distancia entre los mismos,

$$L = \frac{1}{2}\, m_1\, v_1^2 + \frac{1}{2}\, m_2\, v_2^2 \; - V(|\vec{r}|)$$

donde

$$\vec{r} \equiv \vec{r}_1 - \vec{r}_2$$

Como el sistema está aislado podemos escoger un sistema de referencia inercial con origen en el CM. En ese sistema de referencia vale que

$$m_1\, \vec{r}_1 + m_2\, \vec{r}_2 = 0$$

Podemos usar las dos últimas ecuaciones para resolver $\vec{r}_1$ y $\vec{r}_2$ en función de $\vec{r}$:

$$\vec{r}_1 = \frac{m_2\, \vec{r}}{m_1 + m_2}\ , \qquad \vec{r}_2 = -\frac{m_1\, \vec{r}}{m_1 + m_2} \tag{7.1}$$

Por lo tanto

$$\vec{v}_1 = \frac{m_2\, \dot{\vec{r}}}{m_1 + m_2}\ , \qquad \vec{v}_2 = -\frac{m_1\, \dot{\vec{r}}}{m_1 + m_2}$$

Reemplazando en el Lagrangiano,

$$L = \frac{1}{2}\, m_1\, \frac{m_2^2\, |\dot{\vec{r}}|^2}{(m_1 + m_2)^2} + \frac{1}{2}\, m_2\, \frac{m_1^2\, |\dot{\vec{r}}|^2}{(m_1 + m_2)^2} \; - V(|\vec{r}|) = \frac{1}{2}\, \mu\, |\dot{\vec{r}}|^2 \; - V(|\vec{r}|)$$

donde

$$\mu \equiv \frac{m_1\, m_2}{m_1 + m_2}$$

es la **masa reducida** del sistema. Esto muestra que la dinámica del sistema de dos partículas se obtiene del estudio de una única partícula de masa μ en presencia de un **campo central** $V(r)$ ($r = |\vec{r}|$; la fuerza $\vec{F} = -\vec{\nabla} V = -dV/dr\ \hat{r}$ es radial respecto del

origen de coordenadas). Resolviendo la evolución $\vec{r}(t)$ de la partícula μ reconstruiremos las evoluciones de m_1 y m_2 mediante las ecuaciones (7.1).

En un campo central no hay torque respecto del origen de coordenadas O. Por lo tanto se conserva el momento angular respecto de O. En particular, la conservación de la dirección de $\vec{L}_O$ implica que el movimiento se realiza en el plano perpendicular a $\vec{L}_O$. Describiremos la posición de la partícula μ mediante coordenadas polares en ese plano, de modo que

$$\dot{\vec{r}} = \frac{d}{dt}\,(r\ \hat{r}) = \dot{r}\ \hat{r} + r\ \dot{\varphi}\ \hat{\varphi}$$

y el Lagrangiano queda

$$L = \frac{1}{2}\ \mu\ (\dot{r}^2 + r^2\ \dot{\varphi}^2\ )\ \ - V(r)$$

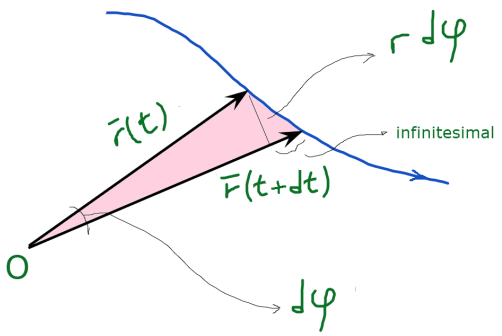
Como la coordenada φ es cíclica, entonces se conserva

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = \mu\ r^2\ \dot{\varphi}$$

que es el módulo del momento angular.

7.2 Segunda Ley de Kepler

En la Figura podemos ver que el área infinitesimal barrida por el vector posición es



$$dA = \tfrac{1}{2}\ base \times altura = \tfrac{r^2\ d\varphi}{2}$$

La "velocidad areolar" es

$$\frac{dA}{dt} = \frac{r^2}{2}\ \dot{\varphi} = \frac{p_\varphi}{2\ \mu} = const.$$

Es decir que la conservación de p_φ implica que la velocidad areolar es constante, que es la Segunda Ley de Kepler (la partícula barre áreas iguales en tiempos iguales).

7.3 Integración

La constante de movimiento p_φ será llamada ℓ en alusión al momento angular:

$$\ell = \mu\ r^2\ \dot{\varphi} \tag{7.2}$$

Además se conserva la energía mecánica

$$E = \frac{1}{2}\ \mu\ (\dot{r}^2 + r^2\ \dot{\varphi}^2\ )\ + V(r)$$

Podemos utilizar la conservación del momento angular para reemplazar $\dot{\varphi}$ en la expresión de E: [1]

$$E = \frac{1}{2}\ \mu\ \dot{r}^2 + \frac{\ell^2}{2\mu r^2}\ + V(r) \tag{7.3}$$

Así obtenemos una relación entre $\dot{r}$ y r que puede ser integrada,

$$\int_{t_o}^{t(r)} dt = \pm \int_{r_o}^{r} \frac{dr}{\sqrt{\frac{2}{\mu}[E - V(r)] - \frac{\ell^2}{\mu^2 r^2}}}\ ,$$

para obtener la evolución radial $r(t)$ (el signo negativo se usa para integrar en los tramos donde $dr/dt < 0$). Conocida la función $r(t)$ se reemplazaría en (7.2) para obtener $\varphi(t)$ mediante una integración. De esa forma, el problema del movimiento de la partícula μ estaría resuelto.

Si en lugar de la evolución temporal nos interesa la trayectoria orbital ($\ell \neq 0$), es decir la función $\varphi(r)$, podemos reemplazar en (7.3)

$$\dot{r} = \frac{dr}{d\varphi}\ \dot{\varphi} = \frac{dr}{d\varphi}\ \frac{\ell}{\mu r^2}\ ,$$

donde usamos (7.2). Así la integración resulta[2]

$$\int_{\varphi_o}^{\varphi(r)} d\varphi = \pm \int_{r_o}^{r} \frac{dr}{\sqrt{2\mu\ell^{-2}\ [E - V(r)]\ r^4 - r^2}} \tag{7.4}$$

1. No sería correcto reemplazar $\dot{\varphi}$ en el Lagrangiano, y derivarlo luego para obtener la dinámica radial, pues reemplazar una relación entre coordenadas y velocidades altera la forma funcional del Lagrangiano y, con ello, las ecuaciones dinámicas. Las constantes de movimiento resultan de una primera integración de las ecuaciones de movimiento; una vez obtenidas, podemos combinarlas para terminar de integrar el problema.

2. Las *integrales elípticas* tienen la forma $\int R(x, P(x)^{1/2})\ dx$ donde R es una función racional de sus argumentos, y $P(x)$ es un polinomio en x de grado 3 o 4. La integral en (7.4) es elíptica si $V(r) = r^n$ con $-4 \leq n \leq 2$.

7.4 Potencial efectivo

En la expresión (7.3) para E vemos que el término asociado a ℓ, que proviene del movimiento orbital, podría pensarse como una contribución al potencial. Podríamos decir que la energía E exhibe un potencial efectivo

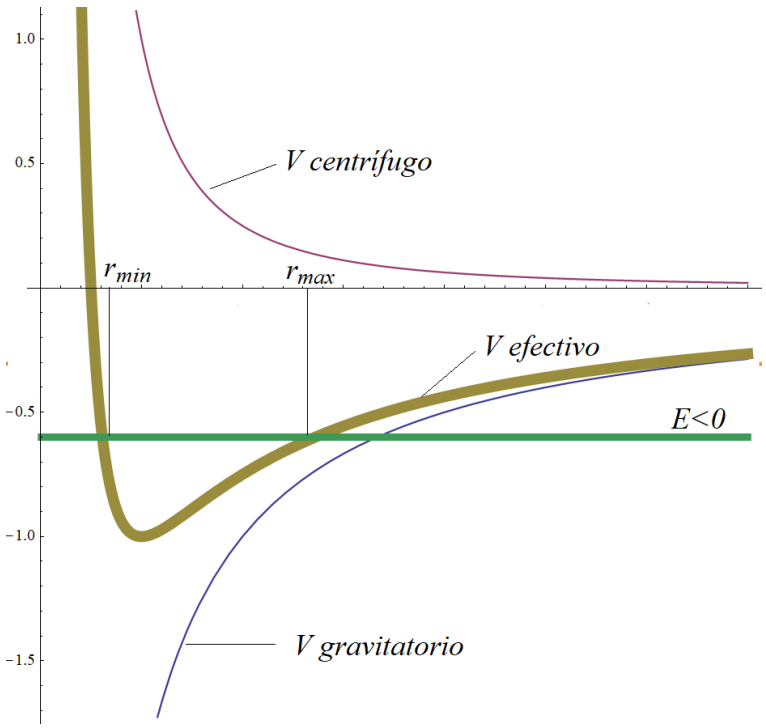
$$V_{efec} = V(r) + \frac{\ell^2}{2\mu r^2}$$

El término orbital se llama **potencial centrífugo**, pues de él se obtiene la fuerza centrífuga,

$$-\vec{\nabla}\left(\frac{\ell^2}{2\mu r^2}\right) = \frac{\ell^2}{\mu\ r^3}\ \hat{r} = \frac{\mu^2\ r^4\ \dot{\varphi}^2}{\mu\ r^3}\ \hat{r} = \mu\ r\ \dot{\varphi}^2\ \hat{r}$$

La energía mecánica escrita como $E = \mu\ \dot{r}^2/2 + V_{efec}(r)$ tiene la forma de la energía de un movimiento unidimensional en un potencial $V_{efec}(r)$. Así, el movimiento radial de la partícula μ podría entenderse en términos de este potencial.

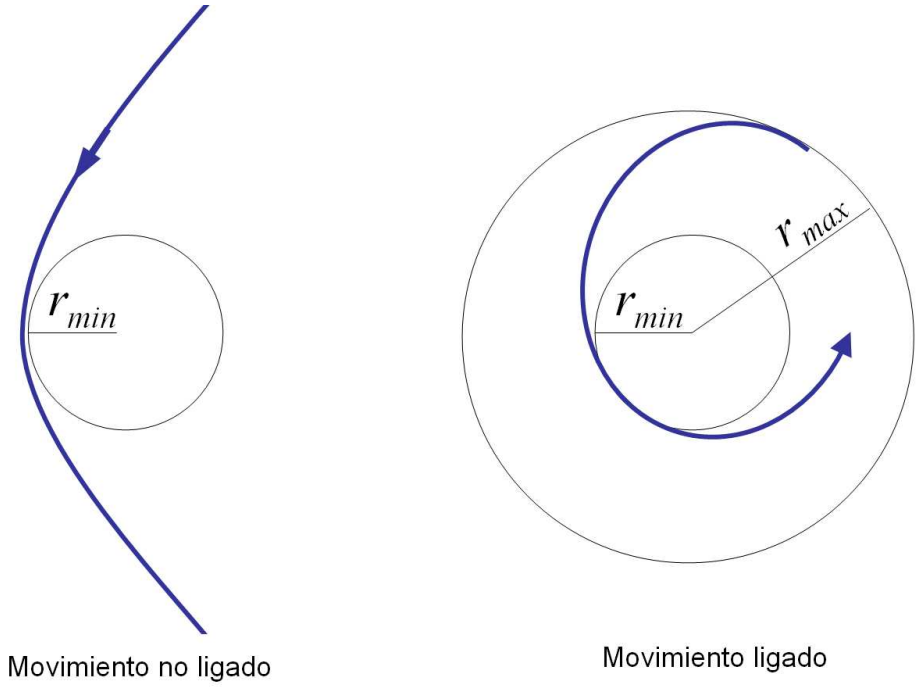
Por ejemplo, si la partícula orbita ($\ell \neq 0$) en un campo gravitatorio el potencial efectivo será como el que muestra la Figura:



Si $E \geq 0$ el radio evolucionará entre un $r_{\min}$ y $r = \infty$ (movimiento no ligado). En cambio, si $E < 0$ el movimiento radial quedará confinado entre un radio mínimo $r_{\min}$ y un radio máximo $r_{\max}$ (movimiento ligado). Los "puntos de retorno" del potencial efectivo no significan que el movimiento se detiene, pues la partícula continúa su movimiento

orbital para conservar el momento angular. Significan que el crecimiento o decrecimiento del radio se detiene ($\dot{r} = 0$); pero el ángulo sigue avanzando.[3]

En el plano orbital, los movimientos ligados y no ligados se ven como en la siguiente Figura:

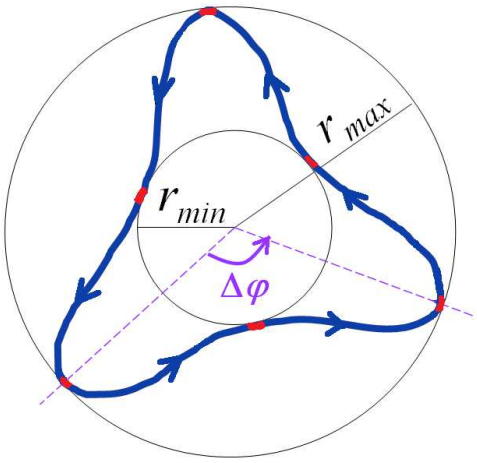


En el caso ligado la órbita se cerrará o no dependiendo del tipo de potencial. En efecto, el ángulo barrido mientras la partícula va desde $r_{\max}$ a $r_{\min}$ y vuelve a $r_{\max}$ es (ver (7.4))

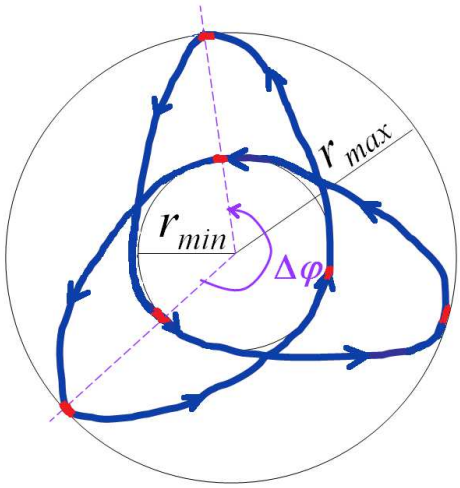
$$\Delta\varphi = -\int_{r_{\max}}^{r_{\min}} \frac{dr}{\sqrt{\frac{2\mu(E-V)}{\ell^2}\, r^4 - r^2}} + \int_{r_{\min}}^{r_{\max}} \frac{dr}{\sqrt{\frac{2\mu(E-V)}{\ell^2}\, r^4 - r^2}} = 2\int_{r_{\min}}^{r_{\max}} \frac{dr}{\sqrt{\frac{2\mu(E-V)}{\ell^2}\, r^4 - r^2}}$$

Si $\Delta\varphi$ es un múltiplo de 2π entonces la órbita se cerrará, pues el radio regresará a su valor máximo al cabo de un número entero de giros. Si $\Delta\varphi$ es un submúltiplo de 2π, $\Delta\varphi = 2\pi/k$, la órbita se cerrará al cabo de k oscilaciones del radio. En un caso general, si al cabo de k oscilaciones del radio se completan m giros, $\Delta\varphi\ k = 2\pi\ m$, entonces la órbita será cerrada (ver Figura). Que este sea o no el resultado de la integral (7.5) depende del potencial $V(r)$. Los puntos donde se anula $\dot{r}$ (marcados en rojo en la Figura)

3. En los puntos de retorno se anula el radicando en la ecuación (7.4). Esta anulación no implica un mal comportamiento de la integral; mediante un desarrollo en serie de Taylor se obtiene que el radicando es lineal en $r - r_{\min}$ o en $r - r_{\max}$ según sea el caso.



$3\ \Delta\varphi = 2\pi$



$3\ \Delta\varphi = 4\pi$

se llaman **puntos absidales**. Como vimos en la ecuación (7.5) el ángulo barrido para ir desde $r_{\min}$ hasta $r_{\max}$ es igual al que se barre para ir desde $r_{\max}$ hasta $r_{\min}$. Lo mismo sucede al ir desde $r_{\min}$ hasta un $r < r_{\max}$ o al revés. Esto significa que las órbitas (sean o no cerradas) son simétricas respecto de los puntos absidales. Si elegimos $\varphi = 0$ en un punto absidal, digamos en $r = r_{\min}$, entonces la ecuación (7.4) dice que $\varphi(r)$ para r entre $r_{\min}$ y el siguiente punto absidal es

$$\varphi(r) = \int_{r_{\min}}^{r} \frac{dr}{\sqrt{2\mu\ell^{-2}\ [E - V(r)]\ r^4 - r^2}} \tag{7.5}$$

Existen dos tipos de potenciales centrales cuyas órbitas ligadas son todas cerradas:

$$V(r) \propto \frac{1}{r}\ , \qquad\qquad V(r) \propto r^2$$

En el caso del potencial elástico isótropo $V(r) = (1/2)kr^2$ es fácil ver que es así. La fuerza es $\vec{F} = -\vec{\nabla}V = -k\ r\ \hat{r} = -k\ \vec{r}$. Las ecuaciones de movimiento para las componentes cartesianas de $\vec{r}(t) = x(t)\ \hat{\imath} + y(t)\ \hat{\jmath}$ son

$$-kx = m\ddot{x}\ , \qquad -ky = m\ddot{y}$$

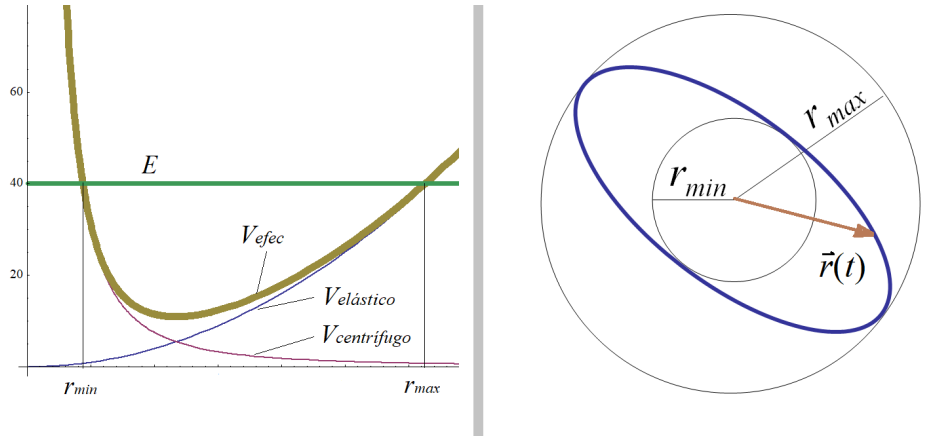
que muestran dos oscilaciones de la misma frecuencia $\omega_o = \sqrt{k/m}$. Por ejemplo, una solución particular es

$$x = A\ \cos\omega_o t\ , \qquad y = B\ \sin\omega_o t$$

que corresponde a una elipse cuyos ejes coinciden con los ejes cartesianos:[4]

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$$

Pero como el potencial es isótropo, dependiendo de las condiciones iniciales obtendremos también elipses con los ejes orientados en otras direcciones (ver Figura).



7.4.1 Órbita circular

Si el potencial efectivo tiene un mínimo, y el valor de E coincide con el valor mínimo de V_{efec} , entonces $r_{\min} = r_{\max} \equiv r_{circ}$. La órbita será circular de radio r_{circ} tal que

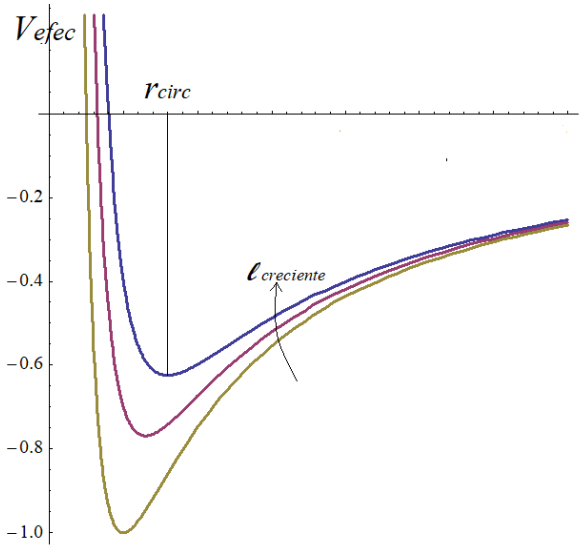
$$\frac{dV_{efec}}{dr}|_{r=r_{circ}} = 0$$

Por cierto, en ese caso tendremos una órbita cerrada independientemente de las demás características del potencial. Debemos recordar que el potencial efectivo depende de las condiciones iniciales a través de ℓ; es decir, para cada valor de ℓ hay un potencial efectivo. Para cada valor de ℓ tal que V_{efec} alcance un valor mínimo habrá una órbita circular.

El potencial gravitatorio tiene la forma $V = -\alpha/r$. El radio de la órbita circular es

$$r_{circ} = \frac{\ell^2}{\alpha\ \mu}$$

La Figura muestra potenciales efectivos del caso gravitatorio para varios valores de ℓ:



7.5 Problema de Kepler

Para resolver la forma de la órbita $\varphi(r)$ en el caso gravitatorio, debemos calcular la integral

$$\varphi(r) = \int \frac{dr}{\sqrt{2\mu\ell^{-2}\,[E - V(r)]\,r^4 - r^2}}$$

con el potencial

$$V(r) = -\frac{\alpha}{r}$$

Como el potencial gravitatorio es atractivo, entonces es $\alpha > 0$. La solución servirá también para la interacción atractiva coulombiana, o para la repulsión coulombiana si cambiamos el signo de α.

Para calcular la integral es conveniente hacer un cambio de variable:

$$u \equiv \frac{1}{r} \qquad \Rightarrow \qquad du = -\frac{dr}{r^2}$$

Llevando r^4 fuera de la raíz, y reemplazando $V(u) = -\alpha u$ obtenemos

$$\varphi(u) = -\int \frac{du}{\sqrt{2\mu\ell^{-2}\,(E + \alpha\,u) - u^2}}$$

Podemos llamar

$$a \equiv 2\mu\ell^{-2}\,E\ , \qquad b \equiv 2\mu\ell^{-2}\,\alpha\ .$$

4. La elipse está centrada en el origen del vector $\vec{r}$, de modo que el ángulo barrido entre dos máximos de r sucesivos es $\Delta\varphi = \pi$.

Así tenemos

$$-\varphi(u) = \int \frac{du}{\sqrt{a + b\ u - u^2}} = \arccos\left(\frac{2\ u - b}{\sqrt{b^2 + 4\ a}}\right) + cte$$

Resolviendo para u:

$$\frac{1}{r} = u = \frac{b}{2} + \frac{1}{2}\sqrt{b^2 + 4\ a}\ \cos(\varphi + cte)$$

Recordemos que hemos elegido el signo de la raíz correspondiente a un tramo de r creciente. Para fijar $\varphi = 0$ cuando $r = r_{\min}$ (es decir, cuando u es máximo), debemos anular la *cte* de integración; entonces, restableciendo los valores de a y b:[5]

$$\frac{\ell^2}{\mu\ \alpha}\ \frac{1}{r} = 1 + \sqrt{1 + \frac{2\ \ell^2 E}{\mu\ \alpha^2}}\ \cos\varphi \tag{7.6}$$

Antes de averiguar qué tipo de trayectoria es esta, veamos que el mismo resultado se obtiene por otra vía.

7.6 Ecuación de Binet

Ya dijimos que la evolución radial de la masa reducida μ se podía entender a la luz del potencial efectivo. Entonces podemos escribir la ecuación[6]

$$\mu\ \ddot{r} = -\frac{dV_{efec}}{dr}$$

Atendiendo a la forma del potencial efectivo, y llamando $f(r) = -dV/dr$,

$$\mu\ \ddot{r} = \frac{\ell^2}{\mu\ r^3} + f(r) \tag{7.7}$$

Como nos interesa la órbita $r(\varphi)$, necesitamos una ecuación en la variable φ. Para ello nos valemos de la conservación de p_φ, que lleva a

$$\dot{\varphi} = \frac{\ell}{\mu\ r^2}$$

Entonces

$$\frac{d}{dt} = \dot{\varphi}\ \frac{d}{d\varphi} = \frac{\ell}{\mu\ r^2}\ \frac{d}{d\varphi}$$

5. b^2 sale fuera de la raíz como $|b|$. Si $\alpha < 0$ (potencial coulombiano repulsivo) es $b < 0$. Entonces resulta

$$\frac{\ell^2}{\mu\ \alpha}\ \frac{1}{r} = 1 - \sqrt{1 + \frac{2\ \ell^2 E}{\mu\ \alpha^2}}\ \cos\varphi$$

6. Por supuesto, la ecuación no es más que la derivada temporal de $E = (1/2)\ \mu\ \dot{r}^2 + V_{efec}(r)$.

Si además optamos por la variable $u \equiv r^{-1}$, entonces

$$\dot{r} = \frac{d}{dt}\,\frac{1}{u} = -\frac{1}{u^2}\,\frac{du}{dt} = -\frac{1}{u^2}\,\frac{\ell}{\mu\, r^2}\,\frac{du}{d\varphi} = -\frac{\ell}{\mu}\,\frac{du}{d\varphi}$$

Por lo tanto

$$\ddot{r} = \frac{\ell}{\mu\, r^2}\,\frac{d}{d\varphi}\,\dot{r} = -\frac{\ell^2 u^2}{\mu^2}\,\frac{d^2u}{d\varphi^2}$$

reemplazando este resultado en la ecuación (7.7) obtenemos la **ecuación de Binet**

$$\frac{d^2u}{d\varphi^2} + u + \frac{\mu}{\ell^2 u^2}\, f\left(\frac{1}{u}\right) = 0$$

Esta ecuación es válida para cualquier potencial central. Si elegimos el origen del ángulo φ en un punto absidal donde r es mínimo (u es máximo), tendremos las siguientes condiciones de contorno

$$u(\varphi = 0) = u_{\max}\ , \qquad \frac{du}{d\varphi}(\varphi = 0) = 0$$

Vemos que tanto la ecuación como las condiciones de contorno son invariantes ante el cambio $\varphi \to -\varphi$. Esto significa que las órbitas son simétricas respecto de los puntos absidales.

En el caso gravitatorio es

$$f = -\frac{\alpha}{r^2} = -\alpha\, u^2$$

y la ecuación de Binet queda

$$\frac{d^2u}{d\varphi^2} + u = \frac{\alpha\,\mu}{\ell^2}$$

cuya solución, con las condiciones de contorno dadas es

$$u = A\cos\varphi + \frac{\alpha\,\mu}{\ell^2}$$

donde la amplitud A de la oscilación radial queda determinada por el valor de la energía. Podemos reescribir la solución en la siguiente forma

$$\frac{\ell^2}{\mu\,\alpha}\,\frac{1}{r} = 1 + e\,\cos\varphi$$

donde e es una constante de integración. Este resultado es igual al obtenido en la ecuación (7.6), siendo [7]

$$e = \sqrt{1 + \frac{2\,\ell^2 E}{\alpha^2 \mu}}$$

7. Si $\alpha < 0$ (repulsión coulombiana), la constante de integración se ajusta al resultado en Nota 5.

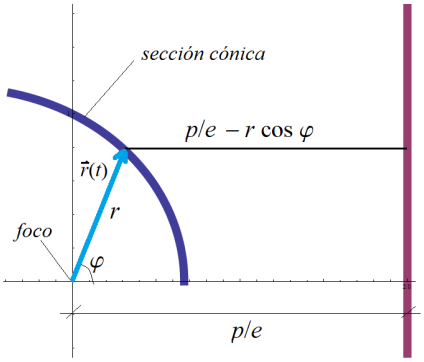
7.7 Secciones cónicas

La solución para la órbita barrida por el vector posición $\vec{r} = r\ \hat{r} = r\ (\cos\varphi\ \hat{\imath} + \sin\varphi\ \hat{\jmath})$ en el caso gravitatorio es una **sección cónica**. Las secciones cónicas son curvas planas que cumplen la ecuación

$$\frac{p}{r} = 1 + e\ \cos\varphi$$

que puede leerse así:

$$\frac{p}{e} - r\ \cos\varphi = \frac{r}{e}$$



Como muestra la Figura, $\frac{p}{e} - r\ \cos\varphi$ es la distancia de cada punto de la curva a una recta fija, mientras que r es la distancia de ese mismo punto a un punto F que llamaremos **foco**. Las secciones cónicas son curvas tales que el cociente entre esas dos distancias es constante a lo largo de la curva; su valor es e^{-1}. Según el valor de e, estas curvas son:

---

1) **circunferencia** $(e = 0)$: $r = p = \frac{\ell^2}{\alpha\ \mu}$, $\quad E = -\frac{\alpha^2 \mu}{2\ \ell^2}$

---

2) **elipse** $(0 < e < 1)$: $-\frac{\alpha^2 \mu}{2\ \ell^2} < E < 0$, e es la excentricidad de la elipse

$r_{\min} = r|_{\varphi=0} = \frac{p}{1+e}$, $r_{\max} = r|_{\varphi=\pi} = \frac{p}{1-e}$ $\quad (\frac{p}{2} < r_{\min} < p < r_{\max} < \infty)$

---

3) **parábola** $(e = 1)$: $\quad E = 0, \quad r_{\min} = r|_{\varphi=0} = \frac{p}{2}$, $r_{\max} = r|_{\varphi=\pi} = \infty$

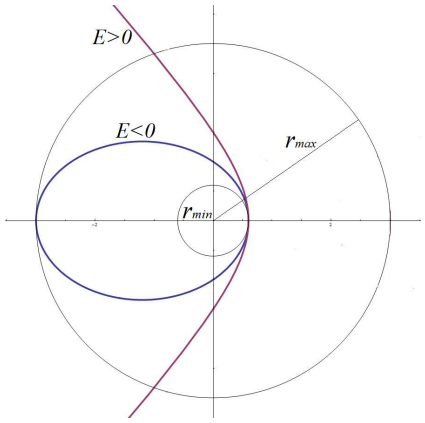
---

4) **hipérbola** $(e > 1)$: $\quad E > 0, \quad r_{\min} = r|_{\varphi=0} = \frac{p}{1+e} < \frac{p}{2}$,[8]

$r_{\max} = \infty$ se alcanza con $\varphi = \pm\varphi_\infty \equiv \pm \arccos[-1/e]$

8. La ecuación para el potencial coulombiano repulsivo, $p/r = 1 - e\ \cos\varphi$ $(e > 1,\ p < 0)$, también genera una trayectoria hiperbólica. La distancia mínima al foco es $r_{\min} = r|_{\varphi=0} = \frac{-p}{e-1}$.

La Figura muestra dos órbitas con el mismo valor de r_{min}; una de energía negativa y otra de energía positiva.[9]



Veamos el caso de la elipse. En la Figura el semieje mayor de la elipse a es igual a

$$a = \frac{1}{2}\,(r_{\max} + r_{\min}) = \frac{1}{2}\left(\frac{p}{1-e} + \frac{p}{1+e}\right) = \frac{p}{1-e^2} = \frac{\frac{\ell^2}{\mu\ \alpha}}{-\frac{2\ \ell^2 E}{\alpha^2 \mu}} = -\frac{\alpha}{2\ E} \tag{7.8}$$

y la distancia entre el centro y el foco es

$$d_{CF} = a - r_{\min} = \frac{1}{2}\,(r_{\max} - r_{\min}) = \frac{1}{2}\left(\frac{p}{1-e} - \frac{p}{1+e}\right) = \frac{e\ p}{1-e^2} = e\ a$$

Por cierto, la elipse resulta más familiar cuando está escrita en coordenadas cartesianas con origen en el centro de la elipse C. El cambio de coordenadas polares con origen en el foco F a cartesianas con origen en C es[10]

$$\begin{aligned} x &= r\ \cos\varphi + e\ a \\ y &= r\ \sin\varphi = r\ \sqrt{1-\cos^2\varphi} \end{aligned}$$

Calculemos la siguiente forma cuadrática:

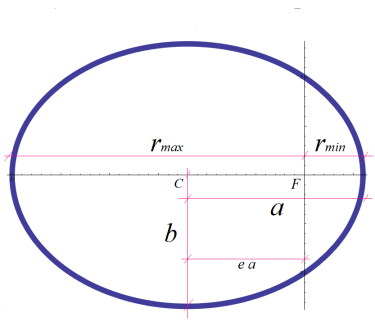
$$\begin{aligned} x^2 + \frac{y^2}{1-e^2} &= r^2\cos^2\varphi + 2\ e\ a\ r\ \cos\varphi + e^2a^2 + \frac{r^2(1-\cos^2\varphi)}{1-e^2} \\ &= -\frac{e^2r^2\cos^2\varphi}{1-e^2} + 2\ e\ a\ r\ \cos\varphi + e^2a^2 + \frac{r^2}{1-e^2} \end{aligned}$$

Como la ecuación de la elipse dice que $p - r = e\ r\ \cos\varphi$, entonces

$$x^2 + \frac{y^2}{1-e^2} = -\frac{(p-r)^2}{1-e^2} + 2\ a\ (p-r) + e^2a^2 + \frac{r^2}{1-e^2}$$

9. Para tener un mismo $r_{min} = p/(1+e)$ con valores distintos de E se deben usar distintos valores de ℓ (se compensa el cambio en e con un cambio en p).

10. El foco F es el origen del vector posición $\vec{r} = r\ \cos\varphi\ \hat{\imath} + r\ \sin\varphi\ \hat{\jmath}$. No deben confundirse las coordenadas x, y con origen en C con las componentes cartesianas del vector posición.



y como $1 - e^2 = p/a$, entonces obtenemos que

$$x^2 + \frac{y^2}{1-e^2} = -\frac{a}{p}\,(p-r)^2 + 2\,a\,(p-r) + \left(1 - \frac{p}{a}\right)\,a^2 + \frac{a\,r^2}{p} = a^2$$

Así recuperamos la elipse de semieje mayor a ; el semieje menor es

$$b = a\,\sqrt{1-e^2} = \frac{p}{\sqrt{1-e^2}} = \sqrt{a\,p} = \frac{\ell}{\sqrt{2|E|\mu}}$$

7.8 Primera Ley de Kepler

En la interacción gravitatoria entre una estrella de masa $m_\star$ y un planeta de masa $m_p << m_\star$, las posiciones de la estrella y el planeta son

$$\vec{r}_P = \frac{m_\star\,\vec{r}}{m_\star + m_P} \approx \vec{r}\,, \qquad\qquad \vec{r}_\star = -\frac{m_P\,\vec{r}}{m_\star + m_P} \approx 0$$

1ra Ley: Las órbitas de los planetas son elipses con el Sol en uno de sus focos.[11]

7.9 Tercera Ley de Kepler

Teniendo en cuenta que el área de una elipse es

$$Área = \pi\,a\,b\,,$$

y que la Segunda Ley de Kepler dice que la velocidad areolar es constante,

$$\frac{dÁrea}{dt} = \frac{1}{2}\,r^2\,\dot{\varphi} = \frac{\ell}{2\,\mu}\,,$$

11. Las órbitas del potencial elástico son elípticas también. En ese caso el centro de la elipse se ubica en el centro de fuerzas, mientras que en el caso gravitatorio el centro de fuerzas coincide con un foco de la elipse. En términos del ángulo $\Delta\varphi$ entre dos $r_{\max}$ consecutivos, es $\Delta\varphi = \pi$ para el potencial elástico, y $\Delta\varphi = 2\pi$ para el potencial gravitatorio.

entonces el área de la elipse se barre en un tiempo T (período de la órbita)

$$T = \frac{Área}{\frac{dÁrea}{dt}} = \frac{\pi\ a\ b}{\frac{\ell}{2\ \mu}} = \frac{\pi\ a^{3/2}\ p^{1/2}}{\frac{\ell}{2\ \mu}} = 2\ \pi\sqrt{\frac{\mu}{\alpha}}\ a^{3/2}$$

Como $\mu = m_1 m_2/(m_1 + m_2)$ y $\alpha = G\ m_1 m_2$ es

$$T = \frac{2\ \pi\ a^{3/2}}{\sqrt{G(m_1 + m_2)}}$$

Cuando se trata de la interacción entre una estrella y un planeta, $m_1 + m_2$ se aproxima por la masa de la estrella. Por lo tanto T^2/a^3 no depende más que de la masa de la estrella.

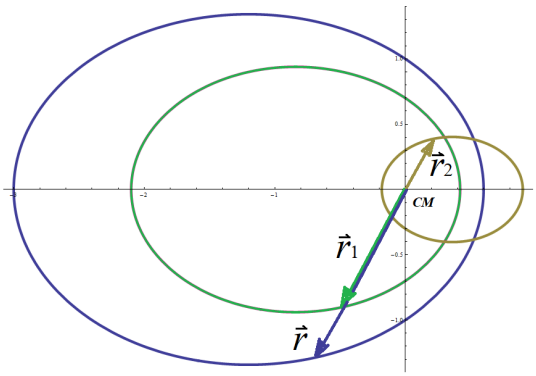
3ra Ley: Los cuadrados de los períodos de las órbitas de los planetas son proporcionales a los cubos de los respectivos semiejes mayores de las órbitas.

7.10 Regreso al problema de dos cuerpos

Habiendo resuelto la evolución $\vec{r}(t)$ de la masa reducida μ podemos recuperar los movimientos de cada una de las dos partículas:

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

Como el extremo de $\vec{r}(t)$ describe una cónica, lo mismo harán los vectores $\vec{r}_1(t)$ y $\vec{r}_2(t)$, cuyo origen está en el CM del sistema.



Material adicional

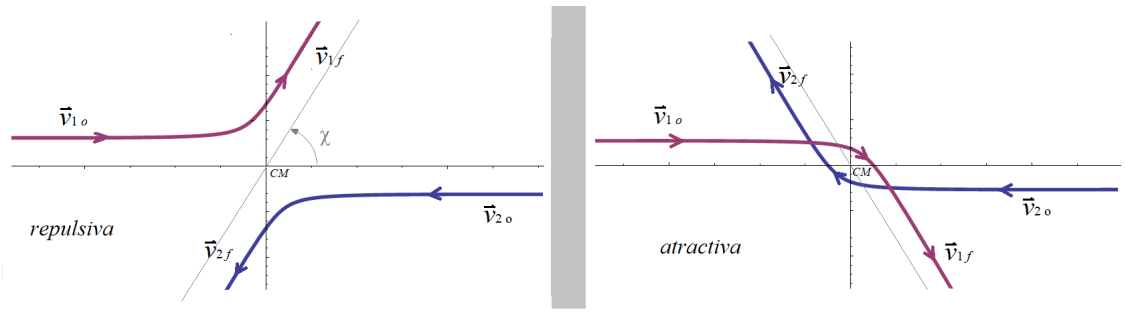
Puntos de Lagrange: https://solarsystem.nasa.gov/resources/754/what-is-a-lagrange-point/

Problema de 3 cuerpos: https://www.youtube.com/watch?v=et7XvBenEo8

8 Choque elástico. Scattering

8.1 Choque elástico

Si el potencial de interacción en un sistema de dos cuerpos tiende a cero en el infinito, entonces existen órbitas no ligadas donde las partículas están infinitamente separadas tanto en la configuración inicial como en la configuración final: $|\vec{r}| \longrightarrow \infty$ cuando $t \longrightarrow \pm\,\infty$. En esas configuraciones inicial y final la energía es puramente cinética, pues la energía potencial tiende a cero. Como la energía mecánica se conserva, entonces la energía cinética inicial es igual a la final. Esto es lo mismo que sucede en un choque elástico, aunque en el caso de un choque no se trate de interacciones a distancia. La igualdad entre energía cinética inicial y final junto con la conservación de la cantidad de movimiento total nos permitirán conocer algunos aspectos que tienen en común este tipo de interacciones. La Figura muestra las trayectorias hiperbólicas de dos partículas cargadas iguales en interacción coulombiana –tanto atractiva como repulsiva– tal como se ve en el sistema CM.



Nótese que las velocidades siempre tendrán direcciones opuestas, como condición necesaria para conservar $\vec{P} = 0$. El ángulo entre las direcciones inicial y final depende de los detalles de la interacción; no resulta de los teoremas de conservación.

La anulación de $\vec{P}$ significa que los valores iniciales y finales de las velocidades en ese sistema CM cumplen que

$$\vec{v}_{2o} = -\frac{m_1}{m_2}\,\vec{v}_{1o}\ , \qquad\qquad \vec{v}_{2f} = -\frac{m_1}{m_2}\,\vec{v}_{1f}$$

Además se conserva la energía cinética,

$$\frac{1}{2}\,m_1\,v_{1o}^2 + \frac{1}{2}\,m_2\,v_{2o}^2 = \frac{1}{2}\,m_1\,v_{1f}^2 + \frac{1}{2}\,m_2\,v_{2f}^2$$

Para valores iniciales dados, los valores finales cumplirán estas relaciones si

$$|\vec{v}_{1f}| \;=\; |\vec{v}_{1o}| \;, \qquad\qquad |\vec{v}_{2f}| \;=\; |\vec{v}_{2o}|$$

Como fue dicho, la dirección de las velocidades finales no queda determinada de esta forma, sino que depende de los detalles de la interacción. En cuanto a los módulos, las velocidades finales tienen los mismos módulos que las respectivas velocidades iniciales.

Una propiedad importante del choque elástico es que el módulo de la velocidad relativa antes y después del choque es el mismo. En el sistema CM el resultado se alcanza en forma muy simple. Debido a que $\vec{P} = 0$ siempre, entonces

$$\vec{v}_{rel} = \vec{v}_1 - \vec{v}_2 = \vec{v}_1 + \frac{m_1}{m_2}\,\vec{v}_1 = \frac{m_1 + m_2}{m_2}\,\vec{v}_1 \tag{8.1}$$

o también

$$\vec{v}_{rel} = \vec{v}_1 - \vec{v}_2 = -\frac{m_2}{m_1}\,\vec{v}_2 - \vec{v}_2 = -\frac{m_1 + m_2}{m_1}\,\vec{v}_2 \tag{8.2}$$

Como los módulos de $\vec{v}_1$ y $\vec{v}_2$ no cambian por efecto del choque, lo mismo sucede con $|\vec{v}_{rel}|$:

$$|\vec{v}_{rel\;o}| = |\vec{v}_{rel\;f}|$$

A diferencia de las velocidades individuales, que están sujetas al teorema de adición de velocidades, las velocidades relativas son **invariantes galileanos** (tienen el mismo valor en todos los sistemas de referencia en traslación relativa). Por lo tanto, $|\vec{v}_{rel\;o}| = |\vec{v}_{rel\;f}|$ es un resultado válido en cualquier sistema de referencia inercial.

Usando las ecuaciones (8.1) y (8.2) las velocidades finales en el sistema CM se pueden escribir así:

$$\vec{v}_{1f} = \frac{m_2}{m_1 + m_2}\,v_{rel}\;\hat{n}\;, \qquad\qquad \vec{v}_{2f} = -\frac{m_1}{m_1 + m_2}\,v_{rel}\;\hat{n}$$

donde $\hat{n}$ es el versor que da la dirección final aún desconocida (es la dirección de las asíntotas de salida en la Figura anterior). En términos de las cantidades de movimiento, este resultado dice que $\vec{p}_{1f} = \mu\;v_{rel}\;\hat{n}$, $\vec{p}_{2f} = -\mu\;v_{rel}\;\hat{n}$.

Ahora transformaremos las velocidades finales a un sistema de referencia S arbitrario. Si $\vec{V}_{CM}$ es la velocidad del CM en el sistema de referencia S, entonces habrá que sumar $\vec{V}_{CM}$ al resultado anterior (adición galileana de velocidades). Si ahora $\vec{v}_{1o}$, $\vec{v}_{2o}$ son las velocidades iniciales medidas en S, entonces

$$\vec{V}_{CM} = \frac{m_1\;\vec{v}_{1o} + m_2\;\vec{v}_{2o}}{m_1 + m_2}$$

Entonces las velocidades finales en S son

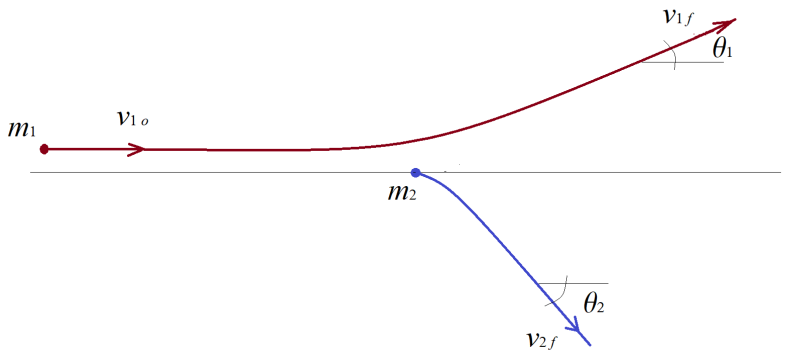
$$\begin{aligned}
\vec{v}_{1f} &= \frac{m_2}{m_1 + m_2}\,v_{rel}\;\hat{n} + \frac{m_1\;\vec{v}_{1o} + m_2\;\vec{v}_{2o}}{m_1 + m_2} \\
\vec{v}_{2f} &= -\frac{m_1}{m_1 + m_2}\,v_{rel}\;\hat{n} + \frac{m_1\;\vec{v}_{1o} + m_2\;\vec{v}_{2o}}{m_1 + m_2}
\end{aligned}$$

Recordemos que el valor de v_{rel} es el mismo en todos los sistemas de referencia inerciales. Agreguemos que $\hat{n}$ es la dirección de $\vec{v}_{rel\ f}$ y tampoco depende del sistema de referencia inercial que se utilice. También es $\hat{n}$ la dirección final de $\vec{r}_1 - \vec{r}_2$ (como es fácil ver en el sistema CM), que también es invariante galileano. Introduciremos el ángulo χ para indicar la dirección de $\hat{n}$:

$$\hat{n} = \cos\chi\ \hat{\imath} + \sin\chi\ \hat{\jmath}$$

8.1.1 El caso $v_{2o} = 0$

En muchos experimentos, una partícula juega el papel de "proyectil" y la otra de "blanco". En ese caso es $\vec{v}_{2o} = 0$ en el sistema del laboratorio, y $\vec{v}_{rel\ o} = \vec{v}_{1o}$.



Los ángulos θ_1 y θ_2 se obtienen de los resultados para las velocidades finales en el sistema del laboratorio:

$$\tan\theta_1 = \frac{v_{1f\ y}}{v_{1f\ x}} = \frac{m_2\ v_{1o}\ \sin\chi}{m_2\ v_{1o}\cos\chi + m_1\ v_{1o}} = \frac{\sin\chi}{\cos\chi + \frac{m_1}{m_2}}$$

$$\tan\theta_2 = -\frac{v_{2f\ y}}{v_{2f\ x}} = \frac{m_1\ v_{1o}\ \sin\chi}{-m_1\ v_{1o}\cos\chi + m_1\ v_{1o}} = \frac{\sin\chi}{1-\cos\chi}$$

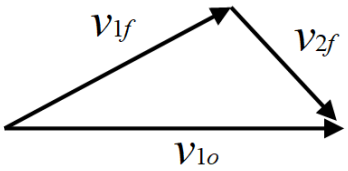
Cuando $m_1 = m_2$ entonces $\vec{v}_{2f}$ es perpendicular a $\vec{v}_{1f}$. En efecto, podemos valernos de una identidad trigonométrica para obtener

$$\tan(\theta_1+\theta_2) = \frac{\tan\theta_1 + \tan\theta_2}{1-\tan\theta_1\tan\theta_2} = \frac{1+\frac{m_1}{m_2}}{-1+\frac{m_1}{m_2}}\ \cot\left(\frac{\chi}{2}\right)$$

1) Si $m_1 = m_2$ entonces $\tan(\theta_1+\theta_2)$ diverge. Luego se tiene que $\theta_1+\theta_2 = \pi/2$.
Podemos entender este resultado así: como $m_1 = m_2$, la conservación de $\vec{P}$ queda

$$\vec{v}_{1o} = \vec{v}_{1f} + \vec{v}_{2f}$$

que puede verse como el siguiente triángulo,



mientras que la conservación de T,

$$v_{1o}^2 = v_{1f}^2 + v_{2f}^2 \ ,$$

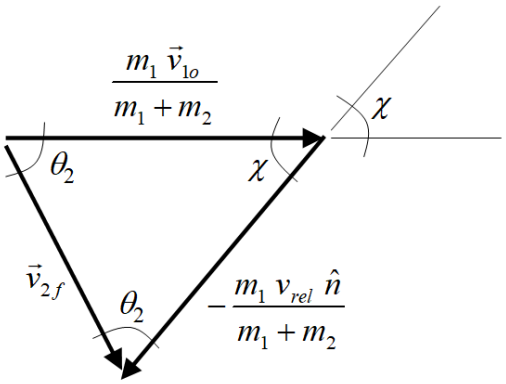
es el teorema de Pitágoras en el triángulo formado por los tres vectores. Por lo tanto $\vec{v}_{1f} \perp \vec{v}_{2f}$.

