

Libros de **Cátedra**

Notas de Mecánica Analítica

Preparadas para el curso
de la Licenciatura en Física

Nicolás Grandi

FACULTAD DE
CIENCIAS EXACTAS

e
exactas


EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA

Notas de Mecánica Analítica

Preparadas para el curso de la Licenciatura en Física

Nicolás Grandi

Facultad de Ciencias Exactas



Dedicado a Héctor Vucetich, y a su pasión por la ciencia y el conocimiento.

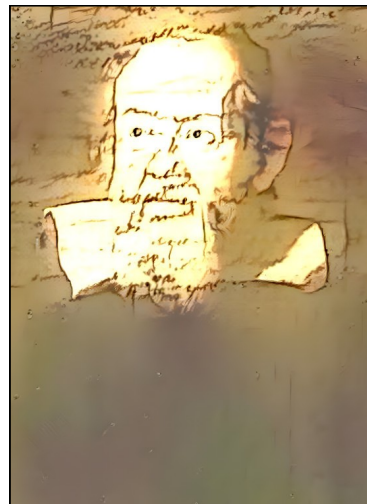
Prólogo

Sobre la Mecánica

La *Mecánica* es la ciencia del movimiento, es decir el conjunto de reglas que nos permiten predecir cómo cambiarán con el tiempo las mediciones realizadas sobre un sistema físico.

Su origen se remonta a los antiguos griegos, principalmente en la figura de Aristóteles y sus ideas de causalidad y de ley natural. En su forma moderna, se funda en el célebre trabajo de Isaac Newton *Philosophiæ naturalis principia mathematica* publicado en 1687. En él, Newton propuso una descripción matemática unificada, englobando las observaciones terrestres que Galileo Galilei había resumido en su *principio de inercia*, junto a las

tres leyes empíricas que había formulado Johannes Kepler para describir el movimiento de los cuerpos celestes. En este sentido, se trató de la primera *unificación* de la física, y sería seguida por otras en los siglos subsiguientes. Puede decirse que el trabajo de Newton se encuentra en la base de la revolución científica, cuyas consecuencias resultaron en una completa reorganización social y una mejora sustancial en la calidad de vida de la humanidad. La Mecánica se desarrolló enormemente durante los siglos XVIII y XIX, siendo de particular interés para los fines de este curso las contribuciones de Joseph-Louis Lagrange y William Rowan Hamilton, quienes dieron forma a lo que hoy conocemos como *Mecánica Analítica*.



Galileo Galilei

En la segunda mitad del siglo XIX, los trabajos de Ludwig Boltzmann demostraron que las leyes de la Mecánica podían combinarse con las ideas de probabilidad, para aplicarlas al comportamiento de un gran número de partículas. Esta construcción predecía correctamente muchas propiedades de los sistemas macroscópicos, motivando la unificación de Mecánica y Termodinámica que hoy conocemos como *Mecánica Estadística*. Casi en paralelo, otra unificación tenía lugar en la física: la de la Óptica, la Electricidad y el Magnetismo, dando lugar a lo que llamamos hoy *Electromagnetismo*. A principios del siglo XX una de las consecuencias del Electromagnetismo, en particular el carácter absoluto de la velocidad de la luz, parecía contradecir los principios de la Mecánica. Esta oposición fue resuelta por Albert Einstein, quien modificó la Mecánica unificándola con el Electromagnetismo en la forma de la *Mecánica Relativista*. Diez años más tarde, Einstein incluiría también a la Gravitación, al formular su célebre teoría de la *Relatividad General*.

Hasta ese momento histórico, la Mecánica Analítica había resistido la unificación con otras áreas de la física sin mayores modificaciones, y parecía explicar la totalidad de los fenómenos. Sin embargo, su némesis llegaría al confrontarla con la Química: una descripción mecánica de la estructura atómica, y en particular de la interacción de los átomos con la luz, parecía requerir de la hipótesis extra de *cuantización*. Según este principio, la energía se intercambiaría entre sistemas físicos en forma de paquetes o *cuantos* de valor fijo. Esta idea resistió su incorporación en las leyes generales de la Mecánica, siendo en cambio compatible con una descripción ondulatoria del movimiento. En consecuencia, para el estudio de sistemas a la escala atómica la Mecánica fue reemplazada por una nueva teoría, basada en principios muy diferentes, que hoy llamamos *Mecánica Cuántica*.

La Mecánica Cuántica incluye a la Mecánica Analítica cuando los cuerpos que describimos son lo suficientemente grandes. Este *límite clásico* hace que el estudio de la Mecánica siga siendo fundamental para la comprensión de una enorme variedad de fenómenos físicos. Más aún, sabemos hoy que varias de las leyes de la Mecánica Analítica contenían desde su formulación algunos rastros de su origen cuántico, por lo que su análisis permitía predecir la necesidad de una ulterior modificación.

Sobre estas Notas

Este libro está basado en los cursos dictados en el Departamento de Física de la Universidad Nacional de La Plata durante los años 2020 y 2021, para estudiantes de la Licenciatura en Física. Durante esos años, la pandemia de COVID19 impuso el formato de clases virtuales a través de videoconferencia. Para las mismas, resultó de enorme utilidad la combinación del software de videoconferencias Jitsi con la pizarra virtual HedgeDoc. Al final de cada clase, lo escrito en la pizarra quedaba disponible en formato Markdown, y se publicaba en Internet para consulta de los estudiantes. Dichas notas constituyeron la semilla de este texto

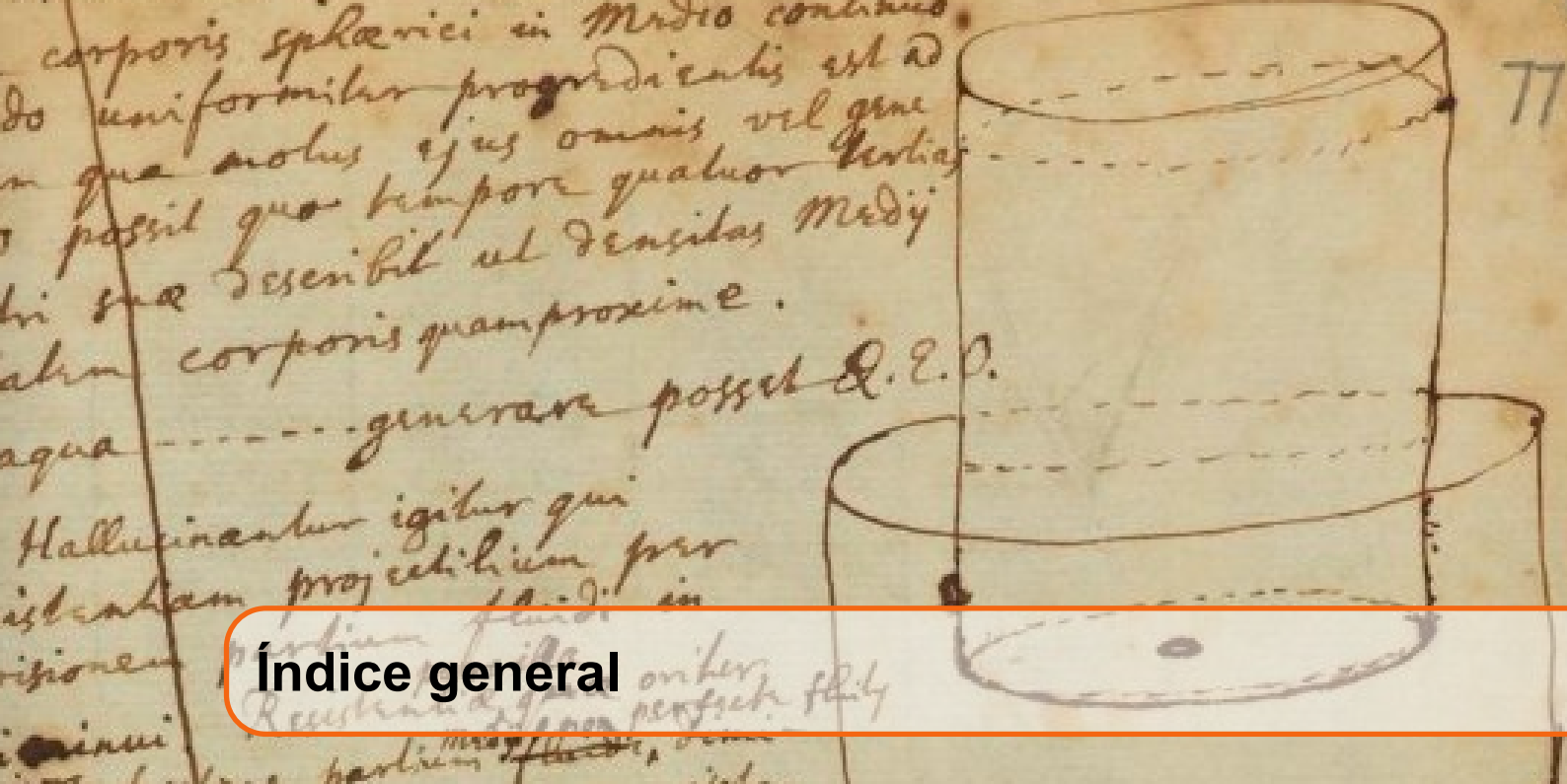
La primera parte se concentra en el formalismo básico. Esto incluye la *formulación newtoniana* de la Mecánica, que es la que se estudia en los cursos básicos de física. En ella, las nociones de *fuerza* y *partícula* cumplen un rol central. Suele estar descrita en términos de un conjunto de vectores cuyas componentes se escriben en un sistema cartesiano de coordenadas, configurando lo que llamamos *mecánica vectorial*. La reescritura de las ecuaciones involucradas en coordenadas curvilíneas generales conduce a la *formulación lagrangiana*, en la que toda la información de la dinámica del sistema está contenida en una sola función o *lagrangiano*, que depende de las *coordenadas generalizadas* y *velocidades generalizadas* de sus partículas.

La segunda parte se dedica a las aplicaciones. El formalismo básico se utiliza en la descripción del movimiento de dos sistemas mecánicos de gran interés: el *sistema de dos cuerpos* con interacción central, y el *cuerpo rígido*. Respecto del primero, se discuten sus propiedades generales y sus consecuencias para el estudio del movimiento planetario y del fenómeno de la dispersión. En cuanto al segundo, se describe la dinámica de un cuerpo rígido libre y el movimiento del trompo. Recorriendo este camino, se adquiere práctica en la descripción del movimiento en términos del formalismo lagrangiano, aprovechando sus principales virtudes.

En la tercera parte se estudian las consecuencias conceptuales de la formulación lagrangiana. En particular, se formaliza la reducción mediante *truncaciones* y *vínculos* del número de variables que describen el movimiento, y luego se aprende a relajar los vínculos permitiendo *pequeñas oscilaciones*. Se formula el *principio de acción estacionaria*, en el cual las ecuaciones que describen el movimiento se deducen minimizando una cierta cantidad. Se explora la relación entre simetrías y cargas conservadas demostrando el *teorema de Noether*, que es uno de los resultados más profundos e importantes de la física teórica.

La cuarta parte desarrolla la *formulación hamiltoniana*. Se exploran las consecuencias que el principio general de *reversibilidad* tiene en la forma de las ecuaciones de movimiento de un sistema mecánico. Se define una manera de combinar magnitudes observables conocida como *paréntesis de Poisson*, que permite reformular la Mecánica Analítica poniendo en evidencia su similitud con la Mecánica Cuántica. Se exploran también las *transformaciones canónicas* y la *ecuación de Hamilton-Jacobi*, que es una huella en la mecánica de las partículas del carácter ondulatorio de la función de onda. Finalmente, se discute la incorporación de vínculos en la formulación hamiltoniana.

Disfruté mucho del dictado de este curso, que se refiere a una de las áreas más elegantes de la Física Teórica. Espero humildemente haber podido retratar en estas notas algo de su belleza.



Índice general

I	Formalismo básico	
Clase 1	Formulación newtoniana	17
1.1	Objetivos	17
1.2	Mecánica vectorial	17
1.2.1	Leyes de Newton	18
1.2.2	Teoremas de conservación	21
1.3	Resumen	25
Clase 2	Coordenadas generalizadas	27
2.1	Objetivos	27
2.2	Coordenadas generalizadas	27
2.2.1	Ecuaciones de Lagrange	28
2.2.2	Fuerzas conservativas y monogénicas	32
2.2.3	Fuerzas dependientes de las velocidades	33

2.2.4	Teoremas de conservación	35
2.3	Resumen	38

II

Aplicaciones

Clase 3	El problema de dos cuerpos I: Definición	41
3.1	Objetivos	41
3.2	El problema de dos cuerpos	41
3.2.1	Problema equivalente de una partícula	44
3.2.2	Solución angular	47
3.2.3	Solución radial	48
3.3	Resumen	51
Clase 4	El problema de dos cuerpos II: Órbitas	53
4.1	Objetivos	53
4.2	Órbitas del problema de dos cuerpos	53
4.2.1	Teorema de Bertrand	56
4.3	Resumen	60
Clase 5	El problema de dos cuerpos III: Kepler	61
5.1	Objetivos	61
5.2	El problema de Kepler	61
5.2.1	Órbitas del problema de Kepler	62
5.2.2	Movimiento en el tiempo	68
5.2.3	Vector de Laplace-Runge-Lenz	69
5.3	Resumen	71

Clase 6 El problema de dos cuerpos VI: Dispersión 73

6.1	Objetivos	73
6.2	El problema de la dispersión	73
6.2.1	Ángulo de dispersión	74
6.2.2	Sección eficaz	78
6.2.3	Experimentos de dispersión	80
6.3	Resumen	86

Clase 7 Cuerpo rígido I: Rotaciones 87

7.1	Objetivos	87
7.2	Rotaciones espaciales	87
7.2.1	Transformaciones ortonogonales	91
7.2.2	Teorema de Euler	95
7.2.3	Rotaciones infinitesimales	99
7.3	Resumen	103

Clase 8 Cuerpo rígido II: Cinemática e inercia 105

8.1	Objetivos	105
8.2	Cinemática del cuerpo rígido	105
8.3	Momento de inercia	108
8.3.1	Tensor momento de inercia	109
8.3.2	Ejes principales	113
8.4	Resumen	117

Clase 9 Cuerpo rígido III: Dinámica 119

9.1	Objetivos	119
------------	------------------	------------

9.2	Dinámica del cuerpo rígido	119
9.2.1	Ecuaciones de Euler	119
9.2.2	Cuerpo rígido libre	121
9.2.3	Trompo	124
9.3	Resumen	129

III

Formulación lagrangiana

Clase 10 Transformaciones de Legendre 133

10.1	Objetivos	133
10.2	Transformación de Legendre	133
10.2.1	Hamiltoniano	138
10.2.2	Ruthiano	141
10.3	Truncaciones	143
10.4	Resumen	146

Clase 11 Vínculos 147

11.1	Objetivos	147
11.2	Origen físico de los vínculos	148
11.2.1	Vínculos holonómicos y coordenadas adaptadas	148
11.2.2	Vínculos no holonómicos	159
11.2.3	Multiplicadores de Lagrange	161
11.3	Resumen	167

Clase 12 Pequeñas oscilaciones 169

12.1	Objetivos	169
12.2	Oscilaciones en sistemas autónomos	169
12.3	Generalización	179

12.4	Resumen	182
------	---------	-----

Clase 13 Principio de acción estacionaria 183

13.1	Objetivos	183
13.2	Funcionales y cálculo variacional	183
13.3	Principio de Hamilton de acción estacionaria	191
13.3.1	Acción para las ecuaciones de Hamilton	196
13.3.2	Generalización de los sistemas mecánicos	197
13.4	Resumen	204

Clase 14 Simetrías 205

14.1	Objetivos	205
14.2	Simetrías	205
14.2.1	Teorema de Noether	207
14.3	Resumen	212

IV

Formulación hamiltoniana

Clase 15 Reversibilidad 215

15.1	Objetivos	215
15.2	Sistemas con configuraciones discretas	216
15.2.1	Leyes de evolución	216
15.2.2	Reversibilidad y disipación	219
15.2.3	Configuraciones y estados	221
15.3	Sistemas con configuraciones continuas	226
15.4	Resumen	233

Clase 16	Paréntesis de Poisson	235
16.1	Objetivos	235
16.2	Paréntesis de Poisson	235
16.2.1	Álgebra de observables	239
16.3	Resumen	243
Clase 17	Transformaciones canónicas	245
17.1	Objetivos	245
17.2	Transformaciones canónicas	245
17.2.1	Transformaciones canónicas infinitesimales	250
17.2.2	Versión hamiltoniana del teorema de Noether	252
17.3	Resumen	255
Clase 18	Ecuación de Hamilton-Jacobi	257
18.1	Objetivos	257
18.2	La ecuación de Hamilton Jacobi	257
18.3	Separación de variables	260
18.4	Resumen	264
Clase 19	Vínculos hamiltonianos	265
19.1	Objetivos	265
19.2	Vínculos hamiltonianos	265
19.2.1	Vínculos primarios y hamiltoniano total	267
19.2.2	Vínculos secundarios y hamiltoniano extendido	269
19.2.3	Transformaciones de gauge	272
19.3	Resumen	274



Formalismo básico

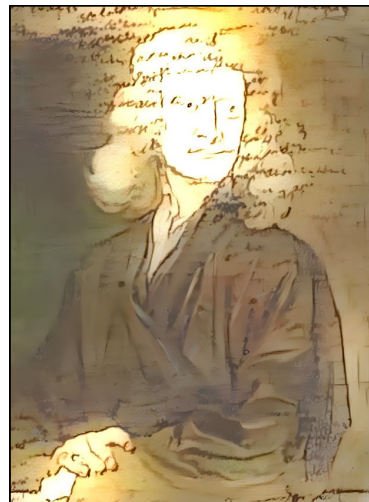
Clase 1	Formulación newtoniana	17
1.1	Objetivos	
1.2	Mecánica vectorial	
1.3	Resumen	
Clase 2	Coordenadas generalizadas	27
2.1	Objetivos	
2.2	Coordenadas generalizadas	
2.3	Resumen	

1. Formulación newtoniana

1.1 Objetivos

En esta clase haremos un breve repaso de las leyes de la mecánica en la forma en las que las aprendimos en los cursos básicos de Física, algo que en adelante referiremos como *formulación newtoniana*. Dado que trabajaremos exclusivamente en coordenadas cartesianas y haremos uso intensivo de vectores, denotaremos este conjunto de conocimientos como *mecánica vectorial*.

Además de recordar algunos elementos básicos que necesitaremos para el resto del curso, nos interesa aquí puntualizar una variedad de aspectos novedosos que serán importantes en las clases subsiguientes. Nos concentraremos por ahora en la descripción de sistemas mecánicos formados por partículas, lo que nos permitirá más adelante explorar sus generalizaciones.



Sir Isaac Newton

1.2 Mecánica vectorial

La mecánica vectorial se basa en la intuición de que las reglas que determinan evolución temporal de los sistemas físicos pueden comprenderse completamente en términos de las que rigen la evolución de cada una de sus partes y las interacciones entre ellas.

Intentemos darle una forma más concreta a esta idea.

Comencemos con un sistema físico cualquiera, del cual siempre se puede dar una descripción macroscópica en términos de un conjunto de cuerpos (objetos que se resisten a la deformación, si bien no es necesario por ahora asumir que sean completamente rígidos) y un conjunto de fluidos (objetos que se deforman fácilmente, aunque tal vez oponiendo alguna resistencia).

Si ahora refinamos nuestra descripción, observando el mismo sistema a través de una lupa o un microscopio, veremos que cada una de sus partes está a su vez formada por un conjunto de cuerpos y un conjunto de fluidos.

Aumentando iterativamente la resolución de nuestra lupa, podemos continuar mejorando nuestra descripción. La idea central de la mecánica vectorial es que este proceso de refinamiento conduce en algún punto a una descripción del sistema hecha completamente en términos de objetos muy simples, que llamamos *partículas*.

Una partícula es un objeto que, a la escala a la que observamos nuestro sistema, está completamente descrito solamente por su posición, por lo que no es necesario hacer consideraciones sobre su orientación ni sobre su estado de deformación. En este sentido, es el sistema mecánico más simple posible, que no requiere ser caracterizado como cuerpo o fluido. Nótese que no estamos asumiendo que las partículas no tengan partes, sino solamente que a la escala a la que llegamos a ver con nuestro microscopio estas partes no son visibles y no juegan ningún rol en la descripción del movimiento del sistema.

1.2.1 Leyes de Newton

Supongamos entonces que el universo entero, es decir el conjunto de todo lo que observamos, está formado por \mathcal{N} partículas, que indexaremos con un índice $n \in \{1, 2, \dots, \mathcal{N}\}$. Su movimiento estará regido por las *leyes de Newton*, que se enuncian como sigue:

1. Existencia de un sistema inercial

A los fines de cualquier experimento, un *sistema inercial* se define tomando una partícula lo suficientemente alejada de todos los demás objetos del universo, y poniendo en ella el origen de un sistema de coordenadas cartesiano, cuyos ejes son perpendiculares entre sí, y están inmóviles respecto de las estrellas lejanas.

Un sistema inercial asume también la existencia de un *reloj*, que es un conjunto de partículas que realizan movimientos periódicos, retornando repetidamente a la misma configuración.

Cualquier otro sistema de coordenadas cartesianas que se mueva con movimiento rectilíneo y uniforme respecto del arriba definido, y cuyos ejes no estén girando respecto de él, también es un sistema inercial. Más aún, cualquier otro sistema inercial estará en movimiento rectilíneo y uniforme respecto del que ya definimos, y sus ejes no estarán girando. Aquí, un movimiento *uniforme* es uno en el cual se recorre la misma distancia en cada período del reloj.

Un detalle muy importante es que esta ley **no es** un caso especial de la segunda ley: no hemos hablado aquí de lo que pasa “en ausencia de fuerzas” porque aún no hemos definido lo que es una fuerza.

2. Ley de fuerza

En un sistema inercial la *posición* de la partícula n -ésima del universo se describe con un vector $\vec{r}_n \in \mathbb{R}^3$, y llamamos *tiempo* al número $t \in \mathbb{R}$ de periodos que ha recorrido el reloj.

Naturalmente \vec{r}_n será una función de t , y podemos denotar a su derivada temporal o *velocidad* como $\dot{\vec{r}}_n = d\vec{r}_n/dt$. Además, para cada una de las partículas del universo existe otro vector *momento lineal* \vec{p}_n que es una función de la velocidad $\dot{\vec{r}}_n$. Este vector satisface la *ley de fuerza*

$$\vec{F}_n = \frac{d\vec{p}_n}{dt}$$

donde \vec{F}_n es la *fuerza* que actúa sobre la partícula n -ésima, y es una función de las posiciones y velocidades de todas las partículas del universo, y eventualmente también del tiempo. Esta función debe cumplir una serie de condiciones de consistencia:

- Primero que nada, si una de las partículas de nuestro sistema está lo bastante lejos de todas las demás, su posición \vec{r}_n es por definición el origen de otro sistema inercial. Pero dijimos en la primera ley que los sistemas inerciales están en movimiento rectilíneo y uniforme entre sí. Esto implica que la velocidad de dicha partícula $\dot{\vec{r}}_n$ debe ser necesariamente constante. Como el momento lineal \vec{p}_n es una función de la velocidad, este también debe ser constante $d\vec{p}_n/dt = 0$. Con esto, concluimos que para una partícula muy alejada de todas las otras, se debe cumplir que $\vec{F}_n = 0$. En otras palabras, la fuerza se anula cuando las partículas se alejan lo suficiente.
- Supongamos ahora que tenemos dos sistemas inerciales en reposo uno respecto del otro, relacionados por una rotación de ángulo θ alrededor de un eje que apunta en la dirección \hat{n} . Si la velocidad de una partícula es $\dot{\vec{r}}_n$ en el primer sistema, en el otro estará dada por un vector rotado un ángulo θ alrededor de \hat{n} . La misma observación debe ser válida para el momento lineal \vec{p}_n , que en el sistema rotado habrá girado un ángulo θ alrededor de \hat{n} . Esto debe cumplirse para cualquier par de sistemas inerciales en reposo relativo. Es fácil convencerse de que esto solo es posible si la relación entre ambos es la proporcionalidad $\vec{p}_n = m_n \dot{\vec{r}}_n$. Aquí m_n es una función de $\dot{\vec{r}}_n$ que tiene que ser invariante frente a rotaciones, es decir que tiene que ser una función de $\dot{\vec{r}}_n^2$.
- La forma precisa de m_n como función de $\dot{\vec{r}}_n^2$ depende de cómo se transforme el tiempo entre los diferentes sistemas inerciales. Si los relojes se sincronizan de modo tal que la velocidad de la luz c sea la misma en todos los sistemas, sabemos que se debe cumplir $m_n = m_n^{\text{rep}} / \sqrt{1 - \dot{\vec{r}}_n^2/c^2}$ donde m_n^{rep} es la *masa en reposo*. En el límite de pequeñas velocidades $\dot{\vec{r}}_n^2 \ll c^2$ tenemos que $m_n \approx m_n^{\text{rep}}$. Buena parte de lo que vamos a discutir a continuación y en el resto del curso se sostiene para una función m_n arbitraria.

3. Ley de acción y reacción

Para enunciar esta ley tenemos que hacer la hipótesis de que las fuerzas satisfacen una suerte de principio de superposición. En efecto, vamos a suponer que la fuerza sobre cualquier partícula se puede descomponer de la siguiente manera

$$\vec{F}_n = \sum_{m=1}^{\mathcal{N}} \vec{F}_{m \rightarrow n}$$

donde la suma recorre todas las partículas del universo, y el símbolo $\vec{F}_{m \rightarrow n}$ representa la fuerza que la partícula m -ésima ejerce sobre la partícula n -ésima. La *ley de acción y reacción* dice que para cualquier par de partículas, estas fuerzas cumplen

$$\vec{F}_{m \rightarrow n} = -\vec{F}_{n \rightarrow m}$$

Nótese que en particular esto implica $\vec{F}_{n \rightarrow n} = -\vec{F}_{n \rightarrow n}$, por lo que la *autofuerza* debe anularse $\vec{F}_{n \rightarrow n} = 0$ y la suma de más arriba corre solamente sobre $m \neq n$.

Si de entre las \mathcal{N} partículas que componen el universo seleccionamos un subconjunto de N de ellas, tendremos un *sistema de N partículas*. La *configuración* dicho sistema estará dada por los N vectores $\vec{r}_n \in \mathbb{R}^3$ que representan la posición de cada una de las partículas en algún sistema inercial en un dado instante de tiempo. Definimos entonces un *espacio de configuración* del sistema de N partículas como $\mathcal{C} = \mathbb{R}^{3N}$.

La posición cambiará a medida que transcurre el tiempo describiendo una curva en el espacio de configuración \mathcal{C} que se denomina la *trayectoria* del sistema. Tanto la forma de tal curva cuanto la velocidad con la que se la recorre a medida que transcurre el tiempo, se obtienen resolviendo un conjunto de *ecuaciones de movimiento* dadas por la ley de fuerza. Nótese de que hay una ecuación vectorial para cada partícula, resultando en un total de $3N$ ecuaciones. En estas ecuaciones, las fuerzas sobre cada partícula \vec{F}_n deben satisfacer la ley de acción y reacción.

Las ecuaciones de movimiento son ecuaciones diferenciales de segundo orden en el tiempo, por lo que cada una de ellas resultará en dos constantes de integración. Dado que tenemos un conjunto de $3N$ ecuaciones, deberemos determinar un total de $6N$ constantes. Para esto, será necesario especificar las *condiciones iniciales*, es decir los valores iniciales para los vectores posición \vec{r}_n y velocidad $\dot{\vec{r}}_n$ de cada una de las partículas. Estos valores iniciales determinan completamente la evolución ulterior, es decir que definen el *estado* del sistema. Podemos hablar entonces de un *espacio de estados* dado por $\mathcal{E} = \mathbb{R}^{6N}$.

Los sistemas físicos difieren entre sí en el número de partículas, y en la forma explícita de la fuerza \vec{F}_n como función de la posición y la velocidad de todas las partículas del universo, y de la masa m_n como función de \dot{r}_n^2 .

Nota:

Escribe Isaac Asimov: sobre *¿Quién fue el científico más grande de la historia?*:

«Si la pregunta fuese *¿Quién fue el segundo científico más grande?* sería imposible de contestar. Hay por lo menos una docena de hombres que, en mi opinión, podrían aspirar a esa segunda plaza. Entre ellos figurarían, por ejemplo, Albert Einstein, Ernest Rutherford, Niels Bohr, Louis Pasteur, Charles Darwin, Galileo Galilei, Clerk Maxwell, Arquímedes y otros.

Incluso es muy probable que ni siquiera exista eso que hemos llamado el segundo científico más grande. Las credenciales de tantos y tantos son tan buenas y la dificultad de distinguir niveles de mérito es tan grande, que al final quizá tendríamos que declarar un empate entre diez o doce.

Pero como la pregunta es «¿Quién es el más grande?», no hay problema alguno. En mi opinión, la mayoría de los historiadores de la ciencia no dudarían en afirmar que Isaac Newton fue el talento científico más grande que jamás haya visto el mundo. Tenía sus faltas, viva el cielo: era un mal conferenciante, tenía algo de cobarde moral y de llorón autocompasivo y de vez en cuando era víctima de serias depresiones. Pero como científico no tenía igual.

Fundó las matemáticas superiores después de elaborar el cálculo. Fundó la óptica moderna mediante sus experimentos de descomponer la luz blanca en los colores del espectro. Fundó la física moderna al establecer las leyes del movimiento y deducir sus consecuencias. Fundó la astronomía moderna estableciendo la ley de la gravitación universal.

Cualquiera de estas cuatro hazañas habría bastado por sí sola para distinguirlo como científico de importancia capital. Las cuatro juntas le colocan en primer lugar de modo incuestionable.»

Las tres leyes arriba enumeradas, cuyas consecuencias exploramos en los cursos básicos de física, nos permitieron construir a partir de ellas una enorme variedad de resultados. En particular, no solo pudimos describir la dinámica de una partícula y de un sistema de partículas, sino también la de los cuerpos rígidos, de los cuerpos elásticos, y de los fluidos. De particular interés para nosotros son los *teoremas de conservación*, que se resumen en lo que sigue.