2) Si $m_1 \gtrless m_2$ entonces $\tan(\theta_1 + \theta_2) \gtrless 0$. Luego se tiene que $\theta_1 + \theta_2 \lesseqgtr \pi/2$.

3) El resultado para $\tan\theta_2$ puede entenderse mejor a partir de la expresión para $\vec{v}_{2f}$; cuando $\vec{v}_{2o} = 0$ queda así

$$\vec{v}_{2f} = \frac{m_1}{m_1 + m_2} \, (\vec{v}_{1o} - v_{rel} \, \hat{n})$$

donde se ve que $\vec{v}_{2f}$ resulta de restar dos vectores del mismo módulo ($v_{rel} = |\vec{v}_{1o}|$ porque $\vec{v}_{2o} = 0$). Entonces se forma el triángulo isósceles que muestra la Figura,



donde vemos que $\quad 2\,\theta_2 + \chi = \pi \quad \Rightarrow \quad \theta_2 = \dfrac{\pi - \chi}{2}$,

que resulta consistente con el resultado previo, pues $\tan\theta_2 = \sin\chi/(1 - \cos\chi) = \cot(\chi/2) = \tan((\pi - \chi)/2)$.

4) Para una dada relación m_1/m_2, el ángulo θ_1 varía con el valor de χ, y tiene un máximo si

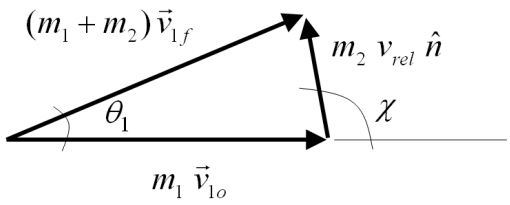
$$0 = \frac{d}{d\chi} \tan\theta_1 = \frac{1 + \frac{m_1}{m_2} \cos\chi}{(\cos\chi + \frac{m_1}{m_2})^2}$$

Entonces θ_1 es máximo cuando $\cos\chi = -m_2/m_1$. Vemos que debe ser $m_1 > m_2$; en caso contrario el proyectil puede salir en cualquier dirección. Reemplazando este valor de χ se obtiene que (ejercicio)

$$\sin\theta_{1\ \max} = \frac{m_2}{m_1} \tag{8.3}$$

Este resultado puede entenderse mejor a partir de la expresión para $\vec{v}_{1f}$; cuando $\vec{v}_{2o} = 0$ queda así

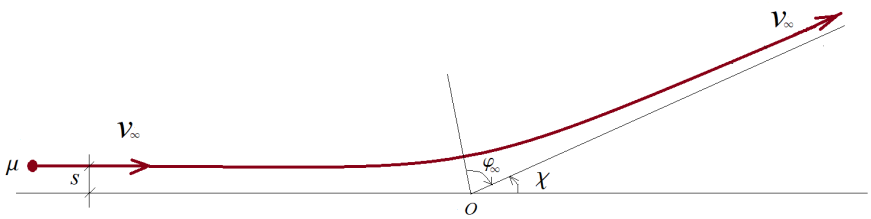
$$(m_1 + m_2)\ \vec{v}_{1f} = m_2\ v_{rel}\ \hat{n} + m_1\ \vec{v}_{1o}$$



Como $v_{rel} = |\vec{v}_{1o}|$ es fijo, lo único que podemos hacer para incrementar θ_1 es elegir χ para que se forme un triángulo rectángulo. Como la relación entre cateto opuesto e hipotenusa es m_2/m_1 llegamos al resultado (8.3). En cambio, si $m_2 > m_1$, el ángulo θ_1 podrá tomar cualquier valor entre 0 y π.

8.2 Scattering

Hasta aquí hemos visto las consecuencias de los teoremas de conservación para la evolución no ligada de un sistema de dos cuerpos. Ahora veremos cómo una interacción específica determina el ángulo χ en función de los datos iniciales. Sabemos que el problema de dos cuerpos se resuelve estudiando una partícula de masa $\mu = m_1 m_2/(m_1 + m_2)$ (la masa reducida) en una posición $\vec{r} = \vec{r}_1 - \vec{r}_2$ respecto del centro de un potencial central. Conocemos la solución en función de las primeras integrales E y ℓ (la energía y el momento angular de los dos cuerpos en el sistema centro de masa, como también la energía de la masa reducida y su momento angular respecto del centro de fuerzas). Pero será conveniente escribir la solución en términos de la velocidad v_∞ de μ en el infinito, y el *parámetro de impacto* s:



Entonces

$$E = \frac{1}{2}\,\mu\, v_\infty^2\ , \qquad\qquad \ell = \mu\ s\ v_\infty \tag{8.4}$$

En la Figura indicamos también el ángulo χ, que es la dirección final del vector $\vec{r}$. Como se ve, χ se relaciona con el ángulo φ_∞ correspondiente a $r \to \infty$, siendo [1]

$$2\ \varphi_\infty + \chi = \pi \tag{8.5}$$

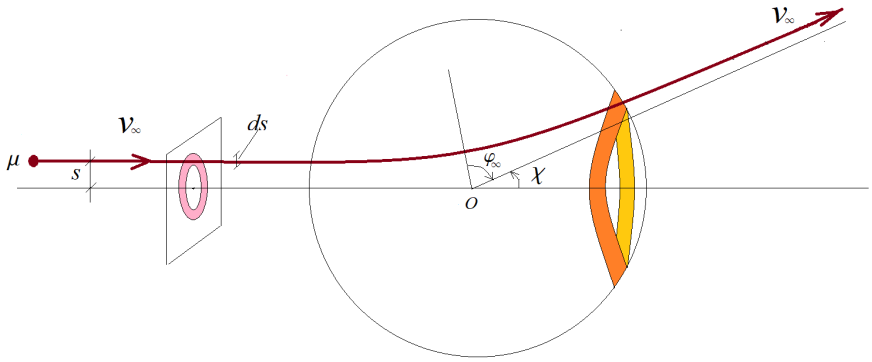
Como este tipo de experimento se suele realizar con haces de partículas que son lanzadas hacia un centro dispersor (*scatter* significa *dispersar* o *esparcir*), lo que interesa conocer es qué fracción de los proyectiles se dispersa con ángulos comprendidos entre θ_1 y $\theta_1 + d\theta_1$. En términos del problema equivalente con la partícula μ, averiguaremos cuántas de estas partículas son dispersadas con ángulos comprendidos entre χ y $\chi + d\chi$; y luego usaremos la relación entre θ_1 y χ ya obtenida para responder la cuestión al nivel del problema de dos cuerpos.

El número dN de partículas que emergen por unidad de tiempo con ángulos comprendidos entre χ y $\chi + d\chi$ contiene la información sobre el potencial dispersor. Como dN es proporcional a la intensidad del haz (supuesta uniforme, o a lo sumo dependiente sólo de s) conviene caracterizar el potencial dispersor mediante el cociente entre dN y la intensidad. Entonces llamaremos

Intensidad I: número de partículas por unidad de tiempo y de superficie atravesada en el haz de entrada.

Sección eficaz de scattering: $\qquad d\sigma \equiv \dfrac{dN}{I}$

$d\sigma$ tiene unidades de superficie. Es igual al área infinitesimal en la sección del haz por donde entraron las dN partículas que luego se dispersaron con ángulos comprendidos entre χ y $\chi + d\chi$ (ver Figura).



En efecto, por el anillo en el haz de entrada pasan $2\pi s\ ds\ I$ partículas. Y estas partículas salen luego con ángulos comprendidos entre χ y $\chi + d\chi$. Por supuesto, la cuestión implica

1. La fórmula vale también para la trayectoria atractiva, donde $\chi < 0$ (ver Figura en §8.1).

conocer la relación entre s y χ; esta relación depende del potencial de interacción. En síntesis,

$$d\sigma = \frac{dN}{I} = \frac{2\pi s\ ds\ I}{I} = 2\pi s\ ds = 2\pi s(\chi)\ \left|\frac{ds}{d\chi}\right|\ d\chi$$

Más adelante calcularemos la función $s(\chi)$ para el potencial coulombiano. La razón para utilizar el valor absoluto de $ds/d\chi$ es que esta derivada puede ser negativa; en efecto $s(\chi)$ es decreciente para el potencial coulombiano repulsivo.

Es común formular la sección eficaz no en términos del ángulo χ sino del ángulo sólido abarcado entre χ y $\chi + d\chi$,

$$d\Omega = 2\pi\ \sin|\chi|\ d\chi\ ,$$

es decir

$$d\sigma = \frac{s(\chi)}{\sin|\chi|}\ \left|\frac{ds}{d\chi}\right|\ d\Omega$$

8.2.1 Potencial coulombiano

En el caso del potencial coulombiano $V(r) = -\alpha/r$ (atractivo o repulsivo: $\alpha > 0$ o $\alpha < 0$) habíamos obtenido (el signo superior corresponde al caso atractivo)

$$\frac{\ell^2}{\mu\ \alpha}\ \frac{1}{r} = 1 \pm e\ \cos\varphi\ , \qquad e = \sqrt{1 + \frac{2\ \ell^2 E}{\alpha^2 \mu}} = \sqrt{1 + \left(\frac{\mu\ s\ v_\infty^2}{\alpha}\right)^2}$$

donde φ se mide desde $r_{\min}$. Por otro lado, de la ecuación de la cónica se obtiene que

$$\cos\varphi_\infty\ = \mp\frac{1}{e} = \mp\frac{1}{\sqrt{1 + \left(\frac{\mu\ s\ v_\infty^2}{\alpha}\right)^2}}$$

La expresión entre paréntesis corresponde a $\tan\varphi_\infty$. φ_∞ se relaciona con χ a través de la ecuación (8.5). Tanto en el caso atractivo como en el repulsivo se obtiene que

$$-\frac{\mu\ s\ v_\infty^2}{\alpha} = \tan\varphi_\infty\ = \tan\frac{\pi - \chi}{2} = \cot\frac{\chi}{2}$$

(si α es positivo, χ es negativo: $-\pi \leq \chi \leq 0$). Entonces

$$s(\chi) = -\frac{\alpha}{\mu\ v_\infty^2}\ \cot\frac{\chi}{2}$$

Por lo tanto

$$\frac{ds}{d\chi} = \frac{\alpha}{2\ \mu\ v_\infty^2}\ \frac{1}{\sin^2\frac{\chi}{2}}$$

Reemplazando $s(\chi)$ y $|ds/d\chi|$ en la sección eficaz [2]

$$d\sigma = 2\pi s(\chi) \left|\frac{ds}{d\chi}\right| d\chi = \pi \left(\frac{\alpha}{\mu\ v_\infty^2}\right)^2 \frac{\cot\frac{|\chi|}{2}}{\sin^2\frac{\chi}{2}}\ d\chi = \pi \left(\frac{\alpha}{\mu\ v_\infty^2}\right)^2 \frac{\cos\frac{\chi}{2}}{\sin^3\frac{|\chi|}{2}}\ d\chi$$

Reemplazando $d\chi$ en términos del ángulo sólido

$$d\Omega = 2\pi\ \sin|\chi|\ d\chi = 2\pi\ \sin\left(\frac{|\chi|}{2} + \frac{|\chi|}{2}\right) = 4\pi\ \sin\frac{|\chi|}{2}\ \cos\frac{\chi}{2}$$

($cos(\chi/2)$ es positivo porque $0 \leq |\chi| \leq \pi$), llegamos a la **fórmula de Rutherford** para la sección eficaz de la interacción coulombiana:

$$d\sigma = \left(\frac{\alpha}{2\ \mu\ v_\infty^2}\right)^2 \frac{1}{\sin^4\frac{\chi}{2}}\ d\Omega$$

Para conocer la sección eficaz de proyectiles y blancos, usaremos las relaciones

$$\tan\theta_1 = \frac{\sin\chi}{\cos\chi + \frac{m_1}{m_2}}\ , \qquad \theta_2 = \frac{\pi - \chi}{2}$$

para reemplazar $\chi(\theta_1)$ y $d\chi = (d\chi/d\theta_1)\ d\theta_1$, y lo mismo para θ_2. Cuando proyectiles y blancos son iguales no tiene sentido distinguirlos después de la dispersión, y deberían sumarse las secciones eficaces respectivas (teniendo en cuenta que si $m_1 = m_2$ es $\theta_1 + \theta_2 = \pi/2$). Por otro lado, la interacción entre los proyectiles dentro del haz ha sido ignorada en este tratamiento.

En 1909 E. Rutherford colaboró con Geiger y Marsden en un experimento de dispersión de un haz de partículas alfa (núcleos de ^{4}He) que atravesaba una lámina de oro. La forma en que las partículas alfa eran dispersadas (algunas con ángulo de deflexión muy grande) indicaban que en la materia existían regiones de carga positiva muy concentrada. Esto le permitió enunciar en 1911 lo que conocemos como el modelo atómico de Rutherford, donde la carga de un signo se concentra en un núcleo que contiene casi toda la masa del átomo, y la carga de signo opuesto se distribuye en electrones de baja masa que orbitan alrededor del núcleo.

Como la masa del núcleo de oro es mucho mayor que la masa de la partícula alfa ($m_2 >> m_1$), entonces $\theta_1 \simeq \chi$, y la fórmula de Rutherford se aplica directamente a la sección eficaz de dispersión de las partículas alfa.

Nótese que la sección eficaz total del potencial Coulombiano diverge:

$$\sigma = \int d\sigma = \pi \left(\frac{\alpha}{\mu\ v_\infty^2}\right)^2 \int_0^\pi \frac{\cos\frac{\chi}{2}}{\sin^3\frac{\chi}{2}}\ d\chi = \infty$$

En efecto, σ corresponde al área de la sección del haz por donde entran las partículas que serán dispersadas. La sección total es infinita porque el potencial de Coulomb dispersa aun aquellos proyectiles cuyo parámetro de impacto tiende a infinito, con ángulo de desviación $\chi \to 0$.

2. Como α y χ tienen signos opuestos vale que $-\alpha\cot(\chi/2) = |\alpha|\cot(|\chi|/2)$.

9 Oscilaciones

9.1 Pequeñas oscilaciones

La aproximación de pequeñas oscilaciones, que se practica en problemas elementales como el del péndulo simple con el objeto de linealizar la ecuación de movimiento, puede aplicarse también en problemas de varios grados de libertad. Consideremos un potencial que depende sólo de las coordenadas generalizadas, $V(q_1, ..., q_n)$, y que posea un **mínimo** en algún punto $(\bar{q}_1, ..., \bar{q}_n)$ del espacio de configuración:

$$\frac{\partial V}{\partial q_\mu}|_{q=\bar{q}} \;=\; 0 \;, \qquad \forall \mu$$

Buscaremos soluciones que se aparten poco de esa posición de equilibrio estable:

$$q_\mu(t) = \bar{q}_\mu + \eta_\mu(t)$$

donde $\eta_\mu(t)$ es una pequeña perturbación; de modo que sólo términos lineales en $\eta_\mu(t)$ serán considerados en las ecuaciones de movimiento. Para ello aproximaremos el Lagrangiano a orden cuadrático en las perturbaciones. El potencial aproximado es

$$V \simeq V(\bar{q}) + \frac{1}{2} \sum_{\mu,\nu} \frac{\partial^2 V}{\partial q_\mu \partial q_\nu}|_{q=\bar{q}} \; (q_\mu - \bar{q}_\mu) \; (q_\nu - \bar{q}_\nu) = V(\bar{q}) + \frac{1}{2} \sum_{\mu,\nu} k_{\mu\nu} \; \eta_\mu \eta_\nu$$

donde

$$k_{\mu\nu} \equiv \frac{\partial^2 V}{\partial q_\mu \partial q_\nu}|_{q=\bar{q}}$$

Las velocidades generalizadas son

$$\dot{q}_\mu(t) = \dot{\eta}_\mu(t)$$

En un sistema holónomo y esclerónomo, la energía cinética es homogénea de grado 2 en las velocidades generalizadas. Como $\dot{q}_\mu = \dot{\eta}_\mu$, la energía cinética aproximada resulta[1]

$$T = \frac{1}{2} \sum_{\mu,\nu} m_{\mu\nu}(q) \; \dot{q}_\mu \; \dot{q}_\nu \simeq \frac{1}{2} \sum_{\mu,\nu} m_{\mu\nu} \; \dot{\eta}_\mu \dot{\eta}_\nu$$

1. $m_{\mu\nu}(q) = \sum_{i=1}^{N} m_i \; \frac{\partial \vec{r}_i}{\partial q_\mu} \cdot \frac{\partial \vec{r}_i}{\partial q_\nu}$

donde

$$m_{\mu\nu} \equiv m_{\mu\nu}(\bar{q})$$

De esta manera, las n ecuaciones de Euler-Lagrange linealizadas resultan

$$\sum_{\nu=1}^{n} (m_{\mu\nu}\ \ddot{\eta}_\nu(t) + k_{\mu\nu}\ \eta_\nu(t)) = 0\ , \qquad \mu = 1, ..., n \tag{9.1}$$

donde $m_{\mu\nu}$, $k_{\mu\nu}$ son dos **matrices reales simétricas** que contienen la información sobre las frecuencias propias del sistema.

Busquemos una solución de la forma[2]

$$\eta_\nu(t) = A_\nu\ e^{i\omega t}\ , \qquad \nu = 1, ..., n$$

donde los coeficientes A_ν son números complejos que contendrán información sobre amplitudes y fases relativas. Reemplazando en las ecuaciones,

$$\sum_{\nu=1}^{n} \left(-\omega^2\ m_{\mu\nu} + k_{\mu\nu}\right) A_\nu = 0\ , \qquad \mu = 1, ..., n \tag{9.2}$$

Vemos que el sistema de ecuaciones diferenciales se ha convertido en un sistema de ecuaciones algebraicas para los coeficientes A_ν.

En un oscilador unidimensional A no queda determinado por la ecuación de movimiento sino por las condiciones iniciales. En un caso con n grados de libertad, veremos que los A_ν quedan determinados en parte por las ecuaciones de movimiento, y en parte por las condiciones iniciales. Las ecuaciones de movimiento determinarán relaciones entre las evoluciones de los distintos grados de libertad en cada posible solución del problema.

Las ecuaciones (9.2) tienen una solución trivial,

$$A_\nu = 0 \qquad \forall\nu \qquad \Rightarrow \qquad q_\nu(t) = \bar{q}_\nu\ ,$$

que corresponde al punto de equilibrio. Para tener oscilaciones en torno al equilibrio necesitamos soluciones no triviales. Pero ocurre que el sistema de ecuaciones lineales (9.2) que deben satisfacer los coeficientes A_ν está igualado a cero. Entonces o todos los A_ν se anulan o la matriz $-\omega^2\ m_{\mu\nu} + k_{\mu\nu}$ es no inversible para permitir soluciones no triviales. Para que la matriz $-\omega^2\ m_{\mu\nu} + k_{\mu\nu}$ sea no inversible, su determinante debe anularse:

$$\det(-\omega^2\ m_{\mu\nu} + k_{\mu\nu}) = 0$$

En ese caso las ecuaciones no son independientes (la anulación del determinante dice que al menos una fila es una combinación de las otras). Es decir que no todos los A_ν podrán ser resueltos, sino que obtendremos relaciones entre los mismos.

2. Como las ecuaciones (9.1) son lineales a coeficientes reales, las partes real e imaginaria de una solución compleja son soluciones reales de las ecuaciones.

La condición $\det(-\omega^2\ m_{\mu\nu} + k_{\mu\nu}) = 0$ debe verse como una forma de seleccionar las frecuencias ω que posibilitan una solución no trivial. La condición corresponde a una ecuación de grado n para la incógnita ω^2, que se caracteriza por sus n raíces.

Vamos a probar que las n raíces ω^2 son positivas, de modo que todas ellas conducen a soluciones que oscilan indefinidamente. Dada una solución no trivial, usaremos su compleja conjugada para multiplicar la ecuación (9.2) por A^*_μ y sumar sobre μ,

$$\sum_{\mu,\nu} \left(-\omega^2\ m_{\mu\nu} + k_{\mu\nu}\right) A_\nu A^*_\mu = 0$$

es decir

$$\omega^2 \sum_{\mu,\nu} m_{\mu\nu}\ A_\nu A^*_\mu = \sum_{\mu,\nu} k_{\mu\nu}\ A_\nu A^*_\mu \tag{9.3}$$

Veamos que cada sumatoria doble es un número real, lo que es consecuencia de la simetría de las respectivas matrices. Los términos con $\mu \neq \nu$ se agrupan en cantidades reales; por ejemplo

$$m_{12}\ A_2 A^*_1 + m_{21}\ A_1 A^*_2 = m_{12}(A_2 A^*_1 + A_1 A^*_2) = 2\ m_{12}\ \mathrm{Re}[A_2 A^*_1]\ ;$$

y los términos diagonales son reales:

$$m_{11}\ A_1 A^*_1 = m_{11}\ \ |A_1|^2$$

Lo mismo sucede con la doble sumatoria que contiene la matriz $k_{\mu\nu}$. Podemos probar que ambas sumatorias son reales en una forma más elegante:

$$(\sum_{\mu,\nu} k_{\mu\nu}\ A_\nu A^*_\mu)^* = \sum_{\mu,\nu} k^*_{\mu\nu}\ A^*_\nu A_\mu = \sum_{\mu,\nu} k_{\mu\nu}\ A^*_\nu A_\mu = \sum_{\mu,\nu} k_{\nu\mu}\ A^*_\mu A_\nu = \sum_{\mu,\nu} k_{\mu\nu}\ A_\nu A^*_\mu$$

donde usamos que $k_{\mu\nu}$ es una matriz real y simétrica.

Como en la ecuación (9.3) ambas sumatorias dobles son reales, entonces concluimos que ω^2 debe ser real. Siendo ω^2 real, entonces todos los coeficientes en las ecuaciones (9.2) son reales; por lo tanto esas ecuaciones admiten soluciones no triviales A_ν reales. Luego es inmediato ver que ambas sumatorias dobles en (9.3) son definidas positivas, pues las matrices $m_{\mu\nu}$ y $k_{\mu\nu}$ definen formas cuadráticas positivas. En efecto,

$$\sum_{\mu,\nu} m_{\mu\nu}\ A_\nu A^*_\mu = \sum_{\mu,\nu} m_{\mu\nu}\ A_\nu A_\mu > 0$$

para valores arbitrarios de A_ν, del mismo modo que $T = (1/2)\sum_{\mu,\nu} m_{\mu\nu}\dot{\eta}_\mu\dot{\eta}_\nu > 0$ para valores arbitrarios (pero no todos nulos) de las velocidades generalizadas. También se concluye que la otra sumatoria doble es positiva pues $\sum_{\mu,\nu} k_{\mu\nu}\eta_\mu\eta_\nu > 0$ como condición para que el punto de equilibrio $\eta_\mu = 0\ \forall\mu$ sea efectivamente un mínimo del potencial (el potencial debe crecer para cualquier desplazamiento del punto de equilibrio). Vale

la pena notar que ambas sumatorias dobles serían positivas aun si los A_ν's fuesen complejos.[3]

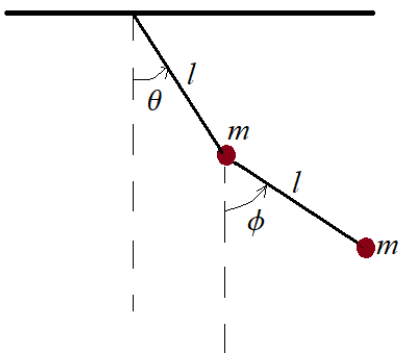
Como en la ecuación (9.3) ambas sumatorias dobles son positivas, entonces ω^2 es positivo; lo que implica que ω es real, y las soluciones oscilan indefinidamente (no hay amortiguación). Una vez que hayamos encontrado las n frecuencias propias del sistema, las identificaremos con una etiqueta s: ω_s , $s = 1, ..., n$. Para cada ω_s el sistema de ecuaciones tendrá una solución real no trivial $A_\nu^{(s)}$, $\nu = 1, ..., n$. Pero, como el sistema de ecuaciones (9.2) es lineal y homogéneo, cada solución es ambigua: multiplicándola por un factor global complejo $C^{(s)}$ tendremos también una solución. Esta ambigüedad sólo dice que la solución correspondiente a la frecuencia ω_s contiene una amplitud y una fase globales que deberán fijarse mediante las condiciones iniciales. Pero las amplitudes *relativas* de las η_ν en cada **modo normal de oscilación** están determinadas por la solución de la ecuación (9.2) correspondiente a la frecuencia ω_s del modo normal correspondiente.

Finalmente, como la ecuación (9.1) es lineal, la solución más general es una combinación de los n modos normales de oscilación:

$$\eta_\nu(t) = \sum_{s=1}^{n} C^{(s)} A_\nu^{(s)} \; e^{i\omega_s t} \; , \qquad \nu = 1, ..., n \tag{9.4}$$

Los coeficientes complejos $C^{(s)}$ indican con qué amplitud y qué fase entra cada modo normal de oscilación en la solución general. Son determinados por las condiciones iniciales.

9.1.1 El péndulo doble plano



Punto de equilibrio: $\qquad \bar{\theta} = 0, \quad \bar{\phi} = 0$

3. Usar que $\sum_{\mu,\nu} m_{\mu\nu} \; A_\nu A_\mu^*$ es real, y que $2\,\mathrm{Re}[A_\nu A_\mu^*] = \mathrm{Re}[A_\nu]\,\mathrm{Re}[A_\mu] + \mathrm{Im}[A_\nu]\,\mathrm{Im}[A_\mu]$.

Coordenadas: $\eta_1 = \theta, \quad \eta_2 = \phi$

Matrices $m_{\mu\nu}$ y $k_{\mu\nu}$: $\mathbf{M} = \begin{pmatrix} 2m\ell^2 & m\ell^2 \\ m\ell^2 & m\ell^2 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 2mg\ell & 0 \\ 0 & mg\ell \end{pmatrix}$

Frecuencias propias: ${\omega_1}^2 = \frac{g}{\ell}(2+\sqrt{2})\ , \qquad {\omega_2}^2 = \frac{g}{\ell}(2-\sqrt{2})$

Modo normal 1: $A_\theta^{(1)} = -A_\phi^{(1)}/\sqrt{2}$ (oscilación a contrafase)

Modo normal 2: $A_\theta^{(2)} = A_\phi^{(2)}/\sqrt{2}$ (oscilación en fase)

Como previmos, la resolución para cada frecuencia no permite obtener todos los coeficientes A_ν sino una relación entre ellos, que nos dice con qué amplitud relativa entra cada coordenada en el modo de oscilación en cuestión. La solución general es una combinación de ambos modos:

$$\begin{aligned} \theta(t) &= C^{(1)} A_\theta^{(1)}\ e^{i\omega_1 t} + C^{(2)} A_\theta^{(2)}\ e^{i\omega_2 t} \\ \phi(t) &= C^{(1)} A_\phi^{(1)}\ e^{i\omega_1 t} + C^{(2)} A_\phi^{(2)}\ e^{i\omega_2 t} \end{aligned}$$

Si absorbemos $A_\phi^{(1)}$, $A_\phi^{(2)}$en las constantes de integración $C^{(1)}$, $C^{(2)}$ respectivamente, el resultado es

$$\begin{aligned} \theta(t) &= -\frac{1}{\sqrt{2}} C_1\ e^{i\omega_1 t} + \frac{1}{\sqrt{2}} C_2\ e^{i\omega_2 t} \\ \phi(t) &= C_1\ e^{i\omega_1 t} + C_2\ e^{i\omega_2 t} \end{aligned}$$

Los complejos C_1, C_2, se determinan con las condiciones iniciales. Por ejemplo, si las condiciones iniciales fueran tales que $C_2 = 0$, entonces el movimiento correspondería al modo normal 1.

9.2 Ortonormalidad

Si bien el método de resolución podría darse por concluido, veremos algunas propiedades de las soluciones $A_\nu^{(s)}$ que serán de utilidad. Consideremos las ecuaciones (9.2) para dos frecuencias distintas ω_s y ω_t:

$$\begin{aligned} \omega_s^2 \sum_{\nu=1}^{n} m_{\mu\nu}\ A_\nu^{(s)} &= \sum_{\nu=1}^{n} k_{\mu\nu}\ A_\nu^{(s)}\ , \\ \omega_t^2 \sum_{\nu=1}^{n} m_{\mu\nu}\ A_\nu^{(t)} &= \sum_{\nu=1}^{n} k_{\mu\nu}\ A_\nu^{(t)} \end{aligned}$$

"Multipliquemos" la primera por $\sum_\mu A_\mu^{(t)}$ y la segunda por $\sum_\mu A_\mu^{(s)}$; usemos la simetría de $m_{\mu\nu}$ y $k_{\mu\nu}$ y restemos:

$$(\omega_s^2 - \omega_t^2) \sum_{\mu,\nu} m_{\mu\nu}\ A_\mu^{(t)} A_\nu^{(s)} = 0$$

Si los autovalores ω_s^2 son todos distintos esto significa que

$$\sum_{\mu,\nu} m_{\mu\nu}\ A_\mu^{(t)} A_\nu^{(s)} = 0\ , \qquad s \neq t$$

Esto es una suerte de ortogonalidad entre vectores. En efecto, para cada valor de s tenemos un "vector" $\vec{A}^{(s)}$ de n componentes $A_\nu^{(s)}$. La relación anterior puede verse como un "producto interno" $\vec{A}^{(s)} \cdot \vec{A}^{(t)}$ en una "métrica" $m_{\mu\nu}$; la relación anterior dice que los "vectores" $\vec{A}^{(s)}$'s son ortogonales entre sí.[4] Aun si hubiese autovalores iguales (degeneración), podrían elegirse ortogonales los autovectores del subespacio correspondiente. Podemos avanzar más sobre esta línea; en el caso $s = t$, estaríamos ante el cuadrado de la norma $\vec{A}^{(s)} \cdot \vec{A}^{(s)}$ de cada vector. Como los $\vec{A}^{(s)}$'s no quedan completamente determinados por las ecuaciones (9.2) podemos valernos de la ambigüedad remanente para fijarla mediante una condición de normalización. Así diremos que los $A_\nu^{(s)}$ satisfacen una condición de **ortonormalidad**:

$$\sum_{\mu,\nu} m_{\mu\nu}\ A_\mu^{(t)} A_\nu^{(s)} = \delta_{st}\ , \qquad \forall s,t \tag{9.5}$$

9.3 Notación matricial

Como hay n coeficientes $A_\nu^{(s)}$ para cada una de las n frecuencias propias ω_s, podemos organizarlos todos en una matriz $\mathbf{A}$ de componentes $A_{\nu s} \equiv A_\nu^{(s)}$ (cada columna corresponde a un modo normal). La relación de ortonormalidad (9.5) significa que[5]

$$\mathbf{A^T\ M\ A = 1}$$

Por ejemplo, en el péndulo doble la matriz $\mathbf{A}$ resulta

$$\mathbf{A} = \begin{pmatrix} A_\theta^{(1)} & A_\theta^{(2)} \\ A_\phi^{(1)} & A_\phi^{(2)} \end{pmatrix} = \begin{pmatrix} -A_\phi^{(1)}/\sqrt{2} & A_\phi^{(2)}/\sqrt{2} \\ A_\phi^{(1)} & A_\phi^{(2)} \end{pmatrix}$$

y la normalización se consigue con

$$A_\phi^{(1)} = \frac{1}{\ell\sqrt{(2-\sqrt{2})\ m}}\ , \qquad A_\phi^{(2)} = \frac{1}{\ell\sqrt{(2+\sqrt{2})\ m}}\ .$$

4. Notar que $m_{\mu\nu}$ en la ecuación $\det(-\omega^2 m_{\mu\nu} + k_{\mu\nu}) = 0$ ocupa el lugar que tendría la identidad en el problema usual de autovalores y autovectores.

5. $\sum_{\mu\nu} m_{\mu\nu}\ A_\mu^{(t)} A_\nu^{(s)} = \sum_{\mu\nu} m_{\mu\nu}\ A_{\mu t}\ A_{\nu s} = \sum_{\mu\nu} (\mathbf{A^T})_{t\mu} (\mathbf{M})_{\mu\nu} (\mathbf{A})_{\nu s}$.

Podemos ver que el formalismo de pequeñas oscilaciones se resume en tres matrices de $n \times n$:
1) La matriz $\mathbf{M}$ de componentes $m_{\mu\nu}$
2) La matriz $\mathbf{K}$ de componentes $k_{\mu\nu}$
3) La matriz $\mathbf{A}$ de componentes $A_{\nu s} \equiv A_\nu^{(s)}$ (cada columna representa un modo normal).

Con ayuda de las matrices $\mathbf{A}$, $\mathbf{M}$, $\mathbf{K}$, las $n \times n$ ecuaciones

$$\sum_{\nu=1}^{n} \left(-\omega_s^2 \; m_{\mu\nu} + k_{\mu\nu}\right) A_\nu^{(s)} = 0 \; , \qquad \mu = 1, ..., n; \qquad s = 1, ..., n$$

se escriben en una única ecuación matricial:

$$\mathbf{M} \; \mathbf{A} \; \mathbf{\Omega}^2 = \mathbf{K} \; \mathbf{A} \tag{9.6}$$

donde $\mathbf{\Omega}^2$ es una matriz diagonal formada por los cuadrados de las frecuencias propias: $\Omega_{ts}^2 = \omega_s^2 \; \delta_{ts}$ o bien

$$\mathbf{\Omega}^2 = \mathrm{diag}(\omega_1^2, \omega_2^2, ..., \omega_n^2)$$

Si multiplicamos por la traspuesta de $\mathbf{A}$,

$$\mathbf{A^T} \; \mathbf{M} \; \mathbf{A} \; \mathbf{\Omega}^2 = \mathbf{A^T} \; \mathbf{K} \; \mathbf{A}$$

Por otro lado, hemos visto que la relación de ortonormalidad significa que

$$\mathbf{A^T} \; \mathbf{M} \; \mathbf{A} = \mathbf{1} \tag{9.7}$$

Entonces

$$\mathbf{\Omega}^2 = \mathbf{A^T} \; \mathbf{K} \; \mathbf{A} \tag{9.8}$$

lo que significa que la matriz $\mathbf{A}$ que resuelve las ecuaciones dinámicas diagonaliza tanto a $\mathbf{M}$ como a $\mathbf{K}$.

9.4 Coordenadas normales

Las coordenadas η_ν podrían no ser las más adecuadas para el sistema físico en consideración. Organicemos las η_ν's en una columna que llamaremos $|\eta >$,

$$|\eta > = \begin{pmatrix} \eta_1 \\ \eta_2 \\ ... \\ ... \\ \eta_n \end{pmatrix} \; ,$$

para escribir las ecuaciones de movimiento

$$\sum_{\nu=1}^{n} (m_{\mu\nu}\ \ddot{\eta}_\nu + k_{\mu\nu}\ \eta_\nu) = 0\ , \qquad \mu = 1, ..., n$$

en notación matricial:

$$\mathbf{M}\ |\ddot{\eta}> + \mathbf{K}\ |\eta> = 0$$

A la luz de las relaciones matriciales anteriores, esta ecuación sugiere introducir **coordenadas normales** ξ_s definidas mediante la siguiente transformación:

$$|\eta> = \ \mathbf{A}\ |\xi> \tag{9.9}$$

En efecto, reemplazando en las ecuaciones de movimiento resulta:

$$\mathbf{M}\ \mathbf{A}\ |\ddot{\xi}> + \mathbf{K}\ \mathbf{A}\ |\xi> = 0$$

Multiplicando a izquierda por $\mathbf{A^T}$, y usando las ecuaciones (9.7) y (9.8),

$$|\ddot{\xi}> + \mathbf{\Omega}^2\ |\xi> = 0$$

que significa

$$\ddot{\xi}_s + \omega_s^2\ \xi_s = 0\ , \qquad s = 1, ..., n \tag{9.10}$$

Vemos que el conjunto de ecuaciones de movimiento luce ahora como un conjunto de osciladores armónicos desacoplados, cada uno de ellos correspondiendo a cada una de las frecuencias propias del sistema. En el modo normal s sólo la coordenada ξ_s oscila; las demás coordenadas permanecen iguales a cero.

El desacople puede verificarse también en el Lagrangiano. Si llamamos $< \eta|$ a la fila formada por las η_ν's ($< \eta|\ = |\eta >^{\mathbf{T}} =< \xi|\ \mathbf{A^T}$), el Lagrangiano linealizado se escribe

$$\begin{aligned} L &= \frac{1}{2} < \dot{\eta}|\ \mathbf{M}\ |\dot{\eta}> - \frac{1}{2} < \eta|\ \mathbf{K}\ |\eta> \\ &= \frac{1}{2} < \dot{\xi}|\ \mathbf{A^T}\ \mathbf{M}\ \mathbf{A}\ |\dot{\xi}> - \frac{1}{2} < \xi|\ \mathbf{A^T}\ \mathbf{K}\ \mathbf{A}\ |\xi> \\ &= \frac{1}{2} < \dot{\xi}|\dot{\xi}> - \frac{1}{2} < \xi|\ \mathbf{\Omega}^2\ |\xi> = \sum_{s=1}^{n} \frac{1}{2}\ (\dot{\xi}_s^2 - \omega_s^2\ \xi_s^2) \end{aligned}$$

Vemos que las coordenadas normales permiten que el Lagrangiano del sistema se muestre como una suma de n osciladores armónicos desacoplados, cada uno con su frecuencia propia ω_s. En otras palabras, al diagonalizar $\mathbf{M}$ y $\mathbf{K}$ (las matrices que representan las energías cinéticas y potencial) la matriz $\mathbf{A}$ está señalando las direcciones privilegiadas en el espacio de configuración respecto de las cuales la descripción del movimiento es más simple.

La solución de las ecuaciones (9.10) es

$$\xi_s(t) = C^{(s)}\ e^{i\omega_s t}\ ;$$

reemplazando en la solución general (9.4),

$$\eta_\nu(t) = \sum_{s=1}^{n} A_\nu^{(s)}\ \xi_s(t)\ , \qquad \nu = 1, ..., n$$

que no es más que la ecuación (9.9), ahora escrita para la evolución temporal del sistema.

Si se desea obtener las coordenadas normales a partir de las coordenadas originales, usaremos la relación de ortonormalidad para invertir la ecuación (9.9):

$$\mathbf{A^T\ M}\ |\eta> =\ \mathbf{A^T\ M\ A}\ |\xi> = |\xi>$$

En particular, para dados valores iniciales de $|\eta>$ y $|\dot{\eta}>$ tendremos los respectivos valores de $C^{(s)}$

$$\mathbf{A^T\ M}\ |\eta_o> = |\xi_o> = |\mathbf{C}>$$

$$\mathbf{A^T\ M}\ |\dot{\eta}_o> =\ |\dot{\xi}_o> = |i\ \omega\ \mathbf{C}> = i\ \mathbf{\Omega}\ \ |\mathbf{C}>$$

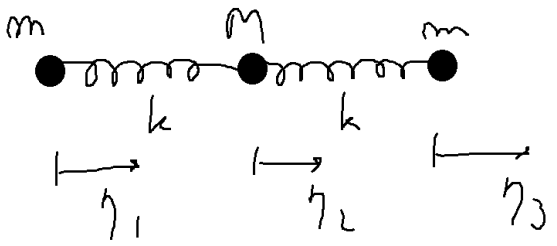
Como finalmente nos interesa la parte real de la solución $|\eta(t)>$, tomaremos parte real en las anteriores ecuaciones,

$$\mathbf{A^T\ M}\ \ \mathrm{Re}\,|\eta_o> = \mathrm{Re}\,|\mathbf{C}>$$

$$-\mathbf{\Omega}^{-1}\mathbf{A^T\ M}\ \ \mathrm{Re}\,|\dot{\eta}_o> = \mathrm{Im}\,|\mathbf{C}>$$

de donde resultan los valores de los complejos $C^{(s)}$ que se precisan para satisfacer las condiciones iniciales de la solución real.

9.4.1 La molécula de CO_2



$$L = \frac{1}{2}\ m\ \dot{\eta}_1^2 + \frac{1}{2}\ M\ \dot{\eta}_2^2 + \frac{1}{2}\ m\ \dot{\eta}_3^2 - \frac{1}{2}\ k\ (\eta_2 - \eta_1)^2 - \frac{1}{2}\ k\ (\eta_3 - \eta_2)^2$$

$$\mathbf{M} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} , \qquad \mathbf{K} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} ,$$

Las frecuencias propias resultan

$$\omega_1 = 0 , \qquad \omega_2 = \sqrt{\frac{k}{m}} , \qquad \omega_3 = \sqrt{\frac{2k}{M} + \frac{k}{m}}$$

El resultado $\omega_1 = 0$ puede sorprendernos (habíamos dicho que todos los ω^2 eran positivos). Sin embargo el punto de equilibrio $\bar{\eta}_1 = \bar{\eta}_2 = \bar{\eta}_3 = 0$ no es estrictamente un mínimo de potencial. Nótese que el potencial no cambia ante traslaciones $\Delta\eta_1 = \Delta\eta_2 = \Delta\eta_3$. Este modo traslatorio podría eliminarse trabajando en el sistema CM. Asimismo, en un movimiento tridimensional convendría eliminar también el modo rotatorio, en el que la molécula rota como un rígido.

El siguiente paso es construir la ecuación (9.6) para resolver los $A_\nu^{(s)}$; el resultado es

$$\mathbf{A} = \begin{pmatrix} A_1^{(1)} & A_1^{(2)} & A_1^{(3)} \\ A_1^{(1)} & 0 & -\frac{2m}{M} A_1^{(3)} \\ A_1^{(1)} & -A_1^{(2)} & A_1^{(3)} \end{pmatrix}$$

y la normalización da

$$\mathbf{A} = \begin{pmatrix} \frac{1}{\sqrt{2m+M}} & \frac{1}{\sqrt{2m}} & \sqrt{\frac{M}{2m(2m+M)}} \\ \frac{1}{\sqrt{2m+M}} & 0 & -\sqrt{\frac{2m}{M(2m+M)}} \\ \frac{1}{\sqrt{2m+M}} & -\frac{1}{\sqrt{2m}} & \sqrt{\frac{M}{2m(2m+M)}} \end{pmatrix}$$

Las coordenadas normales son

$$|\xi> = \mathbf{A^T}\ \mathbf{M}\ |\eta> = \begin{pmatrix} \frac{m}{\sqrt{2m+M}} & \frac{M}{\sqrt{2m+M}} & \frac{m}{\sqrt{2m+M}} \\ \sqrt{\frac{m}{2}} & 0 & -\sqrt{\frac{m}{2}} \\ \sqrt{\frac{mM}{2(2m+M)}} & -\sqrt{\frac{2mM}{2m+M}} & \sqrt{\frac{mM}{2(2m+M)}} \end{pmatrix} |\eta>$$

es decir,

$$\begin{aligned} \xi_1 &= \frac{1}{\sqrt{2m+M}} (m\ \eta_1 + M\ \eta_2 + m\ \eta_3) \\ \xi_2 &= \sqrt{\frac{m}{2}}\ (\ \eta_1 - \ \eta_3) \\ \xi_3 &= \sqrt{\frac{mM}{2(2m+M)}}\ (\eta_1 - 2\ \eta_2 + \ \eta_3) \end{aligned}$$

Modo normal 1 ($\xi_2 = 0 = \xi_3$) : $\quad \eta_1 = \eta_2 = \eta_3$ (traslación)

Modo normal 2 ($\xi_1 = 0 = \xi_3$) : $\quad \eta_2 = 0$, $\eta_1 = -\eta_3$

Modo normal 3 ($\xi_1 = 0 = \xi_2$) : $\quad \eta_1 = -\frac{M}{2m}\,\eta_2 = \eta_3$

Por supuesto la forma de los modos normales de oscilación ya era visible en el primer resultado para la matriz **A**, que contenía las amplitudes relativas de cada modo.

Videos

Comportamiento caótico del péndulo doble fuera del régimen de pequeñas oscilaciones

www.youtube.com/watch?v=d3uOKlIEIoU
www.youtube.com/watch?v=7dXZSn3PV5E
www.youtube.com/watch?v=d0Z8wLLPNE0
www.youtube.com/watch?v=TmlpDO3LcOs

9.5 Oscilador forzado

El oscilador forzado está sometido a la fuerza elástica $\vec{F}_{elast} = -kx\ \hat{\imath}$, una fuerza de rozamiento viscoso $\vec{F}_{visc} = -r\dot{x}\ \hat{\imath}$, y una fuerza externa $\vec{F}_{ext} = F_o\ \cos\omega t\ \hat{\imath}$. La ecuación de movimiento es

$$\ddot{x} + \gamma\ \dot{x} + \omega_o^2\ x = \frac{F_o}{m}\ \mathrm{Re}[e^{i\omega t}]\ , \tag{9.11}$$

donde $\omega_o^2 = k/m$, $\gamma = r/m$. Como es una ecuación lineal inhomogénea, la solución general es la suma de la solución general de la ecuación homogénea más una solución particular:

$$x(t) = x_h(t) + x_p(t)$$

La solución homogénea es amortiguada, debido a la disipación producida por el rozamiento viscoso. Se propone la forma general $x_h(t) \propto e^{\lambda t}$, y se obtiene una ecuación cuadrática para λ:

$$\lambda^2 + \gamma\ \lambda + \omega_o^2\ = 0$$

Entonces las dos raíces (que darán dos soluciones homogéneas independientes) son

$$\lambda_{1,2} = -\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_o^2}$$

El segundo término puede ser real o imaginario. Si es real resulta $\lambda_{1,2} < 0$, lo que significa que la solución se amortigua. Si el segundo término es imaginario, la solución

oscila con la frecuencia $\sqrt{\omega_o^2 - \frac{\gamma^2}{4}}$ mientras se amortigua con el factor $e^{-\gamma t/2}$ (movimiento subamortiguado).[6] En ambos casos la solución homogénea se extingue a tiempos largos:

$$x_h(t) \xrightarrow[t\to\infty]{} 0$$

A tiempos largos sólo sobrevive la solución particular. Como los coeficientes de la ecuación de movimiento son reales, podremos buscar una solución particular compleja para la ecuación

$$\ddot{x} + \gamma\ \dot{x} + \omega_o^2\ x = \frac{F_o}{m}\ e^{i\omega t}\ ;$$

de esa forma la parte real de la solución particular será solución de la ecuación (9.11). Proponemos la solución compleja

$$x_p(t) = A\ e^{i\omega t}\ ,$$

donde $A \in \mathbb{C}$. Reemplazando en la ecuación,

$$-\omega^2\ A + i\ \gamma\ \omega\ A + \omega_o^2\ A = \frac{F_o}{m}$$

Entonces[7]

$$A = \frac{\frac{F_o}{m}}{\omega_o^2 - \omega^2 + i\ \gamma\ \omega}$$

Escribamos el complejo A en la forma $|A|\ e^{i\delta}$ para separar amplitud y fase. Por un lado es

$$|A| = \frac{\frac{F_o}{m}}{\sqrt{(\omega_o^2 - \omega^2)^2 + \gamma^2\ \omega^2}}$$

Por otro lado, multiplicando y dividiendo A por el conjugado del denominador resulta

$$A = \frac{\frac{F_o}{m}}{|\omega_o^2 - \omega^2 + i\ \gamma\ \omega|^2}\ (\omega_o^2 - \omega^2 - i\ \gamma\ \omega)$$

En la expresión para A vemos que $\mathrm{Im}[A] < 0$. Además, si $\omega \lesseqgtr \omega_o$ entonces $\mathrm{Re}[A] \gtreqless 0$. Esto significa que la fase de A está en el cuarto cuadrante si $\omega < \omega_o$, y pertenece al tercer cuadrante si $\omega > \omega_o$. En cualquier caso el movimiento atrasa respecto de la fuerza externa. La fase δ cumple que

$$\tan\delta = \frac{\mathrm{Im}[A]}{\mathrm{Re}[A]} = \frac{-\gamma\ \omega}{\omega_o^2 - \omega^2}$$

6. Si $\omega_o = \gamma/2$, entonces las dos raíces son iguales: $\lambda = -\gamma/2$. En ese caso la segunda solución independiente tiene la forma $t\ e^{-\gamma t/2}$. Este caso se denomina amortiguamiento crítico.

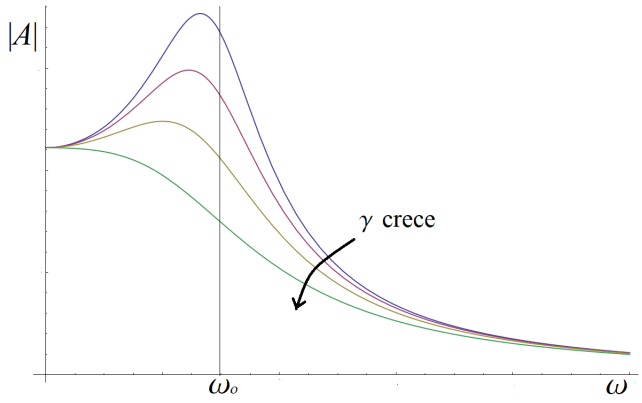
7. Si $\gamma = 0$ y $\omega = \omega_o$, entonces $x_p(t) = F_o t/(2m\omega_o)\ \sin[\omega_o t]$.

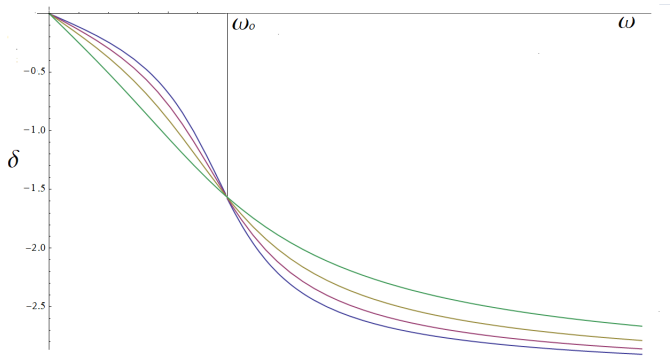
La solución real es

$$x_p(t) = \mathrm{Re}[A\ e^{i\omega t}] = |A|\ \cos(\omega t + \delta)$$

Si la frecuencia externa ω se elige apropiadamente en función de los parámetros del sistema entonces podremos maximizar la amplitud $|A|$:[8]

$$si \quad \omega^2 = \omega_o^2 - \frac{\gamma^2}{2} \quad \Rightarrow \quad |A| = |A|_{\max} = \frac{\frac{F_o}{m}}{\gamma\ \sqrt{\omega_o^2 - \frac{\gamma^2}{4}}}$$





El trabajo realizado por la fuerza externa en cada ciclo se disipa en forma de calor. La potencia media entregada es

$$< P > = \frac{1}{T} \int_0^T F_{ext}(t)\ \dot{x}_P(t)\ dt = \frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} F_o \cos(\omega t)\ |A|\,(-\omega) \sin(\omega t + \delta)\ dt = -\frac{1}{2}\ F_o\, \omega\ |A|\ \sin\delta$$

(recordemos que $\sin\delta < 0$). Entonces

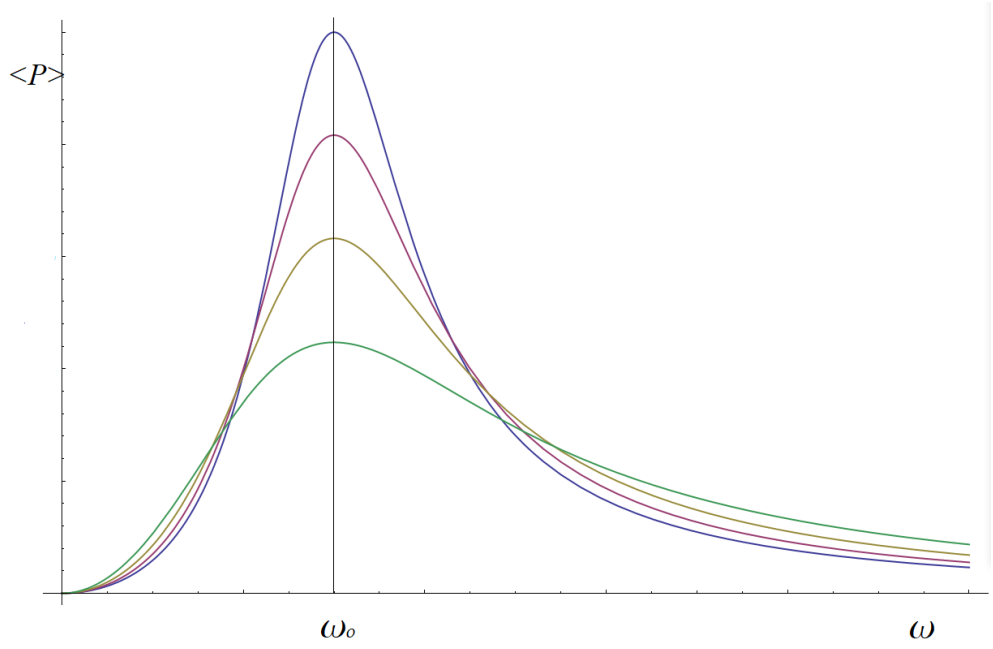
$$< P > = -\frac{1}{2}\ F_o\ \omega\ \ \mathrm{Im}[A] = \frac{F_o^2}{2m}\ \frac{\gamma}{\omega^2(\frac{\omega_o^2}{\omega^2} - 1)^2 + \gamma^2}$$

8. Si $\omega_o < \gamma/\sqrt{2}$, entonces $|A|$ es máximo en $\omega = 0$ (ver Figura).

La potencia media entregada por la fuerza externa al sistema se maximiza cuando la frecuencia externa ω coincide con la frecuencia propia ω_o (frecuencia de **resonancia**); en ese caso la potencia vale

$$< P >_{\max} = \frac{F_o^2}{2m\gamma} = \frac{F_o^2}{2\ r}$$

(en lenguaje de circuitos de corriente alterna, $F_o^2/2$ sería el cuadrado del valor *eficaz* de la fuerza externa).



10 Resonancia paramétrica

Consideremos un sistema oscilante

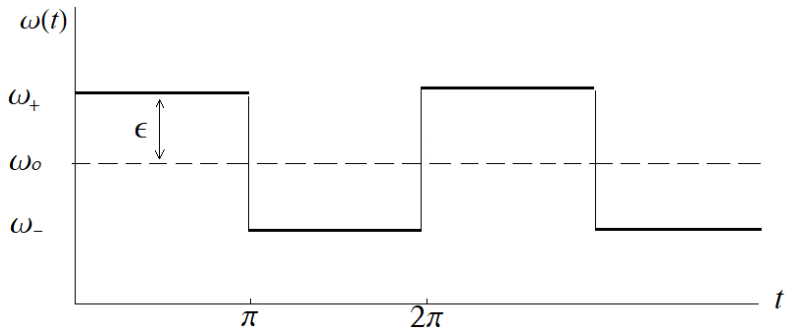
$$\ddot{x} + \omega^2 \; x = 0$$

cuya frecuencia posee una dependencia periódica en el tiempo:

$$\omega^2(t) = \omega_o^2 \; (1 + a(t))$$

donde $a(t)$ es periódica. Hay diferentes situaciones físicas que pueden corresponderse con este tipo de sistema. Por ejemplo las pequeñas oscilaciones de un péndulo cuando su hilo sufre, a su vez, variaciones periódicas de su longitud; como cuando una persona se hamaca sola, moviendo su cuerpo de tal forma que la distancia de su centro de masa al punto de suspensión cambia periódicamente. La llamada *ecuación de Mathieu* es del tipo antedicho, con $a(t) = \kappa \; \cos(2 \; t)$.

Aquí estudiaremos el caso sencillo donde $\omega(t)$ es una onda cuadrada



Llamaremos $T = 2\pi$ al período de la variación de la frecuencia propia (esto corresponde a elegir una unidad de tiempo). De esa forma el sistema es un oscilador armónico en cada semiperíodo de variación de la frecuencia ω, pues ω es constante en cada semiperíodo. Para hacer la transición de un semiperíodo al siguiente habrá que tener en cuenta que el estado final del primer semiperíodo se convierte en el estado inicial del siguiente semiperíodo.

Notemos que en este problema hay dos períodos que interactúan: el período $T = 2\pi$ de la variación de ω, y el período del oscilador $\tau = 2\pi/\omega_o$ en torno al cual se produce la variación. Queremos averiguar si aparecen resonancias entre estos dos períodos.

10.1 El oscilador armónico en el espacio de estados

A los efectos de preparar el abordaje del problema anterior, veamos algunos aspectos del sistema que vamos a perturbar. La solución de la ecuación de oscilador armónico es

$$x(t) = A\ \cos[\omega_o\ t + \delta] \qquad \Rightarrow \qquad v(t) = -A\ \omega_o\ \sin[\omega_o\ t + \delta]$$

Por lo tanto la posición y la velocidad en cada instante satisfacen la ecuación

$$x(t)^2 + \frac{v(t)^2}{\omega_o^2} = A^2$$

Como el par $x(t), v(t)$ caracteriza el estado del sistema, podemos ver a (x, v) como coordenadas en un "espacio de estados".[1] La ecuación anterior dice que el oscilador armónico evoluciona recorriendo una elipse en el espacio de estados; al cabo de un período $\tau = 2\pi/\omega_o$ el oscilador retorna al estado inicial. La elipse se transforma en una circunferencia si usamos las coordenadas $(x, \frac{v}{\omega_o})$. Otras coordenadas interesantes son (ξ, ξ^*) donde

$$\xi \equiv x - i\ \frac{v}{\omega_o}$$

En estas coordenadas la evolución del oscilador armónico está dada por

$$\xi(t) = A\ e^{i(\omega_o t + \delta)} = e^{i\omega_o t}\ \xi(0)$$

que es otra forma de ver que el estado del sistema a tiempo t resulta de rotar un ángulo $\omega_o t$ sobre la circunferencia a partir del estado inicial $\xi(0) = A\ e^{i\delta}$.

En relación al caso que trataremos a continuación, señalamos aquí que la estabilidad del oscilador armónico implica que éste permanece en una región acotada del espacio de estados.

10.2 Variación de la frecuencia propia

Cuando la frecuencia propia varía como una onda cuadrada, tendremos una evolución de oscilador armónico en cada semiperíodo de variación de la frecuencia ω. Para conectar dos semiperíodos contiguos, el estado final del primero de ellos deberá verse como el estado inicial del siguiente semiperíodo:

1) Si $0 < t < \pi$ la frecuencia angular es $\omega = \omega_+$. El estado final de semiperíodo es

$$\xi(\pi) = e^{i\omega_+ \pi}\ \xi(0)$$

1. Este espacio de estados es un remedo del espacio de las fases en el que se realiza el formalismo Hamiltoniano que veremos más adelante.

es decir

$$x(\pi) = \mathrm{Re}[\xi(\pi)] = \cos[\omega_+\ \pi]\ x(0) + \sin[\omega_+\ \pi]\ \frac{v(0)}{\omega_+}$$

$$\frac{v(\pi)}{\omega_+} = -\,\mathrm{Im}[\xi(\pi)] = -\sin[\omega_+\ \pi]\ x(0) + \cos[\omega_+\ \pi]\ \frac{v(0)}{\omega_+}$$

Vemos que hay una transformación lineal que lleva del estado en $t = 0$ al estado en $t = \pi$,

$$\begin{pmatrix} x(\pi) \\ v(\pi) \end{pmatrix} = \begin{pmatrix} \cos[\omega_+\ \pi] & \frac{\sin[\omega_+\ \pi]}{\omega_+} \\ -\omega_+\ \sin[\omega_+\ \pi] & \cos[\omega_+\ \pi] \end{pmatrix} \begin{pmatrix} x(0) \\ v(0) \end{pmatrix}$$

La matriz de transformación sería una rotación de ángulo $\omega_+\ \pi$ en el plano $(x, \frac{v}{\omega_+})$, como muestran las ecuaciones anteriores. Sin embargo preferiremos las coordenadas (x, v) porque no dependen de la frecuencia, que será distinta en el siguiente semiperíodo. Llamaremos M_+ a la matriz que transforma los datos iniciales en datos finales durante el semiperíodo de frecuencia angular ω_+:

$$\begin{pmatrix} x(\pi) \\ v(\pi) \end{pmatrix} = M_+ \begin{pmatrix} x(0) \\ v(0) \end{pmatrix}$$

2) En el semiperíodo siguiente, la matriz que "mapea" el estado inicial en el estado final se construye con la frecuencia angular ω_-. Por lo demás tiene la misma estructura que M_+. Llamaremos M_- a esta matriz. Entonces para el semiperíodo $\pi < t < 2\pi$ la relación entre los estados inicial y final es

$$\begin{pmatrix} x(2\pi) \\ v(2\pi) \end{pmatrix} = M_- \begin{pmatrix} x(\pi) \\ v(\pi) \end{pmatrix}$$

siendo

$$M_\pm \equiv \begin{pmatrix} \cos[\omega_\pm\ \pi] & \frac{\sin[\omega_\pm\ \pi]}{\omega_\pm} \\ -\omega_\pm\ \sin[\omega_\pm\ \pi] & \cos[\omega_\pm\ \pi] \end{pmatrix}$$

En resumen, en un período $T = 2\pi$ de variación de la frecuencia ω, el mapeo entre estado inicial y final viene dado por la matriz $M \equiv M_- M_+$,

$$\begin{pmatrix} x(2\pi) \\ v(2\pi) \end{pmatrix} = M_- \begin{pmatrix} x(\pi) \\ v(\pi) \end{pmatrix} = M_- M_+ \begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = M \begin{pmatrix} x(0) \\ v(0) \end{pmatrix}$$

Podemos entonces entender la evolución del sistema evaluando su estado a tiempos discretos $t = 2k\pi$. En cada período $T = 2\pi$ actúa la matriz M; en k períodos interviene la matriz M^k. Quiere decir que la matriz M contiene información valiosa sobre el carácter de la evolución. Trataremos de extraer de M las condiciones para que se produzcan resonancias. Más concretamente, buscaremos responder si la evolución gobernada por la matriz M deja al sistema en una región acotada del espacio de estados o lleva el sistema hacia regiones de amplitudes crecientes.

Una propiedad importante de la matriz M es que su determinante vale 1 debido a que $\det M_{\pm} = 1$,

$$\det M = \det M_{+} \ \det M_{-} = 1$$

Las transformaciones lineales en un plano que tienen determinante igual a 1 preservan áreas. En efecto, el área del paralelogramo determinado por dos vectores $\vec{A}$, $\vec{B}$, es igual a $\left|\vec{A} \times \vec{B}\right|$,

$$\acute{A}rea = \left|\vec{A} \times \vec{B}\right| = |A_x B_y - A_y B_x| = \left|\det \begin{pmatrix} A_x & B_x \\ A_y & B_y \end{pmatrix}\right|$$

Si aplicamos una transformación lineal Λ a los vectores columna $|A>$, $|B>$,

$$|A'> = \Lambda \ |A> \ , \qquad\qquad |B'> = \Lambda \ |B>$$

Entonces la matriz de los vectores $\vec{A}$, $\vec{B}$ cambia a

$$\begin{pmatrix} A'_x & B'_x \\ A'_y & B'_y \end{pmatrix} = \begin{pmatrix} \Lambda_{xx} & \Lambda_{xy} \\ \Lambda_{yx} & \Lambda_{yy} \end{pmatrix} \begin{pmatrix} A_x & B_x \\ A_y & B_y \end{pmatrix}$$

Tomando el determinante,

$$\acute{A}rea\ ' = \det \Lambda \ \acute{A}rea$$

Si $\det \Lambda = 1$, entonces el área es preservada. En relación a la evolución del sistema bajo consideración, como la matriz M tiene determinante igual a 1 podemos afirmar que los estados iniciales que pertenecen a una determinada región del plano (x, v) evolucionan hacia estados finales que pertenecen a otra región de la misma área.[2]

Pero la pregunta que nos interesa contestar es si la evolución será "estable", en el sentido que el sistema permanece en una región limitada del espacio de estados, o si por el contrario el sistema evoluciona hacia regiones cada vez más alejadas del estado inicial. Para ello nos valdremos de que toda matriz cuadrada –sea o no simétrica– es diagonalizable mediante una transformación de semejanza ($M \to S^{-1} M \ S$, donde S puede no ser ortogonal) siempre que la dimensión del espacio generado por sus autovectores coincida con la dimensión de la matriz; en particular, si los autovalores λ_k son todos diferentes. En tal caso [3]

$$M \longrightarrow \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Como los coeficientes del polinomio característico $P(\lambda) = \det(M - \lambda \ \mathbf{1})$ son reales, sus dos raíces λ_1, λ_2 deben ser reales o complejas conjugadas. Además, como $\det M = 1$, debe cumplirse que

$$\lambda_1 \ \lambda_2 = 1$$

2. Esta propiedad es un caso particular del teorema de Liouville que veremos más adelante.

3. En caso contrario, la matriz admitiría la forma canónica de Jordan $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Entonces hay dos posibilidades:

1) λ_1, λ_2 son complejos conjugados de módulo 1:

$$\lambda_1 = e^{i\alpha}, \qquad \lambda_2 = e^{-i\alpha}$$

Es sabido que estos son los autovalores de la matriz de rotación[4]

$$M = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

lo que significa que en alguna base (es decir, con alguna transformación lineal S de las coordenadas x, v) al cabo de k períodos $T = 2\pi$ el sistema se encontrará en un estado que resulta de rotar un ángulo $k\alpha$ sobre una circunferencia. En este caso el sistema es estable en el sentido que permanece en una región acotada del espacio de estados.

2) λ_1, λ_2 son reales e inversos uno del otro:

$$\lambda_1 = \pm e^{\alpha}, \qquad \lambda_2 = \pm e^{-\alpha}$$

Estos son los autovalores de una transformación hiperbólica[5]

$$M = \pm \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}$$

Mientras que las rotaciones mueven puntos sobre círculos, las transformaciones hiperbólicas mueven puntos sobre hipérbolas.[6] Por lo tanto, en alguna base (es decir, con alguna transformación lineal S de las coordenadas x, v) el sistema evoluciona avanzando sobre una hipérbola al cabo de cada período $T = 2\pi$, lo que lo lleva hacia regiones cada vez más alejadas en el espacio de estados.

---

4. La transformación de semejanza que diagonaliza la matriz de rotación es $S = \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$. A su vez, otra semejanza llevará la matriz de rotación a la forma de la matriz M obtenida en nuestra deducción.

5. La transformación hiperbólica es diagonalizada por la semejanza $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Una segunda semejanza llevará la transformación hiperbólica a la forma de la matriz M obtenida en nuestra deducción.

6. La transformación hiperbólica $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ implica que $x'^2 - y'^2 = x^2 - y^2$.

¿Cómo podemos saber si estamos en el primero o el segundo caso sin calcular los autovalores de $M = M_- M_+$? Vamos a usar que la traza de una matriz cuadrada es invariante ante transformaciones de semejanza ($Tr(S^{-1}M\ S) = Tr\ M$). Cuando la matriz está diagonalizada su traza es $\lambda_1 + \lambda_2$; la idea es comparar este valor con la traza de $M_- M_+$ Notemos que

1) $\lambda_1 = e^{i\alpha}, \quad \lambda_2 = e^{-i\alpha} \quad \Rightarrow \ \lambda_1 + \lambda_2 = 2\ \cos\alpha \quad \Rightarrow \ |Tr\ M| \leqslant 2$

2) $\lambda_1 = \pm e^{\alpha}, \quad \lambda_2 = \pm e^{-\alpha} \quad \Rightarrow \ \lambda_1 + \lambda_2 = \pm 2\ \cosh\alpha \quad \Rightarrow \ |Tr\ M| \geqslant 2$

Esto significa que debemos ver si $|Tr\ M_- M_+|$ es mayor o menor que 2 para así saber si el sistema es estable o resuena. El producto $M_- M_+$ da

$$M_- M_+ = \begin{pmatrix} \cos\omega_-\pi\cos\omega_+\pi - \frac{\omega_+}{\omega_-}\sin\omega_+\pi\sin\omega_-\pi & \ldots \\ \ldots & \cos\omega_+\pi\cos\omega_-\pi - \frac{\omega_-}{\omega_+}\sin\omega_-\pi\sin\omega_+\pi \end{pmatrix}$$

$$\Rightarrow \qquad Tr(M_- M_+) = 2\cos\omega_-\pi\cos\omega_+\pi - \left(\frac{\omega_+}{\omega_-} + \frac{\omega_-}{\omega_+}\right)\sin\omega_+\pi\sin\omega_-\pi$$

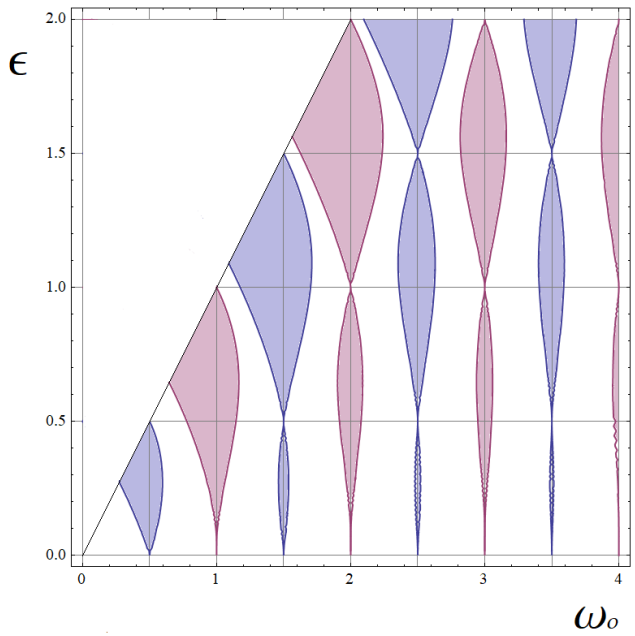
Entonces, si llamamos ω_o al promedio de $\omega_\pm$ (como en la Figura inicial), resulta

$$\omega_\pm = \omega_o \pm \varepsilon$$

Nos interesa saber para qué relación entre ω_o y ε será $|Tr(M_- M_+)| > 2$, y tendremos el fenómeno de **resonancia paramétrica**. Usaremos un programa de cálculo para dibujar las curvas $\varepsilon(\omega_o)$ tales que $Tr(M_- M_+) = \pm 2$. Esas curvas separarán las zonas estables de las de resonancia. Es decir que graficaremos

$$\pm 2 = 2\cos(\omega_o - \varepsilon)\pi\cos(\omega_o + \varepsilon)\pi - \left(\frac{\omega_o + \varepsilon}{\omega_o - \varepsilon} + \frac{\omega_o - \varepsilon}{\omega_o + \varepsilon}\right)\sin(\omega_o + \varepsilon)\pi\sin(\omega_o - \varepsilon)\pi$$

para $\varepsilon < \omega_o$ (pues ω_- debe ser positivo). La Figura muestra el resultado:



Las zonas azules corresponden a $Tr(M_-M_+) \leqslant -2$, y las zonas rojas indican $Tr(M_-M_+) \geqslant 2$. Ambas zonas sombreadas son zonas de resonancia paramétrica (naturalmente, las zonas donde $\varepsilon = 0$ deben ser zonas de estabilidad).[7] Las zonas de resonancia paramétrica están caracterizadas por los valores $\omega_o = k/2$, es decir

$$\frac{2\pi}{\tau} = \frac{k}{2}\,, \qquad k \in \mathbb{N}$$

Si recordamos que la unidad de tiempo fue elegida para que $T = 2\pi$, donde T es el período de la variación de la frecuencia propia del sistema, entonces la expresión anterior dice que

$$\frac{T}{\tau} = \frac{k}{2}\,, \qquad k \in \mathbb{N}$$

Es decir que la resonancia paramétrica aparece cuando la frecuencia propia del sistema oscila con un período T aproximadamente igual a

$$T = \frac{k}{2}\,\tau = \frac{k\pi}{\omega_o}$$

Bibliografía adicional

V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer.

7. En los puntos donde ε, ω_o son ambos enteros o semienteros, los valores de $\omega_\pm$ devienen en enteros de igual paridad o de paridad opuesta. Las $M_\pm$ resultan ser $\pm\mathbf{1}$, con signos iguales u opuestos respectivamente.

10.3 Ecuación de Mathieu

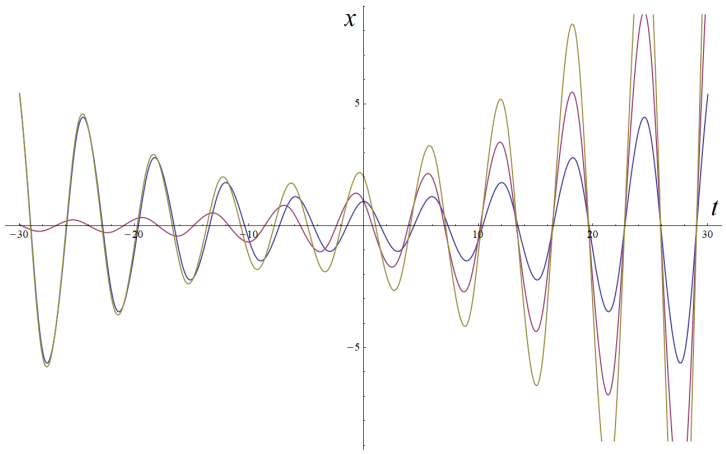
Veamos un ejemplo de resonancia paramétrica en la ecuación de Mathieu

$$\ddot{x} + \omega_o^2 \left[1 + \kappa \ \cos(2t)\right] x = 0$$

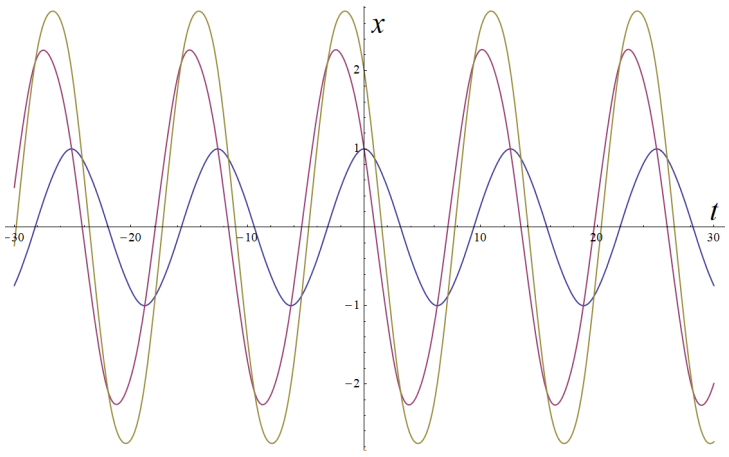
En esta ecuación el período de variación de la frecuencia es $T = \pi$. Para cumplir con la condición de resonancia elegimos $\omega_o = 1$. Vamos a resolver numéricamente la ecuación[8]

$$\ddot{x} + (1 + 0.3 \ \cos(2t)) \ x = 0$$

La siguiente Figura muestra soluciones con distintas condiciones iniciales.



Si en cambio elegimos $\omega_o = .5$, con las mismas condiciones iniciales obtendremos estabilidad:



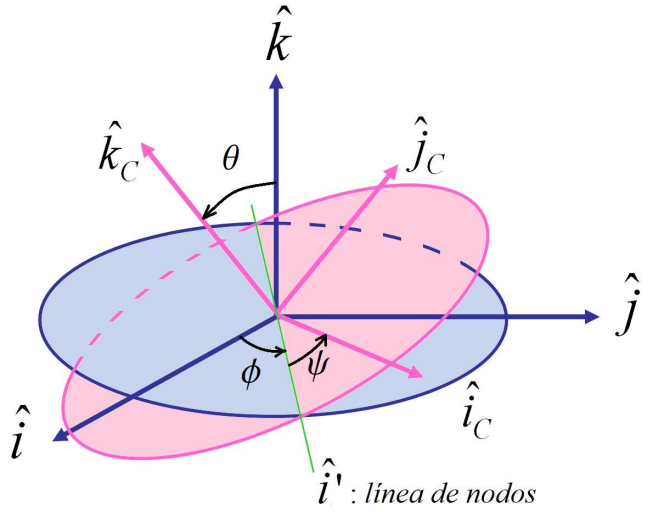
8. En *Mathematica* usar, por ejemplo:

```
NDSolve[{x"[t] + (1 + .3 Cos[2 t]) x[t] == 0, x[0] == 1, x'[0] == 0}, x[t], {t, -30, 30}]
  Plot[Evaluate[x[t] /. %], {t, -30, 30}]
```

11 Cinemática del cuerpo rígido

11.1 Ángulos de Euler

Un cuerpo rígido libre de vínculos tiene seis grados de libertad, que pueden descomponerse en tres grados de libertad traslatorios y tres rotatorios. Podemos usar tres coordenadas que describan la posición de un punto del cuerpo, por ejemplo su centro de masa, y otras tres que describan la orientación del cuerpo respecto de un sistema de referencia. Para este último fin, elijamos una terna $\hat{\imath}_c$, $\hat{\jmath}_c$, $\hat{k}_c$ fija al cuerpo. A medida que el cuerpo **rote**, esta terna cambiará su orientación en el espacio. Introduciremos los **ángulos de Euler** ϕ, θ, ψ para describir la orientación de la terna $\hat{\imath}_c$, $\hat{\jmath}_c$, $\hat{k}_c$.



En la Figura vemos un sistema de referencia en color azul, cuyos versores son $(\hat{\imath}, \hat{\jmath}, \hat{k})$. Por otro lado, en color rojo vemos la terna fija al cuerpo. La **línea de nodos** en color verde corresponde a la intersección de los planos x, y de ambas ternas.

Los ángulos de Euler se definen de la siguiente manera:

ϕ: es el ángulo entre $\hat{\imath}$ y la línea de nodos (que indicaremos con $\hat{\imath}\,'$)

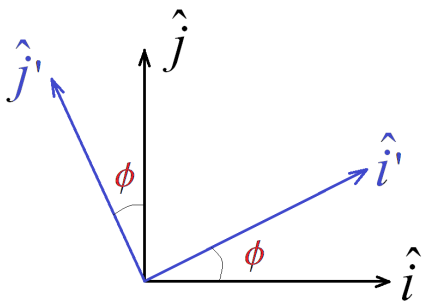
θ: es el ángulo entre $\hat{k}$ y $\hat{k}_c$

ψ: es el ángulo entre la línea de nodos y $\hat{\imath}_c$

Para familiarizarnos con los ángulos de Euler debemos pensar cómo debe girar el cuerpo para que cada uno de los ángulos varíe, y los otros dos permanezcan fijos. Si el cuerpo rota alrededor de $\hat{k}_c$, entonces ψ varía y ϕ, θ permanecen fijos. Para que varíe ϕ (es decir, para que se corra la línea de nodos) sin que varíen θ, ψ el cuerpo debe rotar alrededor de $\hat{k}$. Para que varíe θ sin afectar los otros ángulos, el cuerpo debe girar alrededor de la línea de nodos (es decir, alrededor de $\hat{\imath}'$). Dada una orientación $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ cualquiera, veremos que existe una **matriz de rotación** que conecta $(\hat{\imath}, \hat{\jmath}, \hat{k})$ con $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$.

11.2 Grupo de rotaciones

Repasemos la rotación en un plano



$$\begin{aligned} \hat{\imath} &= \cos\phi\ \hat{\imath}\,' - \sin\phi\ \hat{\jmath}\,' \\ \hat{\jmath} &= \sin\phi\ \hat{\imath}\,' + \cos\phi\ \hat{\jmath}\,' \end{aligned}$$

La transformación de los versores implica una transformación de las componentes de los vectores:

$$\vec{V} = V_x\ \hat{\imath} + V_y\ \hat{\jmath} = (V_x\ \cos\phi + V_y\ \sin\phi)\ \hat{\imath}\,' + (-V_x\ \sin\phi + V_y\ \cos\phi)\ \hat{\jmath}\,'$$

es decir,

$$\begin{aligned} V_x{}' &= V_x\ \cos\phi + V_y\ \sin\phi \\ V_y{}' &= -V_x\ \sin\phi + V_y\ \cos\phi \end{aligned}$$

o bien,

$$\begin{pmatrix} V_x{}' \\ V_y{}' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} V_x \\ V_y \end{pmatrix}$$

que escribiremos en notación compacta

$$|V>' = R_\phi \, |V>$$

Nótese que trasponer R_ϕ es igual a cambiar ϕ por $-\phi$, lo que significa revertir la rotación. Es decir, que $R_\phi^T = R_\phi^{-1}$.

En general las matrices $d \times d$ tales que

$$R^T \; R = \mathbf{1}$$

tienen la propiedad de dejar invariante el producto escalar euclidiano entre vectores d-dimensionales:

$$\begin{aligned} A_x{}' \; B_x{}' + A_y{}' \; B_y{}' + ... \; &= \; '< A|B >' = \; |A >'^{\,T} \;\; |B >' = \; (R \, |A >)^T \; R \, |B > \\ = \; |A >^T \;\; R^{\,T} \; R \, |B > = < A|R^{\,T} R|B > = < A|B > = \; A_x \; B_x + A_y \; B_y + ... \end{aligned}$$

Como los determinantes de una matriz y su traspuesta son iguales, la igualdad $R^T \; R = \mathbf{1}$ implica que

$$\det R = \pm 1$$

Si multiplicamos dos matrices tales que $R^T \; R = \mathbf{1}$ obtendremos otra matriz de ese mismo tipo; en efecto,

$$(R_1 R_2)^T \; (R_1 R_2) = R_2^T \; R_1^T \; R_1 R_2 = R_2^T \; \mathbf{1} \; R_2 = R_2^T \; R_2 = \mathbf{1}$$

Esta propiedad es esencial para que este tipo de matrices constituyan un **grupo**.[1]

1. Un *grupo* $\mathcal{G}$ es un conjunto de transformaciones dotado de una *c*omposición (en nuestro caso el producto de matrices) asociativa y *cerrada* (la composición de elementos de $\mathcal{G}$ está en $\mathcal{G}$), con elemento unidad y tal que cada elemento de $\mathcal{G}$ posee inversa en $\mathcal{G}$.

Las matrices $d \times d$ tales que $R^{\,T} = R^{-1}$ forman el **grupo ortogonal** $O(d)$. El subgrupo con $\det R = 1$ se denomina $SO(d)$ (*special orthogonal group*); si $d = 2$ ó 3 es el grupo de **rotaciones**. Las rotaciones no conmutan, salvo que sean rotaciones alrededor de un mismo eje. Las transformaciones de $O(d)$ con $\det R = -1$ implican cambios de *orientación de la base* de versores (en $d = 3$ la terna derecha cambia a terna izquierda).

11.3 Relación entre $(\hat{\imath}, \hat{\jmath}, \hat{k})$ y $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$

Como dijimos, existe una matriz de rotación de 3×3 que conecta $(\hat{\imath}, \hat{\jmath}, \hat{k})$ con $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$. Para construirla, la descompondremos en tres pasos:

1) $(\hat{\imath}, \hat{\jmath}, \hat{k}) \rightarrow (\hat{\imath}\,', \hat{\jmath}\,', \hat{k}\,')$

Comenzaremos por una rotación de ángulo ϕ alrededor de $\hat{k} = \hat{k}\,'$ para llevar $\hat{\imath}$ hasta la línea de nodos $\hat{\imath}\,'$. Es una rotación que actúa en el plano x, y de las ternas de partida y llegada. Entonces la correspondiente matriz de rotación es

$$R_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$|V>' = R_\phi\ |V>$$

2) $(\hat{\imath}\,', \hat{\jmath}\,', \hat{k}\,') \rightarrow (\hat{\imath}\,'', \hat{\jmath}\,'', \hat{k}\,'')$

Ahora realizaremos una rotación de ángulo θ alrededor de la línea de nodos $\hat{\imath}\,' = \hat{\imath}\,''$ para llevar $\hat{k}\,'$ hasta $\hat{k}\,'' = \hat{k}_c$. Es una rotación en los planos y, z de las ternas de partida y llegada. La matriz de rotación es

$$R_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

$$|V>'' = R_\theta\ |V>' = R_\theta\ R_\phi\ |V>$$

3) $(\hat{\imath}\,'', \hat{\jmath}\,'', \hat{k}\,'') \rightarrow (\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$

Finalmente haremos una rotación de ángulo ψ alrededor de $\hat{k}\,'' = \hat{k}_c$ para llevar $\hat{\imath}\,'', \hat{\jmath}\,''$ a coincidir con $\hat{\imath}_c, \hat{\jmath}_c$. Es una rotación en los planos x, y de las ternas de partida y llegada. La matriz de rotación es

$$R_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$|V>^c = R_\psi\ |V>'' = R_\psi\ R_\theta\ |V>' = R_\psi\ R_\theta\ R_\phi\ |V> \tag{11.1}$$

Por lo tanto, la rotación que transforma las componentes de un vector $\vec{V}$ en la terna $(\hat{\imath}, \hat{\jmath}, \hat{k})$ en las componentes del mismo $\vec{V}$ en la terna $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ es

$$R = R_\psi \; R_\theta \; R_\phi$$

donde ϕ, θ, ψ son los ángulos de Euler de la terna $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ relativos a la terna $(\hat{\imath}, \hat{\jmath}, \hat{k})$.

11.4 Velocidad de rotación

Decimos que un cuerpo rígido **rota** cuando cambia su **orientación** respecto del sistema de referencia. El cambio de la orientación está medido por las derivadas $\dot{\phi}$, $\dot{\theta}$, $\dot{\psi}$. Con ellas podemos definir un vector velocidad de rotación $\vec{\Omega}$. Para ello deberemos tener en cuenta cuál es el eje de rotación correspondiente a la variación de cada ángulo de Euler, según se vio en los tres pasos anteriores. Entonces

$$\vec{\Omega} = \dot{\phi} \; \hat{k} + \dot{\theta} \; \hat{\imath}\,' + \dot{\psi} \; \hat{k}_c \tag{11.2}$$

Tanto las velocidades generalizadas $\dot{\phi}$, $\dot{\theta}$, $\dot{\psi}$ como las direcciones de $\hat{\imath}\,'$, $\hat{k}_c$ pueden variar con el tiempo. De modo que la dirección de $\vec{\Omega}$ también lo hará, definiendo así un **eje instantáneo de rotación**.

Convendrá descomponer $\vec{\Omega}$ en una misma terna, ya sea la $(\hat{\imath}, \hat{\jmath}, \hat{k})$ o la $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$. Es muy fácil descomponer $\hat{\imath}\,'$ ya que la línea de nodos pertenece a los planos x, y de ambas ternas (ver la primera Figura):

$$\hat{\imath}\,' = \cos\phi \; \hat{\imath} + \sin\phi \; \hat{\jmath} \tag{11.3}$$

$$\hat{\imath}\,' = \cos\psi \; \hat{\imath}_c - \sin\psi \; \hat{\jmath}_c \tag{11.4}$$

En cambio, las proyecciones de $\hat{k}$ sobre los versores $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$, así como las proyecciones de $\hat{k}_c$ sobre los versores $(\hat{\imath}, \hat{\jmath}, \hat{k})$ no resultan tan evidentes. Por eso nos valdremos de la matriz R:

i) Las componentes de $\hat{k}$ en la base $(\hat{\imath}, \hat{\jmath}, \hat{k})$ son $(0, 0, 1)$. Para obtener sus componentes en la base $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ recurrimos a la ecuación (11.1):

$$|k>^c = \; R_\psi \; R_\theta \; R_\phi \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \; R_\psi \; R_\theta \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \; R_\psi \begin{pmatrix} 0 \\ \sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} \sin\psi \sin\theta \\ \cos\psi \sin\theta \\ \cos\theta \end{pmatrix}$$

que significa[2]

$$\hat{k} = \sin\psi \sin\theta \; \hat{\imath}_c + \cos\psi \sin\theta \; \hat{\jmath}_c + \cos\theta \; \hat{k}_c \tag{11.5}$$

ii) Para obtener $\hat{k}_c$ descompuesto en la base $(\hat{\imath}, \hat{\jmath}, \hat{k})$ partimos de

$$|k_c >^c = R_\psi \; R_\theta \; R_\phi \; |k_c >$$

En esta ecuación conocemos el lado izquierdo; las componentes de $\hat{k}_c$ en la base $(\hat{\imath}_c, \hat{\jmath}_c, \hat{k}_c)$ son $(0, 0, 1)$. Entonces

$$|k_c > = R_\phi^{-1} \; R_\theta^{-1} \; R_\psi^{-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = R_\phi^T \; R_\theta^T \; R_\psi^T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = R_\phi^T \; R_\theta^T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\phi \sin\theta \\ -\cos\phi \sin\theta \\ \cos\theta \end{pmatrix}$$

El resultado es[3]

$$\hat{k}_c = \sin\phi \sin\theta \; \hat{\imath} - \cos\phi \sin\theta \; \hat{\jmath} + \cos\theta \; \hat{k} \tag{11.6}$$

Nótese el intercambio de roles entre (11.5) y (11.6): $\phi \longleftrightarrow -\psi$, $\theta \longleftrightarrow -\theta$.

Tenemos dos formas de reescribir la velocidad de rotación (11.2):

I) Usamos (11.4) y (11.5) para escribir $\vec{\Omega}$ proyectada en la base del cuerpo:

$$\vec{\Omega} = \Omega_1 \; \hat{\imath}_c + \Omega_2 \; \hat{\jmath}_c + \Omega_3 \; \hat{k}_c$$

resultando

$$\boxed{\begin{array}{lcl} \Omega_1 & = & \dot{\phi} \; \sin\theta \; \sin\psi + \dot{\theta} \; \cos\psi \\ \Omega_2 & = & \dot{\phi} \; \sin\theta \; \cos\psi - \dot{\theta} \; \sin\psi \\ \Omega_3 & = & \dot{\phi} \; \cos\theta + \dot{\psi} \end{array}} \tag{11.7}$$

II) Usamos (11.3) y (11.6) para escribir $\vec{\Omega}$ proyectada en la base del sistema de referencia:

$$\vec{\Omega} = \Omega_x \; \hat{\imath} + \Omega_y \; \hat{\jmath} + \Omega_z \; \hat{k}$$

resultando

2. La proyección $\cos\theta$ sobre el versor $\hat{k}_c$ es evidente en la primera Figura. Las otras dos proyecciones se infieren de la ortogonalidad de $\hat{k}$ con respecto a la línea de nodos (11.4).

3. Se verifica la ortogonalidad respecto de la línea de nodos (11.3).

$$\boxed{\begin{array}{rcl} \Omega_x & = & \dot{\psi}\ \sin\theta\ \sin\phi + \dot{\theta}\ \cos\phi \\ \Omega_y & = & -\dot{\psi}\ \sin\theta\ \cos\phi + \dot{\theta}\ \sin\phi \\ \Omega_z & = & \dot{\psi}\cos\theta + \dot{\phi} \end{array}} \tag{11.8}$$

Nótese que podemos pasar de (11.7) a (11.8) mediante el cambio $\phi \longleftrightarrow -\psi$, $\theta \longleftrightarrow -\theta$, $t \longleftrightarrow -t$.

Ejemplo: en un movimiento plano, como la rotación de un plato alrededor del eje perpendicular al mismo, es posible hacer coincidir las direcciones de $\vec{\Omega}$, $\hat{k}$ y $\hat{k}_c$. En tal caso es $\theta = 0$, y $\vec{\Omega} = (\dot{\phi} + \dot{\psi})\ \hat{k} = (\dot{\phi} + \dot{\psi})\ \hat{k}_c$. Como $\hat{k} = \hat{k}_c$, la línea de nodos no queda determinada; $\phi + \psi$ no es más que el ángulo entre $\hat{\imath}$ y $\hat{\imath}_c$.

11.5 Cinemática del cuerpo rígido

En 1775 Euler demostró que dos configuraciones diferentes de un cuerpo rígido con un punto fijo (articulación) pueden ser siempre conectadas por medio de una única rotación alrededor de un eje que pasa por el punto fijo. La dirección de ese eje es la de aquellos vectores que son invariantes ante tal rotación. El teorema de Euler probaba que las rotaciones forman un grupo, en el sentido que las dos configuraciones podrían conectarse mediante pasos intermedios que también involucrarían rotaciones. Por otra parte, el teorema permitía ver el movimiento de un cuerpo rígido con un punto fijo como una sucesión de rotaciones infinitesimales alrededor de sucesivos ejes instantáneos de rotación.

Recordando que las velocidades de dos puntos O y P de un cuerpo rígido en un movimiento arbitrario siempre cumplen que

$$\vec{v}_P = \vec{v}_O + \vec{\Omega} \times (\vec{r}_P - \vec{r}_O)$$

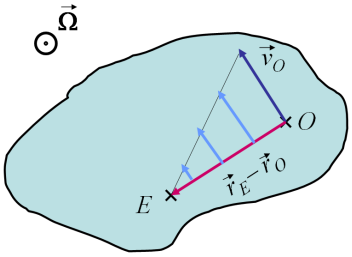
vemos que

$$\vec{\Omega} \cdot \vec{v}_P = \vec{\Omega} \cdot \vec{v}_O$$

de donde concluimos que si un punto O tiene velocidad perpendicular a $\vec{\Omega}$ entonces todos los puntos del cuerpo tienen velocidad perpendicular a $\vec{\Omega}$. En ese caso podemos encontrar un punto E tal que $\vec{v}_E = 0$. En efecto, la posición de E debe cumplir que

$$0 = \vec{v}_O + \vec{\Omega} \times (\vec{r}_E - \vec{r}_O)\ ;$$

como $\vec{v}_O \perp \vec{\Omega}$, la ecuación para $\vec{r}_E$ tiene solución. Una posible solución es tal que $\vec{r}_E - \vec{r}_O \perp \vec{v}_O$.[4] En la Figura vemos una forma de resolver, donde $|\vec{r}_E - \vec{r}_O| = v_O/\Omega$. Pero la solución no es única; cualquier otro punto perteneciente a la recta paralela a $\vec{\Omega}$



que pasa por E también tiene velocidad nula (en el instante considerado). Esa recta es el **eje instantáneo de rotación**.

Si $\vec{\Omega} \cdot \vec{v}_O \neq 0$, entonces podríamos pasar al caso anterior mediante un cambio de sistema de referencia. El nuevo sistema $S\,'$ debería trasladarse con velocidad $\vec{V}$ paralela a $\vec{\Omega}$ para absorber la componente de $\vec{v}_O$ paralela a $\vec{\Omega}$; es decir, $V = \vec{\Omega} \cdot \vec{v}_O\,/\Omega$. Esto enseña que el movimiento más general de un cuerpo rígido se descompone en cada instante en una rotación alrededor del eje instantáneo y una traslación a lo largo del mismo eje (movimiento tipo *tornillo*). Este es el contenido del **Teorema de Chasles** ($\sim$ 1830).

4. El punto E podría estar fuera del cuerpo real; aun así debe ser visto como solidario con el movimiento del cuerpo real.