1.2.2 Teoremas de conservación

Dado un sistema físico formado por N partículas, podemos separar todas las partículas del universo entre aquellas que son internas al sistema $n \in \{1, 2, \dots, N\}$ y aquellas que son externas al mismo $n \in \{N + 1, N + 2, \dots, \mathcal{N}\}$. Una consecuencia de la descomposición de la fuerza que hicimos al

explicar la ley de acción y reacción más arriba es que se puede escribir

$$\vec{F}_n = \sum_{m=1}^N \vec{F}_{m \rightarrow n} + \sum_{m=N+1}^{\mathcal{N}} \vec{F}_{m \rightarrow n}$$

Al segundo término lo llamamos *fuerza externa* sobre la partícula n -ésima y lo denotamos \vec{F}_n^{ext} , lo que nos deja con la expresión

$$\vec{F}_n = \sum_{m=1}^N \vec{F}_{m \rightarrow n} + \vec{F}_n^{\text{ext}}$$

donde ahora la suma corre solamente sobre las partículas internas al sistema. Esta descomposición resulta muy útil para obtener los siguientes resultados

- Conservación del momento lineal

Si usamos la descomposición de la fuerza entre la contribución externa y la interna podemos reescribir la ley de fuerza en la forma

$$\sum_{m=1}^N \vec{F}_{m \rightarrow n} + \vec{F}_n^{\text{ext}} = \frac{d\vec{p}_n}{dt}$$

Ahora podemos sumar a ambos lados sobre todas las partículas de nuestro sistema, para obtener

$$\sum_{n=1}^N \sum_{m=1}^N \vec{F}_{m \rightarrow n} + \sum_{n=1}^N \vec{F}_n^{\text{ext}} = \sum_{n=1}^N \frac{d\vec{p}_n}{dt}$$

Si definimos la *fuerza externa total* $\vec{F}^{\text{ext}} = \sum_{n=1}^N \vec{F}_n^{\text{ext}}$ sobre el sistema y el *momento lineal total* del mismo $\vec{p} = \sum_{n=1}^N \vec{p}_n$, la igualdad de más arriba se puede reescribir como

$$\sum_{n=1}^N \sum_{m=1}^N \vec{F}_{m \rightarrow n} + \vec{F}^{\text{ext}} = \frac{d\vec{p}}{dt}$$

En la suma en el primer término cada par de partículas m y n entra dos veces, una vez como $\vec{F}_{m \rightarrow n}$ y otra como $\vec{F}_{n \rightarrow m}$. Esto nos permite reescribir la suma limitando los índices con la restricción $m > n$, para obtener

$$\sum_{n=1}^N \sum_{m=n+1}^N \left(\vec{F}_{m \rightarrow n} + \vec{F}_{n \rightarrow m} \right) + \vec{F}^{\text{ext}} = \frac{d\vec{p}}{dt}$$

(nótese el cambio de índices en el segundo término dentro de la suma, para que ahora cada par entre una sola vez). Pero entonces la ley de acción y reacción implica inmediatamente que el primer término se anula, dejándonos con

$$\vec{F}^{\text{ext}} = \frac{d\vec{p}}{dt}$$

De aquí deducimos que *si la fuerza externa total sobre un sistema de partículas se anula, entonces el momento lineal total se conserva.*

- Conservación del momento angular

Volviendo a la expresión para la segunda ley de Newton en términos de fuerzas externas e internas

$$\sum_{m=1}^N \vec{F}_{m \rightarrow n} + \vec{F}_n^{\text{ext}} = \frac{d\vec{p}_n}{dt}$$

Podemos tomar a cada lado de la igualdad el producto vectorial con la posición \vec{r}_n de la partícula n -ésima, para obtener

$$\sum_{m=1}^N \vec{r}_n \times \vec{F}_{m \rightarrow n} + \vec{r}_n \times \vec{F}_n^{\text{ext}} = \vec{r}_n \times \frac{d\vec{p}_n}{dt}$$

Integrando por partes el segundo término, y usando que el momento lineal es proporcional a la velocidad y por lo tanto $\dot{\vec{r}}_n \times \vec{p}_n = 0$, obtenemos la expresión

$$\sum_{m=1}^N \vec{r}_n \times \vec{F}_{m \rightarrow n} + \vec{r}_n \times \vec{F}_n^{\text{ext}} = \frac{d(\vec{r}_n \times \vec{p}_n)}{dt}$$

Ahora podemos definir el *momento angular* de cada partícula respecto del origen de coordenadas según $\vec{\ell}_n = \vec{r}_n \times \vec{p}_n$ y el *torque externo* sobre cada partícula respecto del origen como $\vec{\tau}_n^{\text{ext}} = \vec{r}_n \times \vec{F}_n^{\text{ext}}$. Con esto obtenemos

$$\sum_{m=1}^N \vec{r}_n \times \vec{F}_{m \rightarrow n} + \vec{\tau}_n^{\text{ext}} = \frac{d\vec{\ell}_n}{dt}$$

En esta expresión podemos sumar sobre todas las partículas del sistema, resultando en

$$\sum_{n=1}^N \sum_{m=1}^N \vec{r}_n \times \vec{F}_{m \rightarrow n} + \vec{\tau}^{\text{ext}} = \frac{d\vec{\ell}}{dt}$$

donde el *torque total externo* se definió como $\vec{\tau}^{\text{ext}} = \sum_{n=1}^N \vec{\tau}_n^{\text{ext}}$ y el *momento angular total* como $\vec{\ell} = \sum_{n=1}^N \vec{\ell}_n$. De nuevo en el primer término cada par de partículas entra dos veces, lo que nos permite reescribirlo como

$$\sum_{n=1}^N \sum_{m=n+1}^N \left(\vec{r}_n \times \vec{F}_{m \rightarrow n} + \vec{r}_m \times \vec{F}_{n \rightarrow m} \right) + \vec{\tau}^{\text{ext}} = \frac{d\vec{\ell}}{dt}$$

Usando la ley de acción y reacción podemos reordenar el paréntesis

$$\sum_{n=1}^N \sum_{m=n+1}^N (\vec{r}_n - \vec{r}_m) \times \vec{F}_{m \rightarrow n} + \vec{\tau}^{\text{ext}} = \frac{d\vec{\ell}}{dt}$$

Ahora bien, si hacemos la hipótesis adicional de que las fuerzas internas son *centrales*, es decir que la fuerza entre cada par de partículas apunta en la dirección que une las partículas, entonces $(\vec{r}_n - \vec{r}_m) \times \vec{F}_{m \rightarrow n} = 0$ en el primer término, y llegamos a la expresión

$$\vec{\tau}^{\text{ext}} = \frac{d\vec{\ell}}{dt}$$

Con lo que podemos enunciar que *si el torque total externo se anula y las fuerzas internas son centrales, el momento angular total se conserva.*

■ Conservación de la energía

Volvamos a nuestra fórmula inicial para la ley de fuerza, sin separar por ahora entre contribuciones externas e internas

$$\vec{F}_n = \frac{d\vec{p}_n}{dt}$$

Multiplicando esta expresión escalarmente por la velocidad tenemos

$$\vec{F}_n \cdot \dot{\vec{r}}_n = \frac{d\vec{p}_n}{dt} \cdot \dot{\vec{r}}_n$$

Recordemos que $\vec{p}_n = m_n \dot{\vec{r}}_n$ donde m_n es alguna función de \dot{r}_n^2 . Para el caso de una partícula no relativista m_n es simplemente una constante, para el caso relativista tenemos que $m_n = m_n^{\text{rep}} / \sqrt{1 - \dot{r}_n^2/c^2}$. Cualquiera sea la forma de esta función, sabemos que tiene una primitiva $2K_n$ que cumple que $d(2K_n)/d\dot{r}_n^2 = m_n$ (donde hemos introducido un factor 2 por conveniencia). Esta función corresponde a la energía cinética $K_n = m_n \dot{r}_n^2/2$ para el caso no-relativista, y a la expresión $K_n = m_n^{\text{rep}} c^2 \sqrt{1 - \dot{r}_n^2/c^2}$ para el caso relativista. Usando esta primitiva podemos escribir

$$\vec{F}_n \cdot \dot{\vec{r}}_n = \frac{dK_n}{dt}$$

Si ahora integramos entre un instante inicial t_i y uno final t_f obtenemos

$$W_n = \Delta K_n$$

En el lado izquierdo hemos definido el *trabajo* realizado sobre la partícula n -ésima como $W_n = \int \vec{F}_n \cdot \dot{\vec{r}}_n dt$, y la variación de su energía cinética según $\Delta K_n = K_n|_{t_f} - K_n|_{t_i}$.

Nótese que si la fuerza no depende del tiempo ni de las velocidades, entonces el trabajo es una integral de línea $W_n = \int \vec{F}_n \cdot d\vec{r}_n$. Esta integral sólo puede depender de la forma de la fuerza como función de la posición y de la trayectoria que recorre la partícula n -ésima entre su posición inicial y su posición final. Esta última observación nos permite hacer una clasificación de las fuerzas. Llamamos *conservativas* a aquellas fuerzas para las cuales el trabajo realizado en un movimiento cualquiera no depende de la trayectoria sino solamente de los puntos inicial y final. En ese caso escribimos $W_n = -\Delta V_n$ para una cierta función *energía potencial* V_n que es función sólo de la posición \vec{r}_n . Las fuerzas que no cumplen esta propiedad se consideran *no conservativas*, y denotamos el trabajo de las mismas como W_n^{nc} . Entonces tenemos que

$$W_n^{\text{nc}} - \Delta V_n = \Delta K_n$$

Ahora podemos sumar sobre n , obteniendo

$$W^{\text{nc}} - \Delta V = \Delta K$$

con las definiciones obvias para el trabajo total de las fuerzas no conservativas $W^{\text{nc}} = \sum_{n=1}^N W_n^{\text{nc}}$, la energía potencial total $V = \sum_{n=1}^N V_n$ y la energía cinética total $K = \sum_{n=1}^N K_n$. Esto se puede reescribir como

$$W^{\text{nc}} = \Delta E$$

donde la energía *mecánica* total se definió como $E = K + V$.

Entonces podemos decir que *en ausencia de fuerzas no-conservativas internas o externas, la energía mecánica total se conserva*.

Estos teoremas de conservación se pueden reinterpretar en términos de principios de *simetría*, un resultado importante que reformularemos y reobtendremos más adelante en el curso. Por ahora, baste señalar que:

- El momento lineal se conserva cuando la fuerza total externa vale cero. Cuando esto sucede, el movimiento del sistema está completamente determinado por sus fuerzas internas, que dependen de las posiciones y velocidades de las partículas del sistema, y no guardan ninguna relación con las posiciones y velocidades del resto de las partículas del universo. Esto implica que podemos poner el sistema en cualquier lugar del espacio, y su movimiento ulterior no se verá afectado. En otras palabras, *hay invarianza traslacional*. Por lo tanto, la invarianza frente a traslaciones está íntimamente ligada a la conservación del impulso.
- El momento angular se conserva cuando el torque total externo vale cero y las fuerzas internas son centrales. Que el torque total externo se anule implica que la fuerza total externa también es *central*, en el sentido de que apunta hacia o desde el origen de coordenadas. Pero entonces podemos rotar rígidamente el sistema alrededor del origen y su movimiento ulterior no se verá afectado. Es decir que *hay invarianza rotacional*. Esto implica que la invarianza frente a rotaciones está íntimamente relacionada con la conservación del momento angular.
- La energía se conserva cuando todas las fuerzas, tanto internas como externas, son conservativas. Esto requiere en particular que las fuerzas sean independientes del tiempo, ya que solamente de ese modo se puede definir el trabajo como una integral de línea que, en caso de no depender de la trayectoria, implicaría una fuerza conservativa. Por lo tanto podemos decir que *tenemos invarianza frente a traslaciones temporales*. Concluimos que la conservación de la energía está íntimamente relacionada a la invarianza frente a traslaciones temporales.

1.3 Resumen

En esta clase, repasamos las ideas que se adquieren en los cursos básicos de física, en los que la mecánica se escribe en términos de las coordenadas cartesianas de un conjunto de partículas utilizando vectores.

Luego de visitar las tres leyes de Newton analizando cuidadosamente su significado, introducimos los conceptos de espacio de configuración, trayectoria, ecuaciones de movimiento, condiciones iniciales y espacio de estados. Finalmente, recorrimos los teoremas de conservación, y señalamos su relación con diferentes tipos de invarianzas en la descripción de un sistema físico.

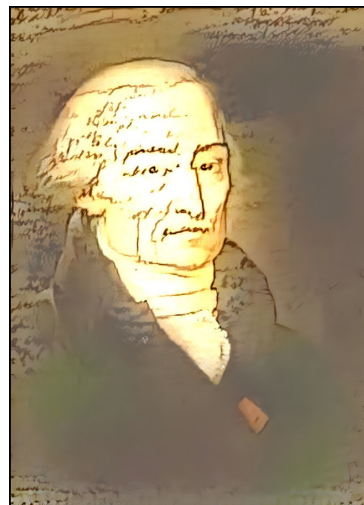
En las clases que siguen, vamos a explorar las consecuencias de estas leyes para estudiar varios sistemas mecánicos de interés. Para poder hacerlo con comodidad, nuestro primer paso será formularlas en una forma que no requiera del uso de sistemas cartesianos y vectores.

2. Coordenadas generalizadas

2.1 Objetivos

En la clase anterior repasamos la formulación newtoniana de la mecánica en su forma de mecánica vectorial, en la cual la posición de cada una de las partículas constituyentes de un dado sistema mecánico está escrita en coordenadas cartesianas.

En esta clase, vamos a reescribir la segunda ley de Newton en coordenadas curvilíneas arbitrarias, lo que nos permitirá más adelante describir una variedad de sistemas mecánicos de interés. En este proceso, descubriremos que para una clase de fuerzas bastante general, las ecuaciones de movimiento pueden escribirse de una forma única que no depende del sistema de coordenadas utilizado, en términos de derivadas de una sola función de las coordenadas curvilíneas y de sus derivadas.



Joseph-Louis Lagrange

2.2 Coordenadas generalizadas

Como vimos en la clase previa, la configuración de un sistema de N partículas está dada por N vectores \vec{r}_n que determinan un punto en el espacio de configuración $\mathcal{C} = \mathbb{R}^{3N}$ del sistema. Quisiéramos ser capaces de describir este espacio reemplazando las $\{\vec{r}_n\}$ con $n \in \{1, \dots, N\}$ por un conjunto cualquiera de $3N$ *coordenadas generalizadas* $\{q_i\}$ con $i \in \{1, \dots, 3N\}$, relacionadas con las primeras por un cambio de variables $\vec{r}_n = \vec{r}_n(q_i, t)$.

Para obtener la trayectoria del sistema, debemos resolver un conjunto de ecuaciones diferenciales dado por la segunda ley de Newton para un sistema de partículas, que en términos de las coordenadas \vec{r}_n se escribe

$$\vec{F}_n = \frac{d\vec{p}_n}{dt}$$

Una manera directa de obtener la forma de estas ecuaciones en términos de las coordenadas generalizadas q_i sería reemplazando directamente el cambio de variables en las ecuaciones. Pero esto nos obligaría a reescribir las derivadas temporales segundas de las variables \vec{r}_n en términos de las nuevas coordenadas q_i y sus derivadas primeras y segundas, usando la regla de la cadena dos veces. Dependiendo del número de partículas involucrado y de la complejidad del cambio de variables, esto puede volverse arbitrariamente complicado. En lo que sigue discutimos una técnica completamente general que nos evita calcular tales derivadas temporales segundas, requiriendo sólo las derivadas primeras.

2.2.1 Ecuaciones de Lagrange

Recordemos que el impulso se escribe como $\vec{p}_n = m_n \dot{\vec{r}}_n$ donde m_n es una función de r_n^2 . Habíamos llamado $2K_n$ a la primitiva de m_n , y vimos que correspondía a la energía cinética para el caso no-relativista. También habíamos definido su suma sobre todas las partículas del sistema como $K = \sum_{n=1}^N K_n$. En este punto nos interesa observar que usando tal *término cinético* siempre podemos escribir el momento lineal de cada partícula en la forma

$$\vec{p}_n = \frac{\partial K}{\partial \dot{\vec{r}}_n}$$

Reemplazando en la segunda ley de Newton, tenemos entonces que

$$\vec{F}_n = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{\vec{r}}_n} \right)$$

Si ahora escribimos $\vec{r}_n = \vec{r}_n(q_i, t)$ como función de un conjunto arbitrario de coordenadas generalizadas $\{q_i\}_{i \in \{1, \dots, 3N\}}$, esta forma de la ley de Newton nos permite considerar a K directamente como una función de las coordenadas generalizadas q_i , las *velocidades generalizadas* \dot{q}_i , y el tiempo. Podemos explotar esto si multiplicamos escalarmente por $\partial \vec{r}_n / \partial q_i$ y sumamos sobre n , para obtener

$$\underbrace{\vec{F}_n \cdot \frac{\partial \vec{r}_n}{\partial q_i}}_{G_i} = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{\vec{r}}_n} \right) \cdot \frac{\partial \vec{r}_n}{\partial q_i}$$

En esta fórmula y en el resto de este texto, el símbolo de sumatoria será sobrentendido cada vez que un índice se repita dos veces, en lo que se conoce como *convención de Einstein*. En este caso la suma que estamos omitiendo corre sobre el índice de partícula n , mientras que el índice de coordenada i aparece una sola vez a cada lado de la igualdad, y por eso no está sumado. En el miembro izquierdo de esta ecuación hemos definido la *fuerza generalizada* G_i asociada a la coordenada generalizada q_i . Reemplazando en ella el cambio de variables, la magnitud G_i queda

escrita como una función de las coordenadas generalizadas, las velocidades generalizadas, y el tiempo. En el miembro derecho podemos integrar por partes, obteniendo

$$G_i = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{\vec{r}}_n} \cdot \frac{\partial \vec{r}_n}{\partial q_i} \right) - \frac{\partial K}{\partial \dot{\vec{r}}_n} \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_n}{\partial q_i} \right)$$

Nótese que debido a la forma del cambio de variables $\vec{r}_n = \vec{r}_n(q_i, t)$ tenemos que las velocidades satisfacen

$$\dot{\vec{r}}_n = \frac{\partial \vec{r}_n}{\partial q_i} \dot{q}_i + \frac{\partial \vec{r}_n}{\partial t}$$

lo que en particular implica la relación entre las derivadas

$$\frac{\partial \vec{r}_n}{\partial q_i} = \frac{\partial \dot{\vec{r}}_n}{\partial \dot{q}_i}$$

Más aún, a partir de la misma fórmula también podemos probar que

$$\frac{\partial \dot{\vec{r}}_n}{\partial q_j} = \frac{\partial^2 \vec{r}_n}{\partial q_j \partial q_i} \dot{q}_i + \frac{\partial^2 \vec{r}_n}{\partial q_j \partial t} = \frac{d}{dt} \left(\frac{\partial \vec{r}_n}{\partial q_j} \right)$$

Usando estas dos últimas fórmulas en el miembro derecho de nuestra segunda ley de Newton en coordenadas generalizadas, podemos reescribirla como

$$G_i = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{\vec{r}}_n} \cdot \frac{\partial \dot{\vec{r}}_n}{\partial \dot{q}_i} \right) - \frac{\partial K}{\partial \dot{\vec{r}}_n} \cdot \frac{\partial \dot{\vec{r}}_n}{\partial q_i}$$

O en otras palabras, usando la regla de la cadena

$$G_i = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_i} \right) - \frac{\partial K}{\partial q_i}$$

En esta forma de la segunda ley de Newton tenemos una ecuación para cada coordenada generalizada q_i , en lo que se conoce como *ecuaciones de Lagrange*.

El punto importante del resultado que hemos obtenido es el siguiente: para escribir las ecuaciones de Lagrange necesitamos la forma de la fuerza generalizada G_i y del término cinético K como función de las coordenadas generalizadas q_i y las velocidades generalizadas \dot{q}_i , y eventualmente del tiempo. Esto sólo involucra calcular derivadas temporales primeras del cambio de variables, con lo que nos hemos ahorrado bastante trabajo.

Podemos usar este resultado para escribir la ley de fuerza en un sistema de coordenadas general, no necesariamente cartesiano. Por ejemplo, podríamos describir la posición de cada partícula en coordenadas esféricas, o cilíndricas. Podríamos incluso usar sistemas de coordenadas en donde cada una de las variables no se refiera a una sola partícula sino a alguna combinación de ellas. Podemos también usar sistemas no necesariamente inerciales, donde el cambio de coordenadas depende del tiempo.

Ejemplo: partícula en coordenadas polares en el plano

Supongamos que dadas las coordenadas cartesianas para una partícula en el plano $\vec{r} = (x, y)$, queremos describir su movimiento en coordenadas polares $\{q_i\} = \{r, \theta\}$. El cambio de variables tiene la forma

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

Tomando una derivada temporal obtenemos para la velocidad

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

Si quisiéramos insertar esto directamente en la ley de fuerza, tendríamos que tomar una derivada más para obtener la aceleración. Se puede ver que cada una de las derivadas segundas resultantes tiene al menos cinco términos, lo que complica bastante la obtención de las ecuaciones de movimiento.

Sin embargo, con lo calculado hasta ahora nos alcanza para escribir el término cinético K . Éste toma la forma

$$K = \frac{1}{2} m \dot{\vec{r}}^2 = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2)$$

Donde usamos que $\cos^2 \theta + \sin^2 \theta = 1$ y por otro lado los productos cruzados conteniendo $\sin \theta \cos \theta$ se cancelaron.

También podemos escribir la fuerza generalizada $G_i = \{G_r, G_\theta\}$, para la que tenemos la formula

$$G_r = \vec{F} \cdot \frac{\partial \vec{r}}{\partial r} = \vec{F} \cdot \underbrace{(\cos \theta, \sin \theta)}_{\vec{r}} \equiv F_r$$

$$G_\theta = \vec{F} \cdot \frac{\partial \vec{r}}{\partial \theta} = \vec{F} \cdot \underbrace{(-r \sin \theta, r \cos \theta)}_{r \vec{\theta}} = x F_y - y F_x \equiv \tau_z$$

En la primera ecuación vemos que la fuerza generalizada en la dirección de r viene dada por el producto escalar entre el vector fuerza y un versor que apunta en la dirección radial. Es decir que corresponde a la componente radial de la fuerza. En la segunda vemos que para la fuerza generalizada en la dirección de θ nos queda el producto vectorial entre el radio y la fuerza, es decir el torque. Esto es una realización concreta de la idea intuitiva de que *el torque es para los ángulos lo que la fuerza es para el vector posición*.

Reemplazando estos resultados en las ecuaciones de Lagrange

$$G_r = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{r}} \right) - \frac{\partial K}{\partial r} \quad G_\theta = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{\theta}} \right) - \frac{\partial K}{\partial \theta}$$

obtenemos explícitamente

$$F_r = m \ddot{r} - m r \dot{\theta}^2 = m \ddot{r} + m a_{\text{cen}} = m \ddot{r} - F_{\text{cen}}$$

$$\tau = \frac{d}{dt} (m r^2 \dot{\theta}) = \frac{d}{dt} (\mathcal{I} \omega) = \frac{d\ell}{dt}$$

Vemos que con este tratamiento aparecen automáticamente todos los conceptos relacionados con el movimiento angular, tales como el torque τ , el momento de inercia $\mathcal{I} = m r^2$, la velocidad angular ω , la aceleración centrípeta $a_{\text{cen}} = -r \dot{\theta}^2$, la *fuerza centrífuga* $F_{\text{cen}} = m r \dot{\theta}^2$, y el momento angular $\ell = m r^2 \dot{\theta}$.

Ejercicio:

Pruebe que para un conjunto de N partículas no relativistas, la energía cinética es una función cuadrática de las velocidades generalizadas. En otras palabras, podemos escribir

$$K = \frac{1}{2} K_{ij}(q_i) \dot{q}_i \dot{q}_j$$

Donde los coeficientes $K_{ij}(q_i)$ son funciones de las coordenadas generalizadas y las masas de las partículas involucradas.

Esta estructura para la energía cinética será muy frecuente en los problemas estudiados en lo que sigue.

Por analogía a la expresión que obtuvimos en coordenadas cartesianas $\vec{p}_n = \partial K / \partial \vec{r}_n$ podemos definir los *momentos generalizados* como

$$p_i = \frac{\partial K}{\partial \dot{q}_i}$$

Dado que la energía cinética es una función de las coordenadas y velocidades generalizadas, los impulsos generalizados también lo serán. Las ecuaciones de Lagrange quedan entonces escritas como

$$G_i + \frac{\partial K}{\partial q_i} = \frac{dp_i}{dt}$$

Esta fórmula se puede expresar en palabras como *la fuerza generalizada más la pseudofuerza generalizada es igual a la derivada del impulso generalizado respecto del tiempo*

Ejemplo: partícula en coordenadas polares en el plano

Volviendo al ejemplo de arriba de una partícula en el plano en coordenadas polares, podemos usar la forma del término cinético K para obtener los impulsos generalizados

$$p_r = \frac{\partial K}{\partial \dot{r}} = m\dot{r} \quad p_\theta = \frac{\partial K}{\partial \dot{\theta}} = mr^2\dot{\theta}$$

Es decir que el impulso generalizado en la dirección r no es más que el impulso radial. Por otro lado, el impulso generalizado en la dirección θ resulta ser el momento angular. En efecto, tenemos que $\ell = |\vec{r} \times \vec{p}| = rp \sin \hat{r}\hat{p} = mrv \sin \hat{r}\hat{p}$. Usando el hecho de que $v \sin \hat{r}\hat{p}$ es la componente tangencial de la velocidad, que sabemos vale $r\dot{\theta}$, podemos reescribir $\ell = mr^2\dot{\theta}$. Esto realiza de manera concreta la idea intuitiva de que “el momento angular es para los ángulos lo que el vector impulso es al vector posición.”

Ahora bien ¿cuál es el orden de las ecuaciones de Lagrange? Observando que tanto la fuerza generalizada G_i como el término cinético K son funciones de las coordenadas generalizadas q_j , las velocidades generalizadas \dot{q}_j , y el tiempo, podemos usar la regla de la cadena para escribir

$$G_i = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_i} \right) - \frac{\partial K}{\partial q_i} = \frac{\partial^2 K}{\partial \dot{q}_i \partial \dot{q}_j} \ddot{q}_j + \frac{\partial^2 K}{\partial \dot{q}_i \partial q_j} \dot{q}_j + \frac{\partial^2 K}{\partial \dot{q}_i \partial t} - \frac{\partial K}{\partial q_i}$$

donde vemos que genéricamente las ecuaciones serán de segundo orden. Dado que tenemos una ecuación para cada una de las coordenadas generalizadas q_i , esto resulta en $3N$ ecuaciones diferenciales de segundo orden. Las $6N$ constantes de integración se fijarán entonces en términos de las condiciones iniciales para q_i y \dot{q}_i , y determinarán completamente la evolución temporal. De este modo, los valores iniciales de q_i y \dot{q}_i reparametrizan nuestro espacio de estados $\mathcal{E} = \mathbb{R}^{6N}$.

Ejercicio:

Usando que $2K_n$ es por definición la primitiva de m_n como función de \dot{r}_n^2 , y aplicando la fórmula que relaciona \dot{r}_n con \dot{q}_i que obtuvimos antes, pruebe que el coeficiente de \ddot{q}_j en la última expresión no puede anularse. Esto demuestra que las ecuaciones que hemos obtenido no pueden ser de orden menor al segundo.

2.2.2 Fuerzas conservativas y monogénicas

Supongamos ahora el caso particular en que la fuerza se deriva de un potencial

$$\vec{F}_n = -\frac{\partial V}{\partial \vec{r}_n}$$

Donde el potencial depende de la posición y eventualmente del tiempo $V(\vec{r}_n, t)$. El caso independiente del tiempo lo conocemos de los cursos previos con el nombre de *fuerza conservativa*. El caso dependiente del tiempo es un tipo particular de fuerza no conservativa que se conoce como *fuerza monogénica*.

La expresión para la fuerza generalizada se puede reescribir para este tipo de fuerzas, según

$$G_i = \vec{F}_n \cdot \frac{\partial \vec{r}_n}{\partial q_i} = -\frac{\partial V}{\partial \vec{r}_n} \cdot \frac{\partial \vec{r}_n}{\partial q_i} \equiv -\frac{\partial V}{\partial q_i}$$

Es decir que la fuerza generalizada viene dada por menos la derivada del potencial respecto de la coordenada generalizada correspondiente. Reemplazando esto en las ecuaciones de Lagrange, escribimos

$$-\frac{\partial V}{\partial q_i} = \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_i} \right) - \frac{\partial K}{\partial q_i}$$

lo que se puede reordenar como

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_i} \right) - \frac{\partial (K - V)}{\partial q_i} = 0$$

Ahora bien, podemos escribir esta ecuación de una manera más simétrica si observamos que, dado que V no depende de \dot{q}_i , en el primer término podemos reemplazar K por $K - V$ obteniendo

$$\frac{d}{dt} \left(\frac{\partial (K - V)}{\partial \dot{q}_i} \right) - \frac{\partial (K - V)}{\partial q_i} = 0$$

En términos de una nueva función $L = K - V$, tenemos que las ecuaciones de Lagrange quedan escritas como

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

Donde L es lo que se conoce como el *lagrangiano* del sistema, y es una función de las coordenadas y velocidades generalizadas, y eventualmente del tiempo $L = L(\dot{q}_i, q_i, t)$.

La ventaja de esta *formulación lagrangiana* es que podemos recuperar el conjunto completo de ecuaciones de movimiento de un sistema mecánico tomando derivadas de una sola función de sus coordenadas generalizadas, sus velocidades generalizadas, y el tiempo. Esta función se puede obtener en términos de las energías cinética y potencial del sistema, lo que implica que para cambiar de coordenadas a partir de la formulación vectorial, sólo tenemos que calcular derivadas primeras respecto del tiempo.

El precio a pagar por esta simplificación fue el de restringirnos a fuerzas que se puedan derivar de un potencial. Podemos pensar que esta es una limitación demasiado fuerte, después de todo quedan excluidas fuerzas tales como el roce, que fue un protagonista importante de los cursos básicos de física. Sin embargo, por un lado y como veremos en la sección siguiente, existe aún una categoría mayor de fuerzas que pueden ser incluidas dentro de la formulación lagrangiana. Por otro, probaremos más adelante bajo consideraciones muy generales que a nivel microscópico sólo pueden existir fuerzas de este tipo.