12 Tensor de inercia. Ecuaciones de Euler

12.1 Momento angular del cuerpo rígido

La velocidad de los puntos de un rígido se escribe $\vec{v}_i = \vec{v}_O + \vec{\Omega} \times (\vec{r}_i - \vec{r}_O)$. Calculemos el momento angular del cuerpo rígido respecto de un punto O que pertenece al mismo:

$$\begin{aligned}
\vec{L}_O &= \sum_i m_i \, (\vec{r}_i - \vec{r}_O) \times \vec{v}_i = \sum_i m_i \, (\vec{r}_i - \vec{r}_O) \times \Big(\vec{v}_O + \vec{\Omega} \times (\vec{r}_i - \vec{r}_O)\Big) \\
&= M \, (\vec{R}_{CM} - \vec{r}_O) \times \vec{v}_O + \sum_i m_i \, \vec{r}_{iO} \times (\vec{\Omega} \times \vec{r}_{iO})
\end{aligned}$$

El primer término se anula si O está fijo o si $O \equiv CM$. En el segundo término preferiremos ver al cuerpo como una distribución continua de materia:

$$\vec{L}_O = \int \delta m \;\; \vec{r} \times (\vec{\Omega} \times \vec{r}) \;\; , \qquad O \text{ es un punto fijo del cuerpo o su } CM \tag{12.1}$$

donde $\vec{r}$ es la posición de δm relativa a O. El doble producto vectorial cumple que

$$\vec{r} \times (\vec{\Omega} \times \vec{r}) = \vec{\Omega} \, r^2 - \vec{r} \, (\vec{\Omega} \cdot \vec{r})$$

En particular, la componente x de $\vec{L}_O$ es

$$\begin{aligned}
L_{O\;x} &= \int \delta m \;\; \big[\Omega_x \, (x^2 + y^2 + z^2) - x \, (\Omega_x \, x + \Omega_y \, y + \Omega_z \, z)\big] \\
&= \int \delta m \;\; \big[\Omega_x \, (y^2 + z^2) - \Omega_y \, x \, y - \Omega_z \, x \, z\big]
\end{aligned}$$

y del mismo modo con las restantes componentes de $\vec{L}_O$. En suma, hay una relación lineal entre las componentes de $\vec{L}_O$ y las de $\vec{\Omega}$, pero $\vec{L}_O$ y $\vec{\Omega}$ no son necesariamente paralelos.

12.2 Tensor de inercia

La relación lineal entre $\vec{L}_O$ y $\vec{\Omega}$ se puede expresar en forma matricial:

$$|L_O> = \mathbf{I}_O\ |\Omega> , \qquad O \text{ es un punto fijo del cuerpo o su } CM \tag{12.2}$$

donde $\mathbf{I}_O$ es el **tensor de inercia**, que contiene información sobre la distribución de la masa del cuerpo,

$$\mathbf{I}_O = \begin{pmatrix} \int \delta m\ (y^2+z^2) & -\int \delta m\ xy & -\int \delta m\ xz \\ -\int \delta m\ xy & \int \delta m\ (x^2+z^2) & -\int \delta m\ yz \\ -\int \delta m\ xz & -\int \delta m\ yz & \int \delta m\ (x^2+y^2) \end{pmatrix}$$

Las componentes del tensor de inercia son

$$I_{O\ \mu\nu} = \int \delta m\ (r^2\ \delta_{\mu\nu} - x_\mu\ x_\nu)$$

donde $r^2 = x^2+y^2+z^2$. El elemento de masa δm se expresa en términos de la densidad de masa $\rho(\vec{r})$ como $\delta m = \rho(\vec{r})\ d\mathcal{V}$; así la integración se realiza sobre el volumen $\mathcal{V}$ del cuerpo.

Propiedades

El tensor de inercia es simétrico ($I_{O\ \mu\nu} = I_{O\ \nu\mu}$), y es aditivo (el tensor de inercia de un cuerpo se puede descomponer en la suma de los tensores de inercia de sus partes).

Las componentes diagonales son positivas y se llaman **momentos de inercia**. Son iguales a la suma de los elementos de masa δm multiplicados por los cuadrados de sus respectivas distancias a los ejes cartesianos que pasan por O (recordemos que x, y, z vienen de las componentes cartesianas de $\vec{r}_{iO}$). Así, por ejemplo, y^2+z^2 en $I_{O\ xx}$ no es más que el cuadrado de la distancia de δm al eje x que pasa por O. Cada momento de inercia no puede superar la suma de los otros dos:[1]

$$I_{O\ xx} + I_{O\ yy} = \int \delta m\ (x^2+y^2+2z^2) \geqslant \int \delta m\ (x^2+y^2) = I_{O\ zz}$$

Las componentes fuera de la diagonal se llaman **productos de inercia**. Si la distribución de masa es simétrica respecto de un plano cartesiano que contiene a O, digamos el plano $x-y$, esto significa que por cada elemento de masa δm ubicado en (x, y, z) habrá un elemento de igual masa ubicado en $(x, y, -z)$ (recordemos que éstas son las posiciones medidas desde O). Esto anulará las sumas $\sum_i m_i\ x_i\ z_i$ y $\sum_i m_i\ y_i\ z_i$; es decir que se obtendrá $I_{O\ xz} = 0 = I_{O\ yz}$.

1. Si todas las partículas del cuerpo pertenecen al plano $z = 0$, entonces vale que $I_{O\ xx}+I_{O\ yy} = I_{O\ zz}$. Si todas las partículas se sitúan sobre el eje z, entonces $I_{O\ xx} = I_{O\ yy} = \int \delta m\ z^2$, $I_{O\ zz} = 0$; el cuerpo se denomina *rotador*, y por ser un segmento recto sólo tiene dos grados de libertad de rotación.

12.3 Ejes principales de inercia

Como sucede con los vectores, las componentes de los tensores dependen de la orientación de los ejes cartesianos. En el caso del tensor de inercia, un cambio en la orientación de los ejes cartesianos modifica los valores de las componentes x, y, z de los vectores $\vec{r}_{iO}$, y con ello modifican las componentes de $\mathbf{I}_O$. En particular se pueden elegir ejes cartesianos tales que $\mathbf{I}_O$ resulte diagonal. En efecto, como $\mathbf{I}_O$ es simétrico entonces posee autovectores ortogonales. Si los ejes cartesianos se eligen en la dirección de los autovectores, entonces $\mathbf{I}_O$ quedará diagonalizado. Los ejes cartesianos que diagonalizan el tensor de inercia se llaman **ejes principales de inercia**.

La existencia de simetría alivia o directamente evita resolver el problema de autovectores para diagonalizar el tensor de inercia. Cuando la distribución de masa posee un plano de simetría que contiene a O, entonces uno de los ejes principales de $\mathbf{I}_O$ es perpendicular al plano de simetría. En efecto, ya vimos que si $x - y$ es el plano de simetría entonces resulta $I_{O\ xz} = 0 = I_{O\ yz}$; la nulidad de las componentes no diagonales del sector z dice que el eje z es un eje principal de $\mathbf{I}_O$. Así quedará por diagonalizar el bloque $x - y$ de $\mathbf{I}_O$. Debe notarse que el CM está contenido en todo plano de simetría. Por lo tanto, en el ejemplo anterior el eje z será también un eje principal de $\mathbf{I}_{CM}$. Si hubiera un segundo plano de simetría tendríamos entonces un segundo eje principal para $\mathbf{I}_{CM}$, y el tercer eje sería perpendicular a los dos anteriores. Veremos enseguida que si O pertenece a un eje principal de $\mathbf{I}_{CM}$, entonces $\mathbf{I}_O$ y $\mathbf{I}_{CM}$ comparten ejes principales (ver Teorema de Steiner). Pero en un caso más general, los ejes principales de $\mathbf{I}_O$ tendrán direcciones diferentes de los de $\mathbf{I}_{CM}$. Cuando un cuerpo tiene simetría de revolución, entonces el eje de la simetría es eje principal de $\mathbf{I}_{CM}$, y los otros ejes son dos ejes perpendiculares cualesquiera en el plano perpendicular al eje de simetría (hay degeneración).

Los autovalores I_1, I_2, I_3 del tensor de inercia se denominan **momentos principales de inercia**. El cuerpo se llama **trompo asimétrico** si $I_1 \neq I_2 \neq I_3$, **trompo simétrico** si $I_1 = I_2 \neq I_3$, y **trompo esférico** si $I_1 = I_2 = I_3$. En el último caso la degeneración es completa, y cualquier dirección es un eje principal.

A continuación vemos una lista de momentos principales de inercia de $\mathbf{I}_{CM}$ para cuerpos de uso frecuente:

esfera (radio R): $I_1 = I_2 = I_3 = \frac{2}{5}MR^2$

cubo (lado L): $I_1 = I_2 = I_3 = \frac{1}{6}ML^2$

cilindro (radio R, longitud L): $I_1 = I_2 = \frac{M}{4}\left(R^2 + \frac{L^2}{3}\right), \quad I_3 = \frac{1}{2}MR^2$

cono (base de radio R, altura L): $I_1 = I_2 = \frac{3M}{20}\left(R^2 + \frac{L^2}{4}\right), \; I_3 = \frac{3}{10}MR^2$

12.4 Teorema de Steiner

$\mathbf{I}_O$ no sólo es sensible a la orientación de los ejes cartesianos; también lo es a la elección de O. Veamos cuál es la relación entre $\mathbf{I}_O$ y $\mathbf{I}_{CM}$ (a igual orientación de los ejes cartesianos). Sea

$\vec{a}$ la posición de O respecto del CM

$\vec{r} = \vec{r}_{i\ O}$ (como hasta aquí)

$\vec{r}\,' = \vec{r}_{i\ CM}$

Entonces $\vec{r} = \vec{r}\,' - \vec{a}$. Si reemplazamos en $\vec{r} \times (\vec{\Omega} \times \vec{r})$, que es la expresión que da lugar a $\mathbf{I}_O$,

$$\begin{aligned} \vec{r} \times (\vec{\Omega} \times \vec{r}) &= (\vec{r}\,' - \vec{a}) \times \Big(\vec{\Omega} \times (\vec{r}\,' - \vec{a})\Big) \\ &= \vec{r}\,' \times (\vec{\Omega} \times \vec{r}\,') - \vec{a} \times (\vec{\Omega} \times \vec{r}\,') - \vec{r}\,' \times (\vec{\Omega} \times \vec{a}) + \vec{a} \times (\vec{\Omega} \times \vec{a}) \end{aligned}$$

Los términos segundo y tercero no contribuyen al tensor de inercia pues

$$\int \delta m\ \vec{r}\,' = 0$$

(los $\vec{r}\,'$ están medidos respecto del CM; por lo tanto la integral resulta igual a M multiplicado por la posición del CM respecto de CM, que es cero). Por otro lado, el primer término da $\mathbf{I}_{CM}$ y el cuarto término es

$$\vec{a} \times (\vec{\Omega} \times \vec{a}) = \vec{\Omega}\ a^2 - \vec{a}\ (\vec{\Omega} \cdot \vec{a})$$

que tiene la estructura ya estudiada. Entonces obtenemos el **Teorema de Steiner**:

$$I_{O\ \mu\nu} = I_{CM\ \mu\nu} + M\ \left(a^2 \delta_{\mu\nu} - a_\mu a_\nu\right)$$

donde usamos que $\int \delta m = M$. El teorema dice que $\mathbf{I}_O$ se descompone en la suma de $\mathbf{I}_{CM}$ más el tensor de inercia de una partícula de masa M ubicada en O (pues $\vec{a}$ es el vector que une el CM con O).

Si los ejes elegidos son los ejes principales de $\mathbf{I}_{CM}$ entonces $\mathbf{I}_O$ permanecerá diagonal siempre que $a_\mu a_\nu$ lo sea; es decir, a_μ debería tener una única componente no nula, lo que significa que O debería estar sobre uno de los ejes principales. En otro caso, los ejes principales de $\mathbf{I}_O$ no coinciden con los de $\mathbf{I}_{CM}$.

12.5 Descomposición del momento angular en ejes principales

La relación lineal (12.2) entre el momento angular $\vec{L}_O$ y la velocidad de rotación $\vec{\Omega}$, cuando O es un punto fijo del cuerpo o su CM, puede ser descompuesta en cualquier

terna de ejes, independientemente del sistema de referencia donde se miden $\vec{L}_O$ y $\vec{\Omega}$. Si elegimos la terna de los ejes principales del cuerpo, que llamaremos $\hat{e}_1, \hat{e}_2, \hat{e}_3$, entonces la ecuación (12.2) queda

$$\begin{pmatrix} L_{O\ 1} \\ L_{O\ 2} \\ L_{O\ 3} \end{pmatrix} = \begin{pmatrix} I_{O1} & 0 & 0 \\ 0 & I_{O2} & 0 \\ 0 & 0 & I_{O3} \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix}$$

o bien

$$\vec{L}_O = I_{O1}\ \Omega_1\ \hat{e}_1 + I_{O2}\ \Omega_2\ \hat{e}_2 + I_{O3}\ \Omega_3\ \hat{e}_3 \tag{12.3}$$

donde O es un punto fijo del cuerpo o su CM. $\vec{L}_O$ y $\vec{\Omega}$ son paralelos sólo si tienen la dirección de un eje principal (es decir, si $\vec{\Omega}$ es autovector de $\mathbf{I}_O$). Por ejemplo, si tienen la dirección de $\hat{e}_1$ entonces es $\vec{L}_O = I_{O1}\ \vec{\Omega}$.

12.6 Dinámica del cuerpo rígido. Ecuaciones de Euler

Hemos visto que para todo sistema de partículas que interactúen con fuerzas $\vec{f}_{ij} \parallel \vec{r}_{ij}$ vale que

$$\vec{N}_O^{ext} = \frac{d\vec{L}_O}{dt}$$

donde O es un punto fijo del sistema de referencia inercial S o el CM del sistema de partículas. Un sistema S_c fijo al cuerpo rota con velocidad $\vec{\Omega}$ respecto de S. Entonces, el teorema de la derivada relativa nos permite escribir

$$\vec{N}_O^{ext} = \frac{d\vec{L}_O}{dt}|_{S_c} + \vec{\Omega} \times \vec{L}_O \tag{12.4}$$

En el sistema fijo al cuerpo la distribución de masa no cambia; por lo tanto no cambian los momentos principales de inercia ni los ejes principales. Entonces, usando (12.3),

$$\frac{d\vec{L}_O}{dt}|_{S_c} = I_{O1}\ \dot{\Omega}_1\ \hat{e}_1 + I_{O2}\ \dot{\Omega}_2\ \hat{e}_2 + I_{O3}\ \dot{\Omega}_3\ \hat{e}_3$$

que vale cuando O es un punto fijo del cuerpo o su CM. En suma, descomponiendo la ecuación (12.4) en los ejes principales del cuerpo obtenemos las **ecuaciones de Euler**

$$\boxed{\begin{array}{rcl} N_{O\ 1}^{ext} & = & I_{O1}\ \dot{\Omega}_1 + (I_{O3} - I_{O2})\ \Omega_2\ \Omega_3 \\ N_{O\ 2}^{ext} & = & I_{O2}\ \dot{\Omega}_2 + (I_{O1} - I_{O3})\ \Omega_1\ \Omega_3 \\ N_{O\ 3}^{ext} & = & I_{O3}\ \dot{\Omega}_3 + (I_{O2} - I_{O1})\ \Omega_1\ \Omega_2 \end{array}} \tag{12.5}$$

válidas cuando O es el CM o es un punto del cuerpo que está fijo en el sistema de referencia inercial S. Estas ecuaciones gobiernan la rotación del cuerpo rígido.

12.7 Energía cinética del cuerpo rígido

$$
\begin{aligned}
T &= \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 = \frac{1}{2}\sum_{i=1}^{N} m_i \left|\vec{v}_O + \vec{\Omega}\times\vec{r}_{iO}\right|^2 \\
&= \frac{1}{2}\, M\, v_O^2 + \sum_{i=1}^{N} m_i\, \vec{v}_O\cdot(\vec{\Omega}\times\vec{r}_{iO}) + \frac{1}{2}\sum_{i=1}^{N} m_i \left|\vec{\Omega}\times\vec{r}_{iO}\right|^2
\end{aligned}
$$

El segundo término se anula si O es un punto fijo del cuerpo o es su CM (pues $\sum m_i\ \vec{r}_{iCM} = 0$). Para el último término usaremos la propiedad cíclica del producto mixto,

$$
\left|\vec{\Omega}\times\vec{r}_{iO}\right|^2 = (\vec{\Omega}\times\vec{r}_{iO})\cdot(\vec{\Omega}\times\vec{r}_{iO}) = \vec{\Omega}\cdot[\vec{r}_{iO}\times(\vec{\Omega}\times\vec{r}_{iO})]
$$

donde reconocemos el factor $\vec{r}\times(\vec{\Omega}\times\vec{r})$ que determina el momento angular (12.1) cuando O es un punto fijo del cuerpo o es su CM. En tal caso la energía cinética resulta

$$
T = \frac{1}{2}\, M\, v_O^2 + \frac{1}{2}\,\vec{\Omega}\cdot\vec{L}_O = \frac{1}{2}\, M\, v_O^2 + \frac{1}{2}\ <\Omega|\ \mathbf{I}_O\ |\Omega>
$$

Vemos que la energía cinética de rotación se escribe en términos de las proyecciones de $\vec{\Omega}$ sobre los ejes principales como

$$
T_{rot} = \frac{1}{2}\, I_{O1}\, \Omega_1^2 + \frac{1}{2}\, I_{O2}\, \Omega_2^2 + \frac{1}{2}\, I_{O3}\, \Omega_3^2
$$

que se puede expresar en función de los ángulos de Euler y sus derivadas.

12.8 Formalismo Lagrangiano

Las ecuaciones de Euler pueden obtenerse también dentro del formalismo Lagrangiano. Sin embargo la comparación entre los sistemas de ecuaciones respectivos se complica porque las fuerzas generalizadas asociadas a los ángulos de Euler, los torques Q_ϕ, Q_θ, Q_ψ, no se relacionan en forma directa con las componentes $N_{O\ 1}^{ext}, N_{O\ 2}^{ext}, N_{O\ 3}^{ext}$ del torque. Esto se debe a que ϕ, θ no corresponden a rotaciones en torno a la terna fija al cuerpo. Sin embargo ψ representa una rotación en torno a $\hat{k}_C = \hat{e}_3$; de modo que $N_{O\ 3}^{ext} = Q_\psi$.[2] Entonces la ecuación de Lagrange para ψ,

$$
\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{\psi}}\right) - \frac{\partial T}{\partial\psi} = Q_\psi
$$

2. Como se vio en §2.1, en un cuerpo rígido sólo las fuerzas externas contribuyen a las fuerzas generalizadas.

debería coincidir con la tercera ecuación de Euler (y las otras dos ecuaciones de Euler se obtendrían permutando ejes). Veamos que efectivamente es así; usando (11.7) tenemos que por un lado es

$$\frac{\partial T_{rot}}{\partial \dot{\psi}} = I_{O3}\ \Omega_3\ \frac{\partial \Omega_3}{\partial \dot{\psi}} = I_{O3}\ \Omega_3$$

y por otro lado es

$$\frac{\partial T_{rot}}{\partial \psi} = I_{O1}\ \Omega_1\ \frac{\partial \Omega_1}{\partial \psi} + I_{O2}\ \Omega_2\ \frac{\partial \Omega_2}{\partial \psi} = I_{O1}\ \Omega_1\ \Omega_2 - I_{O2}\ \Omega_2\ \Omega_1$$

Entonces la ecuación de Lagrange para ψ da lo que esperábamos:

$$I_{O3}\ \dot{\Omega}_3 + (I_{O2} - I_{O1})\ \Omega_1\ \Omega_2 = Q_\psi.$$

13 Trompo simétrico

13.1 Trompo asimétrico libre de torques

Comenzaremos utilizando las ecuaciones de Euler para estudiar cuerpos libres de torque. Recordemos que esas ecuaciones son válidas cuando O es el CM o es un punto del cuerpo que permanece fijo en el sistema de referencia inercial. Pediremos entonces que se anule el torque $\vec{N}_O^{ext}$ respecto de un centro de momentos O con esas características. Por ejemplo, si despreciamos la interacción con el aire, podemos afirmar que un cuerpo lanzado por el aire está actuado sólo por la fuerza peso; en tal caso se anula el torque respecto del CM.[1]

Si un cuerpo libre de torques rota inicialmente sobre un eje principal, digamos el eje 3 (es decir, si $\Omega_1(0) = 0 = \Omega_2(0)$), tendremos que las ecuaciones de Euler dicen que inicialmente se cumple que

$$0 = \dot{\Omega}_1(0) \ , \qquad 0 = \dot{\Omega}_2(0) \ , \qquad 0 = \dot{\Omega}_3(0) \ ;$$

por lo tanto el cuerpo seguirá rotando sobre el eje 3 con velocidad constante Ω_3, conservando el momento angular $\vec{L}_O = I_{O3}\ \Omega_3\ \hat{e}_3$. Nos interesa saber si esta situación es estable o no. Para ello perturbaremos Ω_1, Ω_2,

$$\Omega_1 = \varepsilon_1(t) \ , \qquad \Omega_2 = \varepsilon_2(t) \ ,$$

y desarrollaremos las ecuaciones de Euler a primer orden en las perturbaciones:

$$\begin{aligned} 0 &= I_{O1}\ \dot{\varepsilon}_1 + (I_{O3} - I_{O2})\ \varepsilon_2\ \Omega_3 \\ 0 &= I_{O2}\ \dot{\varepsilon}_2 + (I_{O1} - I_{O3})\ \varepsilon_1\ \Omega_3 \\ 0 &\simeq I_{O3}\ \dot{\Omega}_3 \end{aligned}$$

La tercera ecuación dice que Ω_3 es constante a primer orden en las perturbaciones. Para desacoplar las ecuaciones para $\varepsilon_1(t)$ y $\varepsilon_2(t)$, derivamos la primera y reemplazamos la segunda:

$$0 = \ddot{\varepsilon}_1 + \frac{(I_{O3} - I_{O2})(I_{O3} - I_{O1})\ \Omega_3^2}{I_{O1}\ I_{O2}}\ \varepsilon_1$$

1. El torque del peso respecto de un punto O es $\vec{N}_O^{ext} = \int \vec{r} \times \delta m\ \vec{g} = [\int \delta m\ \vec{r}\,] \times \vec{g} = M\ \vec{R}_{CM-O} \times \vec{g}$, pues los $\vec{r}$ son posiciones relativas a O. El resultado permite ver al peso $M\vec{g}$ como una fuerza aplicada en el CM. Sin embargo, esto sólo es correcto bajo la aproximación de campo gravitatorio uniforme.

Vemos que $\varepsilon_1(t)$ oscila (situación estable) si

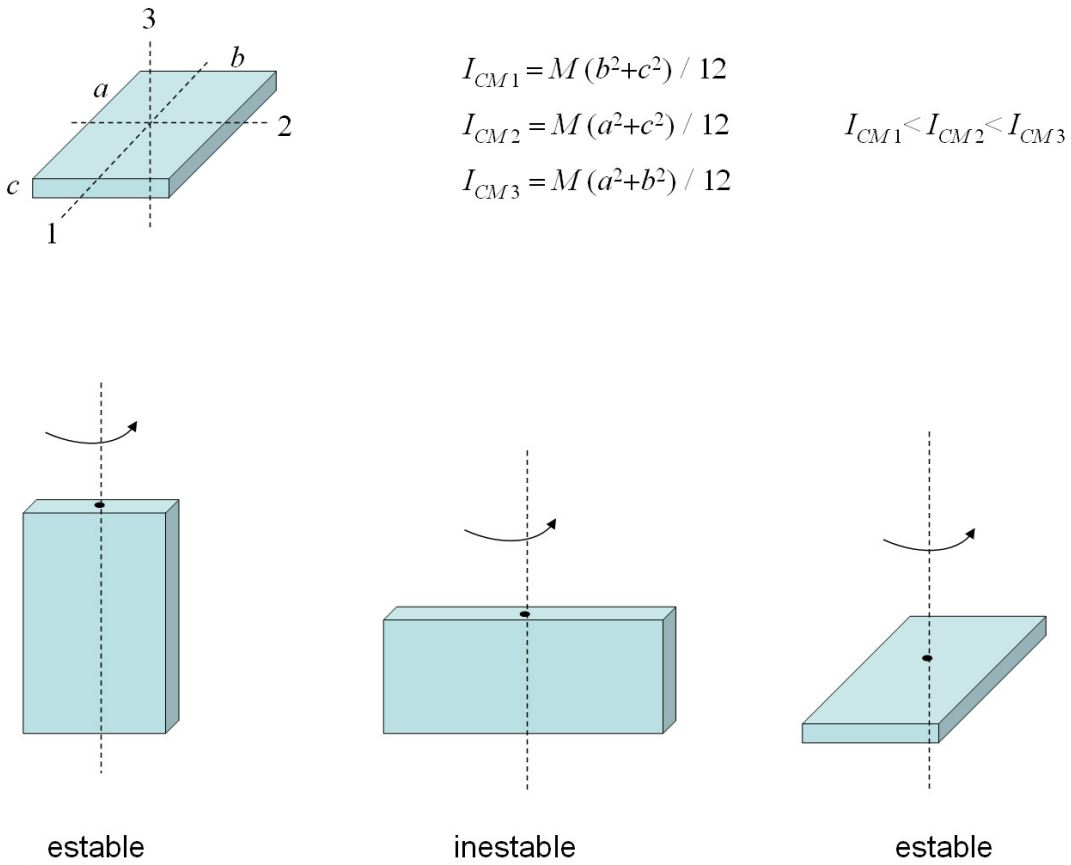
$$(I_{O3} - I_{O2})(I_{O3} - I_{O1}) > 0 \tag{13.1}$$

Lo mismo sucede con $\varepsilon_2(t)$. En ese caso ambas perturbaciones oscilan con una frecuencia dada por

$$\omega^2 = \frac{(I_{O3} - I_{O2})(I_{O3} - I_{O1})\ \Omega_3^2}{I_{O1}\ I_{O2}}$$

En caso contrario tendríamos un crecimiento exponencial de las perturbaciones (mientras valgan las aproximaciones realizadas).

El resultado (13.1) dice que en el trompo simétrico (por ejemplo, una pelota de rugby) la rotación vecina al eje de simetría es siempre estable. En el caso general de un trompo asimétrico hay dos casos estables y uno inestable. Es inestable la rotación alrededor del eje correspondiente al momento principal de inercia de valor intermedio:



Videos

Inestabilidad de la rotación alrededor del "eje intermedio"

https://www.youtube.com/watch?v=gkDHp7I7Cgg

https://www.youtube.com/watch?v=-2cMmwIKTJM

13.2 Trompo simétrico libre de torques

El problema anterior tiene solución exacta si el trompo es simétrico. Si

$$I_{O1} = I_{O2} \equiv I_O$$

las ecuaciones de Euler en ausencia de torque respecto de O resultan

$$\begin{aligned} 0 &= I_O\ \dot{\Omega}_1 + (I_{O3} - I_O)\ \Omega_2\ \Omega_3 \\ 0 &= I_O\ \dot{\Omega}_2 + (I_O - I_{O3})\ \Omega_1\ \Omega_3 \\ 0 &= I_{O3}\ \dot{\Omega}_3 \end{aligned}$$

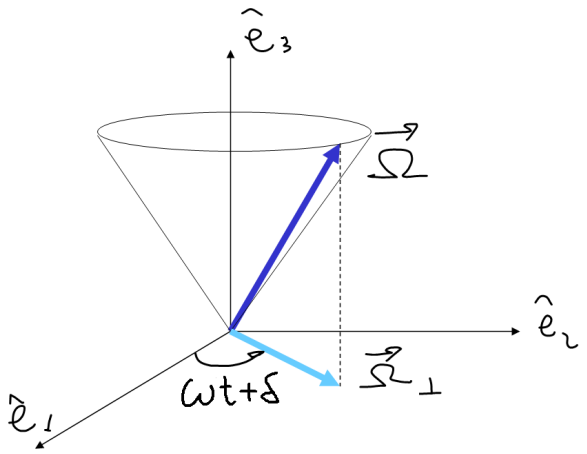
La solución exacta es

$$\begin{aligned} \Omega_1 &= A\ \cos(\omega t + \delta) \\ \Omega_2 &= A\ \sin(\omega t + \delta) \\ \Omega_3 &= cte \end{aligned} \tag{13.2}$$

con

$$\omega = \frac{I_{O3} - I_O}{I_O}\ \Omega_3 \tag{13.3}$$

Nótese que $|\vec{\Omega}|$ es constante. Como también Ω_3 es constante, entonces es constante el ángulo que forma la dirección del eje instantáneo de rotación dada por $\vec{\Omega}$ con el eje principal 3. Como $\vec{\Omega}_\perp = \Omega_1\ \hat{e}_1 + \Omega_2\ \hat{e}_2$ es un vector de módulo constante A que se mueve con velocidad angular constante igual a ω, lo mismo sucede con $\vec{\Omega}$.[2] En el sistema fijo al cuerpo de ejes principales, el vector $\vec{\Omega}$ precede alrededor del eje principal 3; el extremo de $\vec{\Omega}$ describe una circunferencia alrededor de $\hat{e}_3$. El momento angular



$\vec{L}_O = I_O\ \vec{\Omega}_\perp + I_{O3}\ \Omega_3\ \hat{e}_3$, que se conserva en el sistema inercial, está contenido en el plano formado por el eje de simetría del cuerpo y el eje instantáneo de rotación.

2. Nótese que se conserva la energía cinética de rotación $T_{rot} = (1/2)(I_O\ |\vec{\Omega}_\perp|^2 + I_{O3}\ {\Omega_3}^2)$.

A modo de ejemplo, digamos que la precesión de $\vec{\Omega}$ alrededor del eje de simetría del cuerpo es un fenómeno que tiene un efecto en la relación entre nuestro planeta Tierra y la bóveda celeste. La Tierra está sometida a pequeños torques respecto de su CM por parte de la Luna y el Sol, que se pueden ignorar en una primera aproximación. Además, la Tierra es bastante fielmente un trompo simétrico. Entonces vale el resultado obtenido, y podemos decir que el eje instantáneo de rotación de la Tierra precede alrededor de su eje de simetría que une los polos geográficos (en efecto, existe un pequeño ángulo entre ambos ejes). ¿Hay alguna evidencia empírica de esta precesión? Sí la hay pues, como consecuencia, el eje de rotación de la bóveda celeste, que tiene la dirección de $\vec{\Omega}$, precede alrededor de la recta que une los polos geográficos. La *precesión de Chandler* (1891) tiene un período de 427 días.[3] En 1749 Euler había estimado un período de 306 días; como el ángulo entre $\vec{\Omega}$ y $\hat{e}_3$ es muy pequeño, se puede decir que $\Omega_3 \simeq 2\pi/(1\ día)$ y así tenemos el valor de ω:

$$\omega = \frac{I_{CM\ 3} - I_{CM}}{I_{CM}}\ \Omega_3 \simeq \frac{1}{306}\ \frac{2\pi}{1\ día} = \frac{2\pi}{306\ días}$$

donde $(I_{CM\ 3} - I_{CM})/I_{CM} \simeq 1/306$ resulta del achatamiento polar del globo terrestre. La discrepancia entre el valor estimado y el valor observado se atribuye a que la Tierra no es un cuerpo rígido (tiene un núcleo fundido, océanos que son deformados por fuerzas de marea, etc.).

Las evoluciones de los ángulos de Euler deberían surgir de las relaciones entre los mismos y $\Omega_1, \Omega_2, \Omega_3$ (cuyas evoluciones acabamos de obtener); esas relaciones son

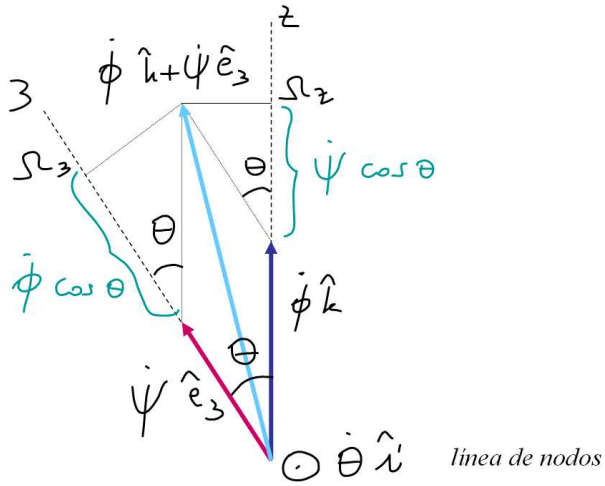
$$\begin{aligned} \Omega_1 &= \dot{\phi}\ \sin\theta\ \sin\psi + \dot{\theta}\ \cos\psi \\ \Omega_2 &= \dot{\phi}\ \sin\theta\ \cos\psi - \dot{\theta}\ \sin\psi \\ \Omega_3 &= \dot{\phi}\ \cos\theta + \dot{\psi} \end{aligned} \tag{13.4}$$

Recordemos que estas relaciones provienen de esta otra:

$$\vec{\Omega} = \dot{\phi}\ \hat{k} + \dot{\theta}\ \hat{\imath}\,' + \dot{\psi}\ \hat{k}_c \tag{13.5}$$

que está representada en la siguiente Figura:

3. No confundir con la *precesión de los equinoccios*, que tiene un período de 25776 años y es producida por los torques ejercidos por la Luna y el Sol.



Ahora bien, los valores de los ángulos de Euler dependen de cómo se eligen la terna fija al cuerpo (en nuestro caso es la terna de ejes principales) y la terna inercial $\hat{\imath}, \hat{\jmath}, \hat{k}$. Haremos una elección de $\hat{k}$ que será la más simple para caracterizar el ángulo θ. Aprovechando que $\vec{L}_O$ se conserva, usaremos un sistema inercial cuyo eje z está orientado como $\vec{L}_O$; entonces $\vec{L}_O = L_O\ \hat{k}$. En ese caso, la relación $|L_O>^C = \ R_\psi R_\theta R_\phi\ |L_O>$ resulta

$$\begin{pmatrix} L_{O1} \\ L_{O2} \\ L_{O3} \end{pmatrix} = R_\psi R_\theta R_\phi \begin{pmatrix} 0 \\ 0 \\ L_O \end{pmatrix} = \begin{pmatrix} L_O \sin\theta \sin\psi \\ L_O \sin\theta \cos\psi \\ L_O \cos\theta \end{pmatrix}$$

y como $L_{O1} = I_O\ \Omega_1$, $L_{O2} = I_O\ \Omega_2$, $L_{O3} = I_{O3}\ \Omega_3$, tendremos que

$$\begin{aligned} I_O\ \Omega_1 &= L_O\ \sin\theta\ \sin\psi \\ I_O\ \Omega_2 &= L_O\ \sin\theta\ \cos\psi \\ I_{O3}\ \Omega_3 &= L_O\ \cos\theta \end{aligned} \tag{13.6}$$

De los resultados obtenidos para $\Omega_1, \Omega_2, \Omega_3$ sabemos que Ω_3 es constante; entonces la última ecuación implica que θ es constante

$$\cos\theta = \frac{I_{O3}\ \Omega_3}{L_O} = cte \tag{13.7}$$

Esto significa que el ángulo entre el eje $\hat{e}_3$ y el vector $\vec{L}_O$ es constante: el eje de simetría del trompo precede alrededor de $\vec{L}_O$.[4]

4. Para un rotador el resultado es $\theta = \pi/2$, pues $I_{O3} = 0$. El rotador libre de torques gira en el plano perpendicular a $\vec{L}_O$. Tanto el rotador como el trompo esférico satisfacen que $\vec{L}_O = I_O\ \vec{\Omega}$. Por lo tanto, en estos dos cuerpos la conservación de $\vec{L}_O$ equivale a la conservación de $\vec{\Omega}$.

Sabiendo que $\dot{\theta} = 0$ [5] y comparando las primeras (o las segundas) ecuaciones en (13.4) y (13.6), se obtiene que la precesión del eje de simetría del trompo $\hat{e}_3$ alrededor de la dirección de $\vec{L}_O$, que viene dado por el corrimiento de la línea de nodos a velocidad $\dot{\phi}$, es uniforme con velocidad

$$\dot{\phi} = \frac{L_O}{I_O} \tag{13.8}$$

Dividiendo la primera y segunda ecuación en (13.6) se obtiene

$$\frac{\Omega_2}{\Omega_1} = \cot\psi$$

y usando la solución (13.2)

$$\tan(\omega t + \delta) = \cot\psi = \tan\left(\frac{\pi}{2} - \psi\right)$$

es decir que el giro alrededor del eje $\hat{e}_3$ se realiza a velocidad constante

$$\dot{\psi} = -\omega$$

Reemplazando estos resultados en (13.5), donde $\hat{k}_c$ es el eje de simetría $\hat{e}_3$, obtenemos

$$\vec{\Omega} = \frac{L_O}{I_O}\ \hat{k} - \omega\ \hat{e}_3 = \frac{\vec{L}_O}{I_O} - \omega\ \hat{e}_3$$

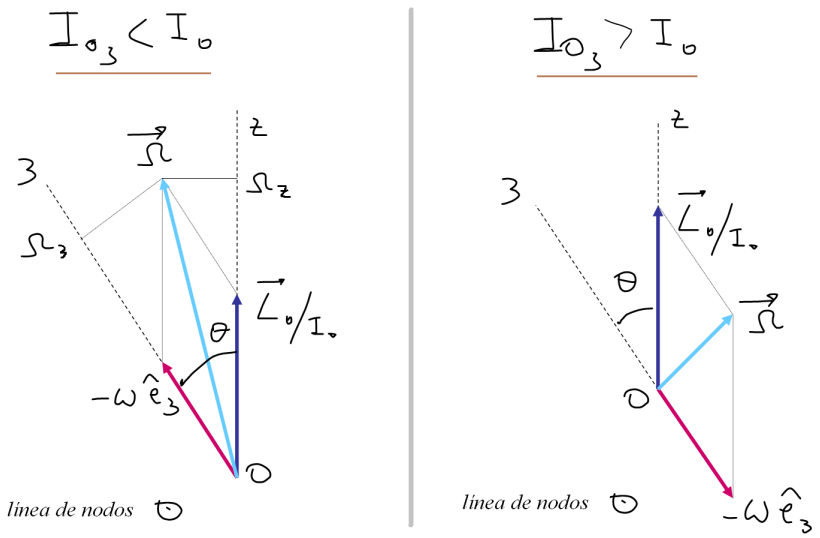
$\vec{\Omega}$ está en el plano formado por $\vec{L}_O$ y el eje de simetría $\hat{e}_3$. A la vez que el eje de simetría precede alrededor de $\vec{L}_O$ con velocidad constante $\dot{\phi} = L_O/I_O$, el cuerpo gira alrededor de $\hat{e}_3$ con velocidad constante $\dot{\psi} = -\omega$. Estas dos velocidades constantes están determinadas por las condiciones iniciales (L_O y ω son dos constantes de integración). El valor de $\dot{\phi}$ se relaciona directamente con el módulo del momento angular, mientras que la relación entre ambas velocidades determina el ángulo θ entre $\vec{L}_O$ y el eje de simetría:

$$\frac{\dot{\psi}}{\dot{\phi}} = \frac{-\omega}{\frac{L_O}{I_O}} = (I_O - I_{O3})\ \frac{\Omega_3}{L_O}\ = \frac{I_O - I_{O3}}{I_{O3}}\ \cos\theta$$

donde usamos (13.3) y (13.6). Por lo tanto, los valores de $\dot{\phi}$ y $\dot{\psi}$ (o los valores L_O y ω) caracterizan completamente al vector $\vec{L}_O$.

En suma, hemos mostrado que el eje instantáneo de rotación precede alrededor del eje de simetría $\hat{e}_3$, y este eje de simetría a su vez precede alrededor de la dirección de $\vec{L}_O$. Esto ocurre de tal forma que $\vec{\Omega}$ permanece en el plano formado por $\vec{L}_O$ y $\hat{e}_3$; este plano, que es perpendicular a la línea de nodos, gira con velocidad $\dot{\phi} = L_O/I_O$. La Figura muestra el comportamiento de los tres vectores en el sistema inercial para los casos de trompo simétrico alargado y achatado (se escogió un valor positivo para Ω_3):

5. Puede sorprender que no haya otra posibilidad para $\dot{\theta}$ que la de anularse. Esto es una consecuencia de haber elegido el eje z en la dirección de $\vec{L}_O$.



La velocidad $\dot{\psi} = -\omega$, el giro alrededor del eje 3, significa que el eje principal 1 (también el eje 2) pasa por la línea de nodos con un período $2\pi/|\omega|$. Como la línea de nodos es perpendicular al plano que contiene a $\vec{L}_O$ y $\vec{\Omega}$, este plano precede alrededor del eje 3 con velocidad ω, como se vio al comienzo de esta sección.

Videos

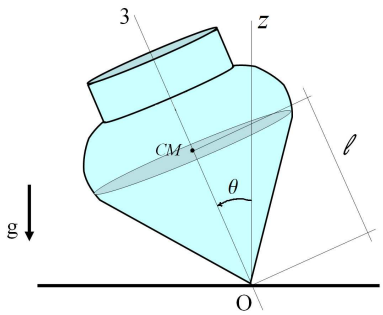
Rotación del trompo simétrico libre de torques (casos alargado y achatado)

https://www.youtube.com/watch?v=s9wiRjUKctU
https://www.youtube.com/watch?v=PDLXVSkDFVk

13.3 Trompo simétrico con punto fijo bajo la acción de la gravedad

Consideremos un trompo simétrico con un punto fijo O en un campo gravitatorio uniforme. Elegiremos el eje z de la terna inercial en la dirección vertical. Su Lagrangiano es

$$
\begin{aligned}
L &= \frac{1}{2} < \Omega|\mathbf{I}_O|\Omega > -mg\ell\cos\theta \\
&= \frac{1}{2}\, I_O\, (\Omega_1^2 + \Omega_2^2) + \frac{1}{2}\, I_{O3}\, \Omega_3^2 - mg\ell\cos\theta \\
&= \frac{1}{2}\, I_O\, (\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \frac{1}{2}\, I_{O3}\, (\dot{\phi}\, \cos\theta + \dot{\psi})^2 - mg\ell\cos\theta
\end{aligned}
$$



Como las coordenadas ϕ, ψ son cíclicas entonces se conservan

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = I_O \ \dot{\phi} \sin^2\theta + I_{O3} \ (\dot{\phi} \ \cos\theta + \dot{\psi}) \ \cos\theta$$

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_{O3} \ (\dot{\phi} \ \cos\theta + \dot{\psi}) = I_{O3} \ \Omega_3 = L_{O3}$$

Así como es $p_\psi = L_{O3}$, también se puede mostrar que $p_\phi = L_{Oz}$.[6] Las conservaciones de L_{Oz} y L_{O3} se deben a que el torque del peso respecto de O tiene la dirección de la línea de nodos, que es perpendicular al plano determinado por $\hat{k}$ y $\hat{e}_3$ (por otro lado, la fuerza de vínculo aplicada en O no realiza torque alguno respecto de O).[7]

Junto a esas dos magnitudes conservadas tenemos además la conservación de la energía

$$E = \frac{1}{2} \ I_O \ (\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \frac{1}{2} \ I_{O3} \ (\dot{\phi} \ \cos\theta + \dot{\psi})^2 + mg\ell \ \cos\theta$$

con lo que logramos tres primeras integrales de movimiento que nos ayudarán a resolver el problema.

Usaremos la conservación de p_ϕ, p_ψ para resolver $\dot{\phi}$, $\dot{\psi}$ en función de θ:

$$\dot{\phi} = \frac{L_{Oz} - L_{O3} \ \cos\theta}{I_O \ \sin^2\theta} \tag{13.9}$$

$$\dot{\psi} = \frac{L_{O3}}{I_{O3}} - \frac{L_{Oz}}{I_O} \ \frac{\cot\theta}{\sin\theta} + \frac{L_{O3}}{I_O} \ \cot^2\theta \tag{13.10}$$

6. Habíamos visto la descomposición $\hat{k} = \sin\psi \sin\theta \ \hat{\imath}_c + \cos\psi \sin\theta \ \hat{\jmath}_c + \cos\theta \ \hat{k}_c$. Estamos usando la terna fija al cuerpo de ejes principales $\hat{e}_1, \hat{e}_2, \hat{e}_3$ donde vale $\vec{L}_O = I_O \ (\Omega_1 \ \hat{e}_1 + \Omega_2 \ \hat{e}_2) + I_{O3} \ \Omega_3 \ \hat{e}_3$. Entonces $L_{Oz} = \hat{k} \cdot \vec{L}_O$ se obtiene fácilmente empleando las expresiones de $\Omega_1, \Omega_2, \Omega_3$ en función de los ángulos de Euler.

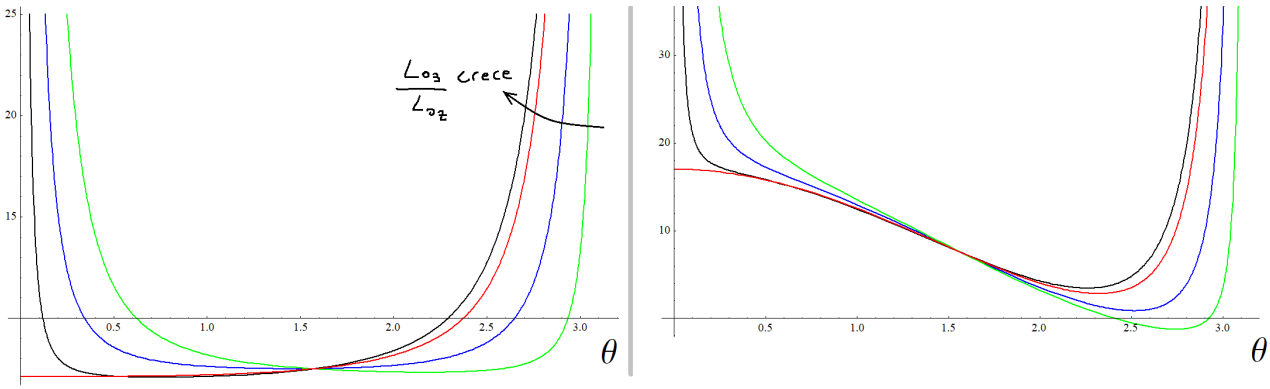
7. De acuerdo al teorema de Steiner, $\mathbf{I}_O$ e $\mathbf{I}_{CM}$ comparten ejes principales. Además es $I_{O3} = I_{CM3}$, mientras que el autovalor degenerado cambia: $I_O = I_{CM} + m \ \ell^2$.

Podemos sustituir estas velocidades generalizadas en la expresión de la energía,[8]

$$E = \frac{1}{2}\, I_O \left[\dot{\theta}^2 + \left(\frac{L_{Oz} - L_{O3}\ \cos\theta}{I_O\ \sin\theta}\right)^2\right] + \frac{1}{2}\,\frac{{L_{O3}}^2}{I_{O3}} + mg\ell\cos\theta$$

que puede verse como la conservación de la energía de un sistema de un único grado de libertad θ, que toma valores en el intervalo $[0, \pi]$, con un potencial efectivo

$$V_{efec} = \frac{1}{2}\,\frac{{L_{Oz}}^2}{I_O\ \sin^2\theta}\left(1 - \frac{L_{O3}}{L_{Oz}}\ \cos\theta\right)^2 + \frac{1}{2}\,\frac{{L_{O3}}^2}{I_{O3}} + mg\ell\cos\theta$$



Las Figuras muestran V_{efec} para dos valores de $mg\ell$ (menor a la izquierda), y distintos valores de L_{O3}/L_{Oz}. Las curvas rojas corresponden a $L_{O3} = L_{Oz}$.[9]

Vemos que V_{efec} posee un mínimo en un valor θ_o que depende de los parámetros del sistema. Si la energía E es próxima al valor mínimo de V_{efec}, entonces podemos tratar la evolución $\theta(t)$ mediante una aproximación de pequeñas oscilaciones. Para ello deberemos identificar la frecuencia propia de las oscilaciones a partir de la derivada segunda del potencial en θ_o. Como I_O juega el papel de "masa" en la energía E tendremos

$$\omega_o^2 = \frac{V''_{efec}(\theta_o)}{I_O}$$

8. No hace falta sustituir $\dot{\psi}$ porque el segundo término de E no es más que $(1/2)\ I_{O3}\ \Omega_3^2 = {L_{O3}}^2/(2\ I_{O3})$.

9. $V_{efec}(\pi/2) = \frac{1}{2}\,\frac{{L_{Oz}}^2}{I_O} + \frac{1}{2}\,\frac{{L_{O3}}^2}{I_{O3}}$. Para realizar la Figura se escogió el mismo valor de $V_{efec}(\pi/2)$ en las distintas curvas.