2.2.3 Fuerzas dependientes de las velocidades

En nuestra deducción de las ecuaciones de Lagrange, usamos el hecho de que el potencial V no depende de las velocidades, por lo que no contribuye al término que contiene la derivada respecto $\dot{\vec{r}}_n$. Sin embargo, una inspección más atenta de ese cálculo nos permite la siguiente generalización. Supongamos que existe algún tipo de fuerza que se puede obtener a partir de una función $V(\dot{\vec{r}}_n, \vec{r}_n, t)$ según la fórmula

$$\vec{F}_n = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{\vec{r}}_n} \right) - \frac{\partial V}{\partial \vec{r}_n}$$

Podemos llamar a $V(\dot{\vec{r}}_n, \vec{r}_n, t)$ un *potencial dependiente de las velocidades*. En el caso particular en el que V no depende de $\dot{\vec{r}}_n$ recuperamos el potencial de una fuerza monogénica, y si además no depende del tiempo estamos en presencia de una fuerza conservativa. Sin embargo, en el caso general aún podemos escribir

$$G_i = \vec{F}_n \cdot \frac{\partial \vec{r}_n}{\partial q_i} = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{\vec{r}}_n} \right) \cdot \frac{\partial \vec{r}_n}{\partial q_i} - \frac{\partial V}{\partial \vec{r}_n} \cdot \frac{\partial \vec{r}_n}{\partial q_i}$$

Integrando por partes en el primer término, y usando las fórmulas que dedujimos más arriba $\partial \vec{r}_n / \partial q_i = \partial \dot{\vec{r}}_n / \partial \dot{q}_i$ y $\partial \dot{\vec{r}}_n / \partial q_j = d(\partial \vec{r}_n / \partial q_i) / dt$, obtenemos

$$G_i = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{\vec{r}}_n} \cdot \frac{\partial \dot{\vec{r}}_n}{\partial \dot{q}_i} \right) - \frac{\partial V}{\partial \dot{\vec{r}}_n} \cdot \frac{\partial \dot{\vec{r}}_n}{\partial q_i} - \frac{\partial V}{\partial \vec{r}_n} \cdot \frac{\partial \vec{r}_n}{\partial q_i} = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{q}_i} \right) - \frac{\partial V}{\partial q_i}$$

Con lo que podemos reconstruir las ecuaciones de Lagrange

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

donde de nuevo hemos definido nuestro lagrangiano como $L = K - V$.

Cuando hay fuerzas dependientes de las velocidades, la definición que dimos más arriba para el momento generalizado debe ser reemplazada por

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

donde entran también las contribuciones provenientes del potencial.

Si bien la generalización aquí definida puede parecer un tanto exótica a primera vista, en los hechos resulta muy útil ya que permite incluir en la formulación lagrangiana una de las fuerzas más importantes de la naturaleza, como se ve en el siguiente ejemplo.

Ejemplo: fuerza de Lorentz

Supongamos como ejemplo sencillo que tenemos una sola partícula, y que la función V es lineal en las velocidades. De manera completamente general podemos entonces escribir

$$V = e \left(\Phi - \frac{1}{c} \dot{\vec{r}} \cdot \vec{A} \right)$$

en términos de una función escalar $\Phi(\vec{r}, t)$ y una función vectorial $\vec{A}(\vec{r}, t)$, donde introducimos las constantes e y c porque serán convenientes más adelante. Si ahora obtenemos la fuerza usando la fórmula anterior

$$F_a = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{r}_a} \right) - \frac{\partial V}{\partial r_a}$$

donde los índices a, b recorren las direcciones cartesianas $\{x, y, z\}$. Esto nos da

$$F_a = -\frac{e}{c} \frac{dA_a}{dt} - e \left(\frac{\partial \Phi}{\partial r_a} - \frac{1}{c} \dot{r}_b \frac{\partial A_b}{\partial r_a} \right)$$

Usando la regla de la cadena en el primer término nos queda

$$F_a = -\frac{e}{c} \left(\frac{\partial A_a}{\partial r_b} \dot{r}_b + \frac{\partial A_a}{\partial t} \right) - e \left(\frac{\partial \Phi}{\partial r_a} - \frac{1}{c} \dot{r}_b \frac{\partial A_b}{\partial r_a} \right)$$

lo que se puede reacomodar según

$$F_a = e \left(-\frac{\partial \Phi}{\partial r_a} - \frac{1}{c} \frac{\partial A_a}{\partial t} \right) - \frac{e}{c} \left(\frac{\partial A_a}{\partial r_b} - \frac{\partial A_b}{\partial r_a} \right) \dot{r}_b$$

Si aquí identificamos Φ con el potencial escalar y \vec{A} con el potencial vector, mientras c será la velocidad de la luz y e la carga eléctrica de nuestra partícula, la expresión que hemos obtenido no es otra cosa que la fuerza de Lorentz

$$\vec{F} = e \left(\vec{E} + \dot{\vec{r}} \times \vec{B} \right)$$

Es decir que la fuerza de Lorentz se puede obtener a partir de un potencial de velocidades, por lo que es un caso particular de las fuerzas dependientes de la velocidad que hemos estudiado en esta sección.

Ejercicio

Podemos escribir el lagrangiano de una partícula relativista cargada en la forma

$$L = K - V = -m_{\text{rep}}c^2\sqrt{1 - \dot{\vec{r}}^2/c^2} - e\left(\Phi - \frac{1}{c}\dot{\vec{r}} \cdot \vec{A}\right)$$

Se puede comprobar lo poderoso que es el método de Lagrange calculando las ecuaciones de movimiento de una partícula cargada en coordenadas polares en el plano, usando la ley de fuerza vectorial de los cursos básicos de física y comparando con el cálculo a partir de este lagrangiano.

Nota:

Joseph-Louis Lagrange fue un matemático y físico italiano nacido en Torino bajo el nombre de Giuseppe Luigi Lagrangia. Considerado por sus contemporáneos «el matemático más grande de Europa», tenía un carácter extremadamente tímido y retraído, y había estudiado derecho en la universidad por considerar el tópico de las matemáticas como demasiado aburrido. Sin embargo, se interesó en la mecánica al leer un artículo del astrónomo, matemático y geólogo inglés Edmond Halley, quién fuera también inspirador de Newton.

Lagrange reformuló la mecánica utilizando el análisis matemático, entre otra enorme cantidad de contribuciones a las matemáticas y la física. Fue también el responsable de la adopción por la Revolución Francesa del sistema métrico decimal.

2.2.4 Teoremas de conservación

Al final de la Clase 1, reinterpretemos los teoremas de conservación como una consecuencia de las invarianzas del sistema. La formulación lagrangiana nos permite formalizar y generalizar esa intuición, como se describe en esta sección.

Según la definición que dimos más arriba, el momento generalizado asociado a la coordenada generalizada q_i se obtiene de la expresión $p_i = \partial L / \partial \dot{q}_i$. Esto nos permite escribir las ecuaciones de Lagrange en la forma

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i}$$

Se puede ver inmediatamente que cada vez que el lagrangiano no dependa de una coordenada q_l el correspondiente momento generalizado p_l se conserva. Una tal coordenada que no aparece explícitamente en el lagrangiano se denomina *coordenada cíclica*. Es importante puntualizar que la velocidad generalizada asociada \dot{q}_l sí aparece en el lagrangiano.

Ejercicio: conservación del momento lineal

Demuestre el teorema de conservación del momento lineal de la mecánica vectorial que vimos la Clase 1 utilizando los resultados de esta sección.

Para hacerlo, primero escriba el lagrangiano de un sistema de partículas en términos de las coordenadas del centro de masas y las de cada partícula en el sistema centro de masas. Luego pruebe que si la coordenada centro de masas es cíclica, para lo que basta que no aparezca en el potencial, entonces el impulso total del sistema se conserva. Convénzase de que si tal coordenada es cíclica, esto implica que la fuerza externa total se anula.

Ejercicio: conservación del momento angular

Use los resultados de esta sección para probar el teorema de conservación del momento angular de la mecánica vectorial que vimos la Clase 1.

Limítese al caso sencillo de una sola partícula en el plano, escribiendo el lagrangiano en términos de coordenadas polares, y probando que si la coordenada angular es cíclica entonces el momento angular se conserva. Pruebe que si el ángulo es cíclico, entonces el torque externo total vale cero.

El hecho de que el lagrangiano sea independiente de la coordenada cíclica q_l implica en particular que nada cambiará si modificamos su valor sumándole una constante arbitraria $q_l \rightarrow q_l + \varepsilon$. Es decir que el problema tiene una *invarianza* frente a cambios en q_l . Una vez más vemos que la existencia de una invarianza está relacionada con la aparición de una cantidad conservada. Más adelante generalizaremos este resultado al caso de las invarianzas que no se manifiestan directamente en una coordenada cíclica.

Vimos en los dos ejemplos previos que cuando hay una coordenada cíclica el momento generalizado correspondiente se conserva, algo que pone en un marco general las leyes de conservación del momento lineal y del momento angular de la mecánica vectorial. Una pregunta pertinente sería entonces cómo se extiende al contexto del formalismo lagrangiano la ley de conservación de la energía. Para ver esto, escribamos las ecuaciones de lagrange en coordenadas generalizadas, según

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

podemos multiplicar esta ecuación por la velocidad generalizada \dot{q}_i y sumar sobre el índice i para obtener

$$\dot{q}_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \dot{q}_i \frac{\partial L}{\partial q_i} = \frac{d}{dt} \left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - \left(\ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial L}{\partial q_i} \right) = 0$$

donde en la segunda igualdad hemos hecho una integración por partes. Sumando y restando $\partial L / \partial t$ dentro del segundo paréntesis reconocemos la derivada total respecto del tiempo dL/dt .

Nos queda

$$\frac{d}{dt} \left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - \left(\frac{dL}{dt} - \frac{\partial L}{\partial t} \right) = 0$$

Esto nos permite reescribir

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

donde hemos definido la energía como una función de las coordenadas y velocidades generalizadas según la fórmula

$$E = \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \dot{q}_i p_i - L$$

De aquí deducimos que cuando el lagrangiano es independiente del tiempo, la energía se conserva. Nótese que esto permite extender la definición de la energía a los sistemas monogénicos y a los que tienen un potencial dependiente de la velocidad. Nuevamente, un teorema de conservación aparece ligado a una invarianza, en este caso la que corresponde a traslaciones temporales $t \rightarrow t + \varepsilon$.

Ejemplo: sistema conservativo no relativista

Para un conjunto de N partículas no relativistas interactuando por medio de un potencial conservativo, tendremos que

$$L = K - V = \frac{1}{2} K_{ij}(q_i) \dot{q}_i \dot{q}_j - V(q_i)$$

De aquí podemos obtener inmediatamente

$$p_i = K_{ij}(q_i) \dot{q}_j$$

con lo cual la energía toma la forma

$$E = K_{ij}(q_i) \dot{q}_i \dot{q}_j - L = \frac{1}{2} K_{ij}(q_i) \dot{q}_i \dot{q}_j + V(q_i) = K + V$$

Es decir que en este caso particular funciona la idea de “cambiar el signo al término potencial” para obtener la energía a partir del lagrangiano y viceversa.

Sin embargo, se puede probar que esto no es verdad en casos más generales, por ejemplo partículas relativistas o con potenciales dependientes de las velocidades.

Ejemplo: partícula relativista

Para una partícula relativista interactuando con un campo electromagnético, tenemos

$$L = K - V = -m_{\text{rep}} c^2 \sqrt{1 - \dot{\vec{r}}^2 / c^2} - e \left(\Phi - \frac{1}{c} \dot{\vec{r}} \cdot \vec{A} \right)$$

de donde se lee el momento lineal en la forma

$$\vec{p} = \frac{m_{\text{rep}} \dot{\vec{r}}}{\sqrt{1 - \dot{\vec{r}}^2 / c^2}} + \frac{e}{c} \vec{A}$$

Lo que implica

$$E = \frac{m_{\text{rep}} c^2}{\sqrt{1 - \dot{\vec{r}}^2 / c^2}} + e \Phi \neq K + V$$

Nótese que en este caso, para pasar del lagrangiano a la energía no funciona la fórmula sencilla de “cambiar el signo al término potencial”.

2.3 Resumen

En esta clase encontramos una manera de reescribir la segunda ley de Newton en coordenadas generalizadas, que no requiere calcular derivadas segundas, y que se aplica a sistemas de partículas relativistas y no relativistas, con fuerzas de tipo lo bastante general como para incluir las gravitacionales y las electromagnéticas.

Por lo tanto en adelante, al menos para este tipo bastante general de sistemas, podemos reemplazar la segunda ley de Newton

$$\vec{F}_n = \frac{d\vec{p}_n}{dt} \quad \text{donde} \quad \vec{p}_n = m_n \vec{\dot{r}}_n$$

para la cual necesitamos saber la masa y las fuerzas que actúan sobre cada una de las N partículas como función de las posiciones y velocidades de todas las demás, por las ecuaciones de Lagrange

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

donde ahora es necesario conocer el lagrangiano como función de cada una de las $3N$ coordenadas y las $3N$ velocidades generalizadas.

Si bien aún necesitamos resolver $3N$ ecuaciones de segundo orden en el tiempo, fuimos capaces de reparametrizar nuestro espacio de configuración $\mathcal{C} = \mathbb{R}^{3N}$ en términos de $3N$ coordenadas curvilíneas arbitrarias $\{q_i\}$ con $i \in \{1, \dots, 3N\}$. Como veremos en los capítulos que siguen, una elección adecuada de tales coordenadas nos permite identificar las coordenadas cíclicas y obtener los correspondientes teoremas de conservación. Las ecuaciones resultantes son de segundo orden en el tiempo, por lo que necesitamos especificar $6N$ condiciones iniciales para las $\{q_i, \dot{q}_i\}$ con $i \in \{1, \dots, 3N\}$, que parametrizan nuestro espacio de estados $\mathcal{E} = \mathbb{R}^{6N}$.

Aplicaciones

Clase 3 El problema de dos cuerpos I: Definición 41

- 3.1 Objetivos
- 3.2 El problema de dos cuerpos
- 3.3 Resumen

Clase 4 El problema de dos cuerpos II: Órbitas 53

- 4.1 Objetivos
- 4.2 Órbitas del problema de dos cuerpos
- 4.3 Resumen

Clase 5 El problema de dos cuerpos III: Kepler 61

- 5.1 Objetivos
- 5.2 El problema de Kepler
- 5.3 Resumen

Clase 6 El problema de dos cuerpos VI: Dispersión 73

- 6.1 Objetivos
- 6.2 El problema de la dispersión
- 6.3 Resumen

Clase 7 Cuerpo rígido I: Rotaciones 87

- 7.1 Objetivos
- 7.2 Rotaciones espaciales
- 7.3 Resumen

Clase 8 Cuerpo rígido II: Cinemática e inercia 105

- 8.1 Objetivos
- 8.2 Cinemática del cuerpo rígido
- 8.3 Momento de inercia
- 8.4 Resumen

Clase 9 Cuerpo rígido III: Dinámica 119

- 9.1 Objetivos
- 9.2 Dinámica del cuerpo rígido
- 9.3 Resumen

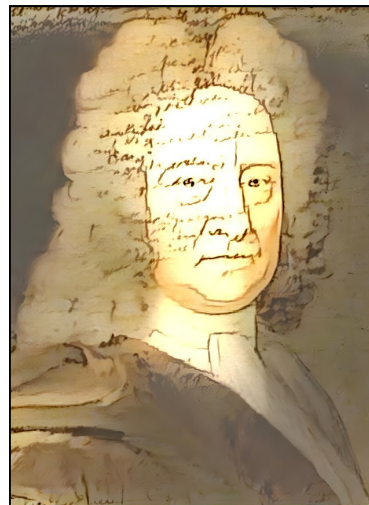
3. El problema de dos cuerpos I: Definición

3.1 Objetivos

En esta clase comenzaremos a aplicar la formulación lagrangiana que obtuvimos en la Clase 2 a ejemplos concretos de interés físico.

En particular, nos dedicaremos al problema de dos cuerpos que interactúan a través de una fuerza central. Por el momento, nos concentraremos en obtener resultados sin especificar la forma explícita de la fuerza central como función de la distancia entre los dos cuerpos.

Estaremos interesados en obtener la evolución temporal de la posición de cada uno de los cuerpos. Veremos que el problema se reduce al de una partícula ficticia que se mueve en una dimensión bajo la influencia de un potencial efectivo.



Edmond Halley

3.2 El problema de dos cuerpos

Supongamos que tenemos dos partículas no relativistas de masa m_1 y m_2 , cuyas posiciones se describen en un sistema inercial a través de los vectores posición \vec{r}_1 y \vec{r}_2 . El espacio de configuración de este sistema es $\mathcal{C} = \mathbb{R}^6$.

Para obtener una descripción lagrangiana de este problema, que nos permitirá luego utilizar coordenadas arbitrarias, podemos comenzar escribiendo su energía cinética como

$$K = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2$$

Supongamos además que las dos partículas interactúan a través de un potencial que no depende de las velocidades ni del tiempo

$$V = V(\vec{r}_1, \vec{r}_2)$$

El lagrangiano correspondiente puede entonces escribirse en la forma

$$L = K - V = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 - V(\vec{r}_1, \vec{r}_2)$$

Al tomar la derivadas del potencial, obtendremos las fuerzas entre ambas partículas, según

$$\vec{F}_1 = \vec{F}_{1 \rightarrow 2} + \vec{F}_1^{\text{ext}} = -\frac{\partial V}{\partial \vec{r}_1}$$

$$\vec{F}_2 = \vec{F}_{2 \rightarrow 1} + \vec{F}_2^{\text{ext}} = -\frac{\partial V}{\partial \vec{r}_2}$$

Si las partículas están alejadas lo suficiente de cualquier otro cuerpo en el universo, entonces no hay fuerzas externas. Por lo tanto tendremos que

$$\vec{F}_{1 \rightarrow 2} = -\frac{\partial V}{\partial \vec{r}_1}$$

$$\vec{F}_{2 \rightarrow 1} = -\frac{\partial V}{\partial \vec{r}_2}$$

Estas fuerzas deben satisfacer la ley de acción y reacción $\vec{F}_{1 \rightarrow 2} = -\vec{F}_{2 \rightarrow 1}$, lo que en particular implica para el potencial

$$\frac{\partial V}{\partial \vec{r}_1} = -\frac{\partial V}{\partial \vec{r}_2}$$

Definiendo dos nuevas variables, la *coordenada relativa* \vec{r} y la posición del centro de masas \vec{r}_{cm} según las fórmulas

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

$$\vec{r}_{\text{cm}} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}$$

podemos reemplazar este cambio de variables en la expresión de más arriba para las derivadas del potencial y usar la regla de la cadena. Con esto tenemos

$$\frac{\partial V}{\partial \vec{r}} + \frac{m_1}{m_1 + m_2} \frac{\partial V}{\partial \vec{r}_{\text{cm}}} = \frac{\partial V}{\partial \vec{r}} - \frac{m_2}{m_1 + m_2} \frac{\partial V}{\partial \vec{r}_{\text{cm}}}$$

Procedemos a cancelar las derivadas respecto de \vec{r} y a reordenar los términos restantes, para obtener

$$\frac{\partial V}{\partial \vec{r}_{\text{cm}}} = 0$$

Esto implica que el potencial es función solamente de la coordenada relativa $\vec{r} = \vec{r}_1 - \vec{r}_2$.

Además, estamos interesados en el problema en el que la fuerza es central, es decir que cumple la condición

$$(\vec{r}_2 - \vec{r}_1) \times \vec{F}_{2 \rightarrow 1} = 0$$

lo que se puede expresar en términos de la coordenada relativa \vec{r}

$$\vec{r} \times \frac{\partial V}{\partial \vec{r}} = 0$$

Esta ecuación implica que $\partial V / \partial \vec{r}$ debe ser proporcional al versor \hat{r} , lo que solo se puede cumplir si escribimos $\partial V / \partial \vec{r} = V'(r)\hat{r}$ donde $V'(r)$ es una función arbitraria de $r = |\vec{r}|$. Esto se puede integrar, para obtener finalmente

$$V = V(|\vec{r}_2 - \vec{r}_1|)$$

Con lo cual el lagrangiano del problema vendrá dado por la expresión

$$L = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 - V(|\vec{r}_1 - \vec{r}_2|)$$

Ejercicio:

Calcule las ecuaciones de Lagrange para este potencial ¿cuántas ecuaciones diferenciales hay? ¿cuántas variables? ¿están acopladas?

Si invertimos la definición de \vec{r} y \vec{r}_{cm} para obtener \vec{r}_1 y \vec{r}_2

$$\vec{r}_1 = \vec{r}_{\text{cm}} + \frac{m_2\vec{r}}{m_1 + m_2}$$

$$\vec{r}_2 = \vec{r}_{\text{cm}} - \frac{m_1\vec{r}}{m_1 + m_2}$$

Las velocidades se obtienen inmediatamente y se pueden reemplazar en la energía cinética, para obtener el lagrangiano en las nuevas coordenadas

$$L = \frac{1}{2}m_1 \left(\dot{\vec{r}}_{\text{cm}} + \frac{m_2\dot{\vec{r}}}{m_1 + m_2} \right)^2 + \frac{1}{2}m_2 \left(\dot{\vec{r}}_{\text{cm}} - \frac{m_1\dot{\vec{r}}}{m_1 + m_2} \right)^2 - V(|\vec{r}|)$$

Expandiendo los cuadrados vemos que los términos proporcionales a $\dot{\vec{r}}_{\text{cm}} \cdot \dot{\vec{r}}$ se cancelan, obteniendo

$$L = \frac{1}{2}m_{\text{tot}}\dot{\vec{r}}_{\text{cm}}^2 + \frac{1}{2}\mu\dot{\vec{r}}^2 - V(r)$$

Donde $m_{\text{tot}} = m_1 + m_2$ es la masa total, y la combinación $\mu = m_1m_2/(m_1 + m_2)$ se conoce como la *masa reducida* del sistema.

Ejercicio:

Escriba las ecuaciones de Lagrange para las coordenadas relativa \vec{r} y centro de masas

\vec{r}_{cm} ¿cuántas son? ¿están acopladas?

Con este cambio de variables, hemos reparametrizado nuestro espacio de configuración $\mathcal{C} = \mathbb{R}^6$ en términos de las coordenadas $\{\vec{r}_{\text{cm}}, \vec{r}\}$.

3.2.1 Problema equivalente de una partícula

La primera observación importante es que las tres componentes del vector posición del centro de masas \vec{r}_{cm} son coordenadas cíclicas, es decir que no aparecen en el lagrangiano. Esto implica que su correspondiente momento generalizado

$$\vec{p}_{\text{cm}} = \frac{\partial L}{\partial \dot{\vec{r}}_{\text{cm}}} = m_{\text{tot}} \dot{\vec{r}}_{\text{cm}}$$

se conserva. Esto es algo que ya sabíamos: *en ausencia de fuerzas externas el momento lineal del centro de masas se conserva*, pero ahora escrito en un nuevo lenguaje. Las ecuaciones de movimiento para \vec{r}_{cm} son entonces

$$m_{\text{tot}} \ddot{\vec{r}}_{\text{cm}} = \vec{p}_{\text{cm}}$$

y se pueden resolver fácilmente como

$$\vec{r}_{\text{cm}} = \frac{\vec{p}_{\text{cm}}}{m_{\text{tot}}}(t - t_0) + \vec{r}_{\text{cm}}^0$$

donde \vec{r}_{cm}^0 es el valor de \vec{r}_{cm} al tiempo $t = t_0$.

En cuanto a las ecuaciones de Lagrange para \vec{r} , estas toman la forma

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\vec{r}}} \right) - \frac{\partial L}{\partial \vec{r}} = \mu \ddot{\vec{r}} + V'(r) \vec{r} = 0$$

donde \vec{r} es el versor radial. En estas ecuaciones no entra ni la variable cíclica \vec{r}_{cm} ni su derivada temporal $\dot{\vec{r}}_{\text{cm}}$, por lo que es evidente que se podrían haber obtenido del lagrangiano *truncado*

$$L_{\text{1part}} = \frac{1}{2} \mu \dot{\vec{r}}^2 - V(r)$$

En otras palabras, hemos transformado el problema de dos cuerpos en el problema equivalente de una sola partícula ficticia de masa μ situada a una distancia \vec{r} del origen, con un potencial que depende sólo de $r = |\vec{r}|$.

Un punto importante de esta *truncación* es que el espacio de configuración resultante $\mathcal{C} = \mathbb{R}^3$ tiene dimensión menor que el original, y lo mismo sucede para el espacio de estados $\mathcal{E} = \mathbb{R}^6$. Otra cuestión a mencionar es que es *consistente*, en el sentido de que las ecuaciones de movimiento para \vec{r} que se obtienen a partir del lagrangiano truncado coinciden con las que se obtendrían usando lagrangiano original.

La dependencia sencilla del potencial nos hace sospechar que el problema será particularmente simple si lo escribimos en coordenadas esféricas. Para esto, definimos

$$x = r \cos \theta \sin \phi \quad y = r \sin \theta \sin \phi \quad z = r \cos \phi$$

y derivamos respecto del tiempo usando la regla de la cadena, para obtener las velocidades en la forma

$$\dot{x} = \dot{r} \cos \theta \sin \phi - r \dot{\theta} \sin \theta \sin \phi + r \dot{\phi} \cos \theta \cos \phi$$

$$\dot{y} = \dot{r} \sin \theta \sin \phi + r \dot{\theta} \cos \theta \sin \phi + r \dot{\phi} \sin \theta \cos \phi$$

$$\dot{z} = \dot{r} \cos \phi - r \dot{\phi} \sin \phi$$

Si reemplazamos estas expresiones en el lagrangiano, obtenemos después de un poco de trabajo

$$L_{1\text{part}} = \frac{1}{2} \mu (\dot{r}^2 + r^2 (\dot{\phi}^2 + \dot{\theta}^2 \sin^2 \phi)) - V(r)$$

donde, al igual que en el ejemplo de las coordenadas polares en el plano que vimos la Clase 2, los productos mixtos se cancelaron. A partir de este lagrangiano, podemos obtener fácilmente las ecuaciones de Lagrange, aplicando la formula general de la clase previa

$$\frac{d}{dt} \left(\frac{\partial L_{1\text{part}}}{\partial \dot{r}} \right) - \frac{\partial L_{1\text{part}}}{\partial r} = \mu \ddot{r} - \mu r (\dot{\phi}^2 + \dot{\theta}^2 \sin^2 \phi) + V'(r) = 0$$

$$\frac{d}{dt} \left(\frac{\partial L_{1\text{part}}}{\partial \dot{\theta}} \right) - \frac{\partial L_{1\text{part}}}{\partial \theta} = \frac{d}{dt} (\mu r^2 \dot{\theta} \sin^2 \phi) = 0$$

$$\frac{d}{dt} \left(\frac{\partial L_{1\text{part}}}{\partial \dot{\phi}} \right) - \frac{\partial L_{1\text{part}}}{\partial \phi} = \frac{d}{dt} (\mu r^2 \dot{\phi}) - \mu r^2 \dot{\theta}^2 \sin \phi \cos \phi = 0$$

Ejercicio:

Escriba el lagrangiano en coordenadas esféricas para el *problema de Hooke*, es decir dos partículas unidas por un resorte. Obtenga las ecuaciones de movimiento.

Este problema provee una descripción clásica para el movimiento de una molécula diatómica. Sin embargo, su importancia fenomenológica es limitada, ya que a la escala molecular los efectos cuánticos comienzan a ser relevantes.

Ejercicio:

Repita lo anterior para el *problema de Kepler*, es decir dos partículas interactuando a través del potencial newtoniano.

Este problema describe el movimiento de los planetas en el sistema solar, y es el que motivó a Newton a desarrollar sus tres leyes de la mecánica y la ley de la gravitación universal.

Fijando primero la atención en la última ecuación, vemos que tiene una solución trivial dada por $\phi = \pi/2$. En esta solución, la partícula se mueve en el plano (x, y) . Las dos ecuaciones restantes pueden escribirse como

$$\mu \ddot{r} - \mu r \dot{\theta}^2 + V'(r) = 0$$

$$\frac{d}{dt} (\mu r^2 \dot{\theta}) = 0$$

Estas ecuaciones pueden obtenerse a partir del lagrangiano

$$L_{2d} = \frac{1}{2} \mu (\dot{r}^2 + r^2 \dot{\theta}^2) - V(r)$$

Se trata de nuevo de un lagrangiano truncado, ya que hemos omitido la variable ϕ , reduciendo una dimensión adicional en el espacio de configuración $\mathcal{C} = \mathbb{R}^2$.

Ejemplo: el movimiento tiene lugar en un plano

La resolución anterior puede hacerse por un camino más corto. Notemos que, dado que la fuerza es central, es decir que apunta hacia el origen de coordenadas en la variable \vec{r} , entonces no habrá torque actuando sobre la partícula ficticia. Esto resulta en que el momento angular asociado $\vec{\ell} = m\vec{r} \times \dot{\vec{r}}$ es constante.

Lo anterior implica que podemos elegir un sistema de coordenadas tal que el eje z apunte en la dirección del momento angular. Ahora bien, por definición el vector \vec{r} es siempre perpendicular a $\vec{\ell}$, por lo que se mantendrá en el plano (x, y) . Con esto, podemos simplificar el cambio de variables, usando coordenadas polares en este plano en lugar de coordenadas esféricas en el espacio. Es decir, escribimos

$$x = r \cos \theta \quad y = r \sin \theta \quad z = 0$$

Este problema corresponde al de una partícula que se mueve en un plano que dimos como ejemplo la Clase 2, lo que simplifica la obtención del lagrangiano

$$L_{2d} = \frac{1}{2} \mu (\dot{r}^2 + r^2 \dot{\theta}^2) - V(r)$$

que coincide con el resultado de más arriba para el lagrangiano truncado.

Lo primero que notamos en estas ecuaciones es que, dado que la coordenada θ es cíclica, el momento generalizado asociado $p_\theta = \partial L_{2d} / \partial \dot{\theta}$ se conserva. Llamamos ℓ a esta constante, y tenemos

$$\ell = \mu r^2 \dot{\theta}$$

Con esto hemos encontrado una *primera integral* del problema. Esto quiere decir que hemos logrado reducir en uno el orden en derivadas de las ecuaciones involucradas. Despejando entonces $\dot{\theta}$ de esta expresión, según

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

obtenemos la ecuación que debemos resolver para obtener la evolución del ángulo θ como función del tiempo. Podemos insertar estos resultados en la ecuación restante para r , lo que resulta en

$$\mu \dot{r}^2 - \frac{\ell^2}{\mu r^3} + V'(r) = 0$$

que es la ecuación que deberíamos resolver para obtener la evolución del radio r como función del tiempo.

Ejercicio:

La tercera ecuación de Lagrange para $L_{1\text{part}}$ tiene otra solución trivial, dada por $\dot{\phi} = 0$. Obtenga la ecuación para r en este caso ¿Qué sucede con la ecuación para θ ? ¿Qué representa este movimiento?

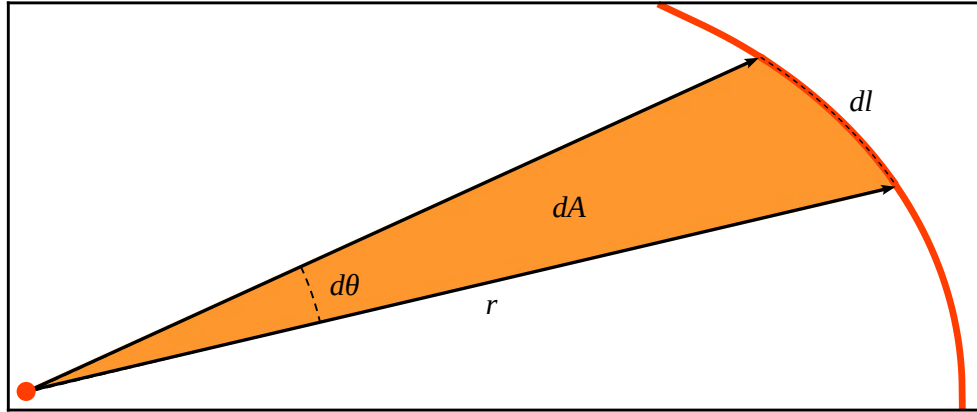


Figura 3.2.1: Ilustración de la segunda ley de Kepler. El área dA formada por el radio r y el diferencial de arco dl se recorre a velocidad uniforme.

3.2.2 Solución angular

Volviendo a la ecuación que obtuvimos más arriba para el movimiento angular

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

vemos que se puede integrar para obtener una solución completa para el ángulo θ como función del tiempo t , según

$$\theta = \theta_0 + \frac{\ell}{\mu} \int_{t_0}^t \frac{dt}{r^2}$$

Por supuesto que para poder hacer explícitamente esta integral, necesitamos conocer la forma de r como función de t . Como veremos en la sección siguiente, dependiendo de la forma del potencial $V(r)$ esta expresión puede o no obtenerse analíticamente.

Podemos reescribir la evolución angular como una relación entre diferenciales

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

por lo tanto el diferencial de arco $dl = r d\theta$ descrito por la partícula ficticia en su movimiento durante un tiempo infinitesimal dt vendrá dado por la expresión

$$dl = \frac{\ell}{\mu r} dt$$

Si desde cada uno de los extremos de este diferencial de arco trazamos una línea hasta el origen, obtenemos un triángulo infinitesimal, que fue barrido por el radio vector \vec{r} durante su evolución durante el tiempo dt , como se puede ver en la figura 3.2.1. El área de este triángulo será $dA = r dl/2$, es decir que

$$dA = \frac{\ell}{2\mu} dt$$

Podemos integrar esto entre dos valores de t para obtener

$$A - A_0 = \frac{\ell}{2\mu} (t - t_0)$$

Con lo que hemos demostrado la *segunda ley de Kepler*: el radio vector \vec{r} barre áreas iguales en tiempos iguales.

3.2.3 Solución radial

La ecuación que tenemos que resolver para el radio es la que habíamos obtenido más arriba, la cual se puede reordenar como

$$\mu \ddot{r} = -V'(r) + \frac{\ell^2}{\mu r^3} = -\frac{d}{dr} \underbrace{\left(V(r) + \frac{\ell^2}{2\mu r^2} \right)}_{V_{1d}}$$

o en otras palabras

$$\mu \ddot{r} = -V'_{1d}(r)$$

donde vemos que hemos transformado el problema en un *problema unidimensional equivalente*, que contiene una partícula que se mueve en una dimensión bajo la influencia de un potencial efectivo de la forma

$$V_{1d} = V + \frac{\ell^2}{2\mu r^2}$$

En esta expresión, el primer término es el potencial del problema, mientras que el segundo se conoce como *barrera centrífuga*, ya que crece sin límite cuando nos acercamos al origen. Se puede ver fácilmente que la ecuación de movimiento de arriba puede obtenerse a partir del lagrangiano efectivo

$$L_{1d} = \frac{1}{2}\mu \dot{r}^2 - V_{1d}$$

Es evidente que hemos reducido aún más el espacio de configuración, realizando una nueva truncación que nos deja con un problema unidimensional cuyo espacio de configuración es $\mathcal{C} = \mathbb{R}$.

Este problema unidimensional equivalente tiene una energía conservada, lo que se puede probar multiplicando la ecuación de movimiento por \dot{r} e integrando. Esta energía toma la forma

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

Naturalmente tenemos que $E \geq V_{1d}$, por lo que el movimiento tendrá lugar en la región de valores de r donde se cumple esta desigualdad. A medida que el radio se acerca a algún borde de dicha región, el valor del potencial efectivo V_{1d} es cada vez más parecido al de la energía E , por lo que el término $\mu \dot{r}^2/2$ es cada vez más pequeño. Cuando el radio alcanza el borde tenemos que $E = V_{1d}$, por lo que $\mu \dot{r}^2/2 = 0$ y el movimiento radial se detiene. Dado que la fuerza radial efectiva $-V'_{1d}$ es no nula, esta situación dura solo un instante y luego el radio comienza a cambiar nuevamente, pero ahora se mueve en la dirección opuesta alejándose del borde. Por esta razón, cualquier punto donde se cumpla $V_{1d} = E$ se denomina *punto de retorno* del potencial, el movimiento radial tendrá lugar rebotando entre los puntos de retorno.

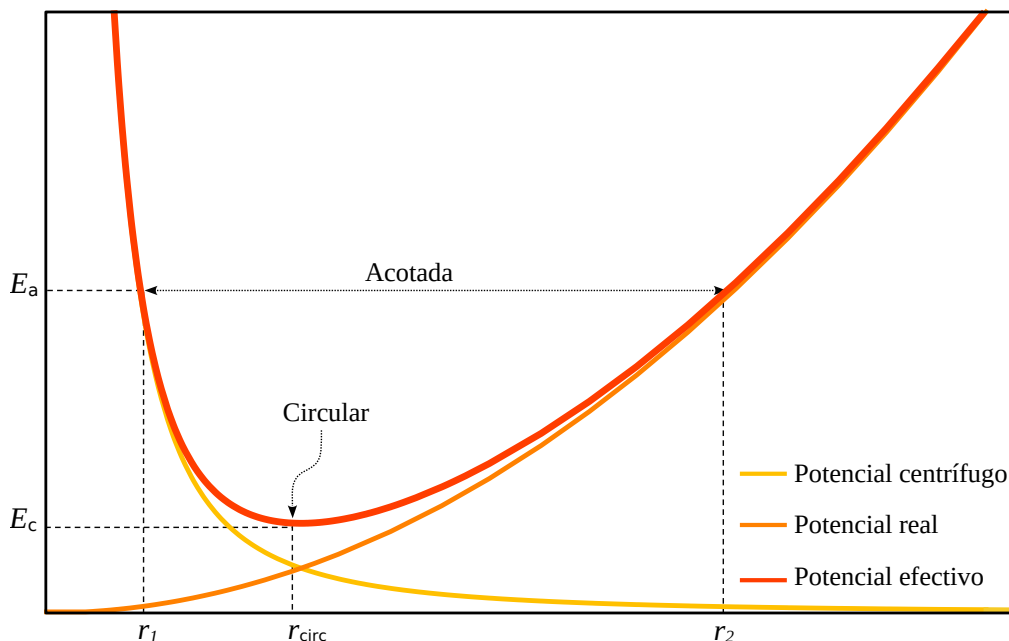


Figura 3.2.2: Potencial efectivo para el problema de Hooke. Cuando la energía del sistema es E_c el sistema recorre órbitas circulares con radio r_{circ} . Cuando la energía es $E_a > E_c$ las órbitas son acotadas. No hay órbitas no acotadas en este sistema. En ningún caso la partícula pasa por el origen.

Ejercicio:

Usando algún software de dibujo como por ejemplo Mathematica o GeoGebra, obtenga gráficas del potencial efectivo para el caso del *potencial de Lennard-Jones* que está dado por

$$V(r) = k \left(\left(\frac{r_{\text{nulo}}}{r} \right)^{12} - \left(\frac{r_{\text{nulo}}}{r} \right)^6 \right)$$

Este potencial describe la interacción entre moléculas neutras como una competencia entre un término repulsivo de corto alcance y uno atractivo de alcance mayor.

De aquí se pueden sacar varias conclusiones cualitativas muy útiles sobre el movimiento, solamente analizando la forma del potencial efectivo

- Si el potencial efectivo crece hacia radios grandes, volviéndose mayor que la energía, entonces habrá un punto de retorno y el movimiento será *acotado*. La partícula ficticia se mantendrá cerca del origen. Si en cambio el potencial está acotado a radios grandes, y si la energía es mayor que la cota del potencial, entonces no habrá un punto de retorno y la órbita será *no acotada*. La partícula ficticia podrá alejarse arbitrariamente del origen.

Las mismas consideraciones se aplican cuando nos movemos hacia radios pequeños: dependiendo de la forma del potencial efectivo y de la energía total, el sistema unidimensional puede tener o no puntos de retorno. Si no hay puntos de retorno, el movimiento llega a $r = 0$. Si en cambio sí los hay, el movimiento radial se detiene en el punto de retorno y cambia de dirección.

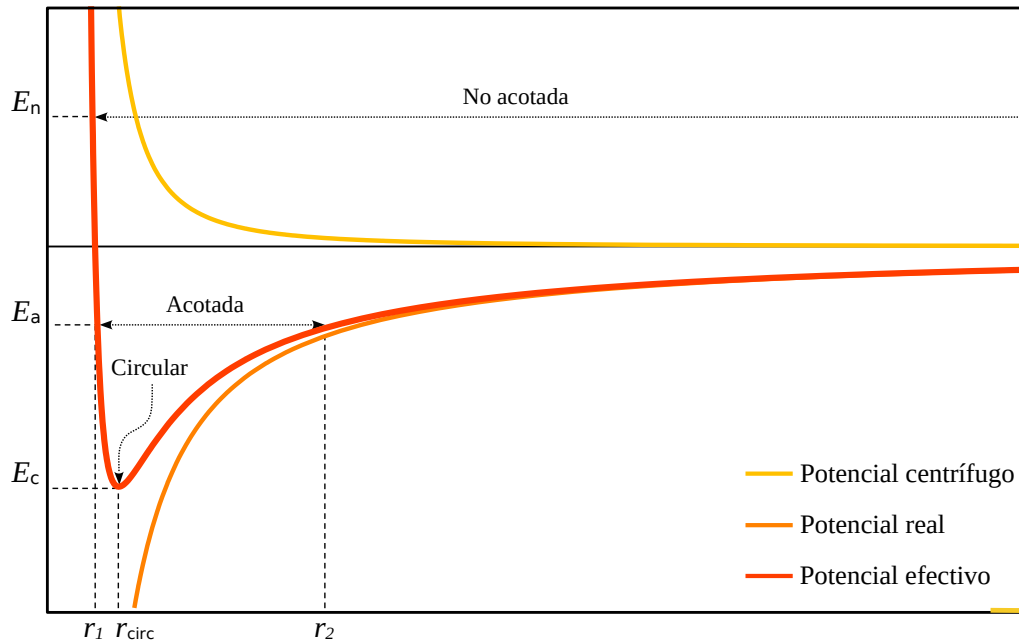


Figura 3.2.3: Potencial efectivo para el problema de Kepler. Cuando la energía del sistema es E_c el sistema recorre órbitas circulares con radio r_{circ} . Cuando la energía es $0 > E_a > E_c$ las órbitas son acotadas. Para energías $E_n > 0$ las órbitas son no acotadas. En ningún caso la partícula pasa por el origen.

- Si el potencial efectivo tiene un mínimo en $r = r_{\text{circ}}$, cualquier movimiento debe tener una energía mayor o igual que el valor mínimo del potencial $V(r_{\text{circ}})$. En particular, si la energía es igual al valor mínimo del potencial $E = V(r_{\text{circ}})$, entonces el sistema unidimensional solo puede quedarse quieto en el punto mínimo $r = r_{\text{circ}}$, ya que no le alcanzaría la energía para ir hacia radios mayores o menores.

En términos del problema inicial, esto no implica no haya movimiento, ya que $\dot{\theta}$ es diferente de cero. Lo que está sucediendo es que el radio se mantiene constante en r_{circ} , por lo que la partícula ficticia realiza una órbita circular alrededor del origen de coordenadas.

Ejercicio:

Obtenga gráficas del potencial efectivo para el caso del problema de Kepler y del problema de Hooke, usando algún software de dibujo como por ejemplo Mathematica o GeoGebra, y analice las correspondientes órbitas.

La energía del problema unidimensional equivalente que hemos encontrado más arriba se puede escribir en términos del potencial real en la forma

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

de donde se puede despejar la velocidad radial según

$$\dot{r} = \sqrt{\frac{2}{\mu} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}$$

Usando esta expresión y escribiendo $\dot{r} = dr/dt$ podemos despejar

$$dt = \frac{dr}{\sqrt{\frac{2}{\mu} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}}$$

Esto permite *reducir el problema a cuadraturas* de acuerdo a

$$t = t_0 + \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{\mu} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}}$$

Junto con la ecuación que escribimos más arriba para θ , esto constituye una solución completa del problema de dos cuerpos con fuerza central.

La primera observación es que hemos obtenido una expresión integral la cual, dependiendo de la forma el potencial, puede no ser posible de integrar explícitamente en términos de funciones elementales. Sin embargo, esta se considera una solución del problema, ya que utilizando métodos numéricos se puede calcular la posición como función del tiempo con precisión arbitraria. Como segundo punto, incluso si pudiéramos resolver explícitamente la integral, obtendríamos t como función de r en lugar de lo opuesto. Esto nos lleva a la necesidad de utilizar una función inversa para obtener r como función de t .

Ejercicio:

Use un software de manipulación algebraica como Mathematica o Maple, o una tabla de integrales para investigar las soluciones radiales del potencial tipo *ley de potencias* $V = -kr^{-n+1}$. ¿Para cuáles valores de n hay una solución explícita en términos de funciones elementales o especiales?

3.3 Resumen

Analizando el problema de dos cuerpos con fuerza central, vimos que se puede transformar en un problema de una sola partícula ficticia que se mueve a en un potencial central. Aplicando entonces la formulación lagrangiana en coordenadas esféricas, fuimos capaces de probar que el movimiento resultante tendrá lugar en un plano.

Usando la conservación del momento angular, pudimos resolver el movimiento angular, y reescribir el movimiento radial como el de un problema unidimensional equivalente con un potencial efectivo. Pudimos reducir este último problema a cuadraturas, escribiendo la dependencia temporal del radio en términos de una sola integral.

Además, demostramos que se cumple la segunda ley de Kepler, que dice que el radio vector que une un planeta al sol barre áreas iguales en tiempos iguales. Lo interesante es que esta propiedad se cumple incluso cuando la interacción central no responde a la ley de cuadrado inverso que caracteriza a la interacción gravitatoria.

4. El problema de dos cuerpos II: Órbitas

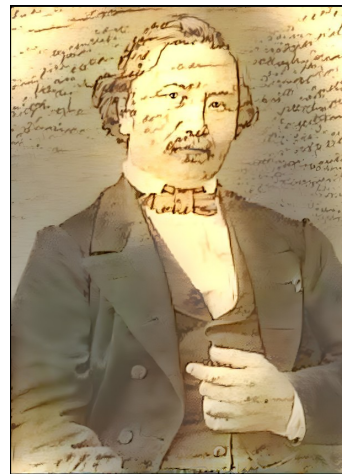
4.1 Objetivos

En esta clase nos concentraremos en las órbitas del problema de dos cuerpos con fuerza central.

En nuestro análisis, sólo estaremos interesados en describir cómo cambia el radio como función del ángulo a medida que se produce el movimiento, sin ocuparnos de la dependencia temporal de ninguna de ambas variables.

Construiremos las ecuaciones diferenciales que satisfacen las órbitas para un potencial general, y estudiaremos las propiedades universales de sus soluciones.

Además, aplicando la técnica de *perturbaciones*, analizaremos qué condiciones debe cumplir el potencial para que el sistema recorra órbitas cerradas.



Joseph Louis François Bertrand

4.2 Órbitas del problema de dos cuerpos

Recordemos las primeras integrales que habíamos obtenido en la Clase 3 para las velocidades angular y radial del problema de dos cuerpos

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

$$\dot{r} = \sqrt{\frac{2}{\mu} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}$$

Con una integral adicional, estas ecuaciones resultaban en una reducción a cuadraturas del problema, lo que es equivalente a una solución completa.

Sin embargo, cuando el movimiento en el tiempo no resulta relevante, podemos dejar de lado estas ecuaciones y concentrarnos en escribir una ecuación que describa exclusivamente las *órbitas* del sistema, es decir el comportamiento de r como función de θ .

Para eliminar la variable t de nuestras ecuaciones de movimiento, dividimos la velocidad radial en la velocidad angular, obteniendo la expresión

$$\frac{dr}{d\theta} = \frac{\dot{r}}{\dot{\theta}} = r^2 \sqrt{\frac{2\mu}{\ell^2} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}$$

Esto permite obtener inmediatamente una solución para la órbita de acuerdo a la integral

$$\theta = \theta_0 + \int_{r_0}^r \frac{dr}{r^2 \sqrt{\frac{2\mu}{\ell^2} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)}}$$

Nuevamente, esto se puede considerar una solución completa de la órbita del problema, si bien dependiendo de la forma del potencial la integral podría no tener una forma explícita en términos de funciones elementales.

Ejercicio:

Use un software de manipulación algebraica como Mathematica o Maple, o una tabla de integrales para investigar las soluciones para las órbitas del potencial tipo *ley de potencias* $V = -kr^{-n+1}$. ¿Para cuáles valores de n hay una solución explícita en términos de funciones elementales o especiales?

Podemos analizar algunas propiedades de la órbita incluso sin obtener una forma explícita para la integral. En efecto, reordenando la ecuación de arriba podemos reescribirla en la forma

$$\frac{1}{r^4} \left(\frac{dr}{d\theta} \right)^2 = \frac{2\mu}{\ell^2} \left(E - \frac{\ell^2}{2\mu r^2} - V(r) \right)$$

Esta ecuación tiene dos simetrías inmediatamente identificables:

- Si $r(\theta)$ es una solución, también lo es $r(\theta + \varepsilon)$ con ε un número real arbitrario. Esto no es sino la invarianza rotacional, que nos dice que podemos rotar el sistema de coordenadas alrededor del origen en el plano del movimiento sin que tenga ninguna consecuencia física.
- Si $r(\theta)$ es una solución, también lo es $r(-\theta)$. Esto sucede porque la derivada aparece al cuadrado, y significa que podemos cambiar el signo del eje z de nuestro sistema de coordenadas sin que haya efectos físicos.

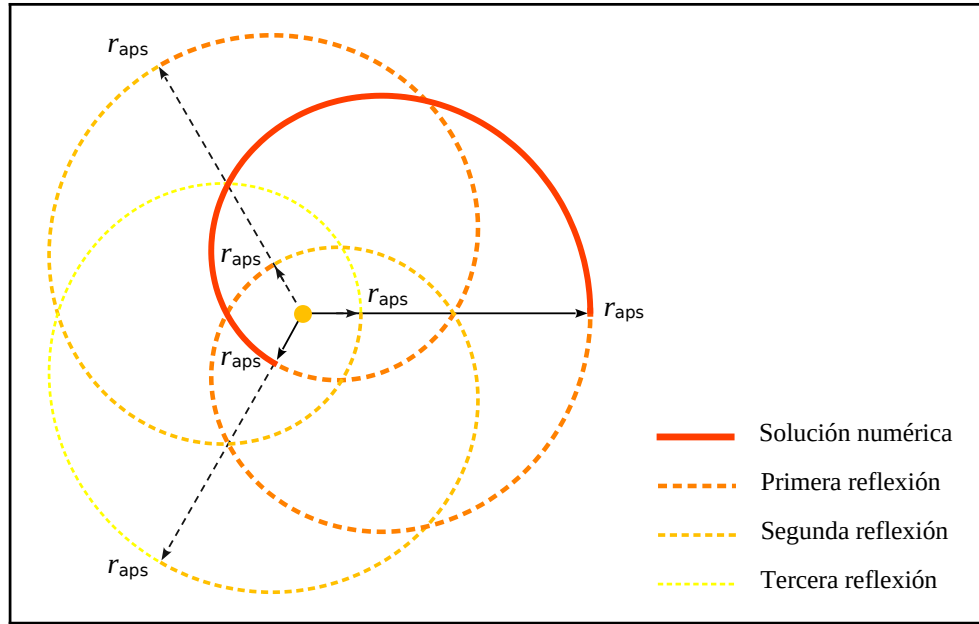


Figura 4.2.1: Las órbitas se reflejan en los ápsides: podemos usar esta propiedad para construir la órbita completa a partir del arco entre dos ápsides. Esto es muy útil al obtener soluciones numéricas, ya que reduce el número de puntos a computar.

Una combinación de ambas transformaciones implica que si $r(\theta)$ es solución, también lo es su forma *reflejada* $r(2\theta_{\text{aps}} - \theta)$ con θ_{aps} una constante cualquiera. Podemos preguntarnos qué valor debería tomar esta constante para que la órbita se refleje sobre sí misma, es decir para que $r(\theta) = r(2\theta_{\text{aps}} - \theta)$. Derivando esta condición y evaluando en $\theta = \theta_{\text{aps}}$ obtenemos $r'(\theta_{\text{aps}}) = -r'(\theta_{\text{aps}})$, lo que implica que $r'(\theta_{\text{aps}}) = 0$. Es decir que θ_{aps} debe ser el *ápside* de la órbita, definido como el punto donde la distancia al origen es máxima o mínima. La conclusión es que *las órbitas se reflejan en sus ápsides*.

La propiedad de reflexión en el ápside resulta muy útil cuando resolvemos las ecuaciones numéricamente, ya que nos permite encontrar la solución para los valores de θ que estén entre dos ápsides cualesquiera, y luego construir el resto aplicando reflexiones.

La ecuación de más arriba puede resultar más simple de resolver si en lugar de escribirla en términos de la variable r la escribimos en términos de una nueva variable $u = 1/r$

$$\left(\frac{du}{d\theta}\right)^2 = \frac{2\mu}{\ell^2} \left(E - \frac{\ell^2}{2\mu} u^2 - V\right)$$

Es posible también escribir una ecuación de segundo orden para la órbita, que a veces simplifica la solución de algunos problemas. Para esto, usamos la ecuación para $\dot{\theta}$ para escribir

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

y lo reemplazamos en la ecuación para r , obteniendo

$$\frac{\ell^2}{\mu r^2} \frac{d}{d\theta} \left(\frac{1}{r^2} \frac{dr}{d\theta} \right) - \frac{\ell^2}{\mu r^3} + V'(r) = 0$$

Escribiendo esto en términos de u y reordenando, nos queda

$$\frac{d^2 u}{d\theta^2} + u + \frac{\mu}{\ell^2} \partial_u V = 0$$

Esta ecuación de segundo orden resulta particularmente útil para estudiar qué sucede cuando deformamos ligeramente una órbita circular, como haremos en la siguiente sección.

Nota:

Como hemos visto, las órbitas y el movimiento en el tiempo para el problema de dos cuerpos con fuerza central se pueden reducir a cuadraturas, lo que es el equivalente a resolverlos exactamente.

El problema de tres cuerpos en cambio es extremadamente difícil de resolver, tanto que se conocen muy pocas soluciones exactas para casos muy particulares, algunas de las cuales fueron obtenidas hace menos de diez años. En el caso general el movimiento es fuertemente *caótico*, lo que significa que cualquier perturbación se amplifica modificando completamente la órbita resultante.

En la novela *El problema de tres cuerpos* el escritor Liu Cixin imagina una civilización que ha evolucionado en el sistema solar de nuestra estrella más cercana Próxima Centauri, que es un sistema triple. Esta especie enfrenta climas extremos cada vez que su planeta se acerca mucho o se aleja demasiado de alguna de las estrellas del sistema, y no conoce una manera de predecir lo que sucederá. Su única certeza es que, en algún momento de su órbita futura, su planeta será finalmente engullido por alguno de los tres soles. De este modo, el único plan posible para la supervivencia es emigrar.

4.2.1 Teorema de Bertrand

Una órbita sencilla que se encuentra como solución para varios potenciales es la órbita circular, que en términos de la variable u está definida como $u(\theta) = u_{\text{circ}}$. Podemos encontrar el valor de u_{circ} usando la ecuación de segundo orden de más arriba, donde al reemplazar la solución se encuentra

$$u_{\text{circ}} = -\frac{\mu}{\ell^2} \partial_u V|_{u_{\text{circ}}}$$

Nos gustaría escribir la ecuación que satisface una pequeña *perturbación* de esta órbita circular, definida según $u(\theta) = u_{\text{circ}} + \varepsilon \delta u(\theta)$, donde ε es un número pequeño y $\delta u(\theta)$ es una función a determinar. Reemplazando en la ecuación de segundo orden para la órbita y desarrollando a primer orden en ε , tenemos

$$\frac{d^2 \delta u}{d\theta^2} + \underbrace{\left(1 + \frac{\mu}{\ell^2} \partial_u^2 V|_{u_{\text{circ}}}\right)}_{\alpha^2} \delta u = 0$$

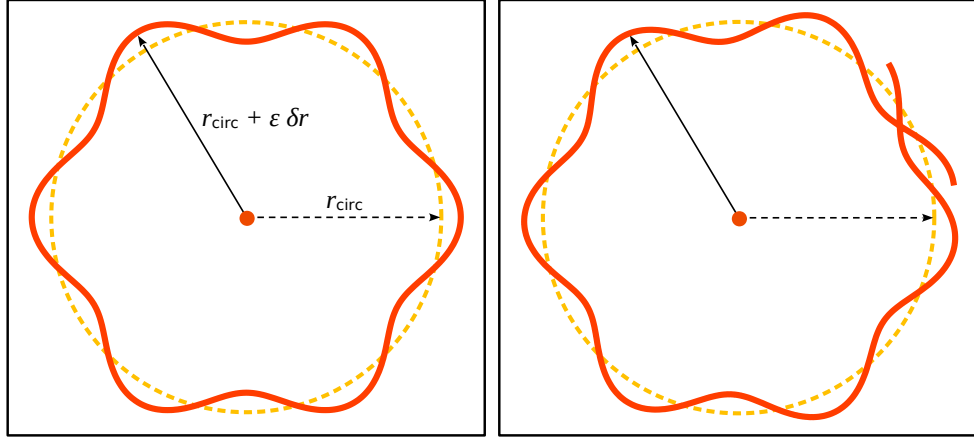


Figura 4.2.2: Teorema de Bertrand: las pequeñas perturbaciones de una órbita circular (izquierda) resultan en una órbita perturbada que se cierra sobre sí misma sólo cuando el radio realiza un número entero de oscilaciones mientras el ángulo da un número entero de vueltas (derecha).

Con la definición de α^2 que hemos hecho en la fórmula, esto puede reescribirse en la forma más compacta

$$\frac{d^2 \delta u}{d\theta^2} + \alpha^2 \delta u = 0$$

que podemos reconocer inmediatamente como la ecuación de un oscilador armónico, cuya solución conocemos

$$\delta u(\theta) = \delta u_0 \cos(\alpha(\theta - \theta_0))$$

siendo δu_0 y θ_0 constantes de integración. Esta perturbación oscila a medida que θ crece recorriendo la órbita. Por lo tanto la órbita perturbada será cerrada si cuando se dieron n vueltas en torno al origen, es decir cuando θ avanzó un ángulo $2\pi n$, la perturbación δu realizó un número m de oscilaciones completas. Para esto necesitamos que $\alpha 2\pi n = 2\pi m$. Esto implica que la órbita será cerrada sólo si α es un número racional $\alpha = m/n$. Si recordamos la expresión para α vemos que esto impone una condición sobre la forma del potencial.

Si ahora queremos resolver el problema a orden cuadrático en la perturbación, escribimos $u = u_{\text{circ}} + \varepsilon \delta u + \varepsilon^2 \delta_{(2)} u$ y volvemos a reemplazar en la ecuación completa, desarrollándola ahora a segundo orden en ε . Nos queda

$$\frac{d^2 \delta_{(2)} u}{d\theta^2} + \alpha^2 \delta_{(2)} u = - \underbrace{\frac{\mu}{2\ell^2} \partial_u^3 V|_{u_{\text{circ}}}}_{2\beta} \delta u^2$$

Insertando explícitamente la solución para δu que habíamos obtenido del desarrollo a primer orden, tenemos

$$\frac{d^2 \delta_{(2)} u}{d\theta^2} + \alpha^2 \delta_{(2)} u = \beta \delta u_0^2 (1 + \cos(2\alpha(\theta - \theta_0)))$$

donde en el segundo término expresamos el cuadrado del coseno según la fórmula $2\cos^2(\dots) = 1 + \cos(2\cdot\dots)$. Esta ecuación corresponde a un oscilador armónico forzado, que según sabemos se

resuelve con

$$\delta_{(2)}u = \frac{\beta}{\alpha^2} \delta u_0^2 \left(1 - \frac{1}{3} \cos(2\alpha(\theta - \theta_0)) \right)$$

Repetiendo el análisis anterior, vemos nuevamente que esta corrección resulta en órbitas cerradas si α es un número racional.

Para ir a tercer orden, escribimos $u = u_{\text{circ}} + \varepsilon \delta u + \varepsilon^2 \delta_{(2)}u + \varepsilon^3 \delta_{(3)}u$ y volvemos a reemplazar en la ecuación, obteniendo esta vez

$$\frac{d^2 \delta_{(3)}u}{d\theta^2} + \alpha^2 \delta_{(3)}u = 4\beta \delta u \delta_{(2)}u - \underbrace{\frac{\mu}{6\ell^2} \partial_u^4 V|_{u_{\text{circ}}}}_{4\gamma} \delta u^3$$

Insertando las formas explícitas de δu y $\delta_{(2)}u$ en el lado derecho, y usando algunas identidades trigonométricas, llegamos a la expresión

$$\frac{d^2 \delta_{(3)}u}{d\theta^2} + \alpha^2 \delta_{(3)}u = \delta u_0^3 \left(\left(\frac{10\beta^2}{3\alpha^2} + 3\gamma \right) \cos(\alpha(\theta - \theta_0)) + \left(\gamma - \frac{2\beta^2}{3\alpha^2} \right) \cos(3\alpha(\theta - \theta_0)) \right)$$

Esto es nuevamente un oscilador forzado, esta vez con dos fuerzas externas. Una de tales fuerzas oscila con la frecuencia natural del oscilador armónico α . Por lo tanto, esta fuerza dará lugar al fenómeno de resonancia, haciendo que $\delta_{(3)}u$ crezca sin límites a medida que crece θ , es decir a medida que la partícula ficticia realiza sus revoluciones en torno al origen.

Hemos encontrado una *inestabilidad*: tenemos una solución para el movimiento del sistema (la órbita circular) cuyas pequeñas perturbaciones crecen arbitrariamente, con lo que dejan de ser pequeñas y destruyen la solución. Para evitar esto necesitamos que la fuerza resonante esté ausente, lo que sucede solamente cuando se cumple la condición

$$10\beta^2 + 9\alpha^2\gamma = 0$$

En este caso, solo queda una fuerza externa que no es resonante, y la solución para el oscilador forzado es

$$\delta_{(3)}u = \delta_{(3)}u_0 \cos(\alpha(\theta - \theta_0)) - \frac{\delta u_0^3}{8\alpha^2} \left(\gamma - \frac{2\beta^2}{3\alpha^2} \right) \cos(3\alpha(\theta - \theta_0))$$

donde $\delta_{(3)}u_0$ es una nueva constante de integración. Esta solución de nuevo da origen a una órbita cerrada cuando α es un número racional.

Este proceso perturbativo se puede continuar a un orden arbitrario en el parámetro infinitesimal ε , lo que nos permite construir una órbita arbitrariamente deformada. Sin embargo, la expansión hasta orden tres es suficiente para nuestros fines, como veremos a continuación.