Una vez identificados los valores de θ_o y ω_o que correspondan a los valores de los parámetros que caracterizan el potencial $-{L_{Oz}}^2/I_O$, L_{O3}/L_{Oz} y $mg\ell$–, entonces la solución de pequeñas oscilaciones será

$$\theta(t) = \theta_o + \eta(t) \simeq \theta_o + \eta_o \ \cos(\omega_o t + \delta)$$

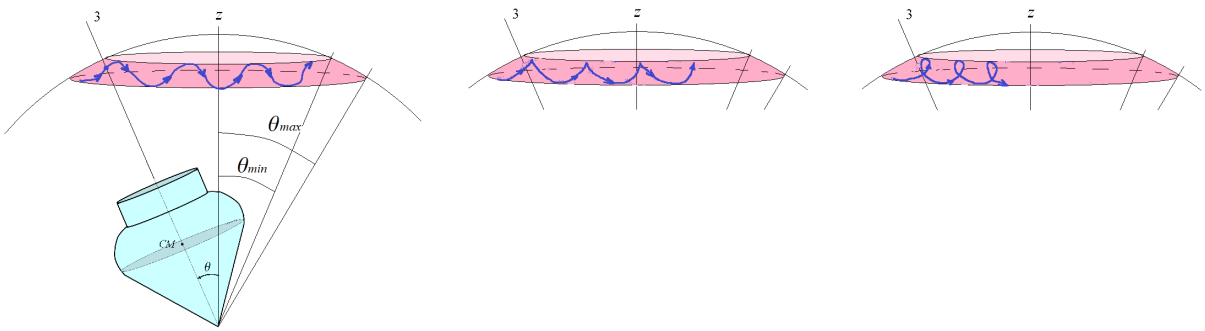
donde η_o es la amplitud de la oscilación, que depende de cuánto se aparta E del mínimo de potencial $V_{efec}(\theta_o)$. Reemplazando esta solución en (13.9) y (13.10) obtendremos las velocidades $\dot{\phi}(t)$ y $\dot{\psi}(t)$. Si la condición inicial para $\theta(t)$ es $\theta(0) = \theta_o$, $\dot{\theta}(0) = 0$ (es decir, $E = V_{efec}(\theta_o)$), entonces el eje de simetría del trompo se mantendrá en equilibrio estable en θ_o. En ese caso las velocidades tendrán valores constantes $\dot{\phi}_o$ y $\dot{\psi}_o$. En cambio, si θ oscila alrededor de θ_o, también oscilarán $\dot{\phi}(t)$ y $\dot{\psi}(t)$ alrededor de $\dot{\phi}_o$ y $\dot{\psi}_o$. Al primer orden perturbativo $\dot{\phi}(t)$ y $\dot{\psi}(t)$ se aproximan por

$$\dot{\phi}(t) \simeq \dot{\phi}_o + \frac{\partial \dot{\phi}}{\partial \theta}|_{\theta_o} \ [\theta(t) - \theta_o] = \dot{\phi}_o + \frac{\partial \dot{\phi}}{\partial \theta}|_{\theta_o} \ \eta(t)$$

$$\dot{\psi}(t) \simeq \dot{\psi}_o + \frac{\partial \dot{\psi}}{\partial \theta}|_{\theta_o} \ [\theta(t) - \theta_o] = \dot{\psi}_o + \frac{\partial \dot{\psi}}{\partial \theta}|_{\theta_o} \ \eta(t)$$

Así ambas velocidades oscilan alrededor de los valores medios $\dot{\phi}_o$ y $\dot{\psi}_o$.

Como θ es la inclinación del eje de simetría del trompo (eje $\hat{e}_3$) respecto de la vertical, mientras que $\dot{\phi}$ y $\dot{\psi}$ representan la velocidad de precesión del eje $\hat{e}_3$ y la velocidad de giro del trompo alrededor del eje $\hat{e}_3$ respectivamente, se concluye que el eje de simetría del trompo "cabecea" (oscilación de θ) realizando un movimiento conocido con el nombre de **nutación**. Mientras esto sucede las velocidades de precesión y de giro alrededor de $\hat{e}_3$ sufren también oscilaciones con la frecuencia de la nutación. La Figura muestra el movimiento del eje de simetría alrededor de la vertical; el movimiento es el resultado de la combinación de nutación y precesión.



i) En el primer caso la velocidad de la precesión $\dot{\phi}$ es siempre positiva; es lo que sucede si $|(\partial\dot{\phi}/\partial\theta)_{\theta_o}\eta_o| < \dot{\phi}_o$.

ii) El segundo caso corresponde a $|(\partial\dot{\phi}/\partial\theta)_{\theta_o}\eta_o| = \dot{\phi}_o$; $\dot{\phi}$ se anula cuando la nutación está en un punto de retorno.

iii) En el tercer caso es $|(\partial\dot{\phi}/\partial\theta)_{\theta_o}\eta_o| > \dot{\phi}_o$; por lo tanto la contribución de la nutación consigue revertir el signo de $\dot{\phi}$, y la precesión por momentos retrocede.[10]

La Tierra está sometida a pequeños torques ejercidos principalmente por la Luna y el Sol. Estos torques producen la nutación del eje terrestre, junto con una precesión lenta del mismo alrededor de la dirección perpendicular al plano de la eclíptica. La precesión tiene un período de unos 26000 años, y es conocida como *precesión de los equinoccios* pues el cambio de la orientación del eje terrestre modifica el punto de la órbita donde el Sol ilumina a plomo sobre el Ecuador.[11]

Valores de θ_o y ω_o^2

El ángulo θ_o minimiza el potencial efectivo. $V_{efec}^{'}(\theta_o) = 0$ implica que

$$(L_{Oz} - L_{O3}\cos\theta_o)\ (L_{O3} - L_{Oz}\cos\theta_o) = mg\ell\ I_O\ \sin^4\theta_o \tag{13.11}$$

Así θ_o depende de dos parámetros, que podemos elegir como L_{O3}/L_{Oz} y $mg\ell\ I_O/{L_{Oz}}^2$.

La frecuencia de pequeñas oscilaciones de θ resulta de calcular $V_{efec}^{''}$ y usar (13.11) para su evaluación en θ_o. Cuando $\theta_o \neq 0$, ω_o^2 se simplifica a

$$\omega_o^2 = \frac{V_{efec}^{''}(\theta_o)}{I_O} = \frac{L_{O3}\ L_{Oz} - mg\ell\ I_O\ (3\ \cos^2\theta_o + 1)}{I_O^2\ \cos\theta_o} \tag{13.12}$$

13.3.1 Trompo vertical

Las curvas rojas en la Figura del potencial efectivo muestran que el caso $L_{O3} = L_{Oz}$ puede tener uno o dos puntos de equilibrio. En ambos casos $\theta_o = 0$ es un punto de equilibrio, que corresponde al **trompo vertical**. Vemos en los gráficos que si el valor de $mg\ell$ es grande, entonces $\theta_o = 0$ será inestable. La transición del caso estable al inestable se reconoce por la anulación de $V_{efec}^{''}(0)$. Si $L_{O3} = L_{Oz} \equiv L_O$, se obtiene que

$$V_{efec}^{''}(\theta) = \frac{L_O^2}{I_O}\frac{2-\cos\theta}{(1+\cos\theta)^2} - mg\ell\ \cos\theta \qquad \Rightarrow \qquad V_{efec}^{''}(0) = \frac{L_O^2}{4\ I_O} - mg\ell$$

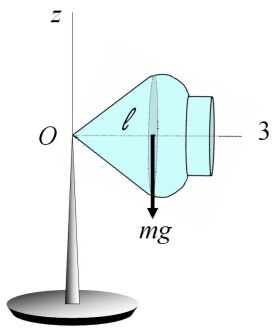
10. $\partial\dot{\phi}/\partial\theta > 0$ para todo valor de θ si $L_{Oz} > L_{Oz} > 0$.

11. El pequeño torque aparece, por un lado, porque el campo gravitatorio del Sol y la Luna no es uniforme sobre la extensión de nuestro planeta. Por el otro lado, la Tierra es achatada y su eje está inclinado respecto de la direcciones Tierra-Sol y Tierra-Luna. De esa forma el "lóbulo" de la Tierra que mira hacia el Sol (o la Luna) recibe una fuerza mayor que el lóbulo opuesto.

Entonces el trompo vertical es estable si $L_O^2 > 4\ mg\ell I_O$. [12]

En el trompo vertical el eje de simetría $\hat{e}_3$ coincide con el eje z del sistema de referencia. Por lo tanto la línea de nodos pierde su significado, y el ángulo azimutal debe verse como $\phi + \psi$. Aunque las expresiones (13.9) y (13.10) para $\dot{\phi}$, $\dot{\psi}$ quedan indeterminadas en $\theta = 0$, su suma $\dot{\phi} + \dot{\psi}$ en $\theta = 0$ vale L_{O3}/I_{O3}; por cierto este resultado refleja el hecho de que $\Omega_3 = \dot{\phi}\ \cos\theta + \dot{\psi}$ es igual a $\dot{\phi} + \dot{\psi}$ si $\theta = 0$.

13.3.2 Trompo horizontal



Las ecuaciones (13.9) y (13.10) dicen que si $\theta = \pi/2$ entonces los valores constantes de $\dot{\phi}$ y $\dot{\psi}$ son:

$$\dot{\phi} = \frac{L_{Oz}}{I_O}\ , \qquad\qquad \dot{\psi} = \frac{L_{O3}}{I_{O3}}$$

El eje de simetría del trompo precede a velocidad constante $\dot{\phi}$, a la vez que el trompo gira alrededor de su eje de simetría a velocidad constante $\dot{\psi}$. Sin embargo estas constantes de movimiento no tienen valor arbitrario pues deben asegurar que $\theta = \pi/2$ sea un punto de equilibrio (mínimo de V_{efec}). Para que $\theta_o = \pi/2$ sea solución de la ecuación (13.11) se debe satisfacer

$$L_{O3}\ L_{Oz} = mg\ell\ I_O \tag{13.13}$$

es decir que la velocidad de precesión está relacionada con la velocidad de giro,

$$\dot{\phi}_o = \frac{mg\ell}{L_{O3}}\ ,$$

que es el resultado que conocemos del curso de mecánica elemental.

Para obtener la frecuencia de las pequeñas oscilaciones alrededor de $\theta_o = \pi/2$, reemplazamos (13.13) en V_{efec} y calculamos nuevamente $V_{efec}^{''}(\theta_o)$; así evitaremos la indeterminación de la expresión (13.12) en $\theta_o = \pi/2$. Se obtiene que

$$\omega_o^2 = \frac{{L_{O3}}^2\ + {L_{Oz}}^2}{I_O^2}$$

12. Normalmente esta condición se pierde por rozamiento en el pivote. Si la condición no se cumple, el mínimo del potencial se halla en $\theta_o = \arccos\left[|L_O|/\sqrt{mg\ell I_O} - 1\right]$. Por otra parte, si $L_{O3} = -L_{Oz}$ habrá equilibrio estable en $\theta_o = \pi$.

13.3.3 Trompo rápido

Se pueden obtener resultados similares a los precedentes para $\theta_o \approx \pi/2$ bajo la aproximación de trompo rápido: $|L_{O3}| \ >> |L_{Oz}|$. En efecto, sea

$$\theta_o = \pi/2 - \varepsilon$$

Entonces

$$\cos\theta_o = \varepsilon + \mathcal{O}(\varepsilon^3) \ , \qquad \sin\theta_o = 1 + \mathcal{O}(\varepsilon^2)$$

Las ecuaciones (13.9) y (13.10) ahora dicen que en ausencia de nutación (es decir, cuando $\theta(t) = \theta_o$) las velocidades de precesión y giro son

$$\dot{\phi}_o \simeq \frac{L_{Oz} - L_{O3}\ \varepsilon}{I_O} \ , \qquad \dot{\psi}_o \simeq \frac{L_{O3}}{I_{O3}} - \frac{L_{Oz}}{I_O}\ \varepsilon \tag{13.14}$$

Para que no haya nutación, $\theta_o = \pi/2 - \varepsilon$ debe cumplir la condición (13.11) para dar un mínimo del V_{efec},

$$L_{O3}\ L_{Oz} - ({L_{O3}}^2 + {L_{Oz}}^2)\ \varepsilon \ \simeq \ mg\ell\ I_O$$

que con la aproximación de trompo rápido se convierte en

$$L_{O3}\ L_{Oz} - {L_{O3}}^2\ \varepsilon \ \simeq \ mg\ell\ I_O \tag{13.15}$$

Reemplazando en (13.14) obtenemos la expresión del ejemplo anterior para la velocidad de precesión:

$$\dot{\phi}_o \simeq \frac{mg\ell}{L_{O3}}$$

También la velocidad de giro tiene la misma expresión que en el ejemplo anterior si el trompo es rápido:

$$\dot{\psi}_o \simeq \frac{L_{O3}}{I_{O3}}$$

La frecuencia (13.12) de las oscilaciones en torno al punto de equilibrio se aproxima por

$$\omega_o^2 \simeq \frac{L_{O3}\ L_{Oz} - mg\ell\ I_O}{I_O^2\ \varepsilon}$$

y usando (13.15) resulta

$$\omega_o \simeq \frac{|L_{O3}|}{I_O} \approx \frac{|L_O|}{I_O}$$

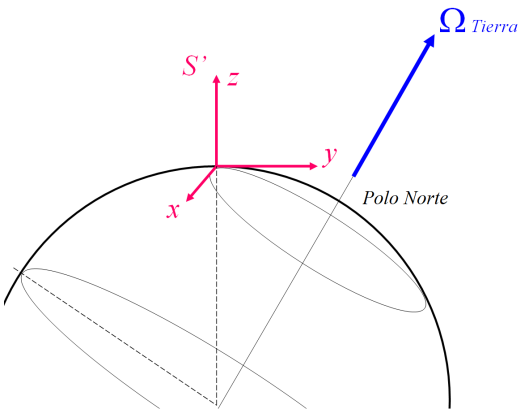
El trompo rápido combina una precesión lenta con una nutación rápida (en la ecuación (13.14) nótese que si ε es suficientemente chico entonces es $\dot{\phi} \approx L_{Oz}/I_O$). La nutación es difícil de observar porque se amortigua rápidamente debido a la fricción en el pivote.

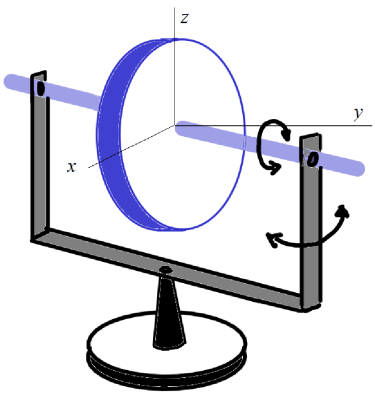
13.4 Giróscopo

Un giróscopo es un disco o rueda cuyo eje se monta en una suspensión cardánica que fija la posición de su CM sin limitar los grados de libertad de rotación. La suspensión cardánica no transmite torques al disco (excepto si dos ejes de la suspensión coinciden), de modo que si éste gira alrededor de su eje, la conservación del momento angular hará que la orientación del eje se mantenga fija, independientemente de los movimientos del soporte. Este comportamiento es aprovechado en sistemas de navegación diversos. En 1852 Foucault construyó el primer giróscopo para ofrecer una nueva demostración de la rotación de la Tierra (además de su famoso péndulo), basada en la persistencia de la orientación del eje de rotación del giróscopo.

13.4.1 Girocompás de Foucault

Foucault se dio cuenta de que una montura diferente permitía usar el giróscopo para señalar la dirección norte-sur. Supongamos que la montura sólo permite el desplazamiento libre del eje de rotación del giróscopo en direcciones horizontales sobre la superficie terrestre (por ejemplo, una montura "semi-cardánica"). Si utilizamos un sistema de referencia S' fijo a la Tierra con el eje z en la dirección vertical, entonces el soporte no ejercerá torques en z para permitir la libre orientación del giróscopo en el plano $x - y$, pero aplicará torques en x e y para evitar que el eje de rotación del giróscopo adquiera una componente vertical (en z) a medida que la Tierra gire.





Suspensión semi-cardánica que permite el libre giro del giróscopo alrededor de su eje, pero limita la libre orientación del eje al plano horizontal.

Para entender la dinámica del girocompás usemos el teorema de la derivada relativa aplicado a $\vec{L}_{CM}$:

$$\frac{d\vec{L}_{CM}}{dt}|_S = \frac{d\vec{L}_{CM}}{dt}|_{S'} + \vec{\Omega}_{Tierra} \times \vec{L}_{CM}$$

En el sistema inercial S, $d\vec{L}_{CM}/dt$ es el torque $\vec{N}^{ext}_{CM}$ provisto por el soporte; entonces

$$\frac{d\vec{L}_{CM}}{dt}|_{S'} = \vec{N}^{ext}_{CM} - \vec{\Omega}_{Tierra} \times \vec{L}_{CM}$$

En el sistema S' fijo a la Tierra, $-\vec{\Omega}_{Tierra} \times \vec{L}_{CM}$ es un torque debido a fuerzas de inercia.[13] En particular, como $N^{ext}_{CM_z} = 0$, para la componente z se tiene que

$$\frac{dL_{CM_z}}{dt}|_{S'} = -(\vec{\Omega}_{Tierra} \times \vec{L}_{CM})_z$$

Si $\vec{L}_{CM}$ está inicialmente en el plano x, y del sistema S', la rotación terrestre generará una componente z de $\vec{L}_{CM}$ en el sistema fijo a la Tierra. Para ello el eje del giróscopo cambiará su orientación en el plano horizontal, agregando así una componente z a su velocidad de rotación. Sin embargo, si el eje del giróscopo tiene inicialmente la orientación norte-sur (es decir que $\vec{L}_{CM}$ sólo tiene componente y en la Figura) entonces se anula la componente z del torque de las fuerzas de inercia, y el giróscopo mantiene su orientación. Analicemos la estabilidad de esa orientación de equilibrio:

$$-(\vec{\Omega}_{Tierra} \times \vec{L}_{CM})_z = -\Omega_{Tierra\ x}\ L_{CMy} + \Omega_{Tierra\ y}\ L_{CMx} = \Omega_{Tierra\ y}\ L_{CMx}$$

que es positivo si $L_{CMx} > 0$, y negativo si $L_{CMx} < 0$, lo que garantiza la estabilidad.[14] Así, el torque de las fuerzas de inercia lleva a orientar el giróscopo hacia el norte. Se requiere un mecanismo de amortiguación para estabilizar la alineación.

Videos

Precesión y nutación
https://www.youtube.com/watch?v=sHnDzGWcqlQ
https://www.youtube.com/watch?v=5Sn2J1Vn4zU

Suspensión cardánica
https://www.youtube.com/watch?v=Y0xkS-fFIy0
https://www.youtube.com/watch?v=ekzwbt3hu2k

13. Para ver todos los torques de inercia en S' hay que escribir $\vec{L}_{CM}$ en términos de $\vec{L}'_{CM}$.
14. $\Omega_{Tierra\ y} > 0$ porque la Tierra gira de oeste a este.

14 Formalismo Hamiltoniano

14.1 Transformada de Legendre

Sea una función $F(u_1,\ ...,\ u_n;\ w_1,\ ...,\ w_m)$. Su diferencial es igual a

$$dF \ = \ \sum_{k=1}^{n} \frac{\partial F}{\partial u_k}\, du_k \ + \ \sum_{j=1}^{m} \frac{\partial F}{\partial w_j}\, dw_j$$

que podemos escribir así:

$$dF \ = \ \sum_{k=1}^{n} d\left(u_k\, \frac{\partial F}{\partial u_k}\right) - \sum_{k=1}^{n} u_k\, d\left(\frac{\partial F}{\partial u_k}\right) + \sum_{j=1}^{m} \frac{\partial F}{\partial w_j}\, dw_j$$

Entonces

$$d\left(\sum_{k=1}^{n} u_k\, \frac{\partial F}{\partial u_k} - F\right) \ = \ \sum_{k=1}^{n} u_k\, d\left(\frac{\partial F}{\partial u_k}\right) - \sum_{j=1}^{m} \frac{\partial F}{\partial w_j}\, dw_j \tag{14.1}$$

Definamos

$$v_k(u,w) \equiv \frac{\partial F}{\partial u_k} \tag{14.2}$$

Si la relación entre las v_k's y las u_k's es inversible, lo que ocurre si la matriz Hessiana de F es inversible:

$$\det\left(\frac{\partial v_i}{\partial u_k}\right) = \det\left(\frac{\partial^2 F}{\partial u_i\, \partial u_k}\right) \neq 0\ ,$$

entonces podremos ver las u_k's como funciones de las v_k's, y definir

$$G(v,w) \equiv \sum_{k=1}^{n} u_k(v,w)\ v_k - F(u(v,w),w) \tag{14.3}$$

De esta forma la ecuación (14.1) dice que

$$dG = \sum_{k=1}^{n} u_k \, dv_k - \sum_{j=1}^{m} \frac{\partial F}{\partial w_j} \, dw_j$$

es decir,

$$\frac{\partial G}{\partial v_k} = u_k \ , \qquad\qquad \frac{\partial G}{\partial w_j} = -\frac{\partial F}{\partial w_j}$$

Si nos abstraemos de las variables w_j's, que no participan en la transformación (son variables *pasivas*), las relaciones

$$v_k = \frac{\partial F}{\partial u_k} \ , \qquad\qquad u_k = \frac{\partial G}{\partial v_k} \ ,$$

exhiben una notable dualidad entre la función F y su **transformada de Legendre** G. Las nuevas variables son las derivadas de la vieja función respecto de las viejas variables; a su vez las viejas variables son las derivadas de la nueva función respecto de las nuevas variables. Pero "viejo" y "nuevo" son intercambiables porque así como G es la transformada de Legendre de F, también F es la transformada de Legendre de G, como resulta de la ecuación (14.3).

Ejemplo. La *energía interna* U de un sistema termodinámico simple puede verse como una función $U(S, V)$ de la entropía S y del volumen V tal que $dU = T \, dS - p \, dV$, es decir

$$\frac{\partial U}{\partial S} = T \ , \qquad\qquad \frac{\partial U}{\partial V} = -p \ .$$

En Termodinámica se define la *energía libre de Helmholtz* $F(T, V)$ como la transformada de Legendre de U respecto de la variable S, cambiada de signo

$$-F(T, V) \equiv TS - U$$

que cumple

$$-\frac{\partial F}{\partial T} = S \ , \qquad\qquad -\frac{\partial F}{\partial V} = p$$

Otros potenciales termodinámicos, como la entalpía o la energía de Gibbs, se definen en forma similar intercambiando distintos pares de variables.

14.2 La función Hamiltoniana

En Mecánica Analítica los sistemas físicos son descriptos mediante una función Lagrangiana $L(q, \dot{q}, t)$ a partir de la que se definen los momentos canónicamente conjugados,

$$p_\mu \equiv \frac{\partial L}{\partial \dot{q}_\mu} \ ,$$

en términos de los cuales las ecuaciones dinámicas de Lagrange se escriben

$$\dot{p}_\mu = \frac{\partial L}{\partial q_\mu} \ . \tag{14.4}$$

La definición de los momentos conjugados puede verse como el primer paso para introducir una transformada de Legendre, esto es la definición de las variables "nuevas".[1] Entonces definimos la función Hamiltoniana $H(q, p, t)$ de un sistema físico como la transformada de Legendre del Lagrangiano $L(q,\ \dot{q},\ t)$ respecto de las velocidades generalizadas $\dot{q}_\mu$,

$$H(q, p, t) \equiv \sum_{\mu=1}^{n} p_\mu \ \dot{q}_\mu(q, p, t) - L(q,\ \dot{q}(q, p, t),\ t) \tag{14.5}$$

Debido a la relación entre variables nuevas y viejas, resulta que

$$\dot{q}_\mu = \frac{\partial H}{\partial p_\mu} \ .$$

Con respecto a las variables pasivas de la transformación (es decir, q, t), tendremos

$$\frac{\partial H}{\partial q_\mu} = -\frac{\partial L}{\partial q_\mu} \ , \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \ . \tag{14.6}$$

Entonces las ecuaciones dinámicas de Lagrange en términos de H se escriben como $\dot{p}_\mu = -\partial H/\partial q_\mu$. Las $2n$ ecuaciones

$$\boxed{\dot{q}_\mu = \frac{\partial H}{\partial p_\mu} \ , \qquad \dot{p}_\mu = -\frac{\partial H}{\partial q_\mu} \ , \qquad \mu = 1,\ ...,\ n} \tag{14.7}$$

se denominan **ecuaciones canónicas de Hamilton** (1834).

Nota. H es la magnitud vista en §6.4 que se conserva cuando el Lagrangiano no depende explícitamente de t (de acuerdo con (14.6), H tampoco dependerá explícitamente de t en tal caso), y que coincide con la energía mecánica en muchas situaciones. Más allá del valor numérico de H en cada instante –que no depende de que se lo calcule usando las $\dot{q}$'s o los p's–, para la formulación de las ecuaciones de Hamilton es fundamental que H esté escrito en función de las p's, que son las variables "nuevas" de la transformación.

1. Recordemos que $\det(\partial^2 L/\partial \dot{q}_\mu \partial \dot{q}_\nu) \neq 0$ para que sea posible escribir $\dot{q}_\mu = \dot{q}_\mu(q, p, t)$.

14.3 Evolución en el espacio de las fases

¿Qué tienen de notable las ecuaciones de Hamilton? Hemos pasado de las n ecuaciones de Lagrange, que son ecuaciones de segundo orden para las incógnitas $q_\mu(t)$, a un sistema de $2n$ ecuaciones de primer orden para las incógnitas $q_\mu(t)$, $p_\mu(t)$. En principio, siempre es posible hacer una conversión de este tipo agregando variables (en este caso las p_μ's). Lo realmente sorprendente y bello es que la conversión de n ecuaciones de segundo orden a $2n$ ecuaciones de primer orden se ha logrado de tal forma que las derivadas de las **variables canónicas** (q_μ, p_μ) están despejadas en las ecuaciones de Hamilton (el Hamiltoniano H no contiene derivadas de q o p).

Si representamos la evolución del sistema en el **espacio de las fases**, que es el espacio de $2n$ dimensiones definido por las variables canónicas (q_μ, p_μ), como una curva $(q_\mu(t), p_\mu(t))$ parametrizada por el tiempo t entonces el "vector" de $2n$ componentes tangente a la curva es $(\dot{q}_\mu, \dot{p}_\mu)$.[2] Pero sucede que ese vector es dado por la simple derivación del Hamiltoniano del sistema, como enseñan las ecuaciones de Hamilton. Esto significa que en cada punto del espacio de las fases (cada punto representa un *estado* del sistema) conocemos hacia dónde evolucionará el sistema porque esa evolución está escrita en los valores de las derivadas de H en ese punto:

$$(\dot{q}_1,\ ...,\ \dot{q}_n,\ \dot{p}_1,\ ...,\ \dot{p}_n) = \left(\frac{\partial H}{\partial p_1},\ ...,\ \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1},\ ...,\ -\frac{\partial H}{\partial q_n}\right)$$

Este vector tangente es único en cada punto, lo que significa que las evoluciones en el espacio de las fases no se cruzan (salvo en puntos singulares, como veremos). Por lo tanto las posibles evoluciones del sistema se ven en el espacio de las fases como si fueran líneas de un fluido cuya "velocidad" en cada punto queda definida por la ecuación anterior.

14.3.1 Sistemas conservativos de 1 grado de libertad

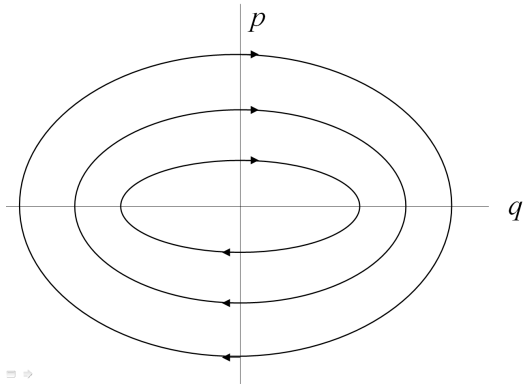
Las evoluciones en el espacio de las fases de los sistemas conservativos de 1 grado de libertad se obtienen inmediatamente como las curvas $H(q,p) = cte.$, ya que esa única integral de movimiento determinará completamente la evolución a partir de cada posible estado inicial $(q(0), p(0))$.

2. Del mismo modo que el vector $\vec{v} = (\dot{x}, \dot{y}, \dot{z})$ es tangente a la trayectoria de una partícula en el espacio físico de tres dimensiones.

Oscilador armónico

$$p = m\dot{q} \ , \qquad\qquad H(q,p) = \frac{p^2}{2m} + \frac{k}{2}\, q^2$$

En el espacio de las fases, las curvas $H(q,p) = cte.$ son elipses que rodean el punto de equilibrio estable $q = 0 = p$:



El vector tangente a las curvas es

$$(\dot{q},\ \dot{p}) = (\frac{p}{m}, -kq)$$

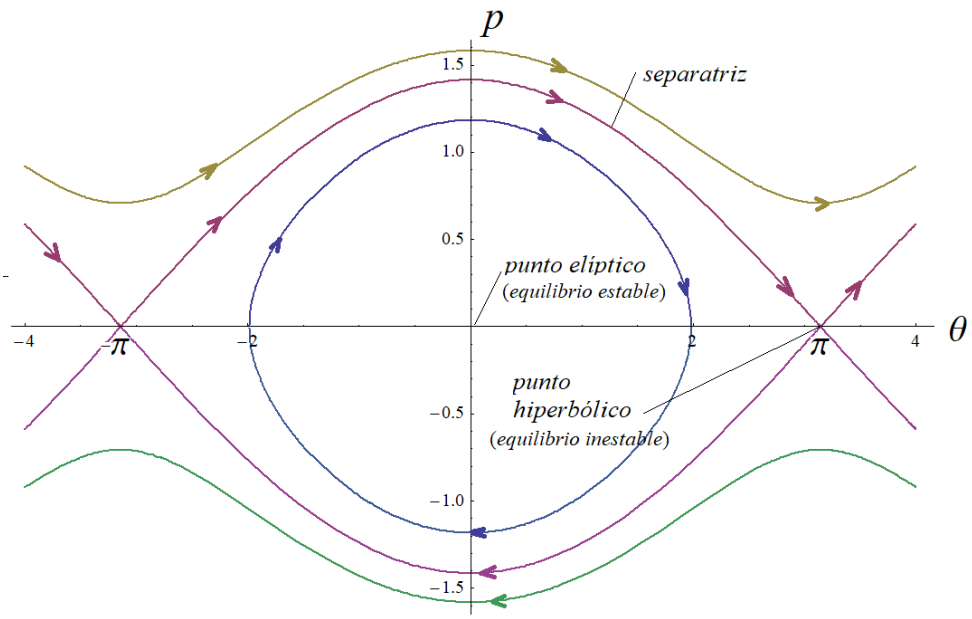
Péndulo simple

$$p = m\ell^2\dot{\theta} \ , \qquad\qquad H(q,p) = \frac{p^2}{2m\ell^2} - mg\ell\cos\theta$$

El vector tangente a las curvas es

$$(\dot{q},\ \dot{p}) = (\frac{p}{m\ell^2}, -mg\ell\sin\theta) \ ,$$

y se anula en los puntos de equilibrio. Las curvas $H(q,p) = cte.$ muestran los dos tipos de movimiento del péndulo: **libración** ($\dot{\theta}$ cambia de signo) en línea azul, y **rotación** ($\dot{\theta}$ no cambia de signo) en línea verde. La *separatriz* entre ambos tipos de comportamiento es la curva $H = mg\ell$.



14.4 Conservación de H

Como mencionamos, el formalismo Lagrangiano nos enseña que H se conserva cuando $\partial L/\partial t = 0$, que según (14.6) significa que H no depende explícitamente del tiempo. Pero veamos la conservación de H a la luz de las ecuaciones de Hamilton:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_{\mu=1}^{n} \left(\frac{\partial H}{\partial q_\mu}\, \dot{q}_\mu + \frac{\partial H}{\partial p_\mu}\, \dot{p}_\mu \right) = \frac{\partial H}{\partial t} + \sum_{\mu=1}^{n} \left(\frac{\partial H}{\partial q_\mu}\, \frac{\partial H}{\partial p_\mu} - \frac{\partial H}{\partial p_\mu}\, \frac{\partial H}{\partial q_\mu} \right) = \frac{\partial H}{\partial t}$$

es decir

$$\dot{H} = \frac{\partial H}{\partial t} \tag{14.8}$$

Aunque esta ecuación no es independiente de las ecuaciones de Hamilton (14.7) sino que es una consecuencia de las mismas, no deja de llamar la atención que t y H quedan relacionadas de la misma manera en que lo hacen los **pares de variables conjugadas** (q_μ, p_μ) en las ecuaciones de Hamilton. Aunque el tiempo t no es una variable dinámica sino el parámetro respecto del cual se describe la evolución del sistema, podría decirse que en algún sentido t y H forman también un par de "variables conjugadas".[3]

3. El tiempo puede incorporarse al conjunto de variables canónicas, al precio de introducir un vínculo entre las variables al momento de variar la acción. Esta estrategia llamada *parametrización* es útil para tratar un sistema no conservativo como si fuese formalmente conservativo (ver Lanczos, Cap. IV, Secc. 10).

14.5 Coordenadas cíclicas

Las ecuaciones de Hamilton muestran que

–Si q_μ es cíclica en el Hamiltoniano $H(q,p,t)$ entonces se conserva p_μ

–Si p_μ es cíclica en el Hamiltoniano $H(q,p,t)$ entonces se conserva q_μ

–Además, en virtud de (14.8): si t es cíclica en $H(q,p,t)$ entonces se conserva H.

14.6 Principio variacional de Hamilton

Las ecuaciones de Hamilton también pueden verse como el resultado de un principio variacional, donde la acción que se hace estacionaria es una funcional $S[q(t),p(t)]$ de $2n$ funciones independientes $q_\mu(t), p_\mu(t)$:

$$0 = \delta \int_{t_1}^{t_2} L\ dt = \delta \int_{t_1}^{t_2} \left(\sum_{\mu=1}^{n} p_\mu\ dq_\mu - H(q,p)\ dt \right) \equiv \delta S[q(t),p(t)]$$

Entonces, usando que la variación de la derivada es la derivada de la variación,

$$\begin{aligned} \delta S[q(t),p(t)] &= \int_{t_1}^{t_2} \left(\sum_{\mu=1}^{n} (\delta p_\mu\ dq_\mu + p_\mu\ d(\delta q_\mu)) - \delta H\ dt \right) \\ &= \sum_{\mu=1}^{n} (p_\mu\ \delta q_\mu)\,|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\sum_{\mu=1}^{n} (\delta p_\mu\ dq_\mu - dp_\mu\ \delta q_\mu) - \delta H\ dt \right) \end{aligned}$$

y como

$$\delta H = \sum_{\mu=1}^{n} \left(\frac{\partial H}{\partial q_\mu} \delta q_\mu + \frac{\partial H}{\partial p_\mu} \delta p_\mu \right)$$

se obtiene que

$$\delta S[q(t),p(t)] = \sum_{\mu=1}^{n} (p_\mu\ \delta q_\mu)\,|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_{\mu=1}^{n} \left[(dq_\mu - \frac{\partial H}{\partial p_\mu}\ dt)\ \delta p_\mu - (dp_\mu + \frac{\partial H}{\partial q_\mu}\ dt)\ \delta q_\mu \right] \tag{14.9}$$

Los dos paréntesis que aparecen dentro de la sumatoria se anulan por separado cuando se cumplen las ecuaciones de Hamilton. Entonces, la validez de las ecuaciones de Hamilton es la condición para que la acción $S[q(t), p(t)]$ sea estacionaria ante variaciones $\delta q_\mu(t)$, $\delta p_\mu(t)$ arbitrarias e independientes que dejen fijos los extremos de integración $q_\mu(t_1)$ y $q_\mu(t_2)$.

14.7 Función principal de Hamilton

Si la acción es evaluada en la evolución real del sistema (es decir, sobre aquellas funciones $q_\mu(t)$, $p_\mu(t)$ que satisfacen las ecuaciones de Hamilton) obtendremos una magnitud que depende de los valores $q_\mu(t_1)$, $q_\mu(t_2)$ que caractericen la solución elegida; es decir, obtendremos una función $S(q_1, t_1; q_2, t_2)$ que llamaremos **función principal de Hamilton**. Por ejemplo, veamos el caso de la partícula libre donde la solución de las ecuaciones de movimiento es

$$q(t) = q_1 + \frac{q_2 - q_1}{t_2 - t_1}(t - t_1), \qquad p = m\,\frac{q_2 - q_1}{t_2 - t_1}\ ;$$

la acción evaluada en esta solución es

$$S(q_1, t_1; q_2, t_2) = \int_{t_1}^{t_2} L\ dt = \int_{t_1}^{t_2} \frac{m}{2}\ \left(\frac{q_2 - q_1}{t_2 - t_1}\right)^2\ dt = \frac{m}{2}\ \frac{(q_2 - q_1)^2}{t_2 - t_1}$$

Una vez que la acción ha sido integrada sobre la evolución real del sistema podemos preguntarnos qué propiedades tiene el resultado como función de los valores de q_μ en los extremos de integración t_1 y t_2. Para ello es suficiente ver cómo cambia la funcional acción ante cambios en los extremos; ese comportamiento está contenido en la ecuación (14.9), de donde se concluye que

$$dS(q_1, t_1; q_2, t_2) = \sum_{\mu=1}^{n} \left(p_{\mu_2}\ dq_{\mu_2} - p_{\mu_1}\ dq_{\mu_1}\right)$$

es decir,

$$\frac{\partial}{\partial q_{\mu_2}} S(q_1, t_1; q_2, t_2) = p_{\mu_2}\ , \qquad \frac{\partial}{\partial q_{\mu_1}} S(q_1, t_1; q_2, t_2) = -p_{\mu_1} \tag{14.10}$$

En el caso de la partícula libre resulta

$$\frac{\partial}{\partial q_2} S(q_1, t_1; q_2, t_2) = m\,\frac{q_2 - q_1}{t_2 - t_1} = p$$

Estas derivadas corresponden al cambio de la acción S cuando se la evalúa sobre dos evoluciones reales entre t_1 y t_2 que empiezan o terminan en puntos vecinos del espacio de configuración.

También podemos considerar dos evoluciones reales entre las mismas configuraciones $q_{\mu 1}$, $q_{\mu 2}$, pero a tiempos inicial o final ligeramente diferentes; para ello calcularemos $\partial S/\partial t$. Por definición vale que

$$\frac{dS}{dt_2} = L(q_2, \dot{q}_2, t_2) \ , \qquad\qquad \frac{dS}{dt_1} = -L(q_1, \dot{q}_1, t_1) \ ,$$

lo que indica que L mide el cambio de S cuando nos mantenemos sobre la misma evolución real pero dejando transcurrir un poco más (o menos) de tiempo. Entonces

$$L(q_2, \dot{q}_2, t_2) = \frac{dS}{dt_2} = \sum_{\mu=1}^{n} \frac{\partial S}{\partial q_{\mu 2}} \ \dot{q}_{\mu 2} + \frac{\partial S}{\partial t_2} = \sum_{\mu=1}^{n} p_{\mu 2} \ \dot{q}_{\mu 2} + \frac{\partial S}{\partial t_2}$$

de donde resulta que

$$\frac{\partial}{\partial t_2} S(q_1, t_1; q_2, t_2) = -H(q_2, p_2, t_2)$$

Del mismo modo

$$\frac{\partial}{\partial t_1} S(q_1, t_1; q_2, t_2) = H(q_1, p_1, t_1)$$

En estas relaciones vemos nuevamente que entre el tiempo y el Hamiltoniano aparece una relación similar a la que observan las variables conjugadas en la ecuación (14.10).

14.8 Notación matricial

Podemos organizar las $2n$ variables canónicas del espacio de las fases en una *columna* de $2n$ componentes η_A,

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

Las ecuaciones de Hamilton relacionan las derivadas temporales de las componentes de $|\eta>$ –es decir, el vector tangente a la evolución en el espacio de las fases– con

las derivadas del Hamiltoniano respecto de las componentes de $|\eta>$. En efecto, las ecuaciones de Hamilton se pueden escribir matricialmente como

$$|\dot{\eta}> = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{-1} & \mathbf{0} \end{pmatrix} |\frac{\partial H}{\partial \eta}> \equiv \omega \, |\frac{\partial H}{\partial \eta}> \tag{14.11}$$

donde cada bloque de la matriz cuadrada que llamamos ω es a su vez una matriz cuadrada de $n \times n$. La relación matricial puede escribirse en notación de índices,

$$\dot{\eta}_A = \sum_{B=1}^{2n} \omega_{AB} \, \frac{\partial H}{\partial \eta_B}$$

La forma simple de las ecuaciones de Hamilton no sólo muestra la evolución de las variables canónicas en cada instante, en términos de las derivadas del Hamiltoniano, sino que permite conocer también cómo evoluciona cualquier función $f(q,p)$ del estado del sistema:

$$\frac{df}{dt} = \sum_{\mu=1}^{n} \left(\frac{\partial f}{\partial q_\mu} \, \dot{q}_\mu + \frac{\partial f}{\partial p_\mu} \, \dot{p}_\mu \right) = \sum_{A=1}^{2n} \frac{\partial f}{\partial \eta_A} \, \dot{\eta}_A = < \frac{\partial f}{\partial \eta} |\dot{\eta}> = < \frac{\partial f}{\partial \eta} |\omega| \frac{\partial H}{\partial \eta} >$$

donde $<|$ denota una *fila* de $2n$ componentes.

14.9 Corchete de Poisson

La matriz antisimétrica ω permite introducir una suerte de "producto" antisimétrico entre funciones en el espacio de las fases conocido como **corchete de Poisson** $\{f,g\}$, y se define así

$$\{f,g\} \equiv < \frac{\partial f}{\partial \eta} |\omega| \frac{\partial g}{\partial \eta} > = \sum_{\mu=1}^{n} \left(\frac{\partial f}{\partial q_\mu} \, \frac{\partial g}{\partial p_\mu} - \frac{\partial f}{\partial p_\mu} \, \frac{\partial g}{\partial q_\mu} \right) \tag{14.12}$$

De acuerdo con lo visto en la Sección anterior, la evolución de cualquier función $f(q,p,t)$ del estado del sistema y del tiempo está dada por

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f,H\} \tag{14.13}$$

Una magnitud $f(q,p)$ se conserva si

$$\{f,H\} = 0 \tag{14.14}$$

Las ecuaciones de Hamilton pueden escribirse como

$$\boxed{\dot{q}_\mu = \{q_\mu, H\}, \qquad \dot{p}_\mu = \{p_\mu, H\}} \tag{14.15}$$

Propiedades

i) $\{f, g\} = -\{g, f\}$ (antisimetría)

ii) $\{f_1 + f_2, g\} = \{f_1, g\} + \{f_2, g\}$ (linealidad)

iii) $\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}$

iv) $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$ (identidad de Jacobi)

La identidad de Jacobi permite mostrar que si $f(q, p)$ y $g(q, p)$ se conservan entonces $\{f, g\}$ también se conserva.

Casos particulares

i) $\{q_\mu, q_\nu\} = 0 = \{p_\mu, p_\nu\}$

ii) $\{q_\mu, p_\nu\} = \delta_{\mu\nu}$

iii) $\{q_\mu, f\} = \frac{\partial f}{\partial p_\mu}$

iv) $\{p_\mu, f\} = -\frac{\partial f}{\partial q_\mu}$

14.10 Transformaciones canónicas. Estructura simpléctica

En la formulación Lagrangiana enfatizamos que la misma es independiente de las coordenadas generalizadas que se elijan. Las ecuaciones de Lagrange son covariantes (no cambian su forma) ante cambios de coordenadas en el espacio de configuración. Ahora nos interesa saber cuáles son las transformaciones de variables en el espacio de las fases que no alteran la forma de las ecuaciones de Hamilton. Como hemos agrupado las $2n$ variables canónicas en una columna $|\eta>$, podemos escribir un cambio de variables general como

$$\eta_A \longrightarrow N_A = N_A(\eta)$$

Entonces

$$\dot{\eta}_A \longrightarrow \dot{N}_A = \sum_{B=1}^{2n} \frac{\partial N_A}{\partial \eta_B}\, \dot{\eta}_B \equiv \sum_{B=1}^{2n} M_{AB}(\eta)\, \dot{\eta}_B$$

o en notación matricial

$$|\dot{\eta}> \longrightarrow |\dot{N}> = \mathbf{M}\, |\dot{\eta}> \tag{14.16}$$

Por otro lado es

$$\frac{\partial f}{\partial \eta_A} = \sum_{B=1}^{2n} \frac{\partial f}{\partial N_B} \frac{\partial N_B}{\partial \eta_A} = \sum_{B=1}^{2n} \frac{\partial f}{\partial N_B} M_{BA}(\eta) \ ,$$

que en notación matricial se escribe[4]

$$< \frac{\partial f}{\partial \eta} | \ = < \frac{\partial f}{\partial N} | \ \mathbf{M} \tag{14.17}$$

o bien

$$| \frac{\partial f}{\partial \eta} > = \ \mathbf{M}^T \ | \frac{\partial f}{\partial N} > \tag{14.18}$$

Queremos saber qué característica debe tener la matriz $\mathbf{M}$, cuyas componentes son las derivadas de las variables nuevas respecto de las variables viejas, para que las ecuaciones de Hamilton no cambien de forma. Si multiplicamos la ecuación (14.11) a izquierda por $\mathbf{M}$ obtenemos

$$\mathbf{M} \ |\dot{\eta} > = \ \mathbf{M} \ \omega \ | \frac{\partial H}{\partial \eta} >$$

es decir,

$$|\dot{N} > = \ \mathbf{M} \ \omega \ \mathbf{M}^T \ | \frac{\partial H}{\partial N} >$$

Entonces, para mantener la forma de las ecuaciones de Hamilton se requiere que

$$\mathbf{M} \ \omega \ \mathbf{M}^T \ = \ \omega \tag{14.19}$$

La matriz antisimétrica ω define una estructura geométrica en el espacio de las fases, llamada **estructura simpléctica** ($\sigma\upsilon\mu\pi\lambda o\kappa o\varsigma$ significa *complejo*), que debe ser preservada por la matriz $\mathbf{M}$. Las matrices de $2n \times 2n$ que cumplen la ecuación (14.19) se denominan **matrices simplécticas**. Las transformaciones de variables canónicas tales que $\mathbf{M}$ resulta simpléctica se denominan **transformaciones canónicas**. Vemos entonces que existen distintos juegos de variables canónicas para un mismo sistema físico, lo que permite elegir el juego de variables que sea más apto para la integración de las ecuaciones de movimiento.[5]

4. La matriz $\mathbf{M}$ va a la izquierda en (14.16), y a la derecha en (14.17), debido a la forma en que se combinan los índices de fila y columna en las sumatorias respectivas. En lenguaje de *geometría diferencial*, estas dos ecuaciones representan la transformación de las componentes de un vector y una *forma diferencial* ante cambio de variables.

5. La ecuación (14.19) puede compararse con la definición de las rotaciones, que podemos escribir $\mathbf{R} \ \mathbf{1} \ \mathbf{R}^T = \mathbf{1}$. Las rotaciones en el espacio físico preservan la estructura euclidiana del mismo, relacionando distintos juegos de coordenadas cartesianas. Al igual que las rotaciones, las matrices simplécticas forman un grupo.

Propiedades

$i)$ El corchete de Poisson es invariante ante transformaciones canónicas:

$$\{f,g\}_{q,p} = < \frac{\partial f}{\partial \eta}|\omega|\frac{\partial g}{\partial \eta} > = < \frac{\partial f}{\partial N}|\mathbf{M}\ \omega\ \mathbf{M}^T|\frac{\partial g}{\partial N} > = < \frac{\partial f}{\partial N}|\omega|\frac{\partial g}{\partial N} > = \ \{f,g\}_{Q,P}$$

Esto significa que el corchete de Poisson tiene el mismo valor ya sea que se lo calcule con las variables viejas (q_μ, p_μ) o con las variables nuevas (Q_μ, P_μ).