Ejemplo: ley de potencias

Podemos testear estos cálculos en el caso de un potencial con la forma de una ley de potencias

$$V = -kr^{-n+1}$$

Primero tenemos que encontrar u_{circ} , mediante la fórmula

$$1 = \frac{\mu}{\ell^2} k(n-1) u_{\text{circ}}^{n-3}$$

luego calculamos α , β y γ , obteniendo

$$\alpha^2 = 3 - n$$

$$\beta = \frac{1}{4} (n-2)(n-3) u_{\text{circ}}^{-1}$$

$$\gamma = \frac{1}{24} (n-2)(n-3)(n-4) u_{\text{circ}}^{-2}$$

Reemplazando en la condición de ausencia de resonancia, esto implica que

$$(n-3)^2 (n-2) (n+1) = 0$$

Con lo cual vemos que o bien $n = -1$ y obtenemos la ley de Hooke, o bien $n = 2$ y obtenemos la ley de Newton. No podemos poner $n = 3$ porque en ese caso α valdría cero y la perturbación de primer orden crecería linealmente.

La condición $10\beta^2 + 9\alpha^2\gamma = 0$ se puede escribir completamente en términos del potencial y sus derivadas, como

$$5u(\partial_u^3 V)^2 + 3(\partial_u V - u\partial_u^2 V)\partial_u^4 V = 0$$

donde hemos reemplazado las expresiones explícitas para α^2 , β y γ , y hemos eliminado μ/ℓ^2 en términos de u_{circ} usando la fórmula que determina la órbita circular. Finalmente reemplazamos u_{circ} en todas partes por u , ya que eligiendo el valor de ℓ^2 podemos hacer que la expresión sea válida en cualquier posición. Con esto hemos obtenido una ecuación diferencial para la fuerza generalizada $G_u = -\partial_u V$, que se puede escribir como.

$$5u(\partial_u^2 G_u)^2 - 3(G_u + u\partial_u G_u)\partial_u^3 G_u = 0$$

Si pudiéramos resolver esta ecuación, obtendríamos aquéllos potenciales para los cuales una pequeña perturbación de la órbita circular resulta en una solución estable que da lugar a una órbita cerrada.

Ejercicio:

Utilizando un software de manipulación algebraica como Mathematica o Maple, reescriba la ecuación diferencial en términos de una nueva variable independiente $\zeta = \log u$ y una nueva variable dependiente $U = \log(-G_u)$. Proponga un desarrollo en serie de la solución, con la forma

$$U = \sum_m a_m \zeta^m$$

y reemplácelo en la ecuación para demostrar que todos los coeficientes a_m deben anularse con la excepción de a_0 y a_1 , el cual puede valer -3 , 0 ó 1 . Yendo hacia atrás

con el cambio de variables, demuestre que el caso $a_1 = -3$ corresponde a la ley de Hooke $V = kr^2$, mientras que el caso $a_1 = 0$ corresponde a la ley de Kepler $V = -k/r$. El caso $a_1 = 1$ debe descartarse ya que conduce a $\alpha = 0$ (ver el ejemplo de más arriba).

Con esto probamos que estas dos leyes de potencias constituyen la solución general al problema, es decir que son los únicos potenciales que dan lugar a órbitas circulares que cuando son perturbadas siguen siendo cerradas. Este resultado se conoce como *teorema de Bertrand*.

4.3 Resumen

Aprendimos que la órbita del problema de dos cuerpos con fuerza central se puede escribir como una cuadratura, sin necesidad de resolver el movimiento como función del tiempo.

También fuimos capaces de escribir una ecuación diferencial de segundo orden para la órbita, y la usamos para estudiar el comportamiento de las perturbaciones alrededor de la órbita circular. Este tipo de cálculos se conocen como *teoría clásica de perturbaciones* y resulta muy útil en cuestiones relacionadas con la mecánica celeste.

A partir de este análisis probamos el teorema de Bertrand, que dice que los únicos potenciales cuyas perturbaciones alrededor de la órbita circular resultan en trayectorias cerradas corresponden al potencial de Kepler y al potencial de Hooke.

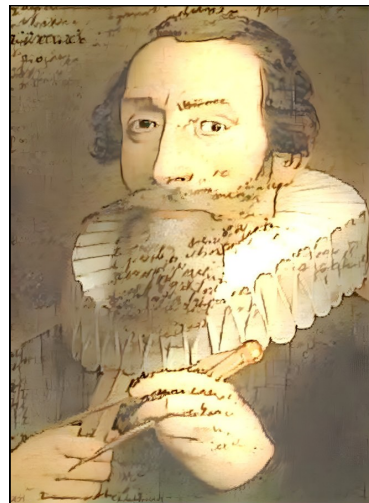
5. El problema de dos cuerpos III: Kepler

5.1 Objetivos

En esta clase nos proponemos estudiar el caso especial del problema de dos cuerpos que interactúan mediante una fuerza central de carácter gravitatorio.

Este problema está en la base de la mecánica celeste, ya que constituye una primera aproximación al movimiento de los planetas en torno al sol, y al de los satélites en torno a cada planeta. Esto significa que entre los resultados de nuestro análisis debemos encontrar las dos leyes de Kepler que aún no hemos demostrado.

Fue la solución de este problema particular lo que llevó a Newton a formular las leyes del movimiento en su célebre *Philosophiæ naturalis principia mathematica*, dando origen a la Mecánica.



Johannes Kepler

5.2 El problema de Kepler

El problema de Kepler corresponde al de dos cuerpos interactuando con una fuerza central, cuando el potencial toma la forma de la interacción newtoniana

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

En esta expresión, la *constante de interacción* k está determinada como $k = G\tilde{m}_1\tilde{m}_2$ donde G es la constante de Newton de la gravitación universal, mientras que \tilde{m}_1 y \tilde{m}_2 son las *masas gravitacionales* de los cuerpos en cuestión. Nótese que en este punto no estamos asumiendo que estas masas coinciden con las masas *inerciales* m_1 y m_2 que entran en la ley de fuerzas.

5.2.1 Órbitas del problema de Kepler

Vamos a comenzar estudiando la forma de las órbitas. Lo primero que podemos hacer para tener una comprensión intuitiva es escribir el potencial efectivo, el cual se lee

$$V_{1d} = -\frac{k}{r} + \frac{\ell^2}{2\mu r^2}$$

De acuerdo a esta fórmula, vemos que se cumplen los siguientes límites

$$\lim_{r \rightarrow 0} V_{1d} = \infty \quad \lim_{r \rightarrow \infty} V_{1d} = 0$$

Más aún, si calculamos su derivada radial

$$V'_{1d} = \frac{k}{r^2} - \frac{\ell^2}{\mu r^3}$$

vemos que es positiva $V'_{1d} > 0$ para radios lo bastante grandes, y negativa para valores lo bastante pequeños. Más aún, existe un solo punto crítico donde $V'_{1d} = 0$, en el radio $r_{\text{circ}} = \ell^2/\mu k$ donde el potencial toma el valor $V_{1d} = -\mu k^2/2\ell^2$.

Estas consideraciones implican que el potencial efectivo es muy grande cerca del origen, luego tiene un mínimo en el cual toma un valor negativo, y luego crece acercándose a cero por debajo cuando r va a infinito. Esto nos permite dibujar un perfil a mano alzada del potencial efectivo, que se puede ver en la figura 3.2.3. Es fácil notar que

- Tendremos órbitas circulares en el mínimo de potencial con radio r_{circ} , para las cuales la energía toma el valor mínimo $E = -\mu k^2/2\ell^2$.
- Cuando la energía es negativa $E < 0$, habrá órbitas acotadas entre dos puntos de retorno: uno cerca del origen donde crece la barrera centrífuga, y otro al alejarnos lo bastante del mismo.
- Cuando la energía es positiva o nula $E \geq 0$, tendremos órbitas no acotadas que nunca pasan por el origen.

Para profundizar el análisis, recordemos que en la Clase 4 encontramos una ecuación de segundo orden para la órbita de un problema de dos cuerpos con fuerza central

$$\frac{d^2 u}{d\theta^2} + u = -\frac{\mu}{\ell^2} \partial_u V$$

En el caso particular del problema de Kepler, el potencial toma la forma $V = -ku$ y por lo tanto su derivada será $\partial_u V = -k$. Con esto, la ecuación para la órbita nos queda escrita como

$$\frac{d^2 u}{d\theta^2} + u = \frac{\mu}{\ell^2} k$$

Esto se puede simplificar mediante el cambio de variables $\zeta = u - \mu k / \ell^2$, con lo que obtenemos

$$\frac{d^2 \zeta}{d\theta^2} + \zeta = 0$$

Esta es la ecuación para un oscilador armónico, cuya solución podemos escribir inmediatamente en la forma $\zeta = \zeta_0 \cos(\theta - \theta_0)$. Volviendo hacia atrás en nuestros cambios de variables para recuperar la forma de la solución en términos de la variable original r , obtenemos

$$\frac{1}{r} = \frac{\mu k}{\ell^2} (1 + e \cos(\theta - \theta_0))$$

Donde el parámetro $e = \zeta_0 \ell^2 / \mu k$ se denomina la *excentricidad* de la órbita. Esta ecuación da la dependencia del radio con el ángulo, y por lo tanto determina completamente las órbitas del problema de Kepler.

Para dar una interpretación física a la excentricidad, busquemos primero la posición de los ápsides θ_{aps} de la órbita, es decir los puntos donde se cumple que $dr/d\theta = 0$. Tomando una derivada de la ecuación anterior y evaluando en θ_{aps} se cumple que

$$\frac{d}{d\theta} \left(\frac{1}{r} \right) = -\frac{\mu k}{\ell^2} e \sin(\theta_{\text{aps}} - \theta_0) = 0$$

Las soluciones de esta ecuación son $\theta_{\text{aps}} = \theta_0$ y $\theta_{\text{aps}} = \theta_0 + \pi$. Reemplazando en la solución para la órbita obtenemos los radios correspondientes a estos dos ápsides, que valen

$$\frac{1}{r_{\text{aps}}^{\pm}} = \frac{\mu k}{\ell^2} (1 \pm e)$$

Nótese que, dado que debe cumplirse $r_{\text{aps}}^{\pm} > 0$, tenemos que el ápside con el signo $-$ solo existe cuando $e < 1$. Dado que en los ápsides también se cumple que $\dot{r} = 0$, podemos reemplazar en la fórmula de la energía, para obtener

$$E = \frac{1}{2} \mu \dot{r}^2 + \frac{\ell^2}{2\mu r^2} - \frac{k}{r} = \frac{\mu k^2}{2\ell^2} (1 \pm e)^2 - \frac{\mu k^2}{\ell^2} (1 \pm e) = \frac{\mu k^2}{2\ell^2} (e^2 - 1)$$

Este resultado demuestra que el parámetro de excentricidad es una medida de la energía. En particular

- La excentricidad vale cero cuando la energía toma el mínimo valor posible (es decir, de acuerdo al análisis de potencial efectivo, cuando la órbita es circular).
- La excentricidad es menor que 1 cuando la energía es negativa, es decir para las órbitas acotadas. Estas órbitas tienen dos ápsides r_{aps}^{\pm} , correspondientes a los dos puntos de retorno del potencial efectivo.
- La excentricidad es mayor que 1 cuando la energía es positiva, es decir para las órbitas abiertas. Estas órbitas tienen un solo ápside r_{aps}^{+} , que corresponde al único punto de retorno del potencial.
- El caso $e = 1$ corresponde a una órbita abierta con energía nula, que también tiene un solo ápside r_{aps}^{+} .

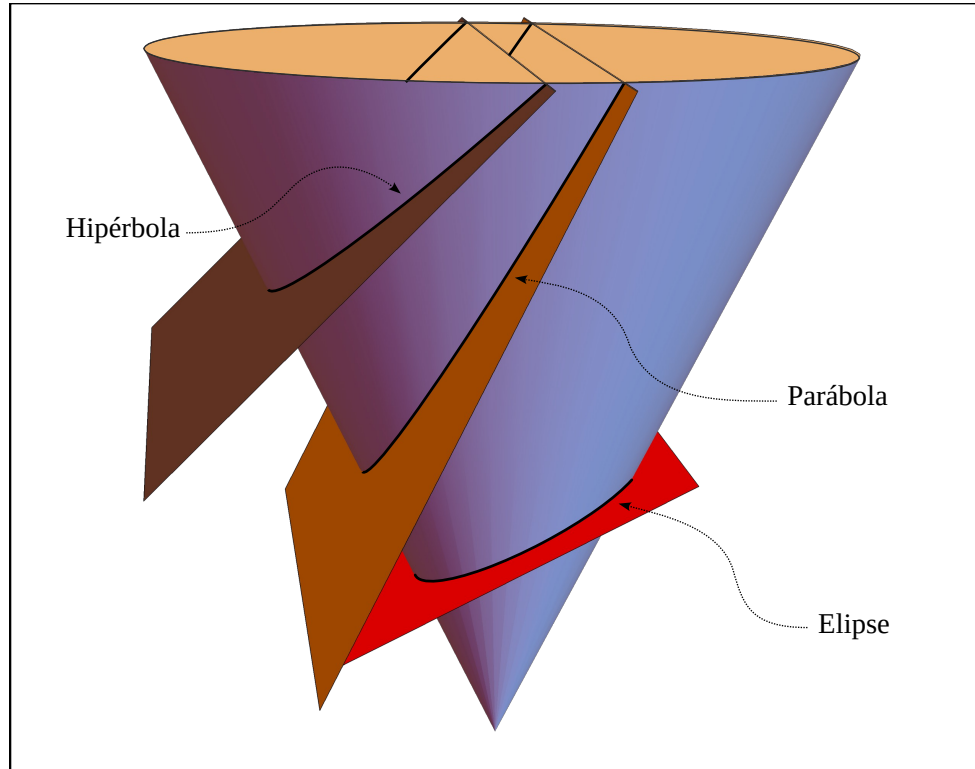


Figura 5.2.1: Secciones cónicas. La intersección de un plano con un cono genera las curvas que realizan las órbitas del problema de Kepler.

Nos gustaría tener una visión más geométrica de la forma de estas órbitas. Para esto, vamos a analizar las *secciones cónicas*, es decir las curvas que se obtienen al cortar un cono con un plano. En el espacio tridimensional \mathbb{R}^3 con coordenadas x, y, z , un cono cuya generatriz forma un ángulo de $\pi/4$ respecto del eje z cumple la fórmula

$$x^2 + y^2 = z^2$$

Por otro lado, un plano que contiene al eje y , formando un ángulo $\arctan(e)$ respecto del eje x al que intersecta en x_0 , se describe con la ecuación

$$z = e(x - x_0)$$

Reemplazando la segunda ecuación en la primera y operando obtenemos

$$x^2 + y^2 = e^2(x - x_0)^2$$

Escribiendo esta fórmula en coordenadas polares en las que se cumple $r^2 = x^2 + y^2$ mientras que $x = r \cos(\theta - \theta_0)$, nos queda

$$\frac{1}{r} = \frac{1}{x_0 e} (1 + e \cos(\theta - \theta_0))$$

Comparando con la solución para la órbita de más arriba, podemos identificar la excentricidad como la tangente del ángulo formado entre el plano y el eje x . Es decir que

- Cuando $e = 0$ el plano es horizontal y la órbita resultante es la intersección de un plano horizontal con un cono, o sea un círculo.
- Cuando la excentricidad es menor que 1, el plano está inclinado en un ángulo menor que $\pi/4$, lo que implica que intersectará el cono en una curva cerrada denominada *elipse*. Los extremos de la elipse corresponden a los dos ápsides r_{aps}^{\pm} .

Esto nos recuerda a la *primera ley de Kepler: los planetas se mueven en órbitas elípticas, con el sol en uno de sus focos*. Solo que el foco de nuestra elipse está en el origen de la coordenada relativa, que en términos de las variables originales \vec{r}_1 y \vec{r}_2 corresponde a la posición \vec{r}_{cm} del centro de masas. Solamente cuando recordamos que la masa del sol m_1 es mucho mayor que la de cualquier planeta m_2 , podemos aproximar $\vec{r}_1 \approx \vec{r}_{\text{cm}}$ y se cumple la primera ley.

- El caso $e = 1$ corresponde a un ángulo de $\pi/4$, es decir una intersección donde el plano es paralelo a la generatriz, la curva resultante es una parábola.
- Para $e > 1$ tenemos que el plano está a más inclinado respecto del eje horizontal que la generatriz, y por lo tanto la intersección es una hipérbola.

En el caso de la órbita elíptica, podemos identificar su semieje mayor a como la media distancia entre los dos ápsides

$$a = \frac{r_{\text{aps}}^+ + r_{\text{aps}}^-}{2} = \frac{1}{2} \frac{\ell^2}{\mu k} \left(\frac{1}{1-e} + \frac{1}{1+e} \right) = \frac{\ell^2}{\mu k} \frac{1}{1-e^2} = -\frac{k}{2E}$$

Por otro lado, el semieje menor b es la media distancia vertical entre los dos puntos que cumplen que $y = r \sin(\theta - \theta_{\text{aps}})$ es estacionario, o sea

$$\frac{dy}{d\theta} = \frac{dr}{d\theta} \sin(\theta - \theta_{\text{aps}}) + r \cos(\theta - \theta_{\text{aps}}) = 0$$

donde podemos usar la solución para la órbita para escribir la derivada en el primer término, resultando en

$$-r^2 \frac{\mu k}{\ell^2} e \sin^2(\theta - \theta_{\text{aps}}) + r \cos(\theta - \theta_{\text{aps}}) = 0$$

Usando que en ese punto $y = b$ y $x = x_{\text{medio}}$, donde x_{medio} es la posición en el eje x del centro de la elipse, tenemos

$$-\frac{\mu k}{\ell^2} e b^2 + x_{\text{medio}} = 0$$

Para determinar x_{medio} usamos el hecho geométrico de que

$$x_{\text{medio}} = \frac{1}{2} (r_{\text{aps}}^- - r_{\text{aps}}^+) = \frac{\ell^2}{2\mu k} \left(\frac{1}{1-e} - \frac{1}{1+e} \right) = \frac{\ell^2 e}{\mu k (1-e^2)}$$

Esto nos permite resolver $b^2 = \ell^4 / \mu^2 k^2 (1-e^2)$ o bien

$$b = \frac{\ell^2}{\mu k} \frac{1}{\sqrt{1-e^2}} = \sqrt{\frac{\ell^2}{\mu k}} \sqrt{a}$$

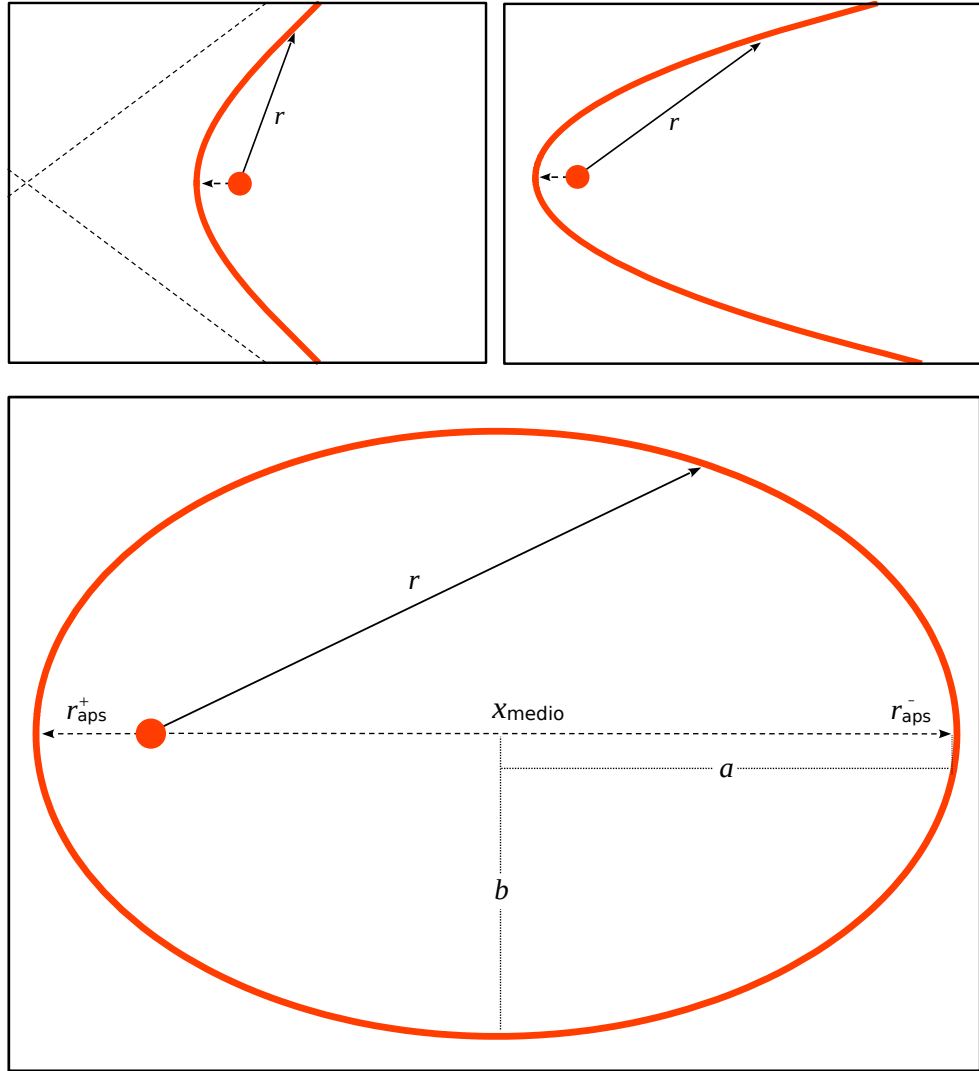


Figura 5.2.2: Órbitas del problema de Kepler: la órbita hiperbólica (arriba a la izquierda) toma asintóticamente la forma de dos rectas, que se consignan en línea punteada, al igual que la órbita parabólica (derecha) tiene un sólo ápside. La órbita elíptica (abajo) tiene dos ápsides, se muestran también sus semiejes mayor a y menor b , y el punto x_{medio} .

Ahora que tenemos a y b podemos usar la fórmula geométrica para el área de una elipse $A = \pi ab$ lo que resulta en $A = \pi a^{3/2} \sqrt{\ell^2 / \mu k}$. Pero si recordamos la segunda ley de Kepler podemos escribir también la expresión $A = T dA/dt = T \ell / 2\mu$ donde T es el tiempo que toma una revolución completa. Eliminando A entre ambas fórmulas, nos queda

$$T = 2\pi \sqrt{\frac{\mu}{k}} a^{3/2}$$

Esto tiene la forma de la *tercera ley de Kepler*: el cuadrado del período de un planeta es proporcional al cubo del semieje mayor de la órbita.

El único problema con este resultado es que la constante de proporcionalidad depende del planeta, y eso no es lo que descubrió Kepler. En efecto, reemplazando explícitamente $k = G\tilde{m}_1\tilde{m}_2$

y $\mu = m_1 m_2 / (m_1 + m_2)$, tenemos que

$$T = 2\pi a^{3/2} \sqrt{\frac{m_1 m_2}{G \tilde{m}_1 \tilde{m}_2 (m_1 + m_2)}}$$

Sin embargo, dado que la masa del sol m_1 es mucho mayor que la masa de cualquier planeta m_2 , podemos despreciar m_2 en esta expresión para obtener

$$T = 2\pi a^{3/2} \sqrt{\frac{m_2}{G \tilde{m}_1 \tilde{m}_2}}$$

En esta expresión, todavía tenemos la masa inercial m_2 y la masa gravitacional \tilde{m}_2 del planeta, por lo que no reproduce correctamente los resultados de Kepler. La observación crucial que hizo Newton es que si ambas masas coinciden entonces la constante de proporcionalidad entre el cuadrado del período y el cubo de la distancia media al sol no depende del planeta, y se cumple la tercera ley de Kepler.

Es interesante señalar que esta coincidencia que observó Newton entre masa inercial y masa gravitacional fue denominada *principio de equivalencia* por Einstein, quien la puso en la base de su *Teoría de la Relatividad General*.

Nota:

Isaac Newton se planteó explicar el movimiento de los planetas codificado en las leyes de Kepler, usando las mismas reglas que rigen el movimiento en la Tierra, que habían sido parcialmente elucidadas por Galileo. A ellos se refería con su célebre frase *parado en los hombros de gigantes*.

Para hacerlo, tuvo que inventar casi completamente el cálculo diferencial. Su libro *Philosophiæ naturalis principia mathematica* no fue publicado por la *Royal Society* de Londres, lo que llevó a Newton a pelearse fuertemente con Hooke, entonces director de la Sociedad. Además Leibniz, quien había visitado de Hooke en Londres, publicó poco después su versión del cálculo diferencial. Esto llevó a Newton a sospechar que Hooke le había mostrado a Leibniz sus manuscritos, lo que lo enojó tanto que se negó a publicar sus descubrimientos de ninguna manera.

En ese momento recibió la visita de Edmund Halley, quien estaba intrigado por una serie de apariciones históricas de cometas cada 76 años. Halley creía que se trataba del mismo cometa, y postuló que los cometas se movían en torno al Sol de modo similar como lo hacen los planetas. Estaba convencido de que podría aplicar los cálculos de Newton a su problema. Para poder hacerlo, tuvo que convencer a Newton de publicar los *Principia* y, como la *Royal Society* se negaba a invertir dinero en ellos, tuvo que pagar la edición de su propio bolsillo.

De este modo, la humanidad está en deuda con Edmund Halley no sólo por identificar su célebre cometa (que volvió a mostrarse 76 años después de su última aparición, como Halley había predicho), sino también por haber hecho posible que no se perdiera

el descubrimiento científico más importante de todos los tiempos, que sentara la semilla de la Revolución Industrial.

5.2.2 Movimiento en el tiempo

En esta sección vamos a investigar cómo se recorren las órbitas del problema de Kepler a medida que transcurre el tiempo. Es decir, queremos obtener explícitamente la evolución de las variables θ y r como función de t . Si recordamos la ecuación para $\dot{\theta} = \ell/\mu r^2$ podemos escribir

$$t = t_0 + \int_{\theta_0}^{\theta} \frac{\mu r^2}{\ell} d\theta$$

Usando la solución explícita para la órbita, tenemos que

$$t = t_0 + \int_{\theta_0}^{\theta} \frac{\ell^3}{\mu k^2 (1 + e \cos(\theta - \theta_0))^2} d\theta$$

La integración explícita depende del valor de la excentricidad e .

- En el caso parabólico $e = 1$ se pueden usar identidades trigonométricas para escribir

$$t = t_0 + \frac{\ell^3}{4\mu k^2} \int_{\theta_0}^{\theta} \sec^4\left(\frac{\theta - \theta_0}{2}\right) d\theta$$

Definiendo una nueva variable de integración $\zeta = \tan\left(\frac{\theta - \theta_0}{2}\right)$ esto se escribe

$$t = t_0 + \frac{\ell^3}{2\mu k^2} \int_0^{\zeta} (1 + \zeta^2) d\zeta$$

lo que se integra fácilmente para dar

$$t = t_0 + \frac{\ell^3}{2\mu k^2} \left(\tan\left(\frac{\theta - \theta_0}{2}\right) + \frac{1}{3} \tan^3\left(\frac{\theta - \theta_0}{2}\right) \right)$$

Esto es un polinomio cúbico en la tangente, que ahora se puede invertir para obtener $\theta(t)$.

- En el caso $e \neq 1$ la integral es más complicada, pero también se puede obtener explícitamente. Haciendo el mismo cambio de variables, obtenemos

$$t = t_0 + \frac{2\ell^3}{\mu k^2} \int_0^{\zeta} \frac{1 + \zeta^2}{(1 + e + (1 - e)\zeta^2)^2} d\zeta$$

lo que se integra a

$$t = t_0 + \frac{2\ell^3}{\mu k^2 (e^2 - 1)} \left(\frac{e \sin(\theta - \theta_0)}{1 + e \cos(\theta - \theta_0)} - \frac{2}{\sqrt{e^2 - 1}} \operatorname{arctanh}\left(\frac{\sqrt{e-1} \tan(\frac{\theta - \theta_0}{2})}{\sqrt{e+1}}\right) \right)$$

Ejercicio:

Utilice tablas de integrales o un software de manipulación algebraica como Mathematica para obtener el movimiento angular como función del tiempo para los dos casos anteriores.

En cuanto al movimiento radial en función del tiempo, recordemos de las Clase 3 que la integral necesaria toma la forma

$$t = t_0 + \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{\mu} \left(E - \frac{\ell^2}{2\mu r^2} + \frac{k}{r} \right)}}$$

En términos de los parámetros orbitales dados por el semieje mayor $a = -k/2E$ y la excentricidad $e^2 = 2\ell^2 E / \mu k^2 + 1$, esta integral se escribe

$$t = t_0 + \sqrt{\frac{\mu a}{k}} \int_{r_0}^r \frac{r dr}{\sqrt{e^2 - (a - r)^2}}$$

Cambiando variables a una nueva variable ζ conocida como *anomalía excéntrica*, según la fórmula $r = a(1 - e \cos \zeta)$, tenemos que la integral se puede calcular de manera explícita, resultando en

$$t = t_0 + \sqrt{\frac{\mu a^3}{k}} (\zeta - e \sin \zeta)$$

Esto no se puede invertir para obtener $r(t)$, pero de todas maneras ofrece una descripción completa del movimiento radial como función del tiempo.

5.2.3 Vector de Laplace-Runge-Lenz

El problema de Kepler tiene una característica esencial que lo diferencia de los demás problemas de dos cuerpos con fuerza central. Para describirla, vamos a volver brevemente a la mecánica vectorial, escribiendo

$$\dot{\vec{p}} = F_r(r) \hat{r}$$

Donde $F_r(r)$ es la componente radial de la fuerza. Tomando el producto vectorial con el momento angular, se puede escribir

$$\dot{\vec{p}} \times \vec{\ell} = F_r(r) \hat{r} \times \vec{\ell} = F_r(r) \frac{\mu}{r} \vec{r} \times (\vec{r} \times \dot{\vec{r}})$$

Usando la identidad $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$ y reordenando

$$\dot{\vec{p}} \times \vec{\ell} = F_r(r) r^2 \mu \left(\frac{\vec{r}}{r^2} \dot{r} - \frac{\dot{\vec{r}}}{r} \right)$$

Nótese que en esta expresión tenemos a ambos lados derivadas totales respecto del tiempo. Esto se puede hacer explícito

$$\frac{d}{dt} (\vec{p} \times \vec{\ell}) + F_r(r) r^2 \mu \frac{d}{dt} \left(\frac{\vec{r}}{r} \right) = 0$$

Para el caso particular del problema de Kepler $F_r(r) = -k/r^2$ con lo cual los dos términos se suman en una sola derivada total, de la forma

$$\frac{d}{dt} \underbrace{(\vec{p} \times \vec{\ell} - k\mu \hat{r})}_{\vec{v}} = 0$$

Por lo tanto, vemos que se conserva el vector $\vec{V} = \vec{p} \times \vec{\ell} - k\mu \vec{r}$, conocido como *vector de Laplace-Runge-Lenz*. En esta prueba hemos usado la forma explícita del potencial newtoniano, por lo que esta conservación es una propiedad exclusiva del problema de Kepler.