$ii)$ $|\det \mathbf{M}| = 1$. En efecto,

$$\det \omega = \det(\mathbf{M}\ \omega\ \mathbf{M}^T) = (\det \mathbf{M})^2 \det \omega \ \Rightarrow\ \det \mathbf{M} = \pm 1$$

Esto significa que el volumen canónico en el espacio de las fases se escribe en la misma forma en cualquier juego de variables canónicas:

$$\begin{aligned} dQ_1...dQ_n\ dP_1...dP_n &= dN_1...dN_{2n} = \left|\det \frac{\partial N_A}{\partial \eta_B}\right|\ d\eta_1...d\eta_{2n} = |\det \mathbf{M}|\ d\eta_1...d\eta_{2n} \\ &= dq_1...dq_n\ dp_1...dp_n \end{aligned}$$

La invariancia del volumen canónico no es más que el último integrante de una secuencia de invariantes conocidos como **invariantes de Poincaré**. En efecto, la independencia de las variables canónicas escogidas también vale para

$$\begin{aligned} \mathcal{J}_1 &\equiv \int\int \sum_\mu dq_\mu dp_\mu \\ \mathcal{J}_2 &\equiv \int\int\ \int\int \sum_{\mu<\nu} dq_\mu dp_\mu\ \ dq_\nu dp_\nu \\ &... \\ \mathcal{J}_n &\equiv \int\int\ ...\ \int\int dq_1 dp_1\ \ ...\ \ dq_n dp_n\ \ , \end{aligned}$$

lo que muestra que la estructura simpléctica habla de una geometría de áreas y volúmenes.

Ejemplo. Veamos qué significa que $\mathbf{M}$ sea simpléctica en un caso de un solo grado de libertad, cuando $\mathbf{M}$ es una matriz de 2×2:

$$\mathbf{M} = \begin{pmatrix} \frac{\partial N_1}{\partial \eta_1} & \frac{\partial N_1}{\partial \eta_2} \\ \frac{\partial N_2}{\partial \eta_1} & \frac{\partial N_2}{\partial \eta_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix}$$

Entonces resulta que

$$\mathbf{M}\ \omega\ \mathbf{M}^T = \begin{pmatrix} 0 & \{Q,P\}_{q,p} \\ \{P,Q\}_{q,p} & 0 \end{pmatrix}$$

Para que $\mathbf{M}$ sea simpléctica la transformación $Q = Q(q,p)$, $P = P(q,p)$ debe cumplir que

$$\{Q,P\}_{q,p} = 1$$

La transformación es canónica si las nuevas variables conjugadas tiene corchete de Poisson igual a 1, como debe ser.

Como ejemplo sencillo, veamos una *transformación de punto*

$$Q = Q(q)\ ,$$

es decir, un cambio de coordenadas en el espacio de configuración. ¿Cómo debe ser la transformación $P = P(q,p)$ para que el cambio de variables resulte canónico? Se debe preservar el corchete de Poisson,

$$1 = \frac{\partial Q}{\partial q}\,\frac{\partial P}{\partial p} - \frac{\partial P}{\partial q}\,\frac{\partial Q}{\partial p} = \frac{dQ}{dq}\,\frac{\partial P}{\partial p}\ ;$$

una posible solución es

$$P(q,p) = \frac{p}{\frac{dQ}{dq}}$$

que satisface que $P\ dQ = p\ dq$.

Lo dicho hasta aquí nos enseña que existen muchos juegos posibles de variables canónicas. También nos da herramientas para verificar si una transformación es o no canónica. Pero no nos da un método para construir este tipo de transformaciones. En el siguiente Capítulo veremos un método para generar transformaciones canónicas.

Material histórico

W.R. Hamilton, *On a General Method in Dynamics; by which the Study of all free Systems of attracting or repelling Points is reduced to the Search and Differentiation of one central Relation, or characteristic Function*, Philosophical Transactions of the Royal Society of London **124**, 247-308 (1834)
https://www.jstor.org/stable/108066?seq=1

15 Generatrices de transformaciones canónicas

Llamamos transformaciones canónicas a los cambios de variables en el espacio de las fases,

$$(q_\mu,\ p_\mu) \quad \longrightarrow \quad (Q_\mu(q_\mu, p_\mu, t),\ P_\mu(q_\mu, p_\mu, t))\,,$$

que mantienen la forma de las ecuaciones de Hamilton. Como ya vimos, este tipo de cambio de variables involucra una matriz simpléctica $\mathbf{M}$ cuyas componentes corresponden a las derivadas de las variables nuevas respecto de las variables viejas. El carácter simpléctico que debe tener la matriz del cambio de variables nos da una manera de verificar si la transformación es canónica o no. Veremos ahora un método para construir transformaciones canónicas.

Para que las ecuaciones de Hamilton tengan la misma forma tanto en las viejas como en las nuevas variables, la acción debe tener la forma canónica en ambos juegos de variables:

$$\sum_{\mu=1}^{n} p_\mu\ dq_\mu - H(q,p,t)\ dt = \sum_{\mu=1}^{n} P_\mu\ dQ_\mu - \overline{H}(Q,P,t)\ dt + dF \tag{15.1}$$

donde dF es un término de borde que puede agregarse sin por ello alterar la forma de las ecuaciones. Nótese que la propia expresión indica que F debe verse como una función $F(q,Q,t)$ pues

$$dF = \sum_{\mu=1}^{n}(p_\mu\ dq_\mu - P_\mu\ dQ_\mu) - (H - \overline{H})dt \tag{15.2}$$

Entonces

$$p_\mu = \frac{\partial F}{\partial q_\mu}\ , \qquad P_\mu = -\frac{\partial F}{\partial Q_\mu}\ , \qquad \overline{H} - H = \frac{\partial F}{\partial t} \tag{15.3}$$

Esto significa que podemos generar una transformación canónica mediante la elección de una **función generatriz** $F(q,Q,t)$; luego la ecuación (15.3) define una transformación $p_\mu = p_\mu(q,Q,t)$; $P_\mu = P_\mu(q,Q,t)$ que automáticamente es canónica. Para que el cambio de variables esté bien definido tiene que ser biyectivo, es decir, tiene que ser inversible para despejar las coordenadas nuevas en función de las viejas y viceversa.

Ejemplo: sea $F(q,Q) = \sum_{\nu=1}^{n} q_\nu Q_\nu$, entonces $p_\mu = Q_\mu$, $P_\mu = -q_\mu$ es una transformación canónica. Evidentemente lo es, ya que es tan sólo un intercambio de los roles entre coordenadas y momentos conjugados. Las ecuaciones de Hamilton no sufrirán cambios de forma, porque coordenadas y momentos aparecen en pie de igualdad en esas ecuaciones.

A diferencia de lo visto en el Capítulo anterior, ahora estamos permitiendo que la transformación canónica dependa del tiempo. Esto no incide en la forma que debe tener la relación entre variables viejas y nuevas para que la transformación sea canónica; simplemente las dos primeras ecuaciones en (15.3) contienen ahora un parámetro t. Por lo tanto, no se alteran las propiedades de la matriz $\mathbf{M}$, que se construye con las derivadas de las coordenadas nuevas respecto de las viejas. En particular no se altera la invariancia del corchete de Poisson, ni la invariancia del volumen canónico (que resulta de $|\det \mathbf{M}| = 1$) o el resto de los invariantes de Poincaré.[1] La presencia de t en la función generatriz F introducirá una dependencia temporal adicional en la evolución de las nuevas variables canónicas que será absorbida por el reemplazo de H por $\overline{H}$, como muestra la tercera ecuación de (15.3), manteniendo así la forma de las ecuaciones de Hamilton.

15.1 Tipos de funciones generatrices

La forma de la ecuación (15.2) nos indujo a considerar a (q,Q) como las variables independientes de la generatriz F. Pero pueden considerarse otros tipos de dependencia funcional. Por eso llamaremos $F_1(q,Q,t)$ a ese primer tipo de función generatriz. Sin embargo podemos considerar otras dependencias funcionales, siempre combinando una variable "vieja" con otra "nueva". Por ejemplo, si reemplazamos

$$P_\mu \; dQ_\mu = d(P_\mu Q_\mu) - Q_\mu \; dP_\mu$$

en la ecuación (15.2) obtenemos

1. Nótese que $\sum(p_\mu dq_\mu - P_\mu dQ_\mu) = \sum(\frac{\partial F(q,Q,t)}{\partial q_\mu} dq_\mu + \frac{\partial F(q,Q,t)}{\partial Q_\mu} dQ_\mu)$ es un diferencial exacto en el espacio de las fases. Por lo tanto, si a un dado t integramos sobre un camino cerrado del espacio de las fases obtendremos una invariancia,

$$\oint \sum_{\mu=1}^{n} p_\mu \; dq_\mu = \oint \sum_{\mu=1}^{n} P_\mu \; dQ_\mu$$

que equivale a la invariancia del primer invariante de Poincaré $\mathcal{J}_1$.

$$d\left(F+\sum_{\mu=1}^{n}P_\mu Q_\mu\right)=\sum_{\mu=1}^{n}(p_\mu\ dq_\mu+Q_\mu\ dP_\mu)-(H-\overline{H})dt$$

La función "F_2" diferenciada en el primer miembro es una transformada de Legendre de $-F$. Como se ve a la derecha, es una función de (q,P,t). Diremos que se trata de una transformación de tipo 2. Así, cada elección de una función $F_2(q,P,t)$ define una transformación canónica a través de las relaciones

$$p_\mu=\frac{\partial F_2}{\partial q_\mu}\ ,\qquad Q_\mu=\frac{\partial F_2}{\partial P_\mu}\ ,\qquad \overline{H}-H=\frac{\partial F_2}{\partial t}$$

Como siempre, las relaciones entre coordenadas nuevas y viejas deben ser inversibles.

Ejemplo:

$$F_2(q,P,t)=\sum_{\nu=1}^{n}\ f_\nu(q,t)\ P_\nu$$

entonces

$$p_\mu=\sum_{\nu=1}^{n}\frac{\partial f_\nu}{\partial q_\mu}\ P_\nu\ ,\qquad Q_\mu=f_\mu(q,t)$$

que es una transformación de punto –las Q's dependen sólo de las q's– dependiente del tiempo. La dependencia temporal explícita de las nuevas variables impone un cambio de Hamiltoniano:

$$\overline{H}=H+\frac{\partial F_2}{\partial t}=H+\sum_{\nu=1}^{n}\frac{\partial f_\nu}{\partial t}\ P_\nu\equiv H+\sum_{\nu=1}^{n}\ g_\nu(Q,t)\ P_\nu$$

Si $f_\nu(q)=q_\nu$ resulta la transformación **identidad**: $p_\mu=P_\mu$, $Q_\mu=q_\mu$. Las transformaciones de punto que no dependen de t satisfacen $\sum p_\mu\ dq_\mu=\sum df_\nu\ P_\nu=\sum P_\nu\ dQ_\nu$.

El truco usado para construir una generatriz tipo 2 se puede repetir para obtener generatrices $F_3(p,Q,t)$ y $F_4(p,P,t)$. Si en (15.2) ahora reemplazamos

$$p_\mu\ dq_\mu=d(p_\mu q_\mu)-q_\mu\ dp_\mu$$

obtendremos

$$d\left(F-\sum_{\mu=1}^{n}p_\mu q_\mu\right)=\sum_{\mu=1}^{n}(-q_\mu\ dp_\mu-P_\mu\ dQ_\mu)-(H-\overline{H})dt$$

La función diferenciada a la izquierda es F_3. Según vemos a la derecha, depende de (p, Q, t), y se cumple

$$-q_\mu = \frac{\partial F_3}{\partial p_\mu} \ , \qquad -P_\mu = \frac{\partial F_3}{\partial Q_\mu} \ , \qquad \overline{H} - H = \frac{\partial F_3}{\partial t}$$

Finalmente, la generatriz de tipo 4 resulta de combinar ambos reemplazos en (15.2):

$$d\left(F - \sum_{\mu=1}^{n}(p_\mu q_\mu - P_\mu Q_\mu)\right) = \sum_{\mu=1}^{n}(-q_\mu \ dp_\mu + Q_\mu \ dP_\mu) - (H - \overline{H})dt$$

La función F_4 diferenciada a la izquierda depende sólo de (p, P, t), como se ve a la derecha. Así resulta que una transformación canónica también puede generarse eligiendo una $F_4(p, P, t)$ y empleando las relaciones

$$-q_\mu = \frac{\partial F_4}{\partial p_\mu} \ , \qquad Q_\mu = \frac{\partial F_4}{\partial P_\mu} \ , \qquad \overline{H} - H = \frac{\partial F_4}{\partial t}$$

15.2 Transformaciones infinitesimales

Si agregamos un término infinitesimal a la generatriz de la transformación identidad tendremos una transformación canónica infinitesimal:

$$F_2(q, P, t) = \sum_{\nu=1}^{n} q_\nu \ P_\nu + \varepsilon \ G(q, P, t)$$

donde ε es un parámetro infinitesimal. La transformación resulta

$$\begin{aligned} Q_\mu &= \frac{\partial F_2}{\partial P_\mu} = q_\mu + \varepsilon \ \frac{\partial G(q, P, t)}{\partial P_\mu} \ , \\ p_\mu &= \frac{\partial F_2}{\partial q_\mu} = P_\mu + \varepsilon \ \frac{\partial G(q, P, t)}{\partial q_\mu} \end{aligned}$$

Al orden más bajo en ε, no cometeremos error si sustituimos P por p como argumento de G. En ese caso, si elegimos una función $G(q, p, t)$ entonces la transformación infinitesimal

$$\begin{aligned} Q_\mu &= q_\mu + \varepsilon \ \frac{\partial G(q, p, t)}{\partial p_\mu} + \mathcal{O}(\varepsilon^2) \\ P_\mu &= p_\mu - \varepsilon \ \frac{\partial G(q, p, t)}{\partial q_\mu} + \mathcal{O}(\varepsilon^2) \end{aligned}$$

es una transformación canónica infinitesimal.

15.3 La evolución del sistema como una transformación canónica

En la transformación canónica infinitesimal que acabamos de ver, elijamos

$$G(q,p,t) = H(q,p,t) \ , \qquad\qquad \varepsilon = \delta t$$

y reemplacemos

$$\begin{aligned} Q_\mu &\simeq q_\mu + \delta t \, \frac{\partial H}{\partial p_\mu} = q_\mu + \delta t \ \dot{q}_\mu \ , \\ P_\mu &\simeq p_\mu - \delta t \, \frac{\partial H}{\partial q_\mu} = p_\mu + \delta t \ \dot{p}_\mu \ , \end{aligned}$$

es decir

$$\begin{aligned} Q_\mu(t) &\simeq q_\mu(t+\delta t) \ , \\ P_\mu(t) &\simeq p_\mu(t+\delta t) \ . \end{aligned}$$

Esto significa que la evolución del sistema puede verse como una sucesión de transformaciones canónicas infinitesimales cuyo generador G es el Hamiltoniano H. En efecto, la transformación canónica $(q,p) \to (Q,P)$ en el instante t nos da por resultado el valor de las variables viejas en el instante $t + \delta t$.

15.3.1 Teorema de Liouville

Como la evolución del sistema en el espacio de las fases puede verse como el resultado de una transformación canónica entonces satisface las propiedades de estas transformaciones. Como el volumen canónico en el espacio de las fases es un invariante de toda transformación canónica, entonces se cumple que si a $t = t_o$ consideramos un conjunto de estados iniciales que llenan una cierta región del espacio de las fases, los respectivos estados finales a $t = t_f$ llenarán una región de igual volumen canónico.

15.4 Relación entre simetría y conservación

Una transformación canónica infinitesimal cambia infinitesimalmente el valor de una magnitud $f(q,p,t)$:

$$\delta f = \sum_{\mu=1}^{n} \left(\frac{\partial f}{\partial q_\mu}\, \delta q_\mu + \frac{\partial f}{\partial p_\mu}\, \delta p_\mu \right) = \varepsilon \sum_{\mu=1}^{n} \left(\frac{\partial f}{\partial q_\mu}\, \frac{\partial G}{\partial p_\mu} - \frac{\partial f}{\partial p_\mu}\, \frac{\partial G}{\partial q_\mu} \right) = \varepsilon\, \{f, G\}$$

Diremos que una transformación canónica infinitesimal es una **simetría** del Hamiltoniano si

$$\delta H = 0 \ ,$$

es decir, si

$$\{H, G\} = 0 \ .$$

Por otro lado sabemos que una magnitud $G(q,p)$ es una constante de movimiento si $\{G, H\} = 0$. Entonces concluimos que las constantes de movimiento generan simetrías del Hamiltoniano.

Ejemplo: si q es una coordenada cíclica entonces la transformación $q \to q + \varepsilon$ es una simetría del Hamiltoniano. El generador de esa simetría es el momento p canónicamente conjugado a q (en efecto, es $\delta q = \varepsilon\, \{q, p\} = \varepsilon$), que como sabemos es una magnitud conservada en tal caso.

16 Ecuación de Hamilton-Jacobi

16.1 Integrabilidad

Hasta ahora vimos varios ejemplos de sistemas cuyas ecuaciones de movimiento pueden ser integradas. Por ejemplo, sistemas conservativos de un único grado de libertad, osciladores lineales acoplados, etc. En el sistema de dos cuerpos vimos cómo valernos de las magnitudes conservadas para resolver la dinámica en términos de *cuadraturas*, donde la solución queda expresada mediante integrales que pueden ser conocidas o no, pero que en el peor de los casos se pueden evaluar numéricamente. Esto puede crear la falsa idea de que todos los sistemas Hamiltonianos son integrables. Pero la integrabilidad de las ecuaciones de Hamilton no está garantizada en general, y no existe un método para chequear la integrabilidad de un dado Hamiltoniano. Aun así, contamos con el siguiente criterio que es satisfecho por todos los sistemas integrables conocidos:

Integrabilidad de Liouville

Si en un sistema Hamiltoniano autónomo (se dice así cuando H no depende explícitamente de t) de n grados de libertad se conocen n primeras integrales[1] $f_\mu(q,p)$ independientes[2] en *involución*, es decir

$$\{f_\mu(q,p)\ ,\ f_\nu(q,p)\} = 0\ , \qquad \forall\, \mu,\nu\,,$$

entonces el sistema es integrable.

Bajo esas condiciones será posible encontrar una transformación canónica $(q,p) \to (Q,P)$ tal que

$$P_\mu = f_\mu(q,p)$$

Como las f_μ's son conservadas, el Hamiltoniano $\overline{H} = H$ no podrá depender de las Q_μ's para así satisfacer que $\dot{P}_\mu = -\partial\overline{H}/\partial Q_\mu = 0$. Por otro lado tendremos que las $\dot{Q}_\mu$'s son constantes de movimiento pues

1. Una *primera integral* o *integral de movimiento* es una magnitud $f(q,p)$ que no depende explícitamente del tiempo, y se conserva durante la evolución del sistema (es decir que $\{f,H\} = 0$). No toda constante de movimiento es una primera integral; por ejemplo en el movimiento de partícula libre se conserva $q - (p/m)t$, que no es una primera integral.
2. Las formas diferenciales df_μ son linealmente independientes.

$$\dot{Q}_\mu = \frac{\partial \overline{H}(P)}{\partial P_\mu} = \dot{Q}_\mu(P) \equiv \nu_\mu$$

Entonces las soluciones $Q_\mu(t)$ son

$$Q_\mu(t) = \nu_\mu \; t + Q_{\mu \; o}$$

(evolución uniforme sobre rectas del espacio de configuración descripto por las Q_μ's).

Por supuesto, la dificultad reside en encontrar las $f_\mu(q,p)$ y la transformación canónica respectiva.

16.2 Ecuación de Hamilton-Jacobi

Una estrategia para resolver las ecuaciones de Hamilton consiste en buscar una transformación canónica $(q,p) \to (Q(q,p,t),\ P(q,p,t))$ que lleve el Hamiltoniano $\overline{H}(Q,P,t)$ a una forma simple. Así podríamos pedir que la transformación canónica anule $\overline{H}$:

$$H(q,p,t) + \frac{\partial F}{\partial t} = 0$$

Si la generatriz es de tipo 1 o 2, se cumple que

$$p_\mu = \frac{\partial F}{\partial q_\mu} \; .$$

Como estas dos ecuaciones recuerdan las derivadas de la función principal de Hamilton, se acostumbra dar el nombre de S a la generatriz. Entonces, combinando ambas ecuaciones, escribimos

$$\boxed{H\left(q_\mu, \frac{\partial S}{\partial q_\mu}, t\right) + \frac{\partial S}{\partial t} = 0} \tag{16.1}$$

que conocemos como **ecuación de Hamilton-Jacobi**. Pero mientras la función principal de Hamilton $S(q_1,t_1;q_2,t_2)$ está sujeta a un doble juego de ecuaciones –un juego para la configuración final y otro para la inicial– la ecuación (16.1) admite una variedad mucho más amplia de soluciones. Por tratarse de una ecuación diferencial en derivadas parciales de $n+1$ variables (q_μ, t), sus soluciones pueden poseer hasta $n+1$ constantes de integración arbitrarias; en ese caso se dice que la solución es **completa**. Una de las

constantes de integración es necesariamente aditiva, pues si $S(q,t)$ es solución de (16.1) también lo es $S(q,t)+cte$. Nos interesa obtener soluciones completas porque las restantes n constantes de integración de la solución completa serán vistas como las variables P_μ's de la generatriz $S(q,P,t)$ de la transformación canónica que lleva a $\overline{H}=0$ (la constante de integración aditiva no juega ningún papel en esa transformación).[3]

Nótese que $\overline{H}=0$ implica que las nuevas variables son constantes de movimiento ($\dot{Q}_\mu = 0 = \dot{P}_\mu$). Por lo tanto la resolución de la ecuación de Hamilton-Jacobi nos llevará a encontrar $2n$ constantes de movimiento que evidentemente no pueden ser sino combinaciones de los $2n$ valores iniciales de las viejas variables, que caracterizan completamente cada solución de las ecuaciones de Hamilton. Una vez hallada la solución completa $S(q,P,t)$, escribiremos la transformación canónica correspondiente,

$$p_\mu = \frac{\partial S}{\partial q_\mu} \ , \qquad Q_\mu = \frac{\partial S}{\partial P_\mu} \ , \qquad \overline{H} = 0 \ ,$$

de donde despejaremos

$$q_\mu = q_\mu(Q,P,t) \ , \qquad p_\mu = p_\mu(Q,P,t) \ .$$

Como (Q_μ, P_μ) son constantes de movimiento de la evolución real del sistema, entonces las relaciones anteriores nos dan directamente las soluciones $q_\mu(t)$, $p_\mu(t)$ de las ecuaciones de Hamilton en las variables originales.

En suma, el problema de resolver las $2n$ ecuaciones de Hamilton puede ser reemplazado por el problema de encontrar una única función $S(q,P,t)$. Sin embargo para que la solución $S(q,P,t)$ de la ecuación de Hamilton-Jacobi pueda usarse como generatriz de una transformación canónica que lleve a $\overline{H}=0$ deberá ser una solución completa, en el sentido que deberá contener n constantes de integración no aditivas que pasarán a jugar el papel de las P_μ's de la transformación canónica.[4] En realidad, la ecuación de Hamilton-Jacobi puede ser tan difícil de resolver como las ecuaciones de Hamilton, pero veremos luego que nos proveerá de herramientas para tratar perturbativamente sistemas que estén "próximos" a un caso integrable. La importancia de las ideas de Hamilton y Jacobi reside en que abrieron nuevas perspectivas para entender el funcionamiento de los sistemas físicos.

3. Llamando P_μ's a las constantes de integración estamos dando el carácter de generatriz de tipo 2 a la solución $S(q,P,t)$. Algunos autores llaman Q_μ's a las constantes de integración; de esa forma la solución $S(q,Q,t)$ es vista como una generatriz de tipo 1.

4. $\det[\partial^2 S/\partial q_\mu \partial P_\nu] \neq 0$ para que el cambio de variables sea inversible.

16.3 Sistema conservativo de 1 grado de libertad

Ejemplificaremos el procedimiento a seguir mediante el oscilador armónico. La ecuación de Hamilton-Jacobi en ese caso es

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{kq^2}{2} + \frac{\partial S}{\partial t} = 0$$

Siempre que el sistema sea autónomo (H no depende explícitamente de t) la solución se puede plantear como[5]

$$S(q_\mu, P_\mu, t) = S_o(q_\mu, P_\mu) - E\ t$$

donde E es una de las n constantes de integración de la solución completa, digamos $P_n \equiv E$, o es una combinación de las mismas $E = E(P_\mu)$ (esto depende de cómo se organice el conjunto de constantes de integración). En los casos de un único grado de libertad, E es la única constante de integración. Entonces la ecuación de Hamilton-Jacobi toma la forma

$$\frac{1}{2m}\left(\frac{dS_o}{dq}\right)^2 + \frac{kq^2}{2} = E$$

que se integra por cuadraturas,

$$S_o = \int \sqrt{2m\left(E - \frac{kq^2}{2}\right)}\ dq$$

(elegimos el signo correspondiente a $p = \partial S/\partial q > 0$).

Ahora haremos la transformación canónica, para lo cual la constante de integración E será vista como una de las variables de la generatriz:

$$p = \frac{\partial S}{\partial q} = \frac{\partial S_o}{\partial q} = \sqrt{2m\left(E - \frac{kq^2}{2}\right)}\ , \qquad\qquad Q = \frac{\partial S}{\partial P} = \frac{\partial S}{\partial E} = \frac{\partial S_o}{\partial E} - t\ ,$$

entonces

$$t{+}Q = \frac{\partial S_o}{\partial E} = \frac{\partial}{\partial E}\int \sqrt{2m(E - \frac{kq^2}{2})}\ dq = m\int \frac{dq}{\sqrt{2m(E - \frac{kq^2}{2})}} = \omega_o^{-1}\arcsin\left(\sqrt{\frac{k}{2E}}\ q\right)$$

de donde resulta

$$q = \sqrt{\frac{2E}{k}}\sin\omega_o(t + Q) \tag{16.2}$$

5. Siempre que una coordenada q sea cíclica en el Hamiltoniano, la dependencia de S en q será lineal. El factor que acompañe a q será una de las constantes de integración P's.

Además se obtiene

$$p = \sqrt{2mE}\cos\omega_o(t+Q) \tag{16.3}$$

Como fue dicho, cuando el sistema evoluciona las nuevas variables Q y $P = E$ permanecen constantes. Por lo tanto las expresiones obtenidas para q, p automáticamente dan las funciones $q(t)$, $p(t)$ que resuelven las ecuaciones de Hamilton.

Invariancia del área

Las transformaciones canónicas preservan el volumen canónico. Si el sistema tiene un único grado de libertad, el espacio de las fases tiene 2 dimensiones y el volumen deviene en un área. Consideremos el área delimitada por la curva $H(q,p) = E$, que es una elipse en coordenadas q, p. El área de una elipse es π por el producto de los semiejes. Como los semiejes corresponden a las amplitudes de q, p, el área es

$$\acute{A}rea = \pi\sqrt{\frac{2E}{k}}\sqrt{2mE} = \frac{2\pi}{\omega_o}\,E$$

En las nuevas variables la elipse $H(q,p) = E$ devino en la recta $P = E$, que no encierra ninguna área. ¿Qué significa la invariancia del área en este caso? Debemos entender que, en cada t, la transformación canónica (16.2)-(16.3) es univaluada si $0 \leq \omega_o(t+Q) < 2\pi$. Es decir que la univaluación de la transformación de variables limita el rango de Q a $\Delta Q = 2\pi/\omega_o$. Entonces, los puntos interiores de la elipse, que corresponden a energías que van desde 0 hasta E ($\Delta P = E$) son mapeados a un rectángulo de altura E y base $2\pi/\omega_o$, cuya área coincide con la de la elipse.[6]

16.4 Separación de variables

En sistemas conservativos de $n > 1$ grados de libertad, el método para resolver la ecuación de Hamilton-Jacobi, cuando esto sea posible, consiste en encontrar coordenadas adecuadas para **separar** la ecuación en n ecuaciones de primer orden que dependan de una **única coordenada**, las que se integrarán por cuadraturas. Por ejemplo, el problema de una partícula en un campo central es separable en coordenadas esféricas; en efecto,

$$T = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right),$$

$$p_r = m\dot{r}\ , \qquad p_\theta = mr^2\dot{\theta}\ , \qquad p_\phi = mr^2\sin^2\theta\,\dot{\phi}$$

6. El centro de la elipse, $q = 0 = p$, es un punto singular de la transformación pues es mapeado a la recta $P = 0$.

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2\theta} \right) + V(r)$$

La ecuación de Hamilton-Jacobi es

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial S}{\partial \theta} \right)^2 + \frac{1}{r^2 \sin^2\theta} \left(\frac{\partial S}{\partial \phi} \right)^2 \right] + V(r) + \frac{\partial S}{\partial t} = 0$$

Para intentar separar la ecuación proponemos la solución (nótese que ϕ, t son cíclicas)

$$S = S_r(r) + S_\theta(\theta) + \alpha_\phi \ \phi - E \ t$$

Entonces reemplazando,

$$\left(\frac{dS_\theta}{d\theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2\theta} = 2mr^2[E - V(r)] - r^2 \left(\frac{dS_r}{dr} \right)^2$$

donde quedan igualadas una función de θ con una función de r (en esto consiste el arte de *separar* la ecuación). La única forma de dar sentido a esta ecuación es que cada miembro sea una misma constante; entonces tendremos dos ecuaciones de primer orden en una sola variable,

$$\left(\frac{dS_\theta}{d\theta} \right)^2 + \frac{\alpha_\phi^2}{\sin^2\theta} = \alpha_\theta^2 \ ,$$

$$2mr^2[E - V(r)] - r^2 \left(\frac{dS_r}{dr} \right)^2 = \alpha_\theta^2 \ ,$$

que se integran por cuadraturas.

Las constantes de integración E, α_θ, α_ϕ, o combinaciones de las mismas, tomarán el rol de variables nuevas cuando se utilice la solución completa S para generar transformaciones canónicas.

17 Variables ángulo-acción

17.1 Variables ángulo-acción

La solución completa de la ecuación de Hamilton-Jacobi $S(q, P, t)$ sirve como generatriz de una transformación canónica que lleva a Hamiltoniano $\overline{H}$ nulo. En tal caso, las nuevas variables Q_μ, P_μ se conservan. Si el sistema es conservativo la solución tiene la forma $S = S_o(q, P) - E(P)\ t$, donde E es una de las n constantes de integración P_μ's de la solución completa (suele elegirse $E \equiv P_n$) o es una combinación de ellas. La función $S_o(q, P)$ debe satisfacer la ecuación

$$H\left(q_\mu, \frac{\partial S_o}{\partial q_\mu}\right) = E\ . \tag{17.1}$$

Si en lugar de S usamos $S_o(q, P)$ como generatriz de una transformación canónica,

$$p_\mu = \frac{\partial S_o}{\partial q_\mu}\ , \qquad Q_\mu = \frac{\partial S_o}{\partial P_\mu}\ , \qquad \overline{H} = H\ , \tag{17.2}$$

entonces la ecuación (17.1) dice que

$$H\left(q_\mu, p_\mu\right) = E(P_1, ..., P_n)$$

y las ecuaciones de Hamilton en las nuevas variables de la transformación generada por $S_o(q, P)$ son

$$\dot{Q}_\mu = \frac{\partial H}{\partial P_\mu} = \frac{\partial E}{\partial P_\mu}\ , \qquad \dot{P}_\mu = -\frac{\partial H}{\partial Q_\mu} = -\frac{\partial E}{\partial Q_\mu} = 0\ .$$

Vemos que las P_μ's se conservan. Entonces las $\dot{Q}_\mu$ también son constantes de movimiento, pues E es una función de las P_μ's conservadas. Las Q_μ evolucionan uniformemente,[1]

1. La diferencia entre las Q_μ's generadas por S y las generadas por S_o resulta de la relación

$$\frac{\partial S_o}{\partial P_\mu} - \frac{\partial S}{\partial P_\mu} = \frac{\partial E}{\partial P_\mu}\ t = \nu_\mu\ t$$

Mientras las Q_μ's generadas por S se conservan, las generadas por S_o evolucionan uniformemente.

$$Q_\mu(t) = \nu_\mu \; t + Q_{\mu\;o} \qquad \text{donde} \qquad \nu_\mu \equiv \frac{\partial E}{\partial P_\mu} \; . \tag{17.3}$$

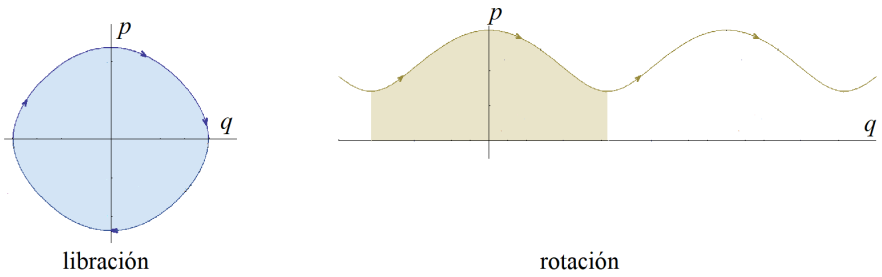
Supongamos ahora que la ecuación (17.1) ha sido integrada mediante separación de variables; es decir, en coordenadas adecuadas la solución completa se escribe

$$S_o = \sum_{\mu=1}^{n} \; S_\mu(q_\mu, P_1, ..., P_{n-1}, P_n) \; ,$$

donde cada término de la suma depende de una única coordenada q, y las funciones S_μ se resuelven por cuadraturas. En ese caso la transformación canónica generada por S_o cumple que el momento conjugado a q_μ,

$$p_\mu = \frac{\partial S_o}{\partial q_\mu} = p_\mu(q_\mu, P_1, ..., P_n), \qquad \mu = 1, ..., n \tag{17.4}$$

depende sólo de q_μ y de las constantes de integración de la ecuación de Hamilton-Jacobi. Así resultan relaciones individuales para cada par de variables conjugadas. Frecuentemente estas relaciones dan curvas cerradas o periódicas en el plano (q, p) correspondiente. Por ejemplo, como vimos en el péndulo simple, podemos tener *libración* o *rotación*:



Si todos los pares de variables conjugadas tienen este tipo de comportamiento podemos definir magnitudes cuyos valores correspondan a las áreas sombreadas:

$$J_\mu \equiv \oint \; p_\mu \; dq_\mu \; , \qquad \mu = 1, ..., n$$

(en el caso de la rotación se integra en uno de los ciclos). Como cada $p_\mu = p_\mu(q_\mu, P_1, ..., P_n)$ depende sólo de su variable conjugada y las constantes de integración, entonces las **variables acción** J_μ's dependen sólo de las constantes de integración:

$$J_\mu = J_\mu(P_1, ..., P_n) \; , \qquad \mu = 1, ..., n \; ,$$

de donde podemos despejar las P_μ's en función de las J_μ's,

$$P_\mu = P_\mu(J_1, ..., J_n) \ , \qquad \mu = 1, ..., n \ ,$$

para escribir la solución completa de la ecuación (17.1) como $S_o(q, J)$. Así escrita, la solución completa $S_o(q, J)$ será usada como generatriz de una transformación canónica $(q_\mu, p_\mu) \to (\mathcal{Q}_\mu, J_\mu)$ que transforme las variables canónicas originales en las variables acción y sus respectivas conjugadas. Es costumbre llamar **variables ángulo** θ_μ a las coordenadas conjugadas a las J_μ multiplicadas por 2π. Como vimos en (17.3), la solución completa de la ecuación (17.1) conduce siempre a variables Q_μ que evolucionan uniformemente, no importa cómo hayamos organizado el conjunto de constantes de integración. En particular esto vale también para las variables ángulo. Pero las variables ángulo gozan de propiedades adicionales, como vamos a ver.

La transformación canónica entrega funciones $\theta_\mu(q, J) = 2\pi\mathcal{Q}_\mu = 2\pi\, \partial S_o(q, J)/\partial J_\mu$. Queremos despejar las coordenadas q's como funciones $q_\mu(\theta, J)$, para así entender cómo evolucionan las q_μ's, dado que sabemos cómo evolucionan las variables ángulo-acción. Nos interesa conocer entonces qué tipo de funciones son las $q_\mu(\theta, J)$. Dados los valores de las J's, los posibles estados del sistema quedan restringidos a los ciclos de la Figura anterior, cuyas áreas son determinadas por las J_μ. Para comprender la relación funcional entre las q's y las θ's, veamos qué cambios experimenta θ_μ cuando una de las q's, digamos q_ν, varía sobre un ciclo mientras las otras q's permanecen fijas:

$$\delta_\nu \theta_\mu \ = \ \frac{\partial \theta_\mu}{\partial q_\nu}\, \delta q_\nu \ = \ 2\pi\, \frac{\partial^2 S_o}{\partial q_\nu\, \partial J_\mu}\, \delta q_\nu \ = \ 2\pi\, \frac{\partial p_\nu}{\partial J_\mu}\, \delta q_\nu$$

donde hemos usado la ecuación (17.2). Entonces en un ciclo de q_ν tenemos que

$$\Delta_\nu \theta_\mu \ = \ 2\pi\, \frac{\partial}{\partial J_\mu} \oint p_\nu\, dq_\nu \ = \ 2\pi\, \frac{\partial J_\nu}{\partial J_\mu} \ = \ 2\pi\, \delta_{\mu\nu} \tag{17.5}$$

donde la derivada respecto de J_μ puede salir fuera de la integral porque ésta se realiza a valores fijos de las J_μ's. El resultado muestra que θ_μ avanza 2π cuando la respectiva q_μ recorre un ciclo completo, permaneciendo fijas las demás coordenadas q. Esto muestra la naturaleza multivaluada de las variables ángulo,[2] a diferencia de las coordenadas q_μ que

2. También la generatriz $S_o(q, J)$ es multivaluada. En cada ciclo de q_μ es

$$\Delta S_o = \oint \frac{\partial S_o}{\partial q_\mu}\, dq_\mu = \oint p_\mu\, dq_\mu = J_\mu \ .$$

$S_o(q, J)$ está determinada a menos de términos que son múltiplos de las J_μ, los que no afectan su condición de ser solución de (17.1).

son univaluadas.[3] Para que las q_μ sean funciones univaluadas de las θ's, las funciones $q_\mu(\theta, J)$ deberán ser multiperiódicas en los ángulos θ con períodos 2π; de esa forma un avance en 2π de cualquiera de los ángulos θ no afectará los valores de las q's. Es decir que, en general, las $q_\mu(\theta, J)$ admiten un desarrollo de Fourier del tipo

$$q_\mu(\theta, J) = \sum_{k_1=-\infty}^{\infty} \ldots \sum_{k_n=-\infty}^{\infty} a_{\mu\ k_1 \ldots k_n}\ e^{i\ (k_1\ \theta_1 + \ldots + k_n\ \theta_n)}$$

donde los coeficientes $a_{\mu\ k_1 \ldots k_n}$ dependen de las variables acción J. Como los momentos p_μ dependen de la respectiva q_μ y de las J's, también ellos son funciones multiperiódicas de las variables ángulo. En general, cualquier función univaluada del estado del sistema $f(q, p)$ admite un desarrollo similar al anterior.

Ejemplo

Según vimos en §16.3, la variable acción del oscilador armónico es

$$J = \oint p\ dq = \textit{Área de la elipse} = \frac{2\pi}{\omega_o}\ E \tag{17.6}$$

y la variable ángulo es

$$\theta\ =\ 2\pi\ \frac{\partial S_o}{\partial J}\ =\ \omega_o\ \frac{\partial S_o}{\partial E}\ =\ \arcsin\left(\sqrt{\frac{k}{2E}}\ q\right)\ =\ \arcsin\left(\sqrt{\frac{\pi m \omega_o}{J}}\ q\right)$$

Por lo tanto q es una función periódica de θ con período 2π:

$$q\ =\ \sqrt{\frac{J}{\pi\ m\ \omega_o}}\ \sin\theta$$

Cuando el sistema evoluciona, las variables ángulo evolucionan uniformemente. Según la ecuación (17.3) resulta $\theta_\mu(t) = 2\pi\nu_\mu\ t + \theta_{\mu_o} = \omega_\mu t + \theta_{\mu_o}$. Por otro lado, las variables acción se conservan. Entonces

$$q_\mu(t) = \sum_{k_1=-\infty}^{\infty} \ldots \sum_{k_n=-\infty}^{\infty} A_{\mu\ k_1 \ldots k_n}\ e^{i\ (k_1\ \omega_1 + \ldots + k_n\ \omega_n)\ t} \tag{17.7}$$

3. Esto vale para la libración. En la rotación, la variable univaluada es $q_\mu - \frac{\theta_\mu}{2\pi}\ q_{\mu o}$, pues q_μ avanza una cantidad fija $q_{\mu o}$ en cada ciclo.

y otro tanto sucede con la evolución de cualquier función univaluada del estado del sistema $f(q,p)$. En particular no sólo las coordenadas que separan la ecuación de Hamilton-Jacobi, sino cualquier tipo de coordenada canónica univaluada evolucionará de esa manera.

Estos resultados son significativos porque la propia transformación canónica nos dice cuánto valen las frecuencias angulares $\omega_\mu = 2\pi\nu_\mu$ involucradas en la evolución del sistema. En efecto la ecuación (17.3) dice que[4]

$$\omega_\mu = 2\pi \, \frac{\partial E(J)}{\partial J_\mu} \tag{17.8}$$

Por ejemplo, en el oscilador armónico la ecuación (17.6) dice que $\omega = 2\pi \, \partial E/\partial J = \omega_o$.

Que las frecuencias del movimiento resulten de calcular las áreas encerradas en los ciclos de las variables canónicas usadas para separar la ecuación de Hamilton-Jacobi, y de escribir la energía en función de esas áreas, es un gran logro de las variables ángulo-acción. Las variables ángulo-acción serán también indispensables para estudiar perturbativamente los sistemas aproximadamente conservativos. Los trabajos de Hamilton (1834) y Jacobi (1836) recibieron el justo reconocimiento cuando Delaunay introdujo las variables ángulo-acción en teoría de perturbaciones aplicada a mecánica celeste (1848), lo que permitió exhibir toda la potencia de la nueva formulación de la Mecánica. Las variables acción fueron utilizadas en los albores de la Mecánica Cuántica para enunciar las reglas de cuantificación de Wilson-Sommerfeld (1915) que condujeron a las *ondas de materia* de de Broglie (1924).

17.1.1 Frecuencias del problema de Kepler

El problema del movimiento ligado de una partícula de masa m en un campo central Kepleriano, $V(r) = -\alpha/r$, se puede separar en coordenadas esféricas (r, θ, ϕ) como ya hemos visto. Una vez calculadas las variables acción, se obtiene que la energía $E(J)$ resulta[5]

$$E = -\frac{2\pi^2 m \, \alpha^2}{(J_r + J_\theta + J_\phi)^2}$$

4. Recordemos que usamos los J's en lugar de los P's, y así las Q's son $\mathcal{Q}_\mu = \theta_\mu/(2\pi)$.

5. Las órbitas que ocurren en el plano $\theta = \pi/2$ corresponden a $J_\theta = 0$ pues el correspondiente ciclo se reduce al punto ($\theta = \pi/2$, $p_\theta = 0$). Cuando la órbita está inclinada un ángulo i, la coordenada θ oscila entre $\pi/2 - i$ y $\pi/2 + i$.

lo que conduce a tres frecuencias iguales:

$$\nu_r = \nu_\theta = \nu_\phi = \frac{4\pi^2 m\ \alpha^2}{(J_r + J_\theta + J_\phi)^3} = \sqrt{\frac{2}{\pi^2 m\ \alpha^2}}\ |E|^{3/2}$$

Siendo las tres frecuencias iguales, entonces la ecuación (17.7) implica que las evoluciones de las coordenadas r, θ, ϕ son periódicas con igual período. La órbita es cerrada, pues al cabo del tiempo $T = \nu^{-1}$ el sistema regresa al estado inicial.

Si recordamos que la relación entre E y a –el semieje mayor de la órbita elíptica– es $|E| = \alpha/(2a)$, obtenemos la 3ra. Ley de Kepler,

$$T^2 = \nu^{-2} = 4\ \pi^2 m\ \alpha^{-1}\ a^3.$$

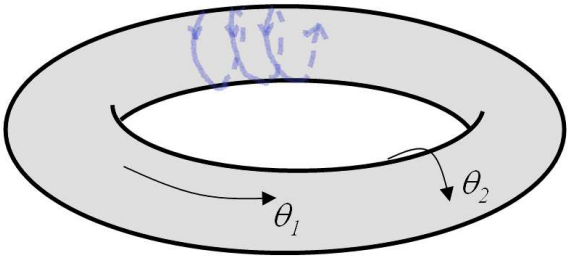
17.2 Sistemas periódicos y condicionalmente periódicos

Las frecuencias ν_μ que caracterizan la evolución del sistema dependen de los valores de las variables acción J_μ. Las variables acción son primeras integrales de movimiento; sus valores quedan determinados por las condiciones iniciales. Si las frecuencias ν_μ son **conmensurables**, es decir que los cocientes entre frecuencias son racionales, entonces las ν_μ se podrán escribir como múltiplos enteros de una frecuencia ν_o (que podemos elegir como el máximo común divisor de las ν_μ's)

$$\nu_\mu = N_\mu\ \nu_o \tag{17.9}$$

es decir que $T_o \equiv \nu_o^{-1}$ es N_μ veces T_μ. Cada período T_μ entra un número entero de veces en T_o. Por lo tanto al cabo del tiempo T_o el sistema regresará a su estado inicial. La evolución del sistema será periódica, y describirá una curva cerrada tanto en el espacio de configuración como en el espacio de las fases. De acuerdo a lo dicho a continuación de la ecuación (17.7), cada función univaluada del sistema evolucionará con una frecuencia que es un múltiplo de ν_o. En el problema de Kepler que acabamos de ver, la conmensurabilidad de las frecuencias se da para cualquier valor de las integrales de movimiento J_μ; pero en un caso general la conmensurabilidad podría darse sólo para ciertas condiciones iniciales. Si las frecuencias no son conmensurables el sistema se dice *múltiple* o *condicionalmente* periódico.

Las n integrales de movimiento J_μ etiquetan un subespacio $n-$dimensional del espacio de las fases en donde la evolución tiene lugar. En ese subespacio evolucionan las variables ángulo, de modo que el subespacio se ve como un toro $n-$dimensional.



A medida que el sistema evoluciona, va trazando una curva sobre el toro. Si las frecuencias son conmensurables, la curva se cerrará. Si las frecuencias son todas inconmensurables, es decir que los cocientes entre frecuencias son todos irracionales, entonces la curva que el sistema desarrolla sobre el toro no se cerrará, e irá cubriendo densamente al toro (pues un irracional se puede aproximar por un racional con la precisión que se desee). Entre estos dos casos existen situaciones intermedias; si entre las frecuencias existen m relaciones de la forma

$$\sum_{\mu=1}^{n} N_{i\mu}\ \nu_\mu = 0\ , \qquad\qquad i = 1, ..., m\ , \tag{17.10}$$

donde los $N_{i\mu}$ son enteros, entonces la evolución del sistema sólo llenará densamente un subespacio del toro de dimensión $n - m$. Si $m = n - 1$, como en (17.9),[6] el subespacio deviene unidimensional (es una curva cerrada).

Cuando las relaciones (17.10) son independientes de los valores de las J_μ se dice que el sistema es m veces **degenerado**. Si $m = n - 1$ el sistema se denomina **completamente degenerado**. Los sistemas completamente degenerados se caracterizan porque la energía E depende de las J_μ's a través de una combinación lineal de las mismas a coeficientes enteros, como sucede en el problema de Kepler y en el oscilador armónico isótropo. Esa característica lleva a que las frecuencias $\nu_\mu = \partial E/\partial J_\mu$ sean conmensurables cualquiera sea el valor de las J's, y las evoluciones siempre describan curvas cerradas. Los sistemas completamente degenerados pueden separarse en más de un tipo de coordenadas.

Bibliografía adicional

M.G. Calkin, *Lagrangian and Hamiltonian Mechanics*, World Scientific.

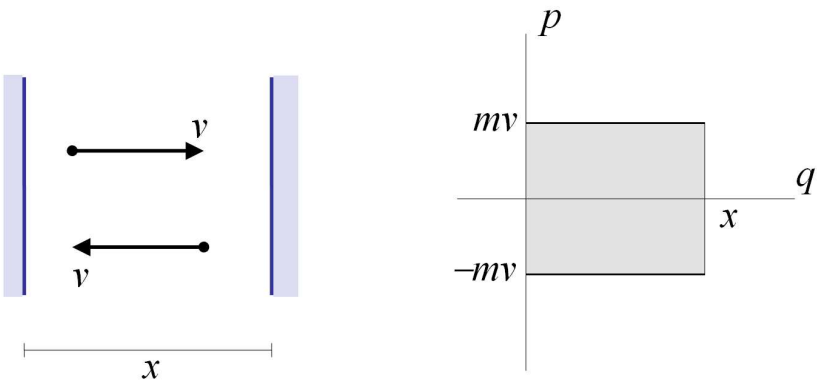
6. En (17.9) hay $m = n$ ecuaciones. Pero una de ellas debe usarse para escribir ν_o en función de alguna de las frecuencias ν_μ del sistema.

18 Invariantes adiabáticos

18.1 Invariancia adiabática

Las soluciones de sistemas conservativos pueden servir para obtener soluciones aproximadas de sistemas no conservativos. Si resolvemos las ecuaciones para el Hamiltoniano conservativo $H(q, p, \lambda)$, donde λ es un parámetro, podremos tener una idea de cómo sería la evolución cuando λ sea una función del tiempo $\lambda(t)$ que varía lentamente. Pero para poder precisar qué significa "lentamente", el sistema conservativo debería contener un tiempo característico. Si el sistema conservativo es periódico con período T entonces $\lambda(t)$ debería cambiar muy poco en el tiempo T; esos cambios lentos se denominan **adiabáticos**. Veremos que las variables acción son casi constantes durante un proceso adiabático.[1]

Para comenzar, mostraremos la **invariancia adiabática** de la variable acción J en un sistema periódico donde una partícula libre rebota elásticamente entre dos paredes separadas por una distancia x, que jugará el papel de un parámetro



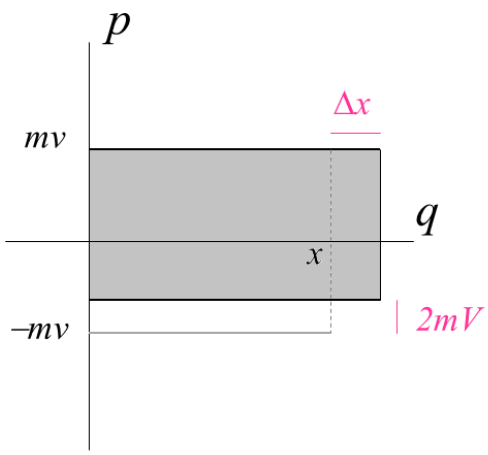
El período de este sistema es $T = 2x/v$, y el valor de la variable acción es $J = 2\ m\ v\ x$. Si ahora dejamos que una de las paredes se mueva a velocidad constante $\dot{x} = V$, entonces la variable acción J resultará afectada, ya que la velocidad de la partícula cambiará

1. No siempre la variación lenta de los parámetros implica que la solución se parece a la del problema conservativo respectivo. Como sabemos, una resonancia paramétrica puede ser el resultado de una variación lenta pero adecuada de un parámetro. Esta cuestión se relaciona con los problemas de convergencia de los desarrollos perturbativos.

al chocar contra la pared móvil, cuya posición habrá cambiado también. Además, la energía de la partícula dejará de conservarse. Sabemos que en un choque elástico se conserva la velocidad relativa, entonces en cada choque con la pared móvil tendremos que

$$v_i - V = v_f + V \qquad \Rightarrow \qquad \Delta v = -2V \qquad \text{ó} \qquad \Delta p = -2mV$$

En la siguiente Figura vemos lo que sucede cuando la pared de la derecha comienza a moverse con velocidad V. La partícula alcanzará la pared en una posición desplazada $x + \Delta x$. Si choca con velocidad v entonces rebotará con velocidad $v - 2V$.



La base del rectángulo aumentará y su altura disminuirá. Si el desplazamiento Δx es pequeño, es decir si $V << v$, podemos estimarlo como $\Delta x \approx VT/2$, siendo $T = 2x/v$ el período del ciclo con paredes fijas:

$$\Delta x \approx V\,\frac{T}{2} = x\,\frac{V}{v} << x$$

Como la variación del parámetro x en un ciclo es $\Delta x << x$, estamos entonces en condiciones adiabáticas. En esta aproximación, la nueva área sombreada (el nuevo valor de J) resulta

$$\begin{aligned} J &= \text{base} \times \text{altura} = (x + \Delta x)\,(2mv - 2mV) \\ &= x\left(1 + \frac{V}{v}\right)\,2mv\left(1 - \frac{V}{v}\right) = 2\,m\,v\,x + \mathcal{O}\left(V^2/v^2\right) \end{aligned}$$

Al orden más bajo en V/v, el valor de J no cambia respecto del que tiene en el caso conservativo. J se comporta como un invariante adiabático pues resulta que $\delta J = \mathcal{O}(\dot{x}^2)$, siendo x el parámetro que varía lentamente en el tiempo; el cambio de J es mucho más

lento que el cambio del parámetro x. Contrariamente, la energía sufre cambios de primer orden en V; en efecto, el cambio de la energía en cada choque es

$$\delta H = \frac{(p-2mV)^2}{2m} - \frac{p^2}{2m} = -2pV + \mathcal{O}(V^2) = \mathcal{O}(\dot{x})$$

Esta situación se reiterará en cada ciclo mientras la pared continúe moviéndose.

18.1.1 Expansión adiabática de un gas ideal

El ejemplo anterior puede aplicarse a la expansión adiabática de un gas de partículas libres en una caja. La invariancia adiabática de $J = 2\ m\ v\ x$ se reproduce en cada una de las tres dimensiones: $v_x x$, $v_y y$, $v_z z$, son aproximadamente constantes de movimiento. Elevando al cuadrado, y promediando sobre las partículas del gas,

$$< v_x^2 >\ x^2 \simeq cte\ , \qquad < v_y^2 >\ y^2 \simeq cte\ , \qquad < v_z^2 >\ z^2 \simeq cte\ .$$

Por isotropía, podemos afirmar que $< v_x^2 >=< v_y^2 >=< v_z^2 >$, que a su vez es proporcional a la temperatura del gas (que es una medida de la energía cinética media del gas). Entonces, multiplicando las tres relaciones anteriores,

$$\text{temperatura}^3\ x^2 y^2 z^2 \simeq cte$$

Es decir

$$\text{temperatura}^3\ \text{volumen}^2 \simeq cte$$

que es la ley de expansión adiabática de los gases ideales monoatómicos.

18.2 Formulación general

Las variables ángulo-acción fueron definidas mediante una transformación canónica generada por la solución S_o de la ecuación de Hamilton-Jacobi independiente del tiempo,

$$H\left(q, \frac{\partial S_o}{\partial q}; \lambda\right) = E(J;\lambda) \tag{18.1}$$

Si el Hamiltoniano adquiriera una dependencia explícita del tiempo a través de la variación lenta del parámetro λ, entonces ya no será cierto que $S = S_o - E\ t$ resuelva la ecuación de Hamilton-Jacobi dependiente del tiempo. Pero nada impide seguir usando la función $S_o(q, J; \lambda(t))$ para generar una transformación canónica. Más aún, $S_o(q, J; \lambda(t))$ seguirá siendo solución de la ecuación (18.1), lo que significa que podremos reemplazar el

Hamiltoniano H por $E(J;\lambda(t))$. Como la generatriz ahora depende del tiempo entonces tendremos un nuevo Hamiltoniano $\overline{H}$ para hacer evolucionar las nuevas variables. La transformación canónica se verá así:

$$p = \frac{\partial S_o(q,J;\lambda)}{\partial q} , \qquad \frac{\theta}{2\pi} = \frac{\partial S_o(q,J;\lambda)}{\partial J} ,$$

$$\overline{H} = H + \frac{\partial S_o(q,J;\lambda)}{\partial t} = E(J;\lambda) + \frac{\partial S_o(q,J;\lambda)}{\partial \lambda}\ \dot{\lambda} \tag{18.2}$$

En las ecuaciones de Hamilton para las variables nuevas, $\overline{H}$ entra como $\overline{H}(\theta,J)$. Aquí debemos cuidar que la derivada $\partial S_o/\partial\lambda$ en el segundo término de $\overline{H}$ se realice antes del reemplazo de q como $q(\theta,J)$. Como la función $q(\theta,J)$ tiene, en general, una dependencia en λ, las operaciones de derivar respecto de λ y reemplazar $q(\theta,J;\lambda)$ no conmutan. En otras palabras, $\partial/\partial\lambda$ deriva la dependencia explícita de S_o en λ, sin que intervenga la dependencia implícita que se introduce a través de q. Las ecuaciones de Hamilton resultan

$$\frac{\dot{\theta}}{2\pi} = \frac{\partial \overline{H}}{\partial J}|_\theta = \frac{\partial E}{\partial J} + \frac{\partial}{\partial J}\left(\frac{\partial S_o(q,J;\lambda)}{\partial \lambda}|_{q=q(\theta,J;\lambda)}\right)\ \dot{\lambda} \tag{18.3}$$

$$\dot{J} = -\frac{\partial \overline{H}}{\partial \frac{\theta}{2\pi}}|_J = -2\pi\ \frac{\partial}{\partial\theta}\left(\frac{\partial S_o(q,J;\lambda)}{\partial \lambda}|_{q=q(\theta,J;\lambda)}\right)\ \dot{\lambda} \tag{18.4}$$

Notemos que J dejó de ser una constante de movimiento; varía lentamente si la variación de λ es lenta. Tampoco vale ya que $\partial E/\partial J$ sea constante en la ecuación para $\dot{\theta}$, ni que θ evolucione uniformemente. Queremos ver que la variación temporal de J, sin embargo, es mucho más lenta que la de λ cuando se la promedia en un ciclo de la variable ángulo. Si suponemos que λ y $\dot{\lambda}$ varían muy poco en un ciclo entonces podremos aproximarlos por sus valores iniciales:

$$\dot{J} \simeq -2\pi\ \ \dot{\lambda}_o\ \left[\frac{\partial}{\partial\theta}\left(\frac{\partial S_o(q,J;\lambda)}{\partial \lambda}|_{q=q(\theta,J;\lambda)}\right)\right]_{\lambda_o}$$

Para promediar $\dot{J}$ en un ciclo aproximaremos J en el integrando por una constante. Esta aproximación está justificada porque $\dot{J}$ ya es de primer orden en $\dot{\lambda}_o$, y basta con calcular a este orden de aproximación. Entonces

$$< \dot{J} > \simeq -\dot{\lambda}_o \int_0^{2\pi} d\theta\ \left[\frac{\partial}{\partial\theta}\left(\frac{\partial S_o(q,J;\lambda)}{\partial \lambda}|_{q=q(\theta,J;\lambda)}\right)\right]_{\lambda_o}$$

Como J y λ son constantes en el integrando, entonces la integral se reduce a evaluar la función entre paréntesis en los extremos de integración. Esa función es del tipo de las

funciones de estado consideradas en sistemas conservativos. Se trata entonces de una función periódica de la variable ángulo θ, lo que implica que $< \dot{J} >$ se anula en primera aproximación.[2] Por lo tanto, $\dot{J}$ no sólo es chica sino que su promedio en un ciclo es aproximadamente cero. Los cambios de J al transcurrir el ciclo tienden a compensarse, evidenciando así el carácter de invariante adiabático de la variable acción J.

Veamos ahora los errores cometidos al aproximar λ y $\dot{\lambda}$ por sus valores iniciales. Para ello usaremos desarrollos en serie de Taylor en $t = 0$:

$$\dot{\lambda}(t) = \dot{\lambda}_o + \ddot{\lambda}_o \, t + ...$$

y [3]

$$\frac{\partial^2 S_o}{\partial q \, \partial \lambda} = \left[\frac{\partial^2 S_o}{\partial q \, \partial \lambda}\right]_{\lambda_o} + \left[\frac{\partial^3 S_o}{\partial q \, \partial \lambda^2}\right]_{\lambda_o} (\lambda - \lambda_o) + ... = \left[\frac{\partial^2 S_o}{\partial q \, \partial \lambda}\right]_{\lambda_o} + \left[\frac{\partial^3 S_o}{\partial q \, \partial \lambda^2}\right]_{\lambda_o} \dot{\lambda}_o \, t + ...$$

donde t corre entre 0 y el período del ciclo. Entonces resulta

$$\frac{\partial^2 S_o}{\partial q \, \partial \lambda} \, \dot{\lambda} = \left[\frac{\partial^2 S_o}{\partial q \, \partial \lambda}\right]_{\lambda_o} \dot{\lambda}_o + \mathcal{O}(\dot{\lambda}_o^2 \, , \, \ddot{\lambda}_o)$$

Por lo tanto

$$< \dot{J} > = \mathcal{O}(\dot{\lambda}_o^2 \, , \, \ddot{\lambda}_o) \, .$$

Comparemos con el ritmo al que cambia la energía $E(J; \lambda)$:

$$\dot{E} = \frac{\partial E}{\partial J} \, \dot{J} + \frac{\partial E}{\partial \lambda} \, \dot{\lambda}$$

Como a orden cero $\partial E / \partial \lambda$ es una constante de movimiento, entonces $< \dot{E} > = \mathcal{O}(\dot{\lambda}_o)$.

Como ejemplo de invariante adiabático mencionemos que la excentricidad de la órbita Kepleriana es un invariante adiabático pues resulta una función de las variables acción, sin la participación de ningún parámetro,[4]

$$e = \sqrt{1 - \frac{J_\phi^2}{(J_r + J_\theta + J_\phi)^2}} \, .$$

En cambio las dimensiones de la órbita serán afectadas por una variación de los parámetros, tales como la masa de la estrella o la constante de gravitación universal.

2. La función $S_o(q, J)$ es multivaluada, pues está determinada a menos de términos que son múltiplos de las J_μ (ver Nota 2 en Capítulo 17). Pero estos términos no contribuyen a $\partial S_o / \partial \lambda$.

3. En el integrando de $< \dot{J} >$ es $\partial^2 S_o / \partial \theta \, \partial \lambda = (\partial^2 S_o / \partial q \, \partial \lambda)(\partial q / \partial \theta)$.

4. Recordemos las expresiones $e = \sqrt{1 + \frac{2 \, \ell^2 E}{\alpha^2 \mu}}$, $E = -\frac{2\pi^2 m \, \alpha^2}{(J_r + J_\theta + J_\phi)^2}$, $J_\phi = \oint p_\phi \, d\phi = 2\pi p_\phi = 2\pi \ell$. No obstante debe mencionarse que los métodos perturbativos para sistemas de más de un grado de libertad presentan aspectos que exceden el marco de lo tratado en esta Sección.

18.3 Oscilador armónico con frecuencia variable

Consideremos un Hamiltoniano de oscilador armónico con frecuencia dependiente del tiempo:

$$H = \frac{p^2}{2m} + \frac{m}{2}\ \omega(t)^2\ q^2$$

Sabemos que la función generatriz que lleva a las variables ángulo-acción es

$$S_o(q, J; \omega) = m\ \omega \int \sqrt{\frac{J}{\pi\ m\ \omega} - q^2}\ dq$$

donde usamos que $E = \omega\ J/(2\pi)$. La relación entre θ y q resulta de

$$\theta = 2\pi\ \frac{\partial S_o}{\partial J}|_q = \arcsin\left(\sqrt{\frac{\pi\ m\ \omega}{J}}\ q\right)$$

es decir

$$q = \sqrt{\frac{J}{\pi\ m\ \omega}}\ \sin\theta \tag{18.5}$$

Además es

$$p = \frac{\partial S_o}{\partial q}|_J = m\ \omega\ \sqrt{\frac{J}{\pi\ m\ \omega} - q^2} = \sqrt{\frac{J\ m\ \omega}{\pi}}\ \cos\theta$$

En las ecuaciones de Hamilton (18.3) y (18.4), $\overline{H}$ debe estar escrito como función de θ, J. Para ello, primero calculamos $\partial S_o(q, J; \omega)/\partial\omega$ y luego reemplazamos q usando (18.5) [5]

$$\overline{H}\ = E(J;\omega) + \frac{\partial S_o}{\partial \omega}|_{q(\theta,J;\omega)}\ \dot{\omega}\ =\ \frac{\omega\ J}{2\pi}\ \left(1 + \frac{\dot{\omega}}{2\omega^2}\ \sin 2\theta\right) \tag{18.6}$$

Entonces las ecuaciones de Hamilton resultan

$$\frac{\dot{\theta}}{2\pi} = \frac{\partial \overline{H}}{\partial J}|_\theta = \frac{\omega}{2\pi}\ \left(1 + \frac{\dot{\omega}}{2\omega^2}\ \sin 2\theta\right)\ , \qquad \dot{J} = -\frac{\partial \overline{H}}{\partial \frac{\theta}{2\pi}}|_J = -\frac{J\dot{\omega}}{\omega}\ \cos 2\theta \tag{18.7}$$

Si $\omega(t)$ y $\dot{\omega}(t)$ varían poco en un ciclo, será suficiente considerar que J, ω y $\dot{\omega}$ son constantes en la expresión de $\dot{J}(t)$ en (18.7).[6] Como vimos en la sección anterior, esta

5. Estos dos pasos no son intercambiables porque q en (18.5) tiene una dependencia en ω.

6. ω varía poco en un ciclo si $\Delta\omega << \omega$, donde $\Delta\omega \simeq \dot{\omega}(0)\ T$, siendo T el período del ciclo. En este ejemplo la propia frecuencia ω es la que da el período del ciclo: $T = 2\pi/\omega$. Entonces la condición de variación lenta de ω se escribe, en este caso, $\dot{\omega} << \omega^2$. Pero al promediar J también decimos que $\dot{\omega}$ varía poco, lo que supone que $\ddot{\omega} << \dot{\omega}\ \omega$.

aproximación introduce errores de orden $\dot{\omega}^2$ y $\ddot{\omega}$ en el cálculo del promedio de $\dot{J}$ en un ciclo, que de esta forma resulta

$$< \dot{J} > \simeq \; -\frac{J\,\dot{\omega}}{2\pi\,\omega} \int_0^{2\pi} d\theta \;\cos 2\theta \;=\; \frac{J\,\dot{\omega}}{4\pi\,\omega} \;\sin 2\theta \;|_0^{2\pi} \;=\; 0$$

Aunque J cambia lentamente mientras transcurre el ciclo, los cambios se promedian a cero en la aproximación más baja. Se dice que J no exhibe cambios *seculares* (es decir, cambios que se acumulen en el tiempo).[7] No sucede así con los cambios de la energía $E = \omega\; J/(2\pi)$; en efecto, $\dot{E}$ es

$$\dot{E} = \frac{1}{2\pi}\;(\dot{\omega}\;J + \omega\;\dot{J}) = \frac{J\,\dot{\omega}}{2\pi}\;(1 - \cos 2\theta)$$

A orden cero, J es una constante de movimiento; entonces $< \dot{E} >= \mathcal{O}(\dot{\omega})$.

Nota. Podemos refinar la forma del invariante adiabático, y lograr una magnitud cuya derivada se anula al orden $\dot{\omega}$ sin necesidad de promediar. En efecto

$$\mathcal{J} \;\doteq\; J(t)\;\left(1 + \frac{\dot{\omega}(t)}{2\omega(t)^2}\right)$$

cumple que

$$\dot{\mathcal{J}} \;=\; J \times \mathcal{O}\left(\frac{\dot{\omega}^2}{\omega^3}, \frac{\ddot{\omega}}{\omega^2}\right)$$

.

La solución para $q(t)$ se obtiene integrando θ en la ecuación (18.7), y reemplazado en (18.5). Al orden más bajo significativo sólo consideraremos el primer término de la expresión para $\dot{\theta}$ en (18.7). La dependencia temporal de ω rompe la evolución temporal uniforme de la variable ángulo; ahora es $\theta(t) \simeq \int \omega(t)\; dt$. Entonces resulta

$$q(t) \;\simeq\; \sqrt{\frac{J(t)}{\pi\; m\; \omega(t)}}\;\sin \int \omega(t)\; dt \;\approx\; \sqrt{\frac{J}{\pi\; m\; \omega(t)}}\;\sin \int \omega(t)\; dt \tag{18.8}$$

7. En la antigua teoría de los cuantos fundada por N. Bohr en 1913, los sistemas multiperiódicos obedecen la regla de cuantificación de Wilson-Sommerfeld (1915), que dice que los invariantes adiabáticos J_μ sólo puede tomar los valores $n_\mu h$, $n_\mu \in \mathbb{N}$. La constante de Planck $h = 6,626 \times 10^{-34}$Js representa el *cuanto* de variable acción. Así, la invariancia adiabática de J es la conservación del número de cuantos ante variación lenta de los parámetros. La idea de invariancia adiabática en la mecánica fue desarrollada por P. Ehrenfest a partir de 1912.

En esta solución aproximada de la ecuación del oscilador armónico con frecuencia lentamente variable, J juega el papel de una constante de integración que caracteriza una amplitud que varía lentamente. Esta forma para $q(t)$ coincide con la de la aproximación WKB en Mecánica Cuántica, aunque los personajes intervinientes tienen otro significado.

Es fácil verificar que el último miembro de (18.8) satisface que

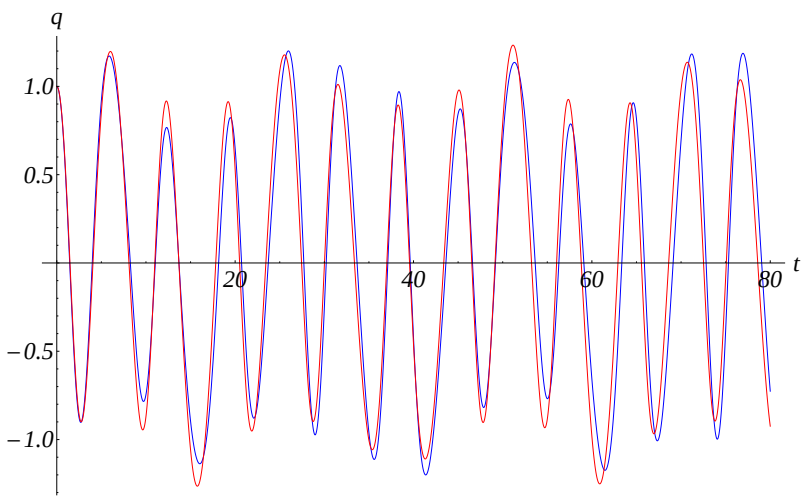
$$\ddot{q}(t) = -\omega(t)^2 \left(1 - \frac{3\,\dot{\omega}(t)^2}{4\,\omega(t)^4} + \frac{\ddot{\omega}(t)}{2\,\omega(t)^3}\right) q(t)$$

lo que confirma el cumplimiento aproximado de la ecuación del oscilador armónico con frecuencia variable en el marco de las aproximaciones detalladas en la Nota 6.

Ejemplo. Consideremos el caso donde ω^2 varía sinusoidalmente alrededor de un valor ${\omega_o}^2$, como sucede en la ecuación de Mathieu tratada en §10.3. Sea entonces

$$\omega(t)^2 = {\omega_o}^2\ (1 + \kappa\ \sin(\sigma\ t))$$

donde $|\kappa| < 1$. Para satisfacer las condiciones de variación lenta, σ deberá ser mucho menor que ω_o. Compararemos la solución numérica de la ecuación del oscilador armónico con frecuencia variable con la solución aproximada obtenida en (18.8), a iguales condiciones iniciales en $t = 0$. Como el acuerdo es excelente cuando $\sigma << \omega_o$, mostraremos un caso marginal para poder apreciar las ligeras discrepancias. Para ello elegimos los parámetros $\omega_o = 1$, $\sigma = 0.7$, $\kappa = 0.6$, que conducen al siguiente resultado (la curva roja es la solución aproximada):



En cambio, la solución aproximada (18.8) no reproduce los comportamientos resonantes.

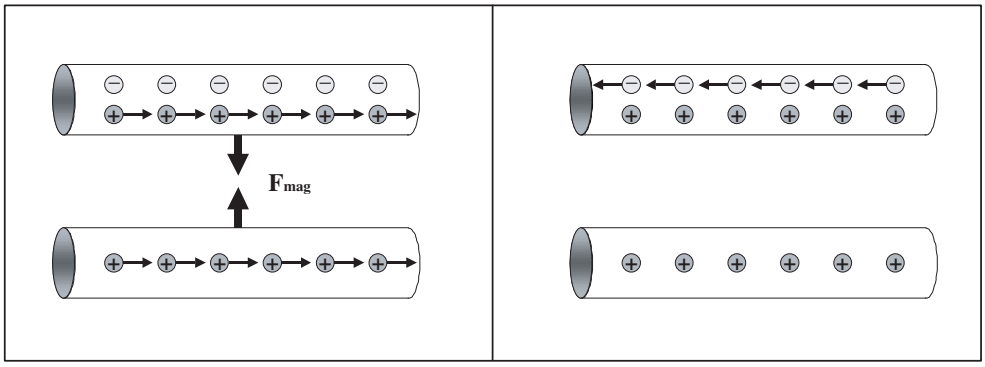
Bibliografía adicional

A.J. Lichtenberg y M.A. Lieberman, *Regular and Chaotic Dynamics*, Springer.

19 Relatividad Especial

19.1 El Principio de relatividad y las leyes de Maxwell

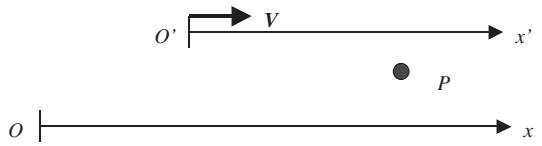
El Principio de relatividad afirma que las mismas leyes fundamentales de la Física son satisfechas en todos los sistemas inerciales. Esto significa que no hay privilegios entre distintos sistemas inerciales, lo que corresponde a la idea de que el estado de movimiento (absoluto) no es detectable; sólo es detectable el movimiento relativo. No hay manera de reconocer el estado de movimiento de un laboratorio mediante un experimento realizado en el mismo (los *cambios* del estado de movimiento serían detectables a través de los efectos inerciales que se generan). El Principio de relatividad se remonta a Galileo, y es satisfecho por las leyes fundamentales de la Mecánica. Queremos saber si las leyes de Maxwell satisfacen o no el Principio de Relatividad. ¿Podemos usar las leyes del electromagnetismo en distintos sistemas inerciales? Para analizar la cuestión consideremos el siguiente ejemplo:



La Figura de la izquierda muestra dos hilos infinitos que portan corrientes iguales; uno es neutro y el otro está cargado. Las leyes del electromagnetismo nos dicen que existe una interacción magnetostática entre ambos hilos. En la Figura de la derecha se representa la misma configuración vista en un sistema de referencia que acompaña a las cargas positivas. En este caso las leyes de Maxwell nos dicen que no existe interacción alguna pues uno de los hilos carece de corriente mientras que el otro hilo carece de carga. Los resultados obtenidos de aplicar las misma leyes en dos sistemas de referencia distintos son claramente contradictorios (la existencia de una interacción debería ser un hecho absoluto, independiente del sistema de referencia). A primera vista esta contradicción parecería indicar que no es posible utilizar las leyes de Maxwell en dos sistemas de referencia distintos. En tal caso las leyes de Maxwell no cumplirían el Principio de relatividad, y sería menester identificar el sistema de referencia donde las leyes de Maxwell

son aplicables. Sin embargo, debe señalarse que el paso de la Figura de la izquierda a la Figura de la derecha no es completamente inocente porque entraña dos suposiciones: que las distancias no cambian y que las cargas no cambian.

La creencia en distancias *absolutas* ha sido una piedra basal en la construcción de la Mecánica de Newton, y debemos admitir que nuestra experiencia cotidiana indica que se trataría de una buena hipótesis sobre la naturaleza del espacio. Sin embargo la experiencia cotidiana es buena consejera sólo en el rango de fenómenos que ella misma abarca. Examinemos cómo entra la creencia en distancias absolutas en las transformaciones de Galileo, que son las transformaciones que usamos para certificar la validez del Principio de relatividad en la Mecánica Clásica.



En el esquema unidimensional de la Figura está claro que $d_{OP} = d_{OO'} + d_{O'P}$. Esta relación entre distancias vale si todas las distancias están medidas en un mismo sistema, ya sea S o S'. Lo que llamamos coordenada x no es más que la distancia d_{OP} medida en S, mientras que x' es la distancia $d_{O'P}$ medida en S'. Por otro lado, $d_{OO'}$ medida en S es igual a Vt. De aquí se concluye que

$$d_{O'P} \; medida \; en \; S = x - Vt \tag{19.1}$$

La creencia en distancias absolutas nos mueve a reemplazar el miembro izquierdo por x', para así obtener las transformaciones de Galileo. Es interesante notar que esta creencia implica que el tiempo debe ser igualmente absoluto. En efecto, para que no existan privilegios entre S y S' la transformación de Galileo inversa debe verse igual a la directa salvo por el cambio de V por $-V$:

$$x' = x - Vt \; , \qquad\qquad x = x' + Vt' \; ;$$

pero esto es consistente sólo si

$$t' = t \; .$$

La transformación de Galileo, basada en la creencia en distancias absolutas o *invariantes* (decimos así de aquellas cantidades cuyo valor no dependen del sistema de referencia), tiene como consecuencia la adición galileana de velocidades: $\vec{u}' = \vec{u} - \vec{V}$, y la invariancia galileana de las aceleraciones: $\vec{a}' = \vec{a}$.

Las leyes fundamentales de la Mecánica cumplen el Principio de relatividad ante transformaciones de Galileo. Por ejemplo, la Segunda Ley de Newton $\vec{F} = m\vec{a}$ se construye con dos miembros que son invariantes galileanos por separado. Por un lado la aceleración lo es; por el otro lado, las fuerzas fundamentales de la Mecánica Clásica dependen de las distancias entre partículas, que son supuestas invariantes. Así, el cumplimiento de la

relación $\vec{F} = m\vec{a}$ en un sistema de referencia es garantía de su cumplimiento en cualquier otro sistema de referencia galileanamente conectado con el primero.

En cambio las leyes de Maxwell no exhiben este comportamiento ante transformaciones de Galileo. Por un lado, en la fuerza de Lorentz no sólo intervienen distancias entre cargas sino también las velocidades de las cargas (que no son invariantes galileanos). Como vimos en el ejemplo, esta característica de la fuerza de Lorentz impide alcanzar resultados consistentes ante un cambio galileano de sistemas de referencia. Por otro lado, las leyes para el campo electromagnético no mantienen su forma ante transformaciones de Galileo; es decir que no satisfacen el Principio de relatividad galileano. Recordemos que las leyes de Maxwell para los campos pueden llevarse a la forma de ecuaciones de onda; en ausencia de fuentes tendremos

$$\mu_o\varepsilon_o \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0 \ , \qquad \mu_o\varepsilon_o \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0 \ ,$$

donde reconocemos la velocidad c de propagación de la onda con un valor que resulta de las constantes de la electrostática y la magnetostática:

$$c^{-2} = \mu_o\varepsilon_o \ .$$

Para sorpresa de Maxwell, el valor de c resultó coincidente con el ya entonces conocido valor de la velocidad de la luz en vacío. Así Maxwell concluyó que la luz es una onda electromagnética. Volviendo a nuestro problema, el hecho que las leyes de Maxwell contengan una velocidad nos dice que estas leyes cambian de forma ante transformaciones de Galileo, porque sabemos que las velocidades finitas no son invariantes galileanos sino que cambian al pasar de un sistema de referencia a otro. Las leyes de Maxwell quedarían excluidas del Principio de relatividad galileano, y habría que establecer cuál es el sistema de referencia donde la luz se propaga con velocidad $c = (\mu_o\varepsilon_o)^{-1/2}$; es decir, cuál es el sistema de referencia donde podemos utilizar las leyes de Maxwell tal como las conocemos. Esto sería una forma de resolver la paradoja de nuestro ejemplo. Maxwell pensaba que ese sistema de referencia privilegiado era, como ocurre con toda ecuación de onda, el sistema de referencia donde el medio de propagación de la onda está en reposo. Maxwell pensaba las ondas electromagnéticas como ondas de materia, perturbaciones de un medio material denominado *éter*. En tal caso el privilegio que adquiere el sistema de referencia donde el medio está en reposo no menoscaba el Principio de relatividad, pues es un privilegio adquirido por razones físicas, y cualquier otro sistema de referencia donde el éter esté en reposo será igualmente privilegiado. Así sucede con las ecuaciones de otras ondas de materia (el sonido, la perturbaciones de la superficie del agua de un estanque, etc.) cuyas velocidades de propagación relativa al medio dependen de las propiedades del medio en cuestión. Sin embargo, a diferencia de lo que sucede con otros medios donde se propagan ondas, el éter se mostraba elusivo a una detección directa. El carácter intangible del éter (la luz se propaga en regiones aparentemente "vacías", el éter no frena el movimiento de los planetas, etc.), movió a los físicos del siglo XIX a intentar detectar no ya el éter mismo, sino nuestro estado de movimiento respecto del éter. Los experimentos buscaban medir diferencias en las velocidades de propagación de

rayos de luz, que revelaran una composición galileana de velocidades entre la velocidad de la luz relativa al éter y la velocidad del laboratorio relativa al éter (Hoek, 1868; Michelson, 1881; Michelson-Morley, 1887; etc.), o alteraciones de la ley de Snell debidas al movimiento relativo al éter del material transparente donde un rayo se refracta (Arago, 1810; Airy, 1871). Los resultados de estos experimentos fueron siempre negativos: el movimiento del laboratorio relativo al éter nunca se evidenció. Se tejieron distintas hipótesis acerca de la interacción del éter con la materia para justificar estos resultados. Estas teorías dinámicas sobre la interacción entre éter y el resto de la materia alcanzaron su forma más elaborada en la Teoría de los electrones de Lorentz (1895).

19.2 La relatividad de Einstein

Mientras la comunidad científica debatía estas cuestiones, en 1905 Einstein cambió el enfoque del problema proponiendo que las leyes de Maxwell son leyes fundamentales que integran, por lo tanto, el conjunto de leyes que satisfacen el Principio de relatividad. Para Einstein el campo electromagnético tiene entidad propia, y no precisa una "materialización" mediante la idea de un éter. Si el éter no existe, entonces no hay nada que privilegie un sistema de referencia frente a otro, y las leyes de Maxwell deben ser válidas en cualquier sistema de referencia inercial. Esto significa que la luz se propaga con la *misma* velocidad c en cualquier sistema de referencia inercial. No existen entonces las diferencias de velocidad buscadas por los experimentadores, y la ley de Snell se cumple en cualquier laboratorio donde el material refractante esté en reposo relativo. Claro que admitir una *velocidad finita invariante* rompe con el teorema de adición de velocidades de Galileo, y supone entonces el abandono de nuestra creencia en distancias y tiempos absolutos. Einstein propuso elevar las leyes de Maxwell al rango de leyes fundamentales, abandonado nuestras nociones intuitivas de espacio y tiempo para subordinarlas a la invariancia de la velocidad de la luz.

Procederemos ahora a replantear la transformación de coordenadas cuidando de no introducir en ellas prejuicio alguno sobre la naturaleza del espacio y el tiempo. En la Figura se muestra una barra recorrida por una partícula; el movimiento relativo barra-partícula está caracterizado por la velocidad relativa V:

En la Figura de la izquierda se describe el movimiento relativo en el sistema fijo a la barra, mientras que la Figura de la derecha lo describe en el sistema fijo a la partícula. Como no estamos dispuestos a prejuzgar sobre la naturaleza del espacio y el tiempo, hemos dibujado la barra con distintas longitudes en cada sistema. Llamamos *longitud propia* L_o a la longitud de la barra en el sistema donde se encuentra en reposo. El tiempo

que demora el recorrido de la partícula a lo largo de la barra también puede depender del sistema de referencia. Llamaremos *tiempo propio* $\Delta\tau$ al tiempo transcurrido entre dos *eventos* medido en el sistema de referencia donde los eventos ocurren en la misma posición (siempre que tal sistema exista). En el caso bajo estudio, los dos eventos son el paso de la partícula por un extremo de la barra y el paso de la partícula por el otro extremo de la barra. Estos dos eventos ocurren en la misma posición en el sistema fijo a la partícula; de modo que el tiempo transcurrido entre los eventos es un tiempo propio en el sistema de la Figura de la derecha. La mera definición de velocidad nos permite escribir:

$$L_o = V\ \Delta t\ , \qquad\qquad L = V\ \Delta\tau\ ,$$

donde Δt es el tiempo del recorrido de la partícula a lo largo de la barra en el sistema de la Figura de la izquierda. De aquí resulta

$$\frac{\Delta t}{\Delta\tau} = \frac{L_o}{L} \tag{19.2}$$

Esta relación nos dice que si estamos dispuestos a renunciar a distancias absolutas ($L_o \neq L$) también habremos de renunciar a tiempos absolutos ($\Delta\tau \neq \Delta t$). Además nos dice que la relación entre la longitud en movimiento y la longitud propia de un mismo cuerpo tiene el mismo carácter que la relación entre tiempo propio y el tiempo entre el mismo par de eventos en otro sistema de referencia. Concretamente, ambas relaciones no pueden depender más que de la velocidad relativa V entre el sistema propio correspondiente y el otro sistema de referencia arbitrario. La ecuación (19.2) dice entonces que la misma función $\gamma(V)$ que expresa la relación entre longitudes ha de expresar también la relación entre tiempos:

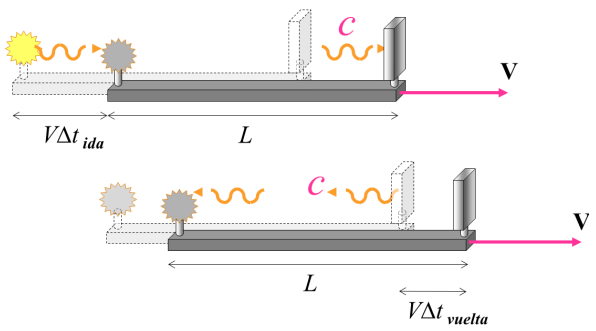
$$\frac{L_o}{L} = \gamma(V)\ , \qquad\qquad \frac{\Delta t}{\Delta\tau} = \gamma(V)\ . \tag{19.3}$$

En realidad las relaciones (19.3) suponen que el espacio y el tiempo son isótropos y homogéneos; de otra forma las relaciones podrían depender del lugar o el instante de ocurrencia de los eventos o de la orientación de la barra. Vamos a admitir, como en la Física clásica, que el espacio está dotado de una geometría euclidiana –por lo tanto es isótropo y homogéneo–, y que las relaciones (19.3) no dependen del tiempo (homogeneidad del tiempo).

La forma de la función $\gamma(V)$ será dictada por la invariancia de la velocidad de la luz a la cual se subordinarán las nociones de espacio y tiempo. Para hallar $\gamma(V)$ reemplacemos la partícula del experimento anterior por un rayo de luz, coloquemos un espejo en un extremo de la barra, y consideremos como par de eventos el paso del rayo de luz por el extremo libre de la barra, y su regreso al mismo luego de reflejarse en el espejo. En este caso el tiempo propio entre los eventos corresponde al sistema fijo a la barra pues ambos eventos ocurren en el mismo extremo de la barra. Como la luz viaja con velocidad c (¡en cualquier sistema!) diremos que

$$c\ \Delta\tau = 2\ L_o\ , \tag{19.4}$$

En el otro sistema calculamos el tiempo Δt entre los eventos con la misma velocidad c para el rayo de luz; teniendo en cuenta el desplazamiento de la barra de longitud L resulta



$$c\ \Delta t_{ida} = L + V\ \Delta t_{ida}\ , \qquad\qquad c\ \Delta t_{vuelta} = L - V\ \Delta t_{vuelta}\ ,$$

entonces

$$\Delta t = \Delta t_{ida} + \Delta t_{vuelta} = \frac{L}{c-V} + \frac{L}{c+V} = \frac{2L}{c}\ \frac{1}{1-\frac{V^2}{c^2}}\ . \tag{19.5}$$

Dividiendo (19.4) y (19.5), y usando las relaciones (19.3) se obtiene

$$\gamma(V) = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}\ . \tag{19.6}$$

Luego, reemplazando en (19.3)

$$\textit{contracción de longitudes} \qquad L = L_o\ \sqrt{1-\frac{V^2}{c^2}} \tag{19.7}$$

$$\textit{dilatación del tiempo} \qquad \Delta t = \frac{\Delta\tau}{\sqrt{1-\frac{V^2}{c^2}}} \tag{19.8}$$

Estas son las relaciones entre longitudes y tiempos que resultan de admitir la existencia de una velocidad finita c invariante (igual en todos los sistemas de referencia); son las dos caras de una misma moneda, como se destaca en la ecuación (19.2). En las transformaciones de Galileo, en cambio, la única velocidad "invariante" es la velocidad infinita. Consistentemente podemos reencontrar las nociones clásicas de espacio y tiempo tomando el límite $c \longrightarrow \infty$ en (19.7) y (19.8). También podemos ver en (19.7) y (19.8) la razón por la cual nuestra experiencia cotidiana no es capaz de revelarnos la relatividad de distancias y tiempos: la velocidad invariante c (la velocidad de la luz) es mucho mayor que las velocidades V involucradas en el rango de fenómenos cotidianos, de modo que el factor $\gamma(V)$ es prácticamente igual a 1 en ese rango. Este comportamiento de distancias y tiempos permite que las densidades de carga y corriente se transformen de manera apropiada para que la leyes de Maxwell sean válidas en todo sistema de referencia inercial. Aunque el efecto sea débil para $V << c$, es suficiente para que el hilo superior derecho de nuestro ejemplo adquiera la densidad de carga negativa necesaria para que la interacción magnetostática de la izquierda se vea como una interacción electrostática en el sistema de la derecha de la Figura.

19.3 Transformaciones de Lorentz

Ahora podemos regresar a la ecuación (19.1) para darle a "$d_{O'P}$ *medida en* S" el tratamiento que corresponde a esta nueva forma de ver el espacio y el tiempo. Como dijimos, la coordenada x' es $d_{O'P}$ *medida en* S', y puede pensarse como la longitud L_o de una regla en reposo en el sistema S'. Entonces $d_{O'P}$ *medida en* S es la longitud contraida de esa regla: $d_{O'P}$ *medida en* $S = \gamma(V)^{-1}$ $d_{O'P}$ *medida en* $S' = \gamma(V)^{-1}\ x'$, y la ecuación (19.1) resulta:

$$x' = \gamma(V)\ (x - Vt) \tag{19.9}$$

Para que los sistemas S y S' estén en pie de igualdad, la transformación inversa de (19.9) debe tener su misma forma salvo por el cambio de V por $-V$,

$$x = \gamma(V)\ (x' + Vt') \tag{19.10}$$

De las ecuaciones (19.9) y (19.10) se pueden despejar t' y t:

$$t' = \gamma(V)\ (t - Vc^{-2}\ x) \tag{19.11}$$

$$t = \gamma(V)\ (t' + Vc^{-2}\ x') \tag{19.12}$$

Mediante un experimento pensado similar al que permitió obtener $\gamma(V)$ se obtiene que las longitudes transversales a la dirección del movimiento no cambian:

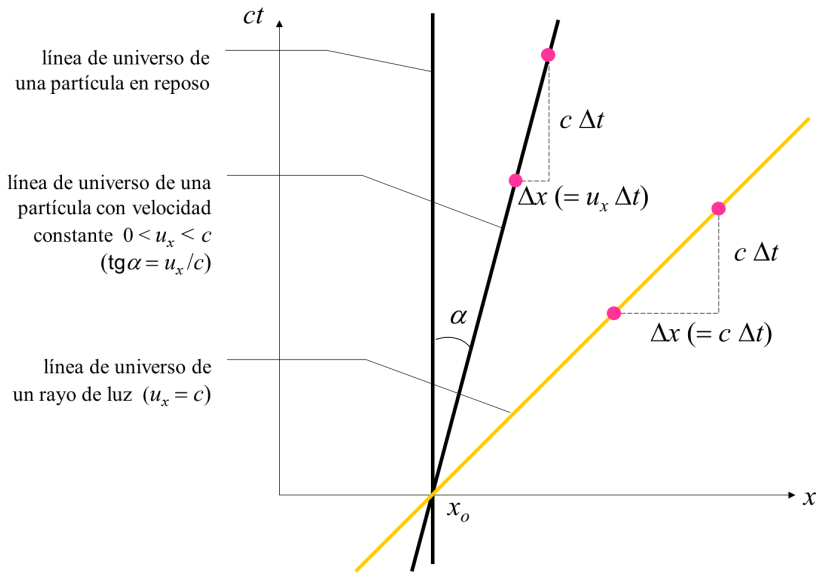
$$y' = y \tag{19.13}$$

$$z' = z \tag{19.14}$$

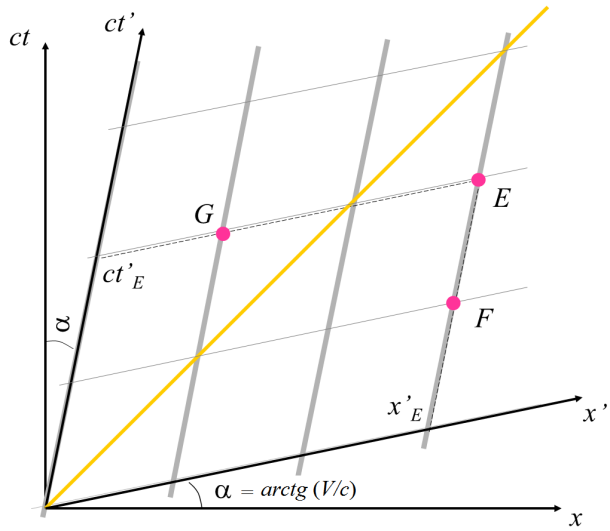
(Ayuda: coloque la barra con el espejo perpendicular a la dirección del eje x, y use el teorema de Pitágoras).

Las transformaciones (19.9-19.14) son las transformaciones de coordenadas que dejan invariante la ecuación de onda con velocidad de propagación c. Se llaman *transformaciones de Lorentz* (Voigt, 1887; Larmor, 1900; Lorentz, 1899, 1904); para $c \longrightarrow \infty$ devienen en las transformaciones de Galileo. Transforman las coordenadas de un evento (un "aquí y ahora", un punto en el espacio-tiempo). Aunque fueron obtenidas antes del trabajo de Einstein, su interpretación era completamente diferente. Mientras que para Einstein la contracción de longitudes y la dilatación del tiempo son efectos puramente cinemáticos que no privilegian sistema de referencia alguno (la barra tendrá longitud L_o en *cualquier* sistema de referencia donde se encuentre fija, y tendrá longitud L en *cualquier* sistema de referencia donde se mueva longitudinalmente con velocidad V), para Lorentz la velocidad V era la velocidad de la barra respecto del éter, y la contracción era un hecho absoluto (verificado en todo sistema de referencia) producido por una interacción con el éter. Según Lorentz la "dilatación del tiempo" no afectaba al tiempo absoluto sino que involucraba a un "tiempo matemáticamente auxiliar".

En Relatividad se utilizan gráficos ct *vs.* x. En tales gráficos un evento es un punto, y un rayo de luz es una recta a 45°. Los movimientos de las partículas se denominan *líneas de universo*. La Figura siguiente muestra un par de líneas de universo de partículas en movimiento uniforme, un rayo de luz y algunos eventos indicados en rojo.



Es común insertar las líneas coordenadas del sistema S' en el gráfico espacio-temporal del sistema S. Para ello usamos las transformaciones (19.9) y (19.11) para determinar el lugar geométrico de los eventos que tienen t' =constante y x' =constante; resultan ser rectas de pendiente V/c y $(V/c)^{-1}$ respectivamente. En particular $t' = 0$ caracteriza a los eventos que forman el eje x', mientras que $x' = 0$ caracteriza a los eventos que forman el eje t':



Relatividad de la simultaneidad

Para extraer resultados cuantitativos de estos gráficos mixtos es necesario *calibrar* previamente los ejes de ambos sistemas de referencia. A diferencia de las rotaciones, las transformaciones de Lorentz no dejan invariantes circunferencias sino hipérbolas (como veremos enseguida); de manera que las unidades de medida en los ejes de S' difieren de las de los ejes de S. En cambio no hay dificultad en extraer resultados cualitativos, y el más evidente de todos es que lo eventos que son simultáneos en S no lo son en S' y viceversa. La simultaneidad de eventos es una noción relativa (mientras que en Física Clásica es absoluta). En efecto, en el sistema S son simultáneos todos los eventos que tengan la misma coordenada t, es decir todos los puntos que estén sobre una misma recta paralela al eje x de la Figura anterior. Por otro lado, en S' son simultáneos todos los eventos que tengan la misma coordenada t', es decir que se trata de puntos dispuestos sobre rectas paralelas al eje x' de la Figura anterior. Evidentemente las nociones de simultaneidad de cada sistema difieren entre sí ($\Delta t = 0$ no implica $\Delta t' = 0$ en las transformaciones de Lorentz). La relatividad de la simultaneidad es el ingrediente necesario para justificar que dos sistemas de referencia no acuerden en las dimensiones de un objeto. La medición de una barra en el sistema S supone la determinación de las posiciones simultáneas de sus extremos. Esta medición no es "buena" para el sistema S' porque éste no comparte la noción de simultaneidad de aquél.