Para entender un poco más lo que significa el vector \vec{V} , probemos primero que es perpendicular al momento angular

$$\vec{V} \cdot \vec{\ell} = (\vec{p} \times \vec{\ell}) \cdot \vec{\ell} - k\mu \vec{r} \cdot \vec{\ell} = 0$$

lo que implica que \vec{V} yace en el plano de la órbita. Por otro lado, tomando su producto escalar con \vec{r} nos queda

$$\vec{V} \cdot \vec{r} = (\vec{p} \times \vec{\ell}) \cdot \vec{r} - k\mu r = \ell^2 - k\mu r$$

donde en la segunda igualdad usamos $(\vec{A} \times \vec{B}) \cdot \vec{C} = (\vec{B} \times \vec{C}) \cdot \vec{A} = (\vec{C} \times \vec{A}) \cdot \vec{B}$. Llamando θ_V el ángulo entre \vec{V} y el eje x , esto se puede reordenar en la forma

$$\frac{1}{r} = \frac{k\mu}{\ell^2} \left(1 + \frac{|\vec{V}|}{\mu k} \cos(\theta - \theta_V) \right)$$

Lo que coincide con la solución para la órbita si identificamos la excentricidad $e = |\vec{V}|/\mu k$ y el ángulo $\theta_V = \theta_{\text{aps}}$. Esta última identificación nos dice que \vec{V} apunta en la dirección del ápside, es decir a lo largo del eje mayor. Por lo tanto, la conservación de \vec{V} implica en particular que el eje mayor de la órbita no rota a medida que transcurre el movimiento.

Nota:

En el siglo XIX, el astrónomo Urbain Le Verrier observó que el eje mayor de la órbita del planeta Urano giraba a medida que transcurría el tiempo, lo que implica que no se conserva el vector de Laplace-Runge-Lenz. Atribuyó esta anomalía a la existencia de un planeta transuránico cuya posición calculó. El planeta Neptuno fue posteriormente descubierto a menos de 1° de donde lo había predicho Le Verrier.

Siguiendo el mismo sistema, Le Verrier propuso la existencia de un planeta aún más cercano al sol que Mercurio, para explicar una anomalía similar observada en el movimiento de éste. Sin embargo, debido a las dificultades asociadas a observar en dirección al Sol, el hipotético planeta Vulcano nunca fue descubierto.

En el siglo XX Albert Einstein explicó las anomalías de la órbita de Mercurio por medio de la teoría de la relatividad general, en la cual el potencial newtoniano no es exactamente $-k/r$ sino que contiene términos adicionales denominados *correcciones postnewtonianas*. Debido a estas correcciones, el vector de Laplace-Runge-Lenz no se conserva cerca del Sol, donde son más importantes. Esto dio fin a la búsqueda del planeta Vulcano.

5.3 Resumen

En esta clase resolvimos el problema de Kepler, es decir el movimiento de un sistema de dos cuerpos con fuerza central gravitacional. Encontramos la forma de sus órbitas, probamos que son de secciones cónicas, y encontramos también su evolución temporal.

En ese camino, demostramos las dos leyes de Kepler que aún nos faltaban. En la demostración de la segunda ley de Kepler, nos encontramos con que la masa gravitacional y la masa inercial de los cuerpos debe coincidir, un ingrediente fundamental que motivó a Einstein en su teoría de la Relatividad General.

Finalmente, descubrimos que el caso particular del problema de Kepler se distingue por tener un vector conservado que apunta en la dirección del eje mayor de la órbita, el vector de Laplace-Runge-Lenz. Nótese que no hemos identificado ninguna invarianza que dé origen a esta cantidad conservada, volveremos a esta cuestión más adelante.

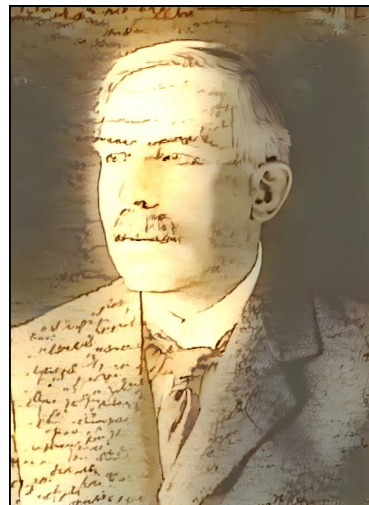
6. El problema de dos cuerpos VI: Dispersión

6.1 Objetivos

El problema de la *dispersión* en Mecánica corresponde a lo que en los cursos básicos de física llamábamos *choque elástico*.

Durante un choque elástico, dos partículas que inicialmente se dirigían la una contra la otra viajando en direcciones opuestas a lo largo de la misma recta, colisionan y son dispersadas, alejándose a lo largo de una nueva recta que forma un cierto ángulo con la dirección de incidencia.

Un punto a recordar es que en los problemas de choque elástico que resolvimos en el pasado, tal ángulo de dispersión siempre se incluía como uno de los datos externos necesarios para resolver el problema. En esta clase vamos a aprender a calcularlo.



Ernest Rutherford

6.2 El problema de la dispersión

En esta sección aprenderemos los conceptos básicos que se utilizan en el análisis de la dispersión. Con el fin de mantener algún grado de intuición, los introduciremos en el caso particular de una colisión de esferas rígidas, para generalizarlos luego a una interacción general.

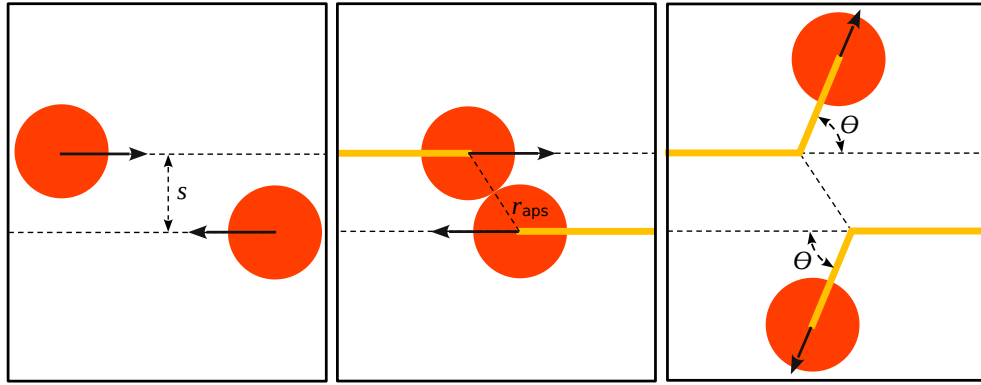


Figura 6.2.1: Problema de dispersión de esferas rígidas. De izquierda a derecha: los dos cuerpos se acercan uno al otro siguiendo rectas paralelas separadas por una distancia s , colisionan cuando la distancia entre sus centros es r_{aps} , y se dispersan siguiendo nuevamente rectas paralelas, que forman un ángulo Θ con la dirección de incidencia.

Primero estudiaremos cómo se obtiene el ángulo de dispersión a partir del potencial, luego introduciremos la idea de sección eficaz como una medida de la dispersión, y finalmente veremos cómo se aplica este conocimiento en los experimentos para obtener información sobre el potencial.

6.2.1 Ángulo de dispersión

Comencemos analizando un problema sencillo de choque, para identificar algunos elementos esenciales que luego podremos generalizar. El experimento consta de tres etapas

1. Inicialmente tenemos dos partículas que viajan en sentidos contrarios siguiendo dos rectas paralelas que están separadas por una cierta distancia transversal. Llamamos *dirección de incidencia* a la dirección de estas rectas, y *parámetro de impacto* a la distancia s que las separa.
2. Estas partículas se van acercando hasta que en algún momento “entran en contacto” y se produce la colisión. La idea intuitiva de “contacto” se puede representar mediante un potencial de *esfera rígida*, que toma la siguiente forma

$$V(|\vec{r}_1 - \vec{r}_2|) = \begin{cases} 0 & |\vec{r}_1 - \vec{r}_2| > r_{\text{aps}} \\ \infty & |\vec{r}_1 - \vec{r}_2| < r_{\text{aps}} \end{cases}$$

donde \vec{r}_1 y \vec{r}_2 denotan la posición de cada una de las partículas, ambas rodeadas por una esfera rígida de radio $r_{\text{aps}}/2$. Naturalmente estas partículas se acercarán la una a la otra, y “chocarán” cuando $|\vec{r}_1 - \vec{r}_2| = r_{\text{aps}}$.

3. Luego del evento, las partículas se alejan de nuevo moviéndose en sentidos opuestos a lo largo de rectas paralelas. Llamamos *dirección de dispersión* a la dirección de estas nuevas rectas, y *ángulo de dispersión* al ángulo Θ que forman con la dirección de incidencia.

Dadas las características del problema descripto, es evidente que se trata de un problema de dos cuerpos con interacción central. Por lo tanto, podemos seguir los pasos de la Clase 3 para describirlo en el sistema de coordenadas en el centro de masas, en términos de una sola partícula ficticia que se mueve en un potencial dado por $V(r)$ dado por

$$V(r) = \begin{cases} 0 & r > r_{\text{aps}} \\ \infty & r < r_{\text{aps}} \end{cases}$$

Esta partícula se acerca inicialmente al origen siguiendo una recta en la dirección de incidencia que pasa a una distancia s del mismo, luego “tiene lugar la colisión”, y finalmente la partícula se aleja del origen siguiendo una recta en la dirección de dispersión.

Analicemos este problema aplicando el método del potencial efectivo que estudiamos en la Clase 3. Para construirlo, necesitamos el valor del momento angular, que vendrá dado por

$$\ell = |\vec{r} \times \vec{p}| = p \underbrace{r \sin(\hat{r}\hat{p})}_s = ps = \sqrt{2\mu E} s$$

donde usamos el hecho de que el momento angular se conserva para evaluarlo en algún instante previo a la colisión, cuando el momento lineal de la partícula cumplía $p^2 = 2\mu E$. Con esto podemos escribir para el potencial efectivo

$$V_{\text{eff}}(r) = \begin{cases} E s^2 / r^2 & r > r_{\text{aps}} \\ \infty & r < r_{\text{aps}} \end{cases}$$

Ahora bien, el movimiento tendrá lugar en la región $V_{\text{eff}} < E$. La primera línea del potencial efectivo nos dice que esta región está determinada por $r > s$ (esto es una propiedad geométrica que es evidente en el diagrama). Por otro lado, la segunda línea nos dice que debe cumplirse $r > r_{\text{aps}}$. Por lo tanto, si $s > r_{\text{aps}}$ tendremos un punto de retorno en $r = s$ (algo que nuevamente se puede ver en el diagrama). Por otro lado, si $s < r_{\text{aps}}$ “se produce la colisión”, es decir que el punto de retorno está en $r = r_{\text{aps}}$.

Como vimos en la Clase 3, en los puntos de retorno se invierte la velocidad radial. Por supuesto nada sucede allí con la componente angular de la velocidad. Esto significa que si $s < r_{\text{aps}}$ entonces en $r = r_{\text{aps}}$ la trayectoria se quiebra, del modo que se ve en el diagrama. El ángulo θ_{aps} formado por el vector \vec{r} con la dirección de incidencia cumplirá

$$\sin \theta_{\text{aps}} = -\frac{s}{r_{\text{aps}}}$$

Dado que la reflexión invierte la componente radial de la velocidad, un análisis geométrico nos permite obtener la dirección de dispersión, la cual forma un ángulo Θ con la de incidencia, dado por

$$\Theta = 2\theta_{\text{aps}} - \pi$$

Es un ejercicio sencillo probar, usando las dos fórmulas anteriores, que se cumple que

$$s = r_{\text{aps}} \sin\left(\frac{\Theta}{2}\right)$$

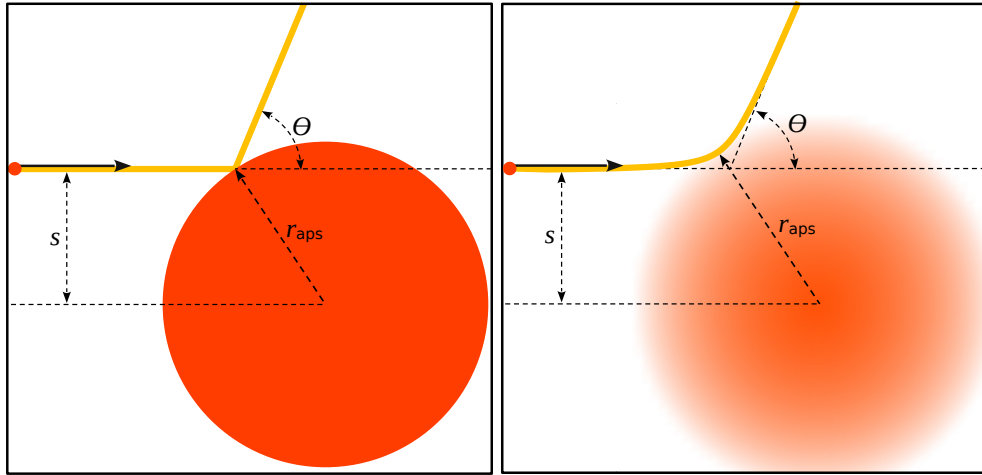


Figura 6.2.2: Problema de dispersión en el centro de masas. A la izquierda vemos el problema de dispersión de esferas rígidas, ahora contra una partícula puntual incidiendo contra una esfera de radio doble. Se acerca siguiendo una recta separada una distancia s del centro, colisiona cuando la distancia es r_{aps} , y se dispersa siguiendo nuevamente una recta, que forma un ángulo Θ con la dirección de incidencia. A la derecha vemos un potencial general, la trayectoria se suaviza pero los conceptos de distancia s y ángulo de dispersión Θ se mantienen en la región asintótica.

Con esto, aprendemos que el ángulo de dispersión depende del parámetro de impacto s y del radio r_{aps} que es una característica del potencial. Más aún, si medimos como varía el ángulo Θ como función del parámetro de impacto s , podemos encontrar el radio r_{aps} del potencial de interacción. Es decir que podemos utilizar un experimento de choque para obtener información acerca del potencial.

En este punto, resulta conveniente notar que los cálculos que realizamos hasta aquí se pueden generalizar sin mayor dificultad para un potencial arbitrario. En otras palabras, en lugar de limitarnos a estudiar partículas que colisionan como si fueran esferas rígidas, nos permitimos también considerar aquéllas que son esferas “blandas”, en el sentido de que su máxima proximidad durante un choque depende de su energía.

En el caso general, para calcular el ángulo de dispersión Θ en función del parámetro de impacto s , utilizaremos la solución para la órbita que obtuvimos en la Clase 4

$$\theta = \theta_0 - \int_{u_0}^u \frac{du}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}}$$

donde la hemos escrito en términos de la variable $u = 1/r$, y hemos reemplazado el momento angular según lo que obtuvimos más arriba $\ell^2 = 2\mu E s^2$. Es fácil ver que si ponemos el instante inicial muy en el pasado tenemos que los valores iniciales satisfacen $u_0 = 0$ y $\theta_0 = \pi$. Entonces

$$\theta = \pi - \int_0^u \frac{du}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}}$$

Podemos separar la integral en el ápside de la órbita y poner

$$\theta = \pi - \int_0^{u_{\text{aps}}} \frac{|du|}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}} + \int_{u_{\text{aps}}}^u \frac{|du|}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}}$$

Donde el + en el segundo término viene del hecho de que u está moviéndose en sentido contrario, por lo que du cambiará de signo. Si ahora ponemos el instante final muy en el futuro, tenemos que $u = 0$ y el ángulo corresponde al ángulo de dispersión $\theta = \Theta$

$$\Theta = \pi - \int_0^{u_{\text{aps}}} \frac{|du|}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}} + \int_{u_{\text{aps}}}^0 \frac{|du|}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}} = \Theta = \pi - 2 \int_0^{u_{\text{aps}}} \frac{du}{\sqrt{\frac{1}{s^2} \left(1 - \frac{V}{E}\right) - u^2}}$$

El ápside u_{aps} que entra en el límite superior de integración, está definido como el punto donde la velocidad radial se anula. Es decir que cumple la fórmula

$$u_{\text{aps}}^2 = \frac{1}{s^2} \left(1 - \frac{V(1/u_{\text{aps}})}{E}\right)$$

Una vez calculada la integral, tendremos una forma explícita para la dependencia en s del ángulo de dispersión Θ .

Ejercicio: esferas rígidas

Probar que estas fórmulas reproducen, en el caso del potencial de esfera rígida, el resultado para el ángulo de dispersión que encontramos más arriba mediante consideraciones geométricas.

Ejemplo: dispersión de Rutherford

En el problema de la dispersión de Rutherford, electrones son disparados sobre un ion negativo, por lo que el potencial en cuestión es un potencial coulombiano repulsivo

$$V(r) = \frac{k}{r}$$

donde $k = Ze^2/4\pi\epsilon_0$ con e la carga de electrón, Ze la carga del ion, y ϵ_0 es la permeabilidad eléctrica del vacío. Las órbitas se pueden resolver de manera completamente análoga a la del problema de Kepler, obteniendo

$$\frac{1}{r} = -\frac{\mu k}{\ell^2} (1 - e \cos(\theta - \theta_{\text{aps}}))$$

Comparando con los resultados de la Clase 5 notamos que, como podría esperarse, hemos cambiando el signo de k respecto de la expresión kepleriana. También hemos cambiado el signo delante de e , esto corresponde a elegir $\theta_0 = \theta_{\text{aps}} - \pi$, de modo tal que en el ápside tenemos $r_{\text{aps}} > 0$ cuando $e > 1$.

Tanto en el pasado lejano cuando la partícula estaba incidiendo, cuanto en el futuro lejano cuando se aleja después de haber sido dispersada, se cumple que $r \rightarrow \infty$ y por lo tanto

$$0 = 1 - e \cos(\theta_{\infty} - \theta_{\text{aps}})$$

esta ecuación tiene dos soluciones $\theta_{\infty}^{\pm} = \theta_{\text{aps}} \pm \text{Arccos}(1/e)$, que corresponden a los valores del ángulo para la partícula incidente y para la partícula dispersada. Esto implica que el ángulo de dispersión vendrá dado por $\Theta = \pi - (\theta_{\infty}^+ - \theta_{\infty}^-) = \pi - 2\text{Arccos}(1/e)$, lo que nos permite reescribir

$$0 = 1 + e \sin\left(\frac{\Theta}{2}\right)$$

De modo que la excentricidad está relacionada con el ángulo de dispersión. Usando nuestros parámetros físicos tenemos que

$$e^2 - 1 = \frac{2E\ell^2}{\mu k^2} = \frac{(2Es)^2}{k^2} = \cot^2\left(\frac{\Theta}{2}\right)$$

Lo que nos permite resolver

$$s = \frac{k}{2E} \cot\left(\frac{\Theta}{2}\right)$$

Nota:

El 19 de octubre de 2017 el astrónomo canadiense Robert Weryk descubrió un objeto que se desplazaba en dirección al Sol siguiendo una órbita hiperbólica. Es decir que no se trataba de un cuerpo menor del Sistema Solar, sino que provenía del espacio interestelar. Se lo bautizó *1I/Oumuamua* donde *1I* indica que se trata del primer objeto interestelar identificado, y *'oumuamua* es una palabra hawaiana que significa “explorador”.

Un descubrimiento similar abre la novela *“Cita con Rama”* del autor inglés Arthur C. Clarke. En el caso de la novela, el objeto era realmente un explorador, es decir una nave estelar de una civilización extraterrestre.

Con los elementos que estudiamos en esta clase, estamos en condiciones de calcular el ángulo con el que fueron desviados *'Oumuamua* o *Rama* al abandonar el sistema solar.

6.2.2 Sección eficaz

En el caso sencillo del potencial de esfera rígida, es evidente que si el parámetro de impacto es mayor al radio de las esferas $s > r_{\text{aps}}$ no se produce ningún choque. En la sección anterior fuimos capaces de probar formalmente este hecho intuitivo hallando los puntos de retorno del potencial efectivo.

Esto significa que si arrojamamos hacia el origen una variedad de partículas con diferentes parámetros de impacto, sólo serán dispersadas aquellas que cumplen $s < r_{\text{aps}}$. Podríamos verificar esto disponiendo del otro lado una pantalla plana perpendicular a la dirección de incidencia, sobre la cual obtendremos una “sombra” circular de radio r_{aps} . La superficie de tal sombra

$$\sigma_T = \pi r_{\text{aps}}^2$$

se denomina *sección eficaz total* y es una medida de qué tan eficiente es el potencial de esfera rígida a la hora de provocar la dispersión de partículas. Podemos reescribir la fórmula anterior de manera trivial como una integral en el parámetro de impacto

$$\sigma_T = \int_0^{r_{\text{aps}}} \int_0^{2\pi} s \, ds \, d\Phi$$

donde hemos agregado también una integral en el ángulo Φ que gira alrededor de la dirección de incidencia. Ahora bien, en la sección anterior aprendimos que el parámetro de impacto s se puede expresar como una función del ángulo de dispersión Θ , lo que nos permite cambiar variables de integración para obtener

$$\sigma_T = \int_0^\pi \int_0^{2\pi} s \left| \frac{ds}{d\Theta} \right| d\Theta d\Phi = \int_\Omega \underbrace{\frac{s}{\sin \Theta} \left| \frac{ds}{d\Theta} \right|}_{\sigma(\Theta)} d\Omega$$

Donde $d\Omega = \sin \Theta d\Theta d\Phi$ es la medida de integración sobre la esfera que rodea el origen. En la última igualdad hemos reescrito la sección eficaz total como la integral sobre tal esfera de una *sección eficaz diferencial*, que está definida como

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

Esta magnitud es una medida de qué tan eficiente resulta este potencial para dispersar partículas con la dirección definida por el ángulo Θ .

Si bien introdujimos los conceptos anteriores motivándolos con el potencial de esfera rígida, lo cierto es que las últimas dos fórmulas pueden utilizarse como una definición general de sección eficaz diferencial y total respectivamente.

La única salvedad es que para expresar la sección eficaz diferencial como una función del ángulo, necesitamos que la expresión para $s(\Theta)$ sea univaluada. Este no es el caso si dos o más parámetros de impacto diferentes resultan en el mismo ángulo de dispersión. Veremos más adelante que en este caso la generalización más útil de la definición de arriba es

$$\sigma(\Theta) = \sum_n \frac{s_n}{\sin \Theta} \left| \frac{ds_n}{d\Theta} \right|$$

donde la suma corre sobre los diferentes valores s_n del parámetro de impacto que resultan en el mismo ángulo de dispersión Θ .

Ejercicio:

Use la fórmula que obtuvimos en la sección previa, que relaciona el ángulo de dispersión con el parámetro de impacto en el caso de la esfera rígida, para calcular la sección eficaz diferencial $\sigma(\Theta)$ del este problema.

Use el resultado para recuperar la sección eficaz total σ_T que escribimos más arriba.

Ejemplo: dispersión de Rutherford

Usando la fórmula que relaciona el ángulo de dispersión con el parámetro de impacto para el problema de la dispersión de Rutherford, podemos calcular la derivada

$$\frac{ds}{d\Theta} = -\frac{k}{4E} \csc^2\left(\frac{\Theta}{2}\right)$$

Esto nos permite obtener para la sección eficaz diferencial

$$\sigma(\theta) = \frac{\frac{k}{2E} \cot\left(\frac{\Theta}{2}\right)}{\sin\Theta} \frac{k}{4E} \csc^2\left(\frac{\Theta}{2}\right) = \left(\frac{k}{4E}\right)^2 \csc^4\left(\frac{\Theta}{2}\right)$$

donde fue necesario utilizar algunas identidades trigonométricas.

Si quisiéramos calcular la sección eficaz total, tendríamos

$$\sigma_T = \int_{\Omega} \sigma(\theta) d\Omega = \left(\frac{k}{4E}\right)^2 \int \csc^4\left(\frac{\Theta}{2}\right) \sin\Theta d\Theta d\Phi$$

lo que se puede reescribir en términos de la variable $\zeta = \sin(\Theta/2)$ como

$$\sigma_T = 2\pi \left(\frac{k}{2E}\right)^2 \int_0^1 \frac{d\zeta}{\zeta^3} \rightarrow \infty$$

con lo que vemos que la sección eficaz total diverge. Para interpretar este resultado, vayamos hacia atrás para escribir la sección eficaz total en su forma original en términos de una integral en el parámetro de impacto

$$\sigma_T = \int_0^\infty \int_0^{2\pi} s ds d\Phi$$

donde para recuperar el resultado infinito tuvimos que integrar hasta valores de s arbitrariamente grandes. Esto significa que todas las partículas son dispersadas sin importar su parámetro de impacto. En otras palabras, la interacción de Coulomb es una *fuerza de largo alcance*.

Nota:

A principios del siglo XX, el descubridor del electrón Joseph John Thomson propuso que el átomo estaba compuesto por un conjunto de electrones, distribuidos dentro de una nube de carga positiva, en lo que se llamó el modelo del *budín de pasas*.

Para testear esta idea, en 1911 Ernest Rutherford hizo pasar partículas alfa a través de una hoja fina de aluminio. Observó ángulos de dispersión muy grandes, que no eran compatibles con el budín de pasas, sino con que toda la carga positiva estuviera acumulada en un solo punto. Descubrió así el núcleo atómico.

6.2.3 Experimentos de dispersión

La razón para definir una sección eficaz diferencial es que resulta ser una magnitud medible con relativa facilidad en un *experimento de dispersión*. Tal experimento consiste en hacer colisionar dos haces de partículas que viajan en direcciones contrarias, y estudiar el producto de la colisión.

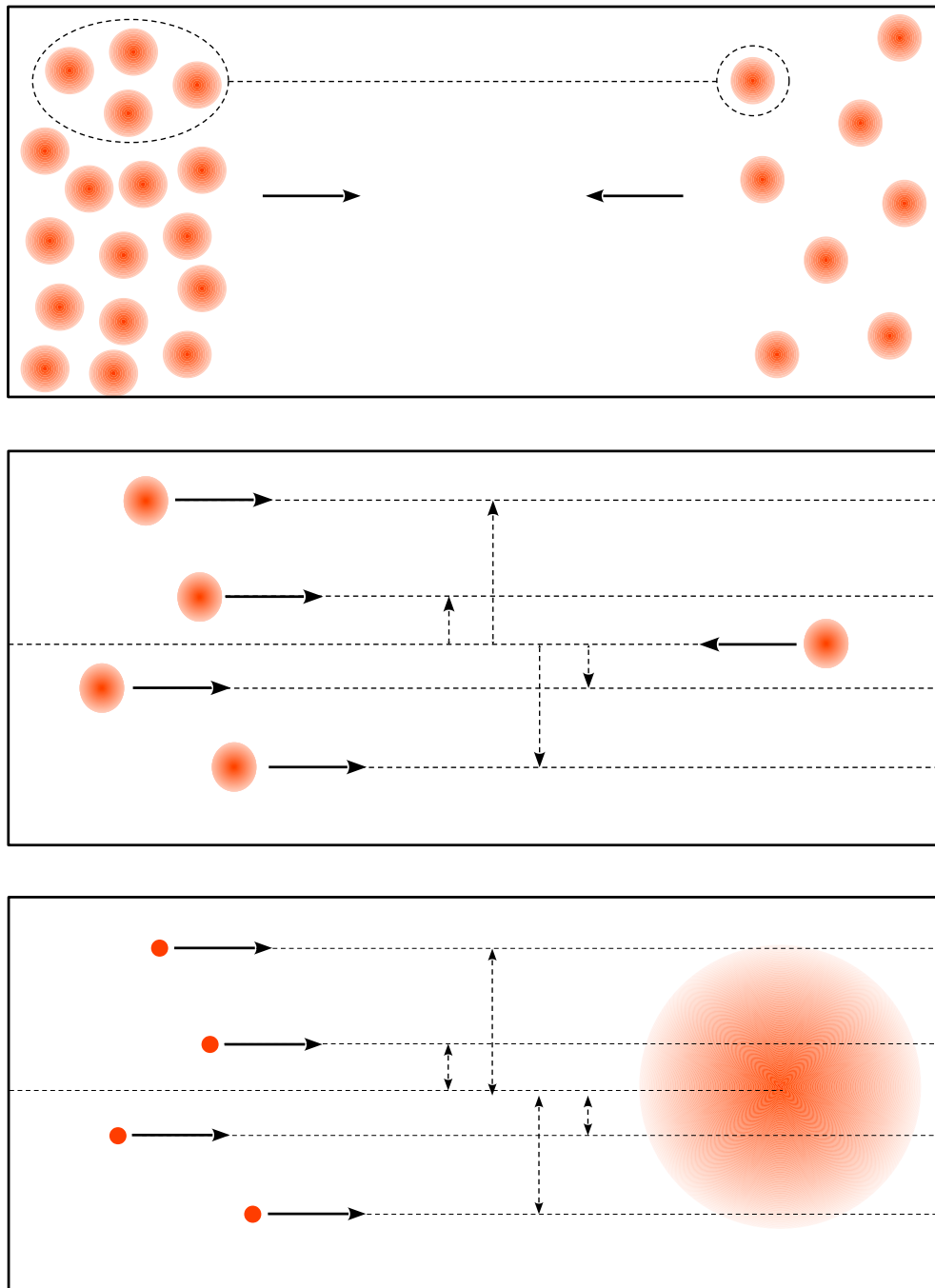


Figura 6.2.3: Experimento de dispersión: dos haces de partículas desplazándose en direcciones contrarias inciden uno contra el otro (primera figura). Un subhaz del haz más concentrado (remarcado con un óvalo en la primera figura) incide contra una partícula del haz más diluido, cada una de sus partículas componentes lo hace con su propio parámetro de impacto (segunda figura). El mismo problema visto desde el centro de masas muestra varias partículas incidiendo contra un centro dispersor (tercera figura).

Para simplificar el problema, notemos lo siguiente

- La primera hipótesis que tenemos que hacer para describir este experimento es que dentro de cada uno de los haces las diferentes partículas no interactúan entre sí.
- La segunda hipótesis es que dada una partícula cualquiera del haz más concentrado, esta interactúa con una sola del haz más diluido.

De esta manera, el problema se descompone en un conjunto de problemas similares, en cada uno de los cuales un haz de partículas (que es un sub-haz del haz más denso) incide sobre una sola partícula (que pertenece al haz más diluido). Tomando ahora una partícula cualquiera del sub-haz incidente, podemos describir su interacción con la partícula que le corresponde del haz diluido en el sistema centro de masas. En este sistema y como sabemos de las secciones previas, el problema corresponde al de una partícula ficticia interactuando con un potencial central en el origen.

Es evidente que el sistema centro de masas estará en una posición distinta para cada una de las partículas del sub-haz incidente. Sin embargo, se puede verificar que el efecto de trasladar el sistema centro de masas de cada una de ellas al mismo punto es solamente el de reescalar la intensidad del sub-haz. Por lo tanto, terminamos reduciendo el problema al de un haz de partículas ficticias que inciden sobre un potencial central en el origen.

Vamos a definir la intensidad I del haz incidente como el número de partículas que atraviesan por unidad de tiempo la unidad de área perpendicular a la dirección de incidencia. Por lo tanto, si tomamos un anillo de radio s y ancho ds , y cortamos un pequeño sector angular $d\Phi$, tendremos el área $dA = s ds d\Phi$ resultante es atravesada durante un tiempo dt por un número de partículas dado por

$$dN = I dA dt = I s ds d\Phi dt$$

Todas estas partículas saldrán dispersadas en la dirección Θ que corresponde al valor de s y con el mismo valor de Φ , por lo que podemos escribir

$$dN = I s \left| \frac{ds}{d\Theta} \right| d\Theta d\Phi dt = I \frac{s}{\sin \Theta} \left| \frac{ds}{d\Theta} \right| d\Omega dt$$

Ahora bien, si hubiera más de un valor del parámetro de impacto que resultara en el mismo ángulo de dispersión, el número total de partículas dispersadas en una dada dirección será

$$dN = I \sum_n \frac{s_n}{\sin \Theta} \left| \frac{ds_n}{d\Theta} \right| d\Omega dt$$

Escrito en términos del diferencial de área $d\tilde{A} = r^2 d\Omega$ de una esfera de radio r , tenemos

$$dN = \frac{I}{r^2} \sigma(\Theta) d\tilde{A} dt$$

Es decir que si nos situamos en el sistema centro de masas a una distancia r del origen y contamos el número total de partículas dN que atraviesan un área $d\tilde{A}$ durante un tiempo dt , podemos obtener con facilidad la sección eficaz diferencial $\sigma(\Theta)$ del choque correspondiente.

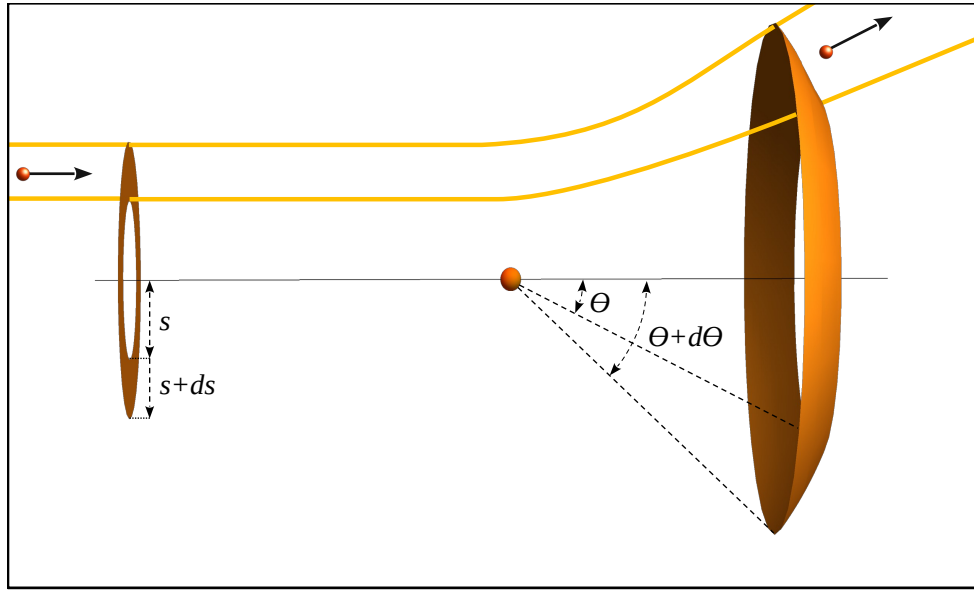


Figura 6.2.4: Experimento de dispersión. Las dN partículas que inciden durante un tiempo dt con parámetros de impacto entre s y $s+ds$ se dispersan con ángulos entre Θ y $\Theta+d\Theta$.

El *problema de dispersión inversa* consiste en obtener información sobre el potencial a partir de la sección eficaz diferencial medida en un experimento de dispersión.

Para ver un primer ejemplo de esto, escribamos la integral para el ángulo de dispersión que habíamos encontrado más arriba en la forma

$$\Theta = \pi - 2 \int_0^{\zeta_{\text{aps}}} \frac{d\zeta}{\sqrt{1 - \frac{V}{E} - \zeta^2}}$$

donde definimos la nueva variable de integración $\zeta = su$ y el límite superior de integración se obtiene de

$$\zeta_{\text{aps}}^2 = 1 - \frac{V(s/\zeta_{\text{aps}})}{E}$$

Si el potencial se anula en el infinito, para s lo bastante grande el límite superior de integración será $\zeta_{\text{aps}} = 1$ y en la integral tenemos V/E pequeño, por lo que podremos escribir

$$\Theta = \pi - 2 \int_0^1 \frac{d\zeta}{\sqrt{1 - \zeta^2}} = 0$$

es decir que un potencial que va a cero lo bastante rápido en el infinito no dispersa partículas con parámetro de impacto grande.

Por otro lado, si s es pequeño y la energía es lo bastante grande, estamos en la situación que se conoce como *dispersión profunda*. En este caso, si el potencial está acotado en el origen podemos escribir

$$\zeta_{\text{aps}}^2 = 1 - \frac{V(0)}{E} = 1$$

donde en la segunda igualdad asumimos que la energía es lo bastante grande $E \gg |V(0)|$. La

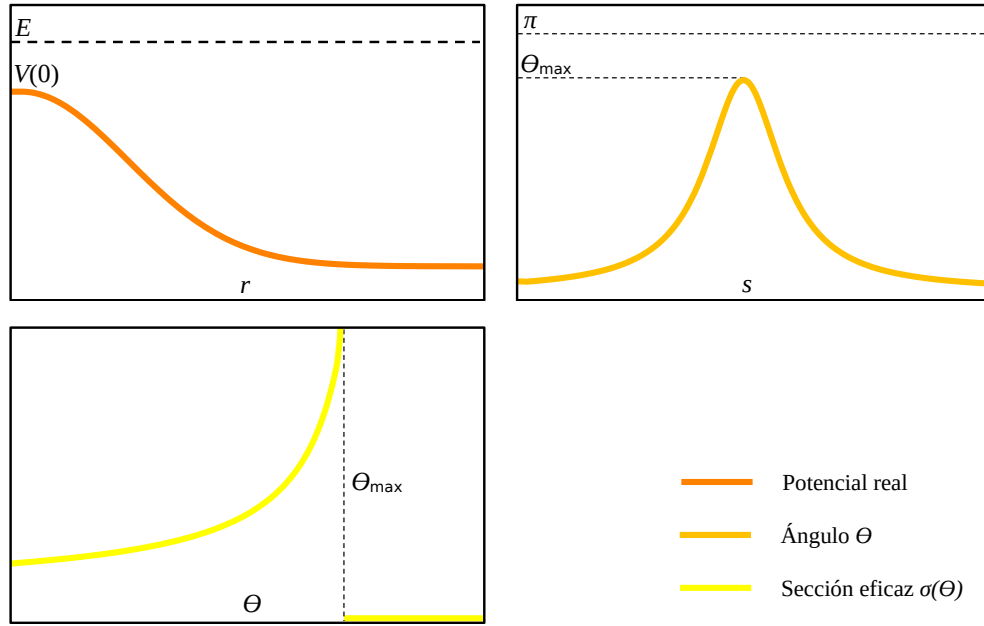


Figura 6.2.5: Dispersión en arcoíris. El potencial V tiene un máximo, lo que implica que partículas con parámetros de impacto s muy grandes o muy pequeños no serán dispersadas. Esto resulta en que el ángulo de dispersión Θ tiene un máximo Θ_{\max} en algún valor intermedio del parámetro de impacto. La sección eficaz diferencial $\sigma(\Theta)$ diverge en Θ_{\max} y cae a cero para ángulos mayores.

integral se puede separar en dos partes, en la forma

$$\Theta = \pi - 2 \int_0^s \frac{d\zeta}{\sqrt{1 - \frac{V}{E} - \zeta^2}} - 2 \int_s^1 \frac{d\zeta}{\sqrt{1 - \frac{V}{E} - \zeta^2}}$$

como s es pequeño podemos descartar la primera integral, y en la segunda reemplazar V por $V(0) \ll E$, para obtener nuevamente

$$\Theta = \pi - 2 \int_0^1 \frac{d\zeta}{\sqrt{1 - \zeta^2}} = 0$$

Es decir que para un potencial acotado en el origen y que se anula rápidamente en infinito, tendremos ángulo de dispersión nulo a energía grande para los dos valores extremos del parámetro de impacto. Esto quiere decir que $\Theta(s)$ debe tener un máximo en algún valor intermedio de s .

Supongamos primero que cuando el ángulo de dispersión alcanza su máximo es aún menor que π . Dado que en el máximo se cumple que $d\Theta/ds = 0$, para ese ángulo la sección eficaz será infinita. Por otro lado siendo ese ángulo el máximo en el cual encontraremos partículas dispersadas, la sección eficaz caerá a cero para ángulos mayores.

Este fenómeno se conoce como *dispersión en arcoíris*. Razonando en sentido inverso, podemos afirmar que cuando un experimento presenta dispersión en arcoíris el potencial va a cero a distancias grandes y está acotado a distancias pequeñas.

También podría suceder que el valor máximo del ángulo de dispersión fuera mayor que π . Esto se explica si para valores intermedios del parámetro de impacto el potencial efectivo tiene un máximo.

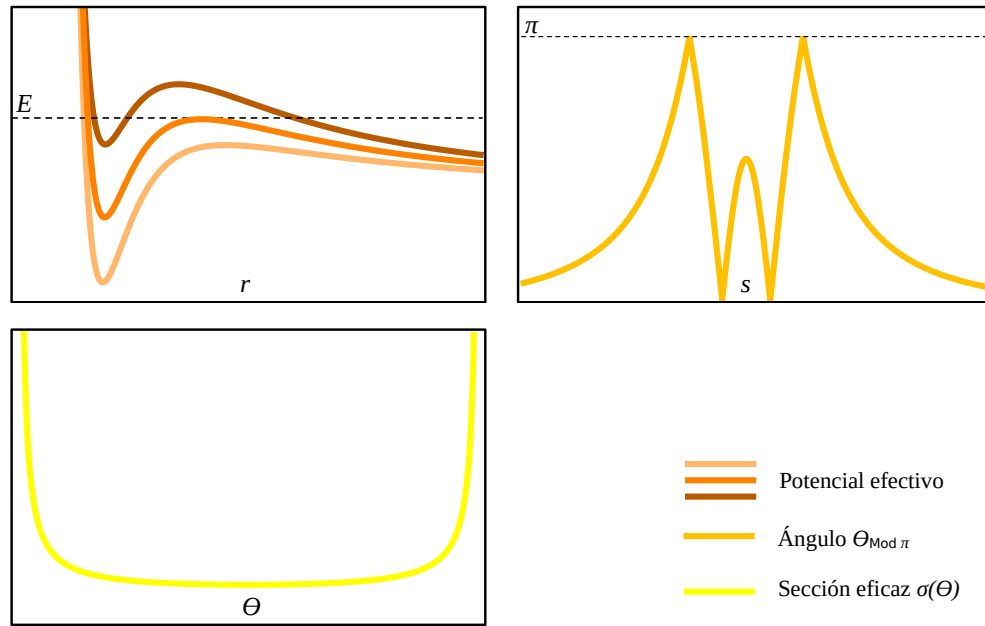


Figura 6.2.6: *Dispersión gloriosa.* El potencial efectivo V_{1d} tiene un máximo, lo que implica que para algún valor del parámetros de impacto s la energía será similar al máximo del potencial, con lo que la partícula se moverá muy lentamente en la dirección radial allí. Esto resulta en que el ángulo de dispersión Θ crecerá mucho, superando π , por lo que debemos corregirlo tomando $\Theta_{\text{Mod}\pi}$. La sección eficaz diferencial $\sigma(\Theta)$ diverge en $\Theta = 0, \pi$.

Una partícula que incide con una energía similar a la energía del máximo permanece mucho tiempo dando vueltas cerca del origen, por lo que el ángulo de dispersión resultante resulta mayor que π . Dado que sólo medimos ángulos entre 0 y π , el ángulo medido sería $\Theta_{\text{Mod}\pi} = \Theta + n\pi < \pi$ con $n \in \mathbb{Z}$. El gráfico resultante para $\Theta_{\text{Mod}\pi}$ como función de s presentaría varias reflexiones en $\Theta_{\text{Mod}\pi} = 0$ y $\Theta_{\text{Mod}\pi} = \pi$.

Para cualquier ángulo entre 0 y π , vemos que hay varios valores de s que contribuyen al mismo ángulo de dispersión. Más aún, en cada uno de los valores de s para los cuales $\Theta_{\text{Mod}\pi} = 0$ o $\Theta_{\text{Mod}\pi} = \pi$, se cumple que $\sin \Theta = 0$ cuando $s \neq 0$, por lo que el cociente en la sección eficaz diferencial diverge y el resultado es infinito.

Este fenómeno se conoce como *dispersión gloriosa*. Invirtiendo el razonamiento, cuando un experimento presenta dispersión gloriosa, sabemos que el potencial efectivo tiene un máximo en los valores intermedios del parámetro de impacto.

6.3 Resumen

En esta clase estudiamos el problema choque de partículas con las herramientas que aprendimos para el problema de dos cuerpos. Pudimos calcular el ángulo de dispersión, que en los problemas de choque de los cursos básicos siempre se nos daba como un dato. Vimos que el resultado depende en detalle de la forma del potencial.

Definimos la sección eficaz diferencial y la sección eficaz total en términos del ángulo de dispersión, y vimos que esas magnitudes pueden medirse mediante experimentos de dispersión de haces de partículas.

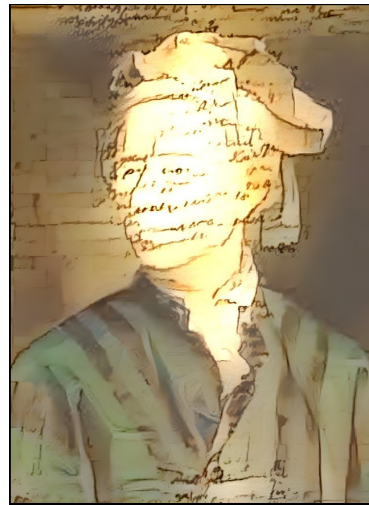
Fuimos capaces de discernir diferentes comportamientos de la sección eficaz diferencial que corresponden a propiedades específicas del potencial, como la de estar acotado a distancias pequeñas y grandes, y tener un máximo para valores intermedios del parámetro de impacto.

7. Cuerpo rígido I: Rotaciones

7.1 Objetivos

En esta clase vamos a estudiar las transformaciones de rotación espacial. Para esto, tendremos que aprender a trabajar con transformaciones lineales representadas en términos de matrices, e identificar cuáles de ellas corresponden a lo que llamamos una rotación.

Si bien aprenderemos conceptos e ideas que son muy útiles en varios ámbitos de la física, nuestro interés inmediato está en su aplicación a la dinámica del cuerpo rígido, que desarrollaremos en las clases subsiguientes.



Leonhard Euler

7.2 Rotaciones espaciales

Durante esta clase vamos a concentrarnos en la descripción vectorial del espacio, donde a cada punto P le corresponde un vector \vec{r} cuyas componentes (x, y, z) en algún sistema cartesiano representan las coordenadas del punto en ese sistema. Supongamos que aplicamos sobre (x, y, z) una transformación que nos da un nuevo conjunto de coordenadas (x', y', z') . Podemos interpretar esto de dos maneras

1. La transformación actúa sobre el espacio, tomando el punto P representado por el vector \vec{r} con coordenadas (x, y, z) , y transformándolo en otro punto P' representado por el vector \vec{r}' con coordenadas (x', y', z') . Esto se conoce como el *punto de vista activo*.

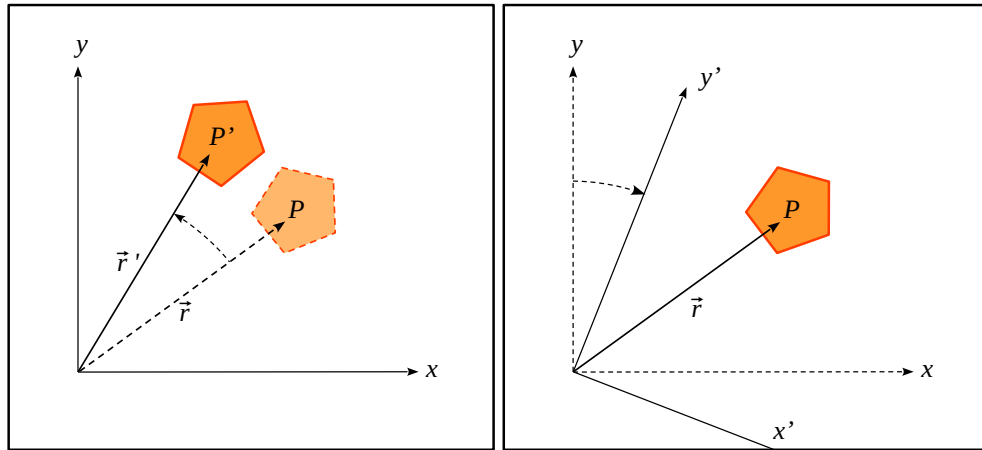


Figura 7.2.1: Punto de vista activo y pasivo de una transformación: en este caso ejemplificado para una rotación en el plano. El punto de vista activo (izquierda) interpreta la transformación como alterando la realidad física, cambiando los vectores. El punto de vista pasivo (derecha) lo entiende como un cambio en la descripción de la misma situación física.

2. La transformación actúa sobre el sistema de coordenadas, tomando el sistema donde el punto P está representado por el vector \vec{r} con coordenadas (x, y, z) y transformándolo en otro sistema donde el mismo punto P está representado por el mismo vector \vec{r} que en el nuevo sistema tiene coordenadas (x', y', z') . Este se denomina *punto de vista pasivo*.

En lo que sigue vamos a usar preferentemente el punto de vista activo, pero todo lo discutido se puede reformular en términos del punto de vista pasivo.

Las transformaciones *lineales* son aquellas que se pueden representar en términos de una matriz invertible M , la cual transforma un vector \vec{r} en otro vector \vec{r}' , según la fórmula

$$\vec{r}' = M \cdot \vec{r}$$

Aquí, para que la expresión de arriba tenga sentido como producto matricial, estamos considerando al vector \vec{r} como una matriz columna

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Si usáramos el punto de vista pasivo, la matriz cambiaría las coordenadas $(\vec{r}) = (x, y, z)$ del vector \vec{r} en el sistema de coordenadas original, por otras coordenadas $(\vec{r})' = (x', y', z')$ las cuales representan al mismo vector, pero en otro sistema de coordenadas. Nótese que usamos la notación en paréntesis (\vec{r}) para referirnos a las coordenadas del vector \vec{r} en un dado sistema. La operación sería exactamente la misma que escribimos arriba, sólo que en términos de la nueva notación

$$(\vec{r})' = M \cdot (\vec{r})$$

El conjunto de todas las transformaciones lineales que actúan sobre el espacio forma una estructura

matemática conocida como *grupo*. Para los fines de este curso, un grupo es un conjunto de operaciones que cumplen las siguientes tres condiciones

1. Propiedad de *clausura*: dadas dos transformaciones del conjunto, su aplicación sucesiva también resulta en una transformación del mismo conjunto.
2. Existencia de *identidad*: hay en el conjunto una transformación tal que al aplicarla antes o después de cualquier otra, no tiene ningún efecto.
3. Existencia de *inversa*: para cualquier transformación del conjunto existe otra tal que, aplicadas ambas en sucesión, no producen ningún efecto.

Esta definición puede parecer un tanto abstracta pero, como veremos inmediatamente, se vuelve muy simple cuando la aplicamos a los ejemplos concretos. Comencemos con el conjunto de las transformaciones lineales antes definidas, y probemos que ese conjunto forma un grupo. Para esto, tendremos que verificar que se cumplen estas tres propiedades

1. Dadas dos transformaciones lineales M_1 y M_2 , la aplicación sucesiva de M_2 y M_1 también es una transformación lineal. Para ver esto, tomamos un vector cualquiera \vec{r} y primero le aplicamos M_1 , definiendo \vec{r}' según

$$\vec{r}' = M_1 \cdot \vec{r}$$

Sobre este resultado \vec{r}' aplicamos subsecuentemente M_2 , definiendo un nuevo vector \vec{r}'' en la forma

$$\vec{r}'' = M_2 \cdot \vec{r}' = M_2 \cdot (M_1 \cdot \vec{r}) = \underbrace{(M_2 \cdot M_1)}_M \cdot \vec{r}$$

donde en la segunda igualdad usamos la asociatividad del producto de matrices para definir una nueva matriz M . En esta expresión, es evidente que si M_2 y M_1 son matrices invertibles, entonces M también lo es. De este modo tenemos una matriz invertible M que toma un vector \vec{r} y lo transforma en un nuevo vector \vec{r}'' , realizando una transformación lineal.

Esto prueba la propiedad de clausura: la aplicación sucesiva de dos transformaciones lineales resulta en una nueva transformación lineal.

2. Existe trivialmente una transformación identidad que deja cualquier vector invariante. Está dada por la matriz identidad I , que es una matriz invertible que aplicada a un vector cualquiera \vec{r} actúa según

$$I \cdot \vec{r} = \vec{r}$$

Es decir que deja el vector invariante. Esto implica que dada otra transformación cualquiera representada por una matriz M tenemos que

$$M \cdot (I \cdot \vec{r}) = M \cdot \vec{r} = I \cdot (M \cdot \vec{r})$$

Donde en la segunda igualdad usamos la propiedad de invarianza de más arriba, pero con el nuevo vector $\vec{r}' = M \cdot \vec{r}$.

Esto prueba que la transformación identidad no tiene ningún efecto aplicada antes o después de cualquier otra transformación.

3. Dada cualquier transformación lineal definida por la matriz M y cualquier vector \vec{r} , definimos

$$\vec{r}' = M \cdot \vec{r}$$

Dado que por definición la matriz M es invertible, podemos usar la matriz inversa M^{-1} para definir una nueva transformación lineal. Si la hacemos actuar sobre \vec{r}' obtenemos

$$M^{-1} \cdot \vec{r}' = M^{-1} \cdot (M \cdot \vec{r}) = \underbrace{(M^{-1} \cdot M)}_I \cdot \vec{r} = \vec{r}$$

donde en la segunda igualdad utilizamos la asociatividad del producto de matrices, y la propiedad de matriz inversa.

Con esto probamos que dada cualquier transformación, existe otra transformación (definida por la matriz inversa), tal que la aplicación sucesiva de ambas no tiene ningún efecto.

Es decir que el conjunto de las transformaciones lineales cumple las propiedades que definen un grupo. Este grupo se denomina *grupo general lineal tridimensional real* y en los textos de física se denota como $GL(3, \mathbb{R})$.

Cualquier subconjunto de transformaciones del espacio que cumpla las tres condiciones de arriba también es un grupo en sí mismo y un *subgrupo* del grupo general lineal tridimensional real.

Nuestra intuición es que las transformaciones que llamamos rotaciones deben encontrarse dentro del grupo $GL(3, \mathbb{R})$, es decir deben ser transformaciones lineales que se realizan multiplicando por una matriz. Para investigar esta cuestión, debemos definir con más precisión a qué llamamos una *rotación*. Utilizaremos las siguientes propiedades que todas las rotaciones cumplen:

- Seleccionamos un *eje de rotación* que pasa por el origen, que podemos representar con un vector unitario \hat{n} que apunta en la dirección del eje.
- Al realizar la rotación, el extremo de cada uno de los vectores \vec{r} del espacio se desplaza siguiendo un círculo que está en el plano perpendicular a \hat{n} a lo largo de un cierto *ángulo de rotación* θ .

Es evidente entonces que cada rotación deja invariante a su vector \hat{n} , y que por otro lado la longitud de cualquier vector \vec{r} no cambia bajo una rotación cualquiera. Esto nos permitirá estudiar en lo que sigue las condiciones que debe cumplir una matriz para representar una rotación.

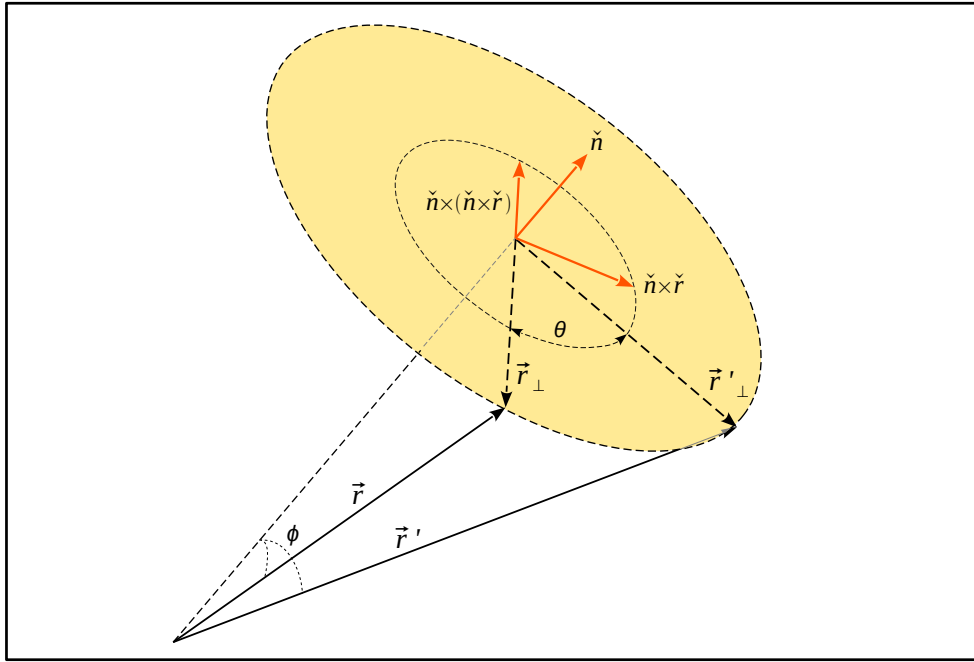


Figura 7.2.2: Rotación: una rotación está definida por un eje cuya dirección está dada por un versor \hat{n} , y un ángulo de rotación θ . Un vector cualquiera gira en torno al eje sin cambiar su módulo. Los versores \hat{n} , $\hat{n} \times \hat{r}$ y $\hat{n} \times (\hat{n} \times \hat{r})$ forman un sistema ortogonal.

Nota:

La teoría de grupos fue fundada por el matemático francés Évariste Galois, quien la escribió apresuradamente durante la noche previa a un duelo que le causó la muerte, con menos de 21 años. Galois fue un matemático brillante y un activista político, que fue expulsado de la École Normale por sus ideas republicanas. Se cree que la causa del duelo fue su amor por una mujer.

7.2.1 Transformaciones ortonogonales

Concentrémonos en las transformaciones que dejan invariante el módulo de los vectores. Sabemos que las rotaciones deben cumplir esta propiedad, por lo que se trata de un buen punto de partida para la discusión. Comencemos definiendo

$$\vec{r}' = R \cdot \vec{r}$$

e intentemos elucidar qué propiedad debe cumplir R para que se verifique que $|\vec{r}'| = |\vec{r}|$. Dado que el módulo de un vector satisface $|\vec{r}|^2 = \vec{r} \cdot \vec{r}$, podemos re-exresar la condición de invarianza como $\vec{r}' \cdot \vec{r}' = \vec{r} \cdot \vec{r}$. Para poder escribir esto en términos de matrices, recordemos que los vectores son matrices columna, con lo que escribiendo el producto escalar en forma matricial tenemos

$$\vec{r}'^t \cdot \vec{r}' = \vec{r}^t \cdot \vec{r}$$

Si ahora usamos que $\vec{r}' = R \cdot \vec{r}$ podemos reescribir el lado izquierdo de esta igualdad según

$$\vec{r}'^t \cdot \vec{r}' = (R \cdot \vec{r})^t \cdot (R \cdot \vec{r}) = \vec{r}^t \cdot (R^t \cdot R) \cdot \vec{r}$$

Con lo cual, para que se cumpla la condición de invarianza necesitamos que $\vec{r}^T \cdot (R^T \cdot R) \cdot \vec{r} = \vec{r}^T \cdot \vec{r}$, es decir que $R^T \cdot R$ no debe tener ningún efecto. Esto implica inmediatamente que

$$R^T \cdot R = I$$

Las matrices que cumplan esta condición dejarán el módulo de cualquier vector invariante, y viceversa, cualquier matriz que deje el módulo de un vector invariante debe cumplirla. Dado que las rotaciones dejan invariante el módulo de cualquier vector, sabemos que deben estar representadas por matrices de este tipo.

La condición de arriba puede reescribirse como $R^T = R^{-1}$, lo que quiere decir que nuestra matriz R es lo que se conoce como una *matriz ortogonal*. El conjunto de todas las matrices ortogonales tridimensionales se denota $O(3)$. Podemos probar que se trata de un grupo, ya que cumple las tres propiedades antes enunciadas, a saber

1. Dadas dos matrices ortogonales $R_1 \in O(3)$ y $R_2 \in O(3)$, se cumple que la composición también es ortogonal $R_2 \cdot R_1 \in O(3)$. En efecto, si escribimos

$$(R_2 \cdot R_1)^T \cdot (R_2 \cdot R_1) = R_1^T \cdot \underbrace{(R_2^T \cdot R_2)}_I \cdot R_1 = R_1^T \cdot R_1 = I$$

con lo que hemos probado que $R_2 \cdot R_1$ también es ortogonal, verificando la propiedad de clausura.

2. La identidad es trivialmente una matriz ortogonal $I \in O(3)$

$$I^T \cdot I = I \cdot I = I$$

Es decir que $O(3)$ contiene una transformación identidad.

3. Dada una matriz ortogonal $R \in O(3)$, se cumple que su inversa también lo es $R^{-1} \in O(3)$. Para verlo, escribimos

$$(R^{-1})^T \cdot R^{-1} = (R^T)^T \cdot R^T = R \cdot R^T = I$$

con lo que probamos que $O(3)$ contiene a la inversa de todas sus transformaciones.

Por lo tanto el conjunto de todas las matrices ortogonales $O(3)$ es un grupo, que se denomina *grupo ortogonal*.

Con lo que hemos discutido hasta este punto, hemos demostrado que las rotaciones deben ser transformaciones ortogonales. La pregunta que podemos hacernos ahora es si todas las transformaciones ortogonales son rotaciones.