19.3.1 Intervalo

Aunque distancias y tiempos no son invariantes por separado ante transformaciones de Lorentz, en cambio sí lo es una combinación de ellos que llamamos intervalo. El intervalo Δs^2 es una forma cuadrática hiperbólica de las diferencias de las coordenadas de un par de eventos:

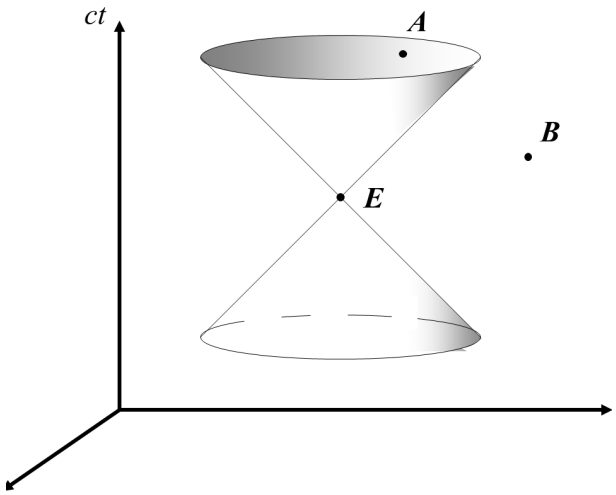
$$\Delta s^2 = c^2 \Delta t^2 - |\Delta \vec{r}|^2 \tag{19.15}$$

donde Δt y $|\Delta \vec{r}|$ son el tiempo transcurrido y la distancia entre dos eventos cualesquiera ($|\Delta \vec{r}|^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$). Para probar la invariancia de Δs^2 usamos que Δy, Δz son de por sí invariantes ante transformaciones de Lorentz a lo largo del eje x; por lo tanto sólo resta probar la invariancia de $c^2 \Delta t^2 - \Delta x^2 = (c\Delta t - \Delta x)(c\Delta t + \Delta x)$ (Ejercicio). Como caso particular, el intervalo entre dos eventos unidos por un rayo de luz ($|\Delta \vec{r}| = c\ \Delta t$) es nulo.

El intervalo tendría el aspecto de una distancia euclidiana en un *espacio-tiempo* de cuatro dimensiones, si no fuera porque las coordenadas espaciales entran con signo opuesto al de la coordenada temporal. Se dice entonces que el espacio-tiempo tiene una geometría *pseudo-euclidiana* (espacio-tiempo de Minkowski).

19.3.2 Cono de luz. Tipos de separación entre eventos

La invariancia del intervalo permite clasificar pares de eventos en forma independiente de los sistemas de referencia. Diremos que un par de eventos tiene separación temporal, espacial o nula según si el intervalo entre ellos es positivo, negativo o cero.[1] Los eventos con separación nula están sobre rayos de luz ($\Delta s^2 = 0 \iff |\Delta\vec{r}| = c\ \Delta t$). Esta clasificación absoluta (independiente del sistema de referencia) introduce en el espacio-tiempo la noción de *cono de luz* de un evento. El cono de luz de un evento E está formado por todos los eventos que tienen separación nula con E. Esto significa que el cono de luz de E está generado por todos los rayos de luz que pasan por E.



Los eventos interiores al cono de E están separados temporalmente de E. Los eventos exteriores al cono de luz de E están separados espacialmente de E. A partir de la Figura anterior no es difícil ver que si dos eventos E y A están separados temporalmente siempre es posible construir un sistema de referencia donde los dos eventos ocurren en la misma posición (de manera que para ellos existe un tiempo propio). En cambio si dos eventos E y B están separados espacialmente siempre existe un sistema de referencia donde los dos eventos son simultáneos; en el resto de los sistemas de referencia o bien ocurre E antes que B, o bien ocurre B antes que E. Por lo tanto no puede hablarse de una relación causa-efecto entre eventos separados espacialmente. En cambio el orden temporal de dos eventos separados temporalmente es el mismo en todos los sistemas de referencia.[2]

1. Algunos autores definen el intervalo con la *signatura* opuesta. En ese caso la separación temporal corresponderá a intervalo negativo, etc. Lo importante de esta clasificación es identificar cuál de los dos términos, el temporal o el espacial, es el que domina en el intervalo entre el par de eventos.

2. La construcción del sistema de referencia donde eventos separados temporalmente ocurran en un mismo lugar ($\Delta\vec{r} = 0$), o donde eventos separados espacialmente ocurran simultáneamente ($\Delta t = 0$), requerirá de un *boost* de Lorentz como el las ecuaciones (19.9-19.14), pero realizado a lo largo de la dirección espacial que une los eventos en cuestión.

En la Figura que precede a la anterior es claro que E ocurre en la misma posición que F en el sistema S', pero no ocurre así en el sistema S. Del mismo modo, E es simultáneo con G en S', pero no lo es en S.

19.4 Composición relativista de movimientos

La transformación de Lorentz de la velocidad de una partícula $\vec{u} = d\vec{r}/dt$ se obtiene diferenciando las transformaciones (19.9-19.14) y tomando los cocientes. De esa forma, las componentes de $\vec{u}' = d\vec{r}'/dt'$ resultan

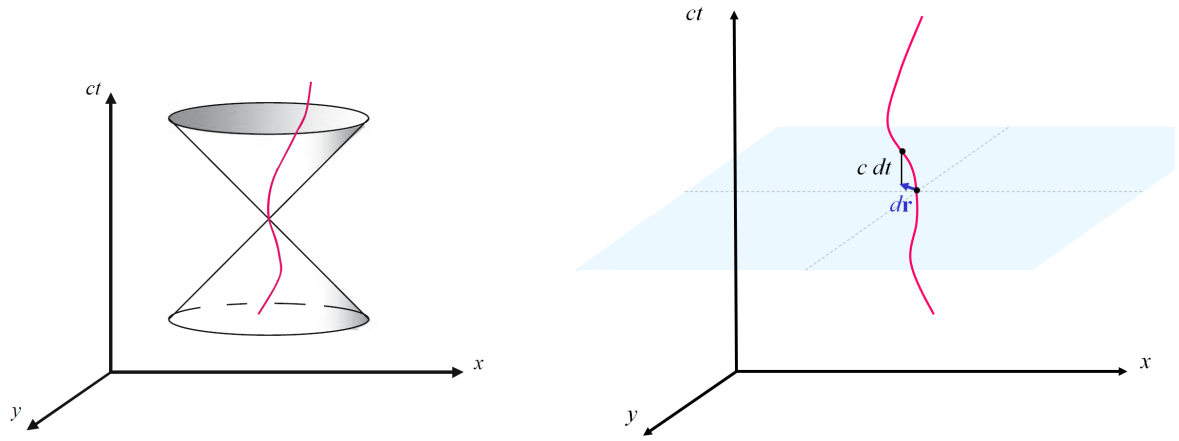
$$u'_x = \frac{u_x - V}{1 - \frac{u_x V}{c^2}} \ , \qquad u'_y = \frac{\sqrt{1 - \frac{V^2}{c^2}}\ u_y}{1 - \frac{u_x V}{c^2}} \ , \qquad u'_z = \frac{\sqrt{1 - \frac{V^2}{c^2}}\ u_z}{1 - \frac{u_x V}{c^2}} \ .$$

La composición relativista de velocidades explica el resultado del experimento realizado por Fizeau en 1851 para medir la velocidad de la luz en una corriente de agua. La composición de las velocidades de la luz y del agua arrojó un resultado difícil de explicar en un contexto histórico pre-relativista. En aquel momento, el resultado se entendió como una consecuencia del *arrastre parcial* del éter propuesto por Fresnel en 1818, quien había tenido que imaginar una singular interacción entre el éter y las sustancias transparentes en movimiento para explicar el resultado del experimento de Arago (1810) que involucraba la refracción de la luz en un prisma móvil.

Mediante un procedimiento similar se obtiene la transformación de Lorentz de las aceleraciones.

19.5 Tiempo propio de la partícula

En la Figura vemos una línea de universo de una partícula con un movimiento inespecífico. Los eventos pertenecientes a la línea de universo de una partícula están causalmente conectados. Por lo tanto el cono de luz de cualquiera de ellos debe contener la línea de universo completa. Para que esto suceda, la velocidad $\vec{u}$ de la partícula debería ser siempre menor que la velocidad de la luz: $|\vec{u}| < c$.



El intervalo entre un par de eventos vecinos de la línea de universo, con coordenadas $(ct,\ \vec{r})$ y $(ct + c\ dt,\ \vec{r} + d\vec{r})$, es

$$ds^2 = c^2dt^2 - |d\vec{r}|^2 = c^2dt^2\ \left(1 - \frac{|d\vec{r}|^2}{c^2dt^2}\right) = c^2dt^2\ \left(1 - \frac{u^2}{c^2}\right) > 0$$

Como el intervalo es invariante, tiene el mismo valor en cualquier sistema de referencia:

$$\sqrt{1 - \frac{u^2}{c^2}}\ c\ dt\ =\ \sqrt{1 - \frac{u'^2}{c^2}}\ c\ dt'\ .$$

En particular, en el sistema de referencia donde la velocidad de la partícula se anule en el instante considerado (**sistema propio** de la partícula), el radicando se hace igual a 1, y el tiempo es el tiempo propio $\Delta\tau$ entre los eventos (en el sistema propio los eventos ocurren en el mismo lugar); entonces la ecuación anterior puede escribirse así

$$d\tau = \gamma(u)^{-1}\ dt\ =\ \gamma(u')^{-1}\ dt'\ . \tag{19.16}$$

Por cierto esto no es más que la dilatación del tiempo, o la relación $\Delta\tau = c^{-1}\Delta s$ para eventos con separación temporal.

El tiempo propio a lo largo de la línea de universo de la partícula se obtiene integrando la expresión anterior,

$$\Delta\tau = \int \gamma(u(t))^{-1}\ dt = \int \sqrt{1 - \frac{u(t)^2}{c^2}}\ dt\,. \tag{19.17}$$

$\Delta\tau$ es el tiempo que mide un **reloj** que viaja junto con la partícula. Los relojes que están en reposo indican el tiempo coordenado Δt del sistema de referencia.

19.5.1 Paradoja de los gemelos

El resultado anterior ilustra la llamada "paradoja" de los gemelos. Dos gemelos se separan porque uno de ellos emprende un viaje espacial. A su regreso, el viajero encuentra que su hermano envejeció más que él. Está claro que la integral $\Delta\tau$ entre dos eventos (en este caso, la partida y la llegada del viajero) depende de la línea de universo utilizada para unir los eventos ($d\tau$ no es un diferencial exacto). Para el gemelo que permaneció en la Tierra es $u = 0$ (entendiendo a la Tierra como un sistema inercial); entonces $\Delta\tau_{Tierra} = t_{llegada} - t_{partida}$. En cambio en $\Delta\tau_{viajero}$ el integrando es siempre menor que 1; así resulta $\Delta\tau_{viajero} < \Delta\tau_{Tierra}$. Lo "paradójico" es que podría pensarse que la situación es simétrica: cualquiera de los dos es un "viajero" visto desde la perspectiva del otro. Lo que rompe la simetría es el hecho de que sólo uno de ellos podría verse como "viajero inercial". Si ambos fuesen inerciales nunca se reencontrarían. El privilegio del viajero inercial es conferido por la propia estructura geométrica del espacio-tiempo de Minkowski.

19.6 Dinámica relativista

Mientras que el Principio de inercia permanece válido en Relatividad Especial, la segunda ley de Newton, y la forma de las fuerzas fundamentales, deben reformularse para satisfacer el Principio de relatividad ante transformaciones de Lorentz. La tercera ley de Newton, que lleva a compensaciones simultáneas a distancia de las variaciones de cantidad de movimiento, tampoco es satisfactoria en el contexto relativista (la simultaneidad no es absoluta en Relatividad).

Para construir la Mecánica relativista partiremos de un principio variacional cuya acción sea un invariante lorentziano. De esa forma, si la acción es estacionaria en un sistema inercial, lo será también en cualquier otro sistema inercial. Así las ecuaciones de Euler-Lagrange cumplirán con el Principio de relatividad ante transformaciones de Lorentz. Por otro lado, la acción relativista debe ser capaz de reproducir la dinámica newtoniana cuando las velocidades sean mucho menores que la velocidad de la luz. Su Lagrangiano debe ser equivalente al Lagrangiano clásico en el límite $|\vec{u}| << c$.

19.6.1 Acción de la partícula libre

El (invariante) tiempo propio a lo largo de la línea de universo de la partícula (19.17) es la elección adecuada para la funcional acción de la partícula libre,

$$S[\vec{r}(t)] \;=\; -m\,c^2 \int d\tau \;=\; -m\,c^2 \int \gamma(u)^{-1}\,dt = \;-m\,c^2 \int \sqrt{1-\frac{|\vec{u}|^2}{c^2}}\;dt\;. \qquad (19.18)$$

El coeficiente $-m\,c^2$ permite obtener el límite clásico correcto:

$$L = -mc^2\,\sqrt{1-\frac{u^2}{c^2}} = -mc^2 + \frac{1}{2}\,m\,u^2 + \mathcal{O}(u^4/c^4)\,.$$

19.6.2 Cantidad de movimiento y energía

El momento canónicamente conjugado $\partial L/\partial \vec{u}$ que resulta del Lagrangiano de partícula libre (19.18) es la **cantidad de movimiento** de la partícula

$$\vec{p} \;\equiv\; \frac{m\,\vec{u}}{\sqrt{1-\frac{u^2}{c^2}}} \;=\; m\,\gamma(u)\,\vec{u} \;=\; m\,\gamma(u)\,\frac{d\vec{r}}{dt} \;=\; m\,\frac{d\vec{r}}{d\tau} \tag{19.19}$$

(en el último paso usamos la ecuación (19.16)), que vale $m\vec{u}$ cuando $|\vec{u}| << c$.

El Hamiltoniano asociado al Lagrangiano de partícula libre,

$$H = \vec{u}\cdot\vec{p} - L = m\,\gamma(u)\,u^2 + mc^2\gamma(u)^{-1} = mc^2\gamma(u)\,\left(\frac{u^2}{c^2} + \gamma(u)^{-2}\right) = mc^2\,\gamma(u)$$

es la **energía** de la partícula

$$E \;\equiv\; m\,\gamma(u)\,c^2\,. \tag{19.20}$$

La energía E es una combinación de **energía en reposo** mc^2 y energía cinética:

$$E \;=\; \frac{m\,c^2}{\sqrt{1-\frac{u^2}{c^2}}} \;=\; mc^2 \;+\; \frac{1}{2}m\,u^2 \;+\; ... \;\equiv\; mc^2 \;+\; T\,, \tag{19.21}$$

donde T es la **energía cinética** de la partícula, que a bajas velocidades coincide con la energía cinética clásica.

Combinando (19.19) y (19.20) obtenemos

$$\vec{p} \;=\; c^{-2}\,E\,\vec{u} \tag{19.22}$$

que es un resultado típicamente relativista: la cantidad de movimiento es un flujo de energía (como sucede en teoría electromagnética).

La ecuación (19.16) puede usarse para reemplazar $\gamma(u)$ en la energía (19.20),

$$\frac{E}{c} = mc\,\frac{dt}{d\tau}\,. \tag{19.23}$$

Así vemos que la energía de la partícula es proporcional a la relación entre el tiempo dt de los relojes del sistema de referencia y el tiempo propio de la partícula $d\tau$.

19.6.3 Transformaciones de E y $\vec{p}$

Como $d\tau$ es invariante, las ecuaciones (19.19) y (19.23) implican que el comportamiento de $(E/c,\ \vec{p})$ ante transformaciones de Lorentz será el mismo que el de $(c\,dt, d\vec{r})$.

Como las transformaciones de Lorentz son lineales, los diferenciales de las coordenadas se transforman como las coordenadas. Entonces

$$E' = \gamma(V)\,(E - V_x\,p_x) = \gamma(V)\,(E - \vec{V}\cdot\vec{p}) \tag{19.24a}$$
$$p'_x = \gamma(V)\,(p_x - V_x\,c^{-2}E) \tag{19.24b}$$
$$p'_y = p_y \tag{19.24c}$$
$$p'_z = p_z \tag{19.24d}$$

Así como $c\ dt$ y $|d\vec{r}|$ se combinan para formar el invariante intervalo, también E^2 y $c^2|\vec{p}|^2$ forman el **invariante energía-cantidad de movimiento** que refleja el valor de la masa de la partícula:

$$E^2 - c^2|\vec{p}|^2 = m^2c^4\gamma(u)^2 - m^2c^2u^2\gamma(u)^2 = m^2c^4\left(1 - \frac{u^2}{c^2}\right)\gamma(u)^2 = m^2c^4\ , \tag{19.25}$$

lo que además muestra que el Hamiltoniano de la partícula libre es

$$H(\vec{p}) = \sqrt{c^2|\vec{p}|^2 + m^2c^4}$$

Diferenciando la ecuación (19.25) se obtiene

$$E\,dE = c^2\,\vec{p}\cdot d\vec{p}\ ,$$

o, usando la ecuación (19.22)

$$dE = \vec{u}\cdot d\vec{p} = d\vec{r}\cdot\frac{d\vec{p}}{dt}\ , \tag{19.26}$$

lo que sugiere que la fuerza en Relatividad Especial está dinámicamente asociada a $d\vec{p}/dt$. En tal caso, el resultado anterior sería un teorema trabajo-energía. Nótese que $\vec{F} = d\vec{p}/dt$ implica que la fuerza no es paralela a la aceleración en general, pues la derivada de $\gamma(u)$ contribuye con un término paralelo a $\vec{u}$. En el límite de trabajo infinito la energía E diverge, y la velocidad de la partícula en (19.20) tiende a c. Así la velocidad de la luz es un límite inalcanzable para la partícula.

19.6.4 Fuerza de Lorentz

Si la partícula no es libre entonces la acción debe completarse con un término de interacción S_{int} que sea invariante lorentziano. En el caso de una carga e interactuando con campos $\vec{E}$, $\vec{B}$ externos descriptos por los potenciales electromagnéticos ϕ y $\vec{A}$, el término de interacción es

$$S_{int} = -e\int(\phi - \vec{u}\cdot\vec{A})\,dt$$

donde reconocemos el potencial dependiente de la velocidad $U = e(\phi - \vec{u} \cdot \vec{A})$ visto en §4.2.[3] La acción $S_{libre} + S_{int}$ corresponde al Lagrangiano $L = -m\gamma(u)^{-1}c^2 - U$. En las ecuaciones de Euler-Lagrange, el primer término de L contribuye con $d\vec{p}/dt$, y el segundo término provee la fuerza de Lorentz, como fue mostrado en §4.2; entonces la ecuación dinámica resulta

$$e\,(\vec{E} + \vec{u} \times \vec{B}) \;=\; \frac{d}{dt}\,(m\;\gamma(u)\;\vec{u})\;,$$

donde $\vec{E} = -\nabla\phi - \partial\vec{A}/\partial t$ y $\vec{B} = \nabla \times \vec{A}$. En 1908 Bucherer confirmó experimentalmente la dinámica relativista, observando el movimiento de un electrón en un campo electrostático. En ese caso, si $\vec{E}$ es uniforme, la ecuación da $(e/m)\vec{E}\,t = \gamma(u)\;\vec{u}$; entonces u tiende a c cuando t tiende a infinito.

19.6.5 Fotones

En las primeras décadas del siglo XX se comprendió que el comportamiento de la luz cuando interactúa con la materia no puede ser descripto por la teoría de Fresnel, que sin embargo describe todos los fenómenos físicos asociados a la propagación de las ondas luminosas. En el intercambio de energía y cantidad de movimiento con la materia, la luz se manifiesta a través de *cuantos* llamados **fotones**, que se comportan como partículas. La energía y cantidad de movimiento de los fotones están, no obstante, asociadas a propiedades ondulatorias (Einstein, 1905; Compton, 1923),

$$E_{foton} = \;\; h\nu\;, \qquad\qquad \vec{p}_{foton} = \;\; \frac{h\nu}{c}\;\hat{n} \quad , \tag{19.27}$$

donde ν es la frecuencia de la luz y $\hat{n}$ es su dirección de propagación ($h = 6,626 \times 10^{-34}$Js es la constante de Planck). Si reemplazamos estos valores en (19.22) obtenemos $\vec{u} = c\;\hat{n}$. Por otro lado, el invariante $E^2 - c^2|\vec{p}|^2$ se anula, lo que dice que la masa del fotón es nula (ver ecuación (19.25)).
Reemplazando en las transformaciones (19.24a) y (19.24b) se obtiene

$$\nu' \;=\; \gamma(V)\,(\,1 - \hat{n} \cdot \vec{V}\,c^{-1})\;\nu\;, \qquad\qquad \cos\theta' = \frac{\cos\theta - \frac{V}{c}}{1 - \frac{V}{c}\;\cos\theta} \tag{19.28}$$

donde $\cos\theta = n_x$. Notablemente estas son las expresiones para el **efecto Doppler** y la **aberración de la luz**[4] que se obtendrían razonando a partir de la teoría ondulatoria.

3. La invariancia lorentziana de S_{int} no será demostrada aquí. Las transformaciones de Lorentz del campo electromagnético mezclan lo eléctrico con lo magnético, así como mezclan espacio y tiempo.

4. Si $V << c$ entonces el ángulo de aberración $\alpha \equiv \theta' - \theta$ es pequeño, y vale que $\cos\theta' = \cos(\theta + \alpha) \approx \cos\theta - \alpha\,\sin\theta$. Aproximando a primer orden en V/c resulta $\alpha \approx (V/c)\,\sin\theta$.

19.7 Energía-cantidad de movimiento. Equivalencia masa-energía

En Relatividad la conservación de la energía y la cantidad de movimiento no pueden ser disociadas. Mientras la conservación de la cantidad de movimiento es una consecuencia de la simetría del Lagrangiano ante traslaciones espaciales, la conservación de energía proviene de la simetría del Lagrangiano ante traslaciones temporales. Pero las transformaciones de Lorentz mezclan el espacio y el tiempo. En consecuencia, la conservación de la cantidad de movimiento en todo sistema de referencia inercial requiere que la energía se conserve también en todo sistema de referencia inercial (y viceversa), lo que es evidente en las transformaciones (19.24) donde E, $\vec{p}$ se combinan para dar E', $\vec{p}\,'$. En Mecánica Clásica, en cambio, la transformación de la cantidad de movimiento es $\vec{p}\,' = \vec{p} - m\,\vec{V}$, que no contiene la energía clásica de la partícula. Así en la transformación de la cantidad de movimiento total de un sistema de partículas, $\vec{P}\,' = \vec{P} - M\,\vec{V}$, sólo la masa total del sistema aparece en el segundo término. En Mecánica Clásica se acepta que la masa total M es una magnitud que se conserva independientemente del resto de las magnitudes (principio clásico de conservación de la masa). De esta forma, la conservación separada de la masa garantiza que la conservación de la cantidad de movimiento total de un sistema aislado no se vea afectada por el cambio de sistema de referencia, y se cumpla en todos los sistemas inerciales ($\vec{P}$ se conserva $\Leftrightarrow$ $\vec{P}\,'$ se conserva). Pero en la transformación lorentziana de la cantidad de movimiento

$$p'_x \;=\; \gamma(V)\,(p_x - V_x\,c^{-2}E) = \gamma(V)\,(p_x - mV_x\; - V_x\,c^{-2}\;T)$$

el clásico término galileano de masa es tan sólo el primer término del desarrollo de la energía (energía en reposo). En Relatividad la energía, y no la masa, debe conservarse junto con la cantidad de movimiento para que la conservación de la cantidad de movimiento de un sistema aislado valga en todo sistema inercial. En suma, los sistemas aislados conservan la energía-cantidad de movimiento.

19.7.1 Equivalencia masa-energía

Ahora que vimos que no hay razón para sostener la conservación de la masa, sino que debemos utilizar en su lugar la conservación de la energía, queda abierta la posibilidad de que la masa, como una forma más de energía, pueda convertirse en otra forma de energía y viceversa. Lo veremos en un par de ejemplos.

Colisión plástica de dos partículas de igual masa

En el sistema **centro de momento**, donde se anula la cantidad de movimiento total, las partículas tienen velocidades iguales y opuestas u antes de la colisión. Cuando colisionan forman una única partícula que permanece en reposo para conservar la cantidad

de movimiento total. Si no se emite ningún tipo de energía, entonces toda la energía queda depositada en la energía en reposo de la partícula final; la conservación de la energía dice que

$$2\, m\, \gamma(u)\, c^2 \;=\; M\, c^2 \;,$$

es decir que la partícula resultante tiene una masa M mayor que $2m$ ($\gamma(u) > 1$). Esto es así porque la partícula resultante no sólo contiene las masas de las partículas que chocaron sino también contiene sus energías cinéticas.

Defecto de masa

La masa (energía en reposo) de un sistema compuesto incluye las masas de los constituyentes y la energía de ligadura medida en el sistema centro de momento. Por ejemplo, un deuterón D está constituido por un protón y un neutrón. La masa del deuterón es inferior a la suma de las masas de un protón y un neutrón libres, lo que significa que la energía de ligadura de los constituyentes es negativa. El **defecto de masa** (la diferencia entre la masa del sistema compuesto y las de sus constituyentes libres) es $(m_D - m_p - m_n)c^2 = -2,22\text{MeV}$. En general, cuando núclidos livianos se fusionan se libera cierta cantidad de energía para conservar la energía total (**fusión nuclear**). En el caso de núclidos pesados sucede lo contrario porque más allá del Fe la energía (negativa) de ligadura por nucleón comienza a disminuir con el aumento del número másico; en ese caso se obtendrá energía dividiendo núclidos (**fisión nuclear**).

Creación (y aniquilación) de partículas

La energía cinética puede ser usada para crear partículas. Por ejemplo un pión neutro π^0 puede ser creado en una colisión entre protones si la energía es suficientemente alta. La reacción debe superar una **energía umbral** para dar cuenta de la partícula creada. En el sistema centro de momento la energía umbral de la reacción

$$p + p \rightarrow p + p + \pi^0$$

es la necesaria para crear el pión y que los productos queden en reposo:

$$E_{umbral} = 2m_p c^2 + m_{\pi^0} c^2 = 1876,54\text{MeV} + 134,98\text{MeV}$$

Para alcanzar esta energía los dos protones deberían colisionar a la velocidad de $u_p = 0,36\; c$ en el sistema centro de momento.

Colisión elástica

El ejemplo anterior es un caso de colisión inelástica. Una colisión se dice **elástica** si las partículas antes y después de la colisión son las mismas. En ese caso la energía en reposo es la misma antes y después, y la conservación de la energía corresponderá a la conservación de la energía cinética relativista total.

Emisión o absorción de energía electromagnética

La interacción entre partículas cargadas puede involucrar radiación electromagnética. En los estados inicial o final, cuando las partículas están muy separadas, las cargas son libres y la radiación emitida o absorbida entra en el balance de energía-cantidad de movimiento a través de fotones. Por ejemplo, un par electrón-positrón se aniquila para dar lugar a dos fotones (se necesitan al menos dos para conservar la cantidad de movimiento). En el sistema centro de momento donde el electrón y el positrón tienen velocidades u_e iguales y opuestas el balance de energía es

$$2\, m_e\, \gamma(u_e)\, c^2 \;=\; 2\, h\, \nu.$$

A la inversa, dos fotones pueden crear un par electrón-positrón si tienen la energía umbral necesaria. La frecuencia ν está en el rango de los rayos γ.

El **efecto Compton** es la interacción de un fotón con un electrón libre. El fotón cede parte de su energía al electrón, y emerge de la colisión con una frecuencia menor que la inicial. El efecto es significativo si la longitud de onda del fotón es del orden o menor que $\lambda_C \equiv h/(m_e c) = 0,00243\text{nm}$ (*longitud de onda Compton*).

19.7.2 Interacciones "a distancia"

La Mecánica Clásica admite la existencia de interacciones a distancia, como las interacciones gravitatorias. En ese caso, la tercera ley de Newton garantiza la conservación de la cantidad de movimiento total como resultado de la compensación de las variaciones simultáneas de las cantidades de movimiento de los cuerpos que interactúan (en *c*ada instante las fuerzas son iguales y opuestas). En Relatividad la simultaneidad no es absoluta; por lo tanto, el enunciado mismo de la tercera ley de Newton no es independiente del sistema de referencia inercial. Así la relatividad de la simultaneidad entra en conflicto con la idea de interacciones a distancia. Este conflicto desaparece en las interacciones **locales**, porque la simultaneidad de dos hechos que suceden en el mismo lugar es absoluta. Por ejemplo, en un problema de colisiones entre partículas libres las interacciones son locales: suceden sólo en los instantes y lugares donde se producen los choques. En cada choque se conserva la cantidad de movimiento, y no hay conflicto con la relatividad de la simultaneidad porque dos partículas chocan cuando están en el mismo lugar. En una interacción local el intercambio de energía-cantidad de movimiento ocurre en un mismo evento, a diferencia de lo que sucedería en una interacción a distancia donde se involucrarían dos eventos cuya simultaneidad dependería del sistema de referencia elegido.

Las observaciones anteriores aún no enseñan cómo formular teorías de interacciones "a distancia" en Relatividad. Las interacciones a distancia de la Mecánica Clásica se describen mediante potenciales dependientes de las distancias entre las partículas. ¿Qué magnitudes deberían tomar el papel de estos potenciales clásicos? Para responder esta pregunta acudiremos a la interacción paradigmática de la Relatividad, que es la

interacción electromagnética. Como vimos en §19.6.4, las cargas eléctricas no interactúan directamente entre sí, sino que lo hacen (localmente) con los *c*ampos $\phi(\vec{r},t)$, $\vec{A}(\vec{r},t)$. A su vez estos campos tienen su propia dinámica, gobernada por una acción electromagnética de donde se derivan las ecuaciones de Maxwell. Estos campos poseen energía y cantidad de movimiento, que resultan de su propio Lagrangiano electromagnético. De esta forma cada carga puede interactuar localmente con el campo vecino, intercambiando energía-cantidad de movimiento. El campo, a su vez, transporta energía-cantidad de movimiento cuando se propaga (ondas electromagnéticas), pudiendo transferirlo parcialmente a otra carga en otro lugar. Las interacciones entre cargas distantes son entonces mediadas por un campo que se propaga a velocidad finita c, lo que supone la existencia de un tiempo de retardo entre la "acción" de una carga y la "reacción" de la otra. La conservación se satisface en cada evento del espacio-tiempo, y no hay conflicto con la relatividad de la simultaneidad. El sistema aislado que conserva la energía-cantidad de movimiento está formado por las cargas y el campo.

Material histórico

A. Einstein, Zur Elektrodynamik bewegter Körper (trad: Sobre la electrodinámica de los cuerpos en movimiento), Annalen der Physik **17**, 891-921 (1905).
https://www.fourmilab.ch/etexts/einstein/specrel/www/

Bibliografía adicional

R. Ferraro, *El espacio-tiempo de Einstein*, Ediciones Cooperativas (Buenos Aires).

R. Ferraro, *From æther theory to Special Relativity*, en Springer Handbook of Spacetime, A. Ashtekar y V.Petkov (eds.), Springer-Verlag GmbH (Heidelberg, 2014), doi:10.1007/978-3-642-41992-8_1.
Versión preliminar: http://arxiv.org/pdf/1302.6965.pdf

20 Acción de Jacobi. Analogía con la óptica geométrica

Consideremos un sistema conservativo holónomo y esclerónomo,

$$T = \frac{1}{2} \sum_{\lambda,\rho} m_{\lambda\rho}(q)\ \dot{q}_\lambda\ \dot{q}_\rho\ , \qquad\qquad V = V(q)$$

Notemos que

$$\sqrt{2T} dt = \sqrt{\sum_{\lambda,\rho} m_{\lambda\rho}(q)\ dq_\lambda\ dq_\rho} \tag{20.1}$$

entonces, como $T = E - V$,

$$dt = \sqrt{\frac{\sum_{\lambda,\rho} m_{\lambda\rho}(q)\ dq_\lambda\ dq_\rho}{2\ (E - V)}} \tag{20.2}$$

Esto significa que el tiempo infinitesimal que el sistema demora en evolucionar entre dos configuraciones vecinas –q_μ y $q_\mu + dq_\mu$– puede calcularse si se conoce con qué energía el sistema pasa de una configuración a la otra. Este es un nuevo ejemplo de que tiempo y energía son magnitudes que juegan juntas.

La curva que la evolución del sistema describe en el espacio de configuración puede ser obtenida de un principio variacional debido a Jacobi. La acción de Jacobi,

$$A[q] = \int_{q_{\mu\ i}}^{q_{\mu\ f}} \sqrt{2(E - V)\ \sum_{\lambda,\rho} m_{\lambda\rho}(q)\ dq_\lambda\ dq_\rho}\ , \tag{20.3}$$

no contiene el tiempo que demora el sistema en evolucionar desde la configuración inicial $q_{\mu\ i}$ hasta la configuración final $q_{\mu\ f}$. En su lugar, A contiene el valor de la energía E con la que esa evolución se realiza. El principio variacional busca en este caso la curva entre dos configuraciones fijas que hace estacionaria la funcional A para un valor determinado de E. La acción de Jacobi no sólo no depende del tiempo sino que es independiente de cualquier parametrización que se elija para la evolución en el espacio de configuración

(invariancia ante reparametrizaciones). Podemos elegir un parámetro cualquiera σ; si multiplicamos y dividimos el integrando por $d\sigma$ tendremos

$$A = \int_{q_{\mu\ i}}^{q_{\mu\ f}} \sqrt{2[E - V(q)] \sum_{\lambda,\rho} m_{\lambda\rho}(q) \frac{dq_\lambda}{d\sigma} \frac{dq_\rho}{d\sigma}}\, d\sigma$$

Así podemos ver el integrando como una suerte de Lagrangiano y escribir las ecuaciones de Lagrange que hacen estacionaria la acción A:

$$0 = \frac{d}{d\sigma}\left[\sqrt{\frac{2[E - V(q)]}{\sum_{\lambda,\rho} m_{\lambda\rho}(q) \frac{dq_\lambda}{d\sigma}\frac{dq_\rho}{d\sigma}}} \sum_\nu m_{\mu\nu}(q) \frac{dq_\nu}{d\sigma}\right] - \frac{\partial}{\partial q_\mu}\left[\sqrt{2[E - V(q)] \sum_{\lambda,\rho} m_{\lambda\rho}(q) \frac{dq_\lambda}{d\sigma}\frac{dq_\rho}{d\sigma}}\right] \tag{20.4}$$

Para verificar que estas ecuaciones efectivamente conducen a la evolución real del sistema en el espacio de configuración, reintroduciremos el tiempo t como parámetro; esto nos permitirá comparar con las conocidas ecuaciones de Lagrange. Recordemos que $d/d\sigma$ en (20.4) es una derivación a lo largo de la evolución del sistema. Si $\sigma = t$ entonces la raíz cuadrada en el primer término de (20.4) puede reemplazarse por 1, pues $E - V = T$ a lo largo de la evolución del sistema. Entonces la ecuación (20.4) toma la forma

$$0 = \frac{d}{dt}\left[\sum_\nu m_{\mu\nu} \frac{dq_\nu}{dt}\right] - \frac{\partial}{\partial q_\mu}\left[2\sqrt{(E - V)\, T}\right] \tag{20.5}$$

donde el corchete del primer término es $\partial T/\partial \dot{q}_\mu$. Por otro lado, las derivadas $\partial/\partial q_\mu$ deben realizarse antes del reemplazo $E - V = T$:

$$\frac{\partial}{\partial q_\mu}\left[2\sqrt{(E - V)\, T}\right] = \left[\sqrt{\frac{E - V}{T}} \frac{\partial T}{\partial q_\mu} - \sqrt{\frac{T}{E - V}} \frac{\partial V}{\partial q_\mu}\right] \tag{20.6}$$

Usando que $E - V = T$ en el resultado (20.6), la ecuación (20.5) queda

$$\frac{d}{dt}\left[\frac{\partial T}{\partial \dot{q}_\mu}\right] - \frac{\partial (T - V)}{\partial q_\mu} = 0$$

que es la forma habitual de las ecuaciones para la evolución temporal del sistema.

20.1 Analogía con la óptica geométrica

La óptica geométrica describe la propagación de la luz mediante trayectorias que llamamos **rayos**. Esas trayectorias pueden obtenerse del principio de Fermat (1658), que es el primer principio variacional de la Historia.[1]

20.1.1 Principio de Fermat

Los rayos de luz se propagan sobre trayectorias que hacen estacionario el camino óptico $\int n\, d\ell$, donde n es el índice de refracción y $d\ell$ es la longitud de arco sobre la trayectoria.

Es notable la analogía entre el principio de Fermat y el principio variacional de Jacobi. La acción de Jacobi podría escribirse como

$$A = \int_{q_{\mu\, i}}^{q_{\mu\, f}} \sqrt{2(E-V)}\ dl\ , \tag{20.7}$$

para una "longitud" dl definida en el espacio de configuración como

$$dl \equiv \sqrt{\sum_{\lambda,\rho} m_{\lambda\rho}(q)\ dq_\lambda\ dq_\rho} \tag{20.8}$$

donde $m_{\lambda\rho}(q)$ juega el papel de un tensor métrico en ese espacio. En el caso de una sola partícula de masa m es mejor usar $d\ell$, la longitud de arco sobre la trayectoria de la partícula en el espacio físico, y escribir la masa dentro de la raíz. Además, como la acción de Jacobi se varía con E constante, entonces podemos escribir el principio variacional de Jacobi para una única partícula como

$$\delta \int_{q_i}^{q_f} \frac{\sqrt{2m(E-V)}}{E}\ d\ell = 0\ .$$

A su vez el principio de Fermat dice que

$$\delta \int_{q_i}^{q_f} \frac{n}{c}\ d\ell = 0\ ,$$

donde incluimos c para ver el integrando como la inversa de la velocidad de fase $w_f = c/n$ de la onda en un medio con índice de refracción n. La velocidad de fase se relaciona con la longitud de onda λ y la frecuencia ν : $w_f = \lambda\nu$. Entonces, comparando ambos principios variacionales vemos que

$$\frac{n}{c} = \frac{\lambda^{-1}}{\nu} \qquad \text{juega el papel de} \qquad \frac{\sqrt{2m(E-V)}}{E} = \frac{p}{E}$$

1. En la Antigüedad, Herón de Alejandría ya había notado que el rayo de luz reflejado en un espejo toma el camino más corto entre el objeto y el observador.

Esta identificación muestra el sentido de las ideas de L. de Broglie, quien en 1924 postuló que las partículas poseen propiedades de onda cuya relación con sus propiedades de partícula son similares a las de los fotones: $E = h\nu$, $p = h\lambda^{-1}$ (h es la constante de Planck).[2] Así la regla de cuantificación de Wilson-Sommerfeld, $\oint p\ dq = Nh$, aplicada a la órbita del electrón en el átomo implica que el perímetro de la órbita $\oint dq$ es igual a $N\lambda$; o sea que la longitud de onda λ de la *onda piloto* de de Broglie cabe un número entero de veces en la órbita.

20.1.2 Ecuación de la eikonal

La óptica geométrica del principio de Fermat no es más que una aproximación de la óptica física, que es válida esencialmente en aquellas regiones del espacio donde la onda luminosa pueda aproximarse por una onda plana.

En la ecuación de onda

$$\nabla^2\varphi - \frac{n^2}{c^2}\,\frac{\partial^2\varphi}{\partial t^2} = 0\ ,$$

donde n es el índice de refracción, consideremos una onda monocromática

$$\varphi(\vec{r},t) = A(\vec{r})\ e^{i\,[2\pi\nu t\ -\ \phi(\vec{r})]}$$

donde $A(\vec{r})$ es una amplitud que puede depender de la posición. Entonces la ecuación de onda queda

$$\nabla^2\varphi = -\left(\frac{2\pi\nu n}{c}\right)^2\ \varphi \tag{20.9}$$

Por otro lado es

$$\begin{aligned}\nabla^2\varphi &= \nabla\cdot\vec{\nabla}\varphi = \nabla\cdot\left[\left(\vec{\nabla}A - iA\ \vec{\nabla}\phi\right)\ e^{i\,[2\pi\nu t\ -\ \phi(\vec{r})]}\right]\\ &= \left(\nabla^2 A - i\ 2\vec{\nabla}A\cdot\vec{\nabla}\phi - iA\ \nabla^2\phi - A\ |\vec{\nabla}\phi|^2\right) e^{i\,[2\pi\nu t\ -\ \phi(\vec{r})]}\end{aligned}$$

que sustituiremos en (20.9) descomponiendo en partes real e imaginaria:

$$\nabla^2 A - A\ |\vec{\nabla}\phi|^2 = -\left(\frac{2\pi\nu n}{c}\right)^2 A \tag{20.10}$$

$$2\vec{\nabla}A\cdot\vec{\nabla}\phi + A\ \nabla^2\phi = 0 \tag{20.11}$$

2. La velocidad de grupo $w_g = \partial\nu/\partial(\lambda^{-1})\ |_{\lambda_o}$ de un paquete de ondas formado por ondas monocromáticas de longitudes de onda similares, que varían poco alrededor de un valor λ_o, se identifica entonces con $\partial E/\partial p = \dot{q}$. En un medio no dispersivo, donde el índice de refracción no depende de λ, es $w_g = \partial(c\ \lambda^{-1}/n)/\partial(\lambda^{-1}) = c/n = w_f$. Para relaciones de dispersión no triviales, w_g (la velocidad del rayo o de la energía) no coincide con w_f (la velocidad de los frentes de onda, medida sobre la perpendicular a los mismos). Para la relación de dispersión relativista $E = c\sqrt{p^2 + m^2c^2}$ se obtiene que $w_g = c^2/w_f$.

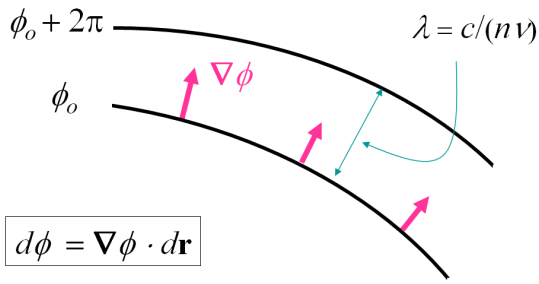
En general, para una magnitud A cualquiera, es $|\nabla^2 A/A| \sim d^{-2}$ donde d es una distancia característica del cambio de A. En la óptica geométrica están excluidos los fenómenos de interferencia, donde la amplitud A varía sensiblemente en dimensiones del orden de la longitud de onda $\lambda = c/(n\nu)$; por lo tanto se requiere que $d >> \lambda$, es decir $|\nabla^2 A/A| << \lambda^{-2}$. Haciendo esta aproximación en la ecuación (20.10) queda la **ecuación de la eikonal**,

$$|\vec{\nabla}\phi|^2 \simeq \left(\frac{2\pi\nu n}{c}\right)^2 \tag{20.12}$$

A su vez, la ecuación (20.11) dice que

$$\frac{\nabla^2\phi}{|\vec{\nabla}\phi|} = -2\,\frac{|\vec{\nabla}A|}{A}\,\cos\alpha\ ,$$

donde α es el ángulo entre $\vec{\nabla}\phi$ y $\vec{\nabla}A$. El miembro derecho tiene valor absoluto mucho menor que λ^{-1}, lo que significa que los cambios de $\vec{\nabla}\phi$ ocurren en dimensiones mucho mayores que la longitud de onda. Si en una región la dirección de $\vec{\nabla}\phi$ es aproximadamente constante entonces los frentes de onda (las superficies de igual fase ϕ a cada t) se parecerán a planos.[3] Según la ecuación (20.12), para que $|\vec{\nabla}\phi|$ sea aproximadamente constante deberá también serlo el índice de refracción n; entonces la longitud de onda $\lambda = c/(n\nu)$ –la distancia entre las crestas de la onda– será aproximadamente constante y los frentes de onda serán aproximadamente paralelos:



La analogía entre la óptica geométrica y la mecánica puede verse ahora como la analogía entre la ecuación de la eikonal (20.12) y la ecuación de Hamilton-Jacobi

$$H\left(q, \frac{\partial S_o}{\partial q}\right) = E\ ,$$

que para una partícula es

$$|\vec{\nabla}S_o|^2 = 2m(E - V)\ .$$

3. En un medio isótropo los rayos tienen dirección perpendicular a los frentes de onda, que es la dirección de $\vec{\nabla}\phi$.

Para que tenga sentido asociar $\vec{p} = \vec{\nabla} S_o$ con rayos, y S_o con la fase de los frentes de onda, deberíamos imaginar en el espacio una superficie inicial donde un conjunto de partículas de igual energía E comienza su movimiento en dirección perpendicular a la misma. Naturalmente este cuidado no es necesario con la luz porque los rayos de luz no son entes individuales sino manifestaciones de un mismo campo.

Como la eikonal ϕ es adimensional, dividiremos S_o por la la constante de Planck $\hbar \equiv h/(2\pi)$ para escribir la ecuación de Hamilton-Jacobi como

$$\left| \vec{\nabla} \, \frac{S_o}{\hbar} \right|^2 = \frac{2m(E-V)}{\hbar^2}$$

La analogía entonces es

$$\frac{S_o}{\hbar} \quad \longleftrightarrow \quad \phi$$

$$\frac{\sqrt{2m(E-V)}}{\hbar} \quad \longleftrightarrow \quad \frac{2\pi\nu n}{c} = \frac{2\pi}{\lambda}$$

La última ecuación también dice que

$$\lambda = \frac{h}{p}$$

que es la longitud de onda de la onda piloto de de Broglie.

20.2 Ecuación de Schrödinger

En 1926 Schrödinger propuso darle un tratamiento ondulatorio a la Mecánica, reemplazando la ecuación de Hamilton-Jacobi (el análogo de la ecuación de la eikonal) por una ecuación de ondas como la (20.9):

$$\nabla^2 \psi = -\frac{2m(E-V)}{\hbar^2} \; \psi$$

es decir

$$-\frac{\hbar^2}{2m} \; \nabla^2 \psi + V \; \psi = E \; \psi$$

que es la ecuación de Schrödinger independiente del tiempo. En lugar de las reglas de cuantificación primitivas, los valores permitidos de la energía E, y otras magnitudes

compatibles, son determinados por las soluciones de esta ecuación. Inmediatamente Schrödinger propuso una ecuación más general,

$$\hat{H}\left(q_\nu,\ \frac{\hbar}{i}\frac{\partial}{\partial q_\nu}\right)\Psi(q_\mu,t) = i\hbar\,\frac{\partial}{\partial t}\Psi(q_\mu,t)$$

donde el operador $\hat{H}$ es el Hamiltoniano con los operadores $(\hbar/i)(\partial/\partial q_\nu)$ reemplazando a los momentos p_ν.[4] El valor de $|\Psi(q_\mu,t)|^2$, convenientemente normalizado, describe la probabilidad de encontrar el sistema en una determinada configuración $\{q_\mu\}$ a tiempo t. En el caso de una partícula, la última ecuación se reduce a la anterior si la función de onda es *estacionaria* $\Psi(\vec{r},t) = \psi(\vec{r})\ e^{-i\frac{E}{\hbar}t}$, donde vemos la asociación entre energía y frecuencia propuesta por de Broglie. Esta descripción *ondulatoria* de la mecánica –mediante una densidad de probabilidades– implica que las partículas manifiesten fenómenos de interferencia del tipo de los de la óptica física.

Material histórico

E. Schrödinger, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, Physical Review **28**, 1049 (1926). http://people.isy.liu.se/jalar/kurser/QF/references/Schrodinger1926c.pdf

4. La construcción de $\hat{H}$, sin embargo, puede estar sujeta a ambigüedades debidas a la no conmutatividad entre q_ν y $\partial/\partial q_\nu$.