Utilizando la propiedad definitoria de las matrices ortogonales, podemos escribir

$$\text{Det}(R^T \cdot R) = \text{Det} I = 1$$

Usando que el determinante de un producto de matrices es el producto de los determinantes de cada factor, y que el determinante de una matriz es igual al de su transpuesta, tenemos que

$$\text{Det}(R^T \cdot R) = \text{Det} R^T \text{Det} R = (\text{Det} R)^2$$

Por lo que se debe cumplir que $(\text{Det}R)^2 = 1$ o en otras palabras

$$\text{Det}R = \pm 1$$

Por lo tanto, el conjunto de las matrices ortogonales se separa en dos conjuntos disjuntos, aquél que contiene a las matrices con $\text{Det}R = 1$, que se conocen como *transformaciones propias o especiales*, y el que contiene a las matrices que tienen $\text{Det}R = -1$, que se denominan *transformaciones impropias*. Se puede probar que el conjunto de las matrices ortogonales propias o especiales forma un grupo, que se denota como $SO(3)$.

Ejercicio:

Probar que las transformaciones ortogonales propias o especiales forman un grupo, es decir que verifican que

1. Dadas dos transformaciones ortogonales propias o especiales, su producto también lo es
2. La identidad es una matriz ortogonal propia o especial.
3. Dada una matriz ortogonal propia o especial, su inversa también lo es.

Ejercicio:

Probar que las transformaciones ortogonales impropias no forman un grupo, porque no se cumple la primera de las tres propiedades enumeradas en el ejercicio anterior.

Una particularidad de esta división de $O(3)$ en dos subconjuntos, es que no podemos movernos continuamente de un sector al otro. En efecto, el determinante de una matriz R cualquiera es una suma de productos de los elementos de matriz, es decir un polinomio que tiene como variables las entradas de la matriz. Todo polinomio es una función continua de sus variables, que cambia suavemente al modificarlas. Por esta razón, no es posible deformar suavemente los elementos de matriz de R para provocar el necesario salto discreto que nos llevaría de una matriz propia $\text{Det}R = 1$ a una impropia $\text{Det}R = -1$.

Una consecuencia inmediata de lo anterior es que las rotaciones son transformaciones propias o especiales. Esto sucede porque, como comentamos más arriba, cualquier rotación está definida con un eje \tilde{n} y un ángulo de rotación θ . En particular, la rotación de ángulo nulo corresponde a la transformación identidad. Dado que el ángulo de rotación se puede deformar continuamente llevándolo desde cualquier valor finito hasta cero, cualquier rotación está continuamente *conectada con la identidad*. Como la identidad es una transformación propia, esto implica que cualquier rotación lo es.

Vamos a estudiar las transformaciones impropias. Un ejemplo de transformación impropia es la reflexión a lo largo del eje x , en la cual reemplazamos el vector $\vec{r} = (x, y, z)$ por un nuevo vector

$\vec{r}' = (x', y', z') = (-x, y, z)$. Esta transformación se conoce como *transformación de paridad en x*, y se representa según

$$\vec{r}' = P_x \cdot \vec{r}$$

Recordando que los vectores son matrices columna, tenemos que

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = P_x \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Por lo que vemos que P_x corresponde a la matriz

$$\begin{pmatrix} -x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P_x} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Es evidente de aquí que se cumple $\text{Det } P_x = -1$ por lo que P_x es una transformación impropia, como habíamos adelantado. Por otro lado, es fácil verificar que se trata en efecto de una transformación ortogonal, según

$$P_x^t \cdot P_x = P_x^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Podemos definir paridades en los otros ejes: P_y representa la reflexión a lo largo del eje y , y P_z la reflexión a lo largo del eje z . Estarán dadas por las matrices

$$P_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad P_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Nótese que cualquier producto binario $P_x \cdot P_y$ o $P_y \cdot P_z$ o $P_z \cdot P_x$ (y sus permutaciones) tendrá determinante 1, por lo que será una transformación propia. Por otro lado el producto triple $P_x \cdot P_y \cdot P_z = -I$ tiene determinante -1 y representa la transformación impropia.

Una transformación impropia cualquiera, es decir una que cumpla $\text{Det } R = -1$, se puede escribir en términos de una transformación propia R' con $\text{Det } R' = 1$ y la transformación de paridad P_x según $R = P_x \cdot R'$. Para demostrar esto, aprovechamos el hecho de que $P_x^2 = I$ para escribir

$$\text{Det } R = \text{Det}(P_x \cdot P_x \cdot R) = \text{Det}(P_x) \text{Det}(P_x \cdot R) = -\text{Det}(P_x \cdot R)$$

Dado que $\text{Det } R = -1$ tenemos de la ecuación de arriba que $\text{Det}(P_x \cdot R) = 1$. Por lo tanto si definimos $R' = P_x \cdot R$ podemos escribir

$$R = P_x \cdot R'$$

donde hemos aprovechado de nuevo el hecho de que $P_x^2 = I$. Es decir que cualquier transformación impropia se puede escribir como la paridad en x actuando sobre una transformación propia. Este último resultado implica que para estudiar el grupo de matrices ortogonales $O(3)$ podemos concentrarnos en el grupo de transformaciones propias o especiales $SO(3)$.

Nota:

Las transformaciones de paridad invierten izquierda y derecha. Eso es fácil de entender si contemplamos nuestras manos: la mano derecha es lo que se obtiene a partir de la mano izquierda si cambiamos el signo del eje que va a lo largo del pulgar.

Una pregunta que podemos hacernos es si la naturaleza es *invariante de paridad*. En otras palabras ¿qué sucede si realizamos dos experimentos idénticos, pero que son imágenes especulares el uno del otro? ¿obtendremos resultados idénticos, que son también imágenes reflejadas? Hoy sabemos que eso no sucede para las interacciones nucleares débiles. En otras palabras, la paridad no es una simetría de la naturaleza, que puede distinguir izquierda y derecha. Se dice que la simetría de paridad está *explícitamente rota*.

En otro contexto, tanto la ecuación de Schrödinger para los electrones como las ecuaciones de Maxwell para los campos electromagnéticos son invariantes de paridad. Esto implica en particular que toda la química, que obedece a las mencionadas ecuaciones, no distingue izquierda y derecha. Aquéllas moléculas que son diferentes de su imagen en un espejo deben producirse en ambas versiones (o *enantiómeros*) en cantidades iguales. Sin embargo, las reacciones químicas que originaron los primeros procesos biológicos se iniciaron, por azar, en moléculas que tenían alguna lateralidad preferida. Podrían haberse iniciado con igual probabilidad en moléculas del tipo reflejado, pero simplemente no fue así. Eso hace que la biología, que evolucionó a partir de ese acontecimiento contingente, distinga izquierda y derecha. Esto corresponde a una simetría de paridad *espontáneamente rota*.

El cuerpo humano transforma nuestros alimentos en *D-glucosa* o *dextrosa*, uno de los enantiómeros de la glucosa. El otro enantiómero, imagen en un espejo de la dextrosa, no tiene ningún valor nutricional. Es decir que al pasar “*A través del espejo*” Alicia hubiera encontrado, con gran probabilidad, la inanición y la muerte. Eso sucedió precisamente al protagonista del relato “*Error técnico*”, de Arthur C. Clarke.

7.2.2 Teorema de Euler

Hasta ahora, hemos probado que toda rotación es un elemento de $O(3)$, pero que no todo elemento de $O(3)$ es una rotación, dado que ese conjunto también contiene a las transformaciones impropias, que no están continuamente conectadas con la identidad. Ahora bien, si nos limitamos a $SO(3)$, compuesto por las matrices ortogonales propias o especiales ¿es correcto decir que todas ellas son rotaciones?

De acuerdo a nuestra definición de más arriba, además de no cambiar el módulo de cualquier vector, una rotación deja invariante la dirección del eje alrededor del cual estamos rotando. Es decir que para cualquier rotación R debe existir un vector \hat{n} tal que

$$R \cdot \hat{n} = \hat{n}$$

Vamos a probar que tal vector existe para todas las matrices de $SO(3)$. Tenemos que una matriz ortogonal cualquiera cumple la propiedad

$$(R - I) \cdot R^t = I - R^t$$

Tomando el determinante a ambos lados, y usando que el determinante de un producto es el producto de los determinantes y que el determinante de una matriz es igual al de su transpuesta, tenemos que para cualquier matriz especial

$$\text{Det}(R - I) = \text{Det}(I - R^t)$$

Usando ahora que $\text{Det}(I - R^t) = \text{Det}((I - R)^t) = \text{Det}(I - R)$ podemos escribir

$$\text{Det}(R - I) = -\text{Det}(R - I)$$

Hemos usado aquí que el determinante de una matriz es $(-1)^d$ veces el determinante de su opuesta, donde d es la dimensión de la matriz, por lo que en nuestro caso tenemos $d = 3$. Esto implica que

$$\text{Det}(R - I) = 0$$

En otras palabras la matriz $R - I$ es singular, por lo que debe existir un vector \hat{n} que cumpla

$$(R - I) \cdot \hat{n} = 0$$

con lo que hemos demostrado que para cualquier matriz ortogonal propia existe un vector invariante $R \cdot \hat{n} = \hat{n}$.

Con esto, hemos demostrado que cualquier transformación de $SO(3)$ deja invariante el módulo de todos los vectores, y tiene un vector invariante. En otras palabras, toda matriz de $SO(3)$ es una rotación.

La pregunta que podemos hacernos ahora es ¿se puede escribir la matriz que representa una dada rotación en términos del eje de rotación \hat{n} y del ángulo de rotación θ ? Para responderla, necesitamos un sistema de ejes ortogonales en el que descomponer nuestros vectores. Notando que para cualquier vector \vec{r} se cumple que \hat{n} es perpendicular a $\hat{n} \times \vec{r}$ y que ambos son perpendiculares a $\hat{n} \times (\hat{n} \times \vec{r})$, podemos construir un sistema de ejes perpendiculares en la dirección de esos tres vectores. Con esto, podemos dibujar el diagrama de la figura 7.2.2.

En el dibujo vemos varios vectores que tienen el mismo módulo, a saber $|\vec{r}_\perp| = |\vec{r}'_\perp| = |\hat{n} \times \vec{r}|r = |\hat{n} \times (\hat{n} \times \vec{r})|r$. Por otro lado, estos vectores son todos perpendiculares a \hat{n} . Esto implica que es

correcto dibujarlos en el mismo círculo. Si descomponemos entonces el vector transformado \vec{r}' en nuestros tres ejes, según

$$\vec{r}' = \alpha \vec{n} + \underbrace{\beta (\vec{n} \times \vec{r}) + \gamma \vec{n} \times (\vec{n} \times \vec{r})}_{\vec{r}'_{\perp}}$$

en términos de coeficientes α , β y γ , vemos que podemos leer la forma de estos coeficientes directamente del dibujo, según

$$\alpha = \vec{r} \cdot \vec{n} = \vec{r}' \cdot \vec{n} \quad \beta = \sin \theta \quad \gamma = -\cos \theta$$

Con esto, podemos escribir explícitamente

$$\vec{r}' = (\vec{r} \cdot \vec{n}) \vec{n} + \sin \theta \vec{n} \times \vec{r} - \cos \theta \vec{n} \times (\vec{n} \times \vec{r})$$

Ahora usamos la identidad vectorial $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{C}(\vec{A} \cdot \vec{B}) - \vec{B}(\vec{A} \cdot \vec{C})$ que nos permite reescribir el último término en esta fórmula para obtener

$$\vec{r}' = (\vec{r} \cdot \vec{n}) \vec{n} + \sin \theta \vec{n} \times \vec{r} + \cos \theta (\vec{r} - \vec{n}(\vec{n} \cdot \vec{r}))$$

Si llamamos eje z a la dirección del versor \vec{n} , tenemos que $\vec{n} = \hat{k}$. Escribiendo $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ nos queda

$$\vec{r}' = z\hat{k} + (\cos \theta x + \sin \theta y)\hat{i} + (-\sin \theta x + \cos \theta y)\hat{j} = R_z \cdot \vec{r}$$

lo que puede ser expresado matricialmente como

$$\vec{r}' = \begin{pmatrix} \cos \theta x + \sin \theta y \\ -\sin \theta x + \cos \theta y \\ z \end{pmatrix} = R_z \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Esto nos permite leer la matriz que representa la rotación de ángulo θ alrededor del eje z en la forma

$$R_z = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Es fácil probar que $\text{Det } R_z = \cos^2 \theta + \sin^2 \theta = 1$ y también que $R_z^t \cdot R_z = I$, confirmando que esta rotación es una matriz de $SO(3)$.

Ejercicio:

Pruebe que para rotaciones alrededor del eje x y del eje y las matrices son R_x y R_y respectivamente, dadas por

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \quad R_y = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}$$

Verifique que en ambos casos las matrices son ortogonales y su determinante vale 1.

Ejercicio:

Demuestre que la matriz $R_{\vec{n}}$ definida según

$$R_{\vec{n}} = \begin{pmatrix} \cos \theta + n_x^2(1 - \cos \theta) & n_x n_y(1 - \cos \theta) + n_z \sin \theta & n_x n_z(1 - \cos \theta) - n_y \sin \theta \\ n_x n_y(1 - \cos \theta) - n_z \sin \theta & \cos \theta + n_y^2(1 - \cos \theta) & n_y n_z(1 - \cos \theta) + n_x \sin \theta \\ n_x n_z(1 - \cos \theta) + n_y \sin \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta & \cos \theta + n_z^2(1 - \cos \theta) \end{pmatrix}$$

representa la rotación en ángulo θ alrededor del eje \vec{n} dada por

$$\vec{r}' = (\vec{r} \cdot \vec{n})\vec{n} + \sin \theta \vec{n} \times \vec{r} + \cos \theta (\vec{r} - \vec{n}(\vec{n} \cdot \vec{r})) = R_{\vec{n}} \cdot \vec{r}$$

y verifique que $\text{Det } R_{\vec{n}} = 1$ y que $R_{\vec{n}}^t \cdot R_{\vec{n}} = I$.

Sin embargo, la forma general de una rotación de eje \vec{n} que hemos presentado en el ejercicio anterior no es la única manera de representar una rotación cualquiera. Una forma alternativa que resulta muy útil para nuestros fines es la que se obtiene según la siguiente construcción:

1. Primero se realiza una rotación en un ángulo ϑ alrededor del eje z . Con esto, los ejes x, y, z se mueven, yendo a parar a un nuevo conjunto de ejes x', y', z . Nótese que, como se trató de una rotación alrededor del eje z , este quedó inalterado.
2. Luego se rota un ángulo ϕ alrededor del eje x' resultante de la rotación anterior, obteniendo un nuevo conjunto de ejes x', y'', z'' .
3. Finalmente se rota un ángulo ψ alrededor del eje z'' resultante de ambas rotaciones.

Esto implica que hemos descompuesto la rotación R en términos de tres rotaciones elementales alrededor de z , de x' , y de z'' . En términos matriciales

$$R = R_{z''} \cdot R_{x'} \cdot R_z$$

Usando la forma explícita de las rotaciones alrededor de cada eje que escribimos más arriba, tenemos que

$$R = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix} \cdot \begin{pmatrix} \cos \vartheta & \sin \vartheta & 0 \\ -\sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Es fácil convencerse que mediante este proceso se puede descomponer cualquier rotación. Los ángulos ϑ, ϕ, ψ se conocen como los *ángulos de Euler* de la rotación.

Nota:

Los ángulos de Euler resultan muy útiles en navegación marítima y aérea, ya que permiten describir la posición de un vehículo respecto del giróscopo de a bordo cuya orientación se mantiene fija.

Como se puede ver en la figura 7.2.4, el ángulo ϑ representa la *dirección*, entendida como hacia adónde apunta el vehículo en el mapa, el ángulo ϕ especifica el *cabeceo*,

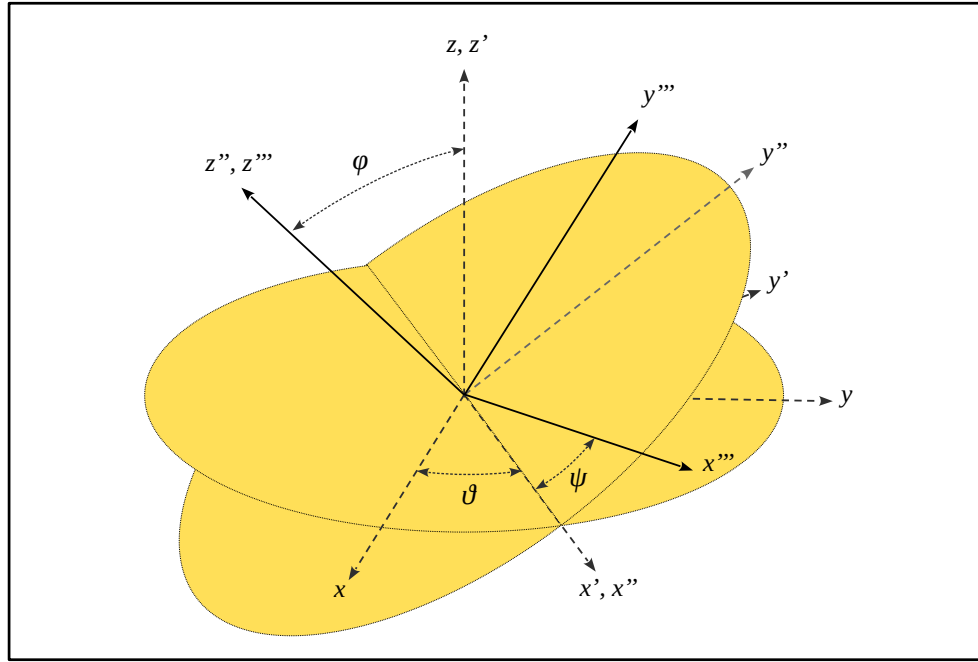


Figura 7.2.3: Ángulos de Euler: una rotación de ángulo ϑ en torno al eje z transforma el sistema de coordenadas x, y, z en el nuevo sistema x', y', z' . A partir de éste, rotamos alrededor del eje x' un ángulo φ para ir al nuevo sistema x'', y'', z'' . Finalmente, una rotación de ángulo ψ en torno al eje z'' nos lleva al sistema de coordenadas rotado final x''', y''', z''' .

es decir si el vehículo se encuentra subiendo o bajando, y finalmente el ángulo ψ corresponde a la *guiñada*, que se refiere a su inclinación lateral respecto de la línea del horizonte.

7.2.3 Rotaciones infinitesimales

Dijimos más arriba que toda rotación está continuamente conectada a la identidad, es decir que podemos acercarla arbitrariamente a la misma tomando el ángulo de rotación cada vez más pequeño. Imaginemos entonces una rotación con un ángulo muy pequeño alrededor de un eje \tilde{n} cualquiera. La podemos escribir como una matriz que está infinitesimalmente cerca de la identidad, en la forma

$$R = I + \varepsilon t_{\tilde{n}}$$

donde $t_{\tilde{n}}$ es una matriz a determinar, y ε es un número muy pequeño que nos asegura que la desviación de la identidad es infinitesimal. Entonces podemos escribir su determinante en potencias de los elementos de la matriz $\varepsilon t_{\tilde{n}}$, según

$$\text{Det } R = 1 + \varepsilon \text{Tr } t_{\tilde{n}} + \mathcal{O}(\varepsilon^2)$$

Ejercicio:

Pruebe la fórmula de arriba escribiendo explícitamente la matriz y calculando su determinante, tirando cuadrados de ε .

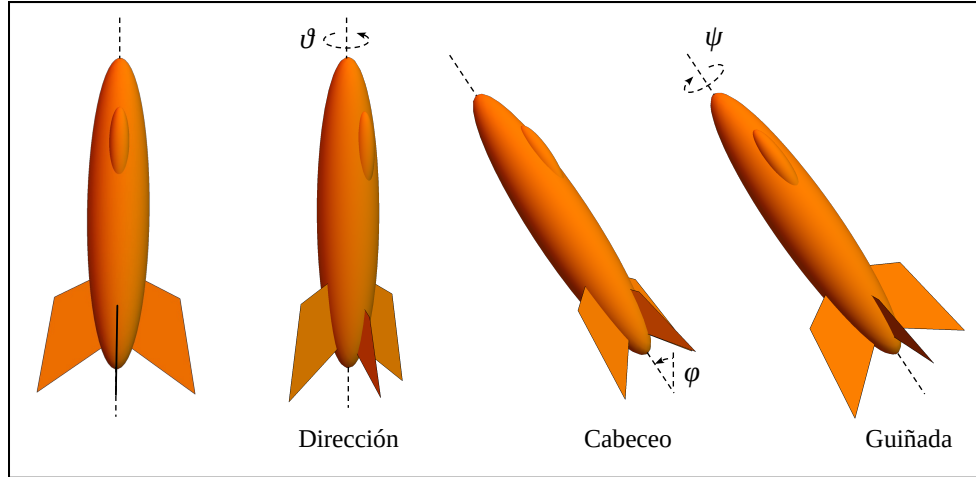


Figura 7.2.4: Ángulos de Euler en aeronáutica. El primer ángulo de Euler ϑ especifica la dirección de navegación, el segundo φ nos dice si la aeronave sube o baja, mientras que el tercero ψ nos habla de la inclinación de sus alas.

Por lo tanto la condición de que la matriz sea especial $\text{Det } R = 1$ implica para que matriz infinitesimal debe cumplir $\text{Tr } t_{\tilde{n}} = 0$. En otras palabras, las matrices especiales cercanas a la identidad difieren de la ella en una matriz de traza nula.

Por otro lado, sabemos que las rotaciones son matrices ortogonales, es decir que cumplen que $R^t \cdot R = I$. Para nuestra rotación cercana a la identidad esto implica

$$(I + \varepsilon t_{\tilde{n}})^t \cdot (I + \varepsilon t_{\tilde{n}}) = I$$

Escribiendo $(I + \varepsilon t_{\tilde{n}})^t = I + \varepsilon t_{\tilde{n}}^t$ y expandiendo la última expresión tenemos

$$(I + \varepsilon t_{\tilde{n}}^t) \cdot (I + \varepsilon t_{\tilde{n}}) = I + \varepsilon (t_{\tilde{n}}^t + t_{\tilde{n}}) + \mathcal{O}(\varepsilon^2) = I$$

Por lo tanto la condición $R^t \cdot R = I$ implica que $t_{\tilde{n}}^t + t_{\tilde{n}} = 0$, o en otras palabras

$$t_{\tilde{n}}^t = -t_{\tilde{n}}$$

Es decir que $t_{\tilde{n}}$ es una matriz antisimétrica. Las matrices ortogonales cercanas a la identidad difieren de ella en una matriz antisimétrica.

Una matriz antisimétrica general se puede escribir como

$$t_{\tilde{n}} = \begin{pmatrix} 0 & -\gamma & \nu \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} = \underbrace{\alpha \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{t_x} + \underbrace{\beta \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}}_{t_y} + \underbrace{\gamma \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{t_z}$$

En términos de coeficientes arbitrarios α, β, γ y las matrices t_a con $a \in \{x, y, z\}$, que se denominan *generadores del grupo de rotaciones*. Por lo tanto tenemos que

$$R = I + \alpha t_x + \beta t_y + \gamma t_z$$

Esta propiedad se resume diciendo que el grupo de rotaciones es un *grupo de Lie*, es decir un grupo que contiene transformaciones infinitesimales arbitrariamente cercanas a la identidad que se pueden escribir en términos de un conjunto de generadores.

Ejercicio: reglas de conmutación

Probar, escribiendo explícitamente las matrices y haciendo el cálculo, que se cumplen las *reglas de conmutación*

$$[t_x, t_y] = t_x \cdot t_y - t_y \cdot t_x = t_z$$

$$[t_y, t_z] = t_y \cdot t_z - t_z \cdot t_y = t_x$$

$$[t_z, t_x] = t_z \cdot t_x - t_x \cdot t_z = t_y$$

Las reglas de conmutación de este ejercicio se pueden resumir en la expresión

$$[t_a, t_b] = \varepsilon_{abc} t_c$$

donde ε_{abc} es el *símbolo de Levi-Civita completamente antisimétrico*, que vale $\varepsilon_{xyz} = 1$ y los demás se obtienen intercambiando los índices y multiplicando por -1 en cada intercambio. La expresión del conmutador se conoce como el *álgebra de Lie* del grupo de rotaciones. Se puede verificar que las componentes de cada uno de los generadores cumplen

$$(t_a)_{bc} = -\varepsilon_{abc}$$

Podríamos preguntarnos cómo se relaciona todo esto con nuestra expresión general para las rotaciones con un ángulo θ alrededor de un eje \hat{n} que habíamos escrito más arriba, es decir

$$\vec{r}' = (\vec{r} \cdot \hat{n}) \hat{n} + \sin \theta \hat{n} \times \vec{r} + \cos \theta (\vec{r} - \hat{n}(\hat{n} \cdot \vec{r}))$$

Si imaginamos una rotación infinitesimal con un ángulo ε , podemos desarrollar $\sin(\varepsilon) = \varepsilon$ y $\cos(\varepsilon) = 1 + \mathcal{O}(\varepsilon^2)$, con lo que si ahora reemplazamos en la ecuación anterior, obtenemos

$$\vec{r}' = \vec{r} + \varepsilon \hat{n} \times \vec{r}$$

o bien, escribiendo esta expresión en componentes,

$$r'_a = r_a - \varepsilon n_b \varepsilon_{bac} r_c = r_a + \varepsilon n_b (t_b)_{ac} r_c = (\delta_{ac} + \varepsilon n_b (t_b)_{ac}) r_c$$

Lo que en términos matriciales se puede resumir como

$$\vec{r}' = (I + \varepsilon \hat{n} \cdot \vec{t}) \vec{r} = (I + \varepsilon t_{\hat{n}}) \cdot \vec{r}$$

donde $\vec{t} = (t_x, t_y, t_z)$ es un vector cuyas componentes son matrices (esto es simplemente una notación para no tener que escribir índices). Si ahora definimos $\vec{\varepsilon} = \hat{n} \varepsilon$ obtenemos la fórmula

$$d\vec{r} = \vec{\varepsilon} \times \vec{r} = (\vec{\varepsilon} \cdot \vec{t}) \cdot \vec{r} = t_{\hat{n}} \cdot \vec{r}$$

La primera igualdad toma una forma conocida cuando escribimos $d\theta = \varepsilon$ y dividimos por dt

$$\dot{\vec{r}} = \vec{\omega} \times \vec{r}$$

Vamos a aplicar las expresiones obtenidas al cálculo de la aceleración de Coriolis en un sistema de coordenadas rotante. Para esto, es natural trabajar usando el punto de vista pasivo de las rotaciones, donde lo que está cambiando no es el vector que las coordenadas representan, sino el sistema de coordenadas en el cual lo descomponemos.

Supongamos que las componentes de un vector cualquiera \vec{G} en un sistema de coordenadas fijo respecto de las estrellas lejanas se llaman $(\vec{G})_S$, mientras que sus componentes en un sistema de coordenadas que rota solidario con la Tierra son $(\vec{G})_T$. El origen de ambos sistemas está en el centro de la Tierra, y suponemos que inicialmente los ejes de ambos sistemas están alineados. Tenemos entonces que, en el instante inicial

$$(\vec{G})_S = (\vec{G})_T$$

Al dejar pasar un pequeño intervalo de tiempo dt , el sistema solidario con la Tierra ha rotado un ángulo $-d\Theta$ respecto del sistema alineado con las estrellas lejanas, y además el vector \vec{G} ha cambiado una cantidad $d\vec{G}$ debido a su movimiento propio, por lo que en lugar de la igualdad de más arriba tendremos que

$$(\vec{G} + d\vec{G})_S = (I + d\vec{\Theta} \cdot \vec{r}) \cdot (\vec{G} + d\vec{G})_T$$

De donde podemos escribir

$$(\vec{G})_S + d(\vec{G})_S = (\vec{G})_T + d(\vec{G})_T + (d\vec{\Theta} \cdot \vec{r}) \cdot (\vec{G})_T + \dots$$

donde hemos escrito $d(\vec{G})_T = d\vec{G}_T$ descartando términos de orden cuadrático en las variaciones. Usando una de las identidades de más arriba, podemos poner

$$d(\vec{G})_S = d(\vec{G})_T + d\vec{\Theta} \times (\vec{G})_T$$

donde el primer término representa el cambio intrínseco en las componentes del vector, mientras que el segundo se refiere al cambio debido a la rotación del sistema de coordenadas fijo a la Tierra. Dividiendo por dt obtenemos

$$\frac{d(\vec{G})_S}{dt} = \frac{d(\vec{G})_T}{dt} + \vec{\Omega} \times (\vec{G})_T$$

Nótese que $d(\vec{G})_T/dt \neq (d\vec{G}/dt)_T$ debido a la rotación propia del sistema solidario con la Tierra.

Aplicando este resultado al vector posición $\vec{G} = \vec{r}$, tenemos

$$(\vec{v})_S = \frac{d(\vec{r})_S}{dt} = \frac{d(\vec{r})_T}{dt} + \vec{\Omega} \times (\vec{r})_T$$

Nótese que $d(\vec{r})_T/dt \neq (\vec{v})_T$. Si ahora pensamos que $\vec{G} = \vec{v}$ podemos escribir

$$(\vec{a})_S = \frac{d(\vec{v})_S}{dt} = \frac{d}{dt} \left(\frac{d(\vec{r})_T}{dt} + \vec{\Omega} \times (\vec{r})_T \right)_S$$

donde hemos puesto un subíndice S para recordar que la magnitud entre paréntesis corresponde a las componentes $(\vec{v})_S$ del vector velocidad en el sistema en reposo respecto de las estrellas lejanas. Aplicando nuestra fórmula para la derivada temporal

$$(\vec{a})_S = \frac{d}{dt} \left(\frac{d(\vec{r})_T}{dt} + \vec{\Omega} \times (\vec{r})_T \right)_T + \vec{\Omega} \times \left(\frac{d(\vec{r})_T}{dt} + \vec{\Omega} \times (\vec{r})_T \right)_T$$

y reordenando obtenemos

$$(\vec{a})_S = \frac{d^2(\vec{r})_T}{dt^2} + 2\vec{\Omega} \times \frac{d(\vec{r})_T}{dt} + \vec{\Omega} \times (\vec{\Omega} \times (\vec{r})_T)$$

Aquí el primer término es la aceleración en el sistema solidario con la Tierra, el segundo término nos da la *aceleración de Coriolis*, mientras que el tercero corresponde a la aceleración centrípeta.

7.3 Resumen

En esta clase introdujimos la idea de grupo, comenzando con el grupo $SL(3, \mathbb{R})$ de transformaciones generales lineales en el espacio. Identificamos su subgrupo ortogonal $O(3)$ dado por las transformaciones que preservan el módulo de los vectores. Vimos que en términos de matrices, corresponden a aquellas cuya transpuesta es a la vez su inversa. Identificamos el subgrupo de matrices especiales $SO(3)$ que tienen determinante unidad, y demostramos que corresponden a las rotaciones.

Aprendimos también a descomponer una rotación en términos de rotaciones sucesivas alrededor de diferentes ejes, con lo que pudimos definir los ángulos de Euler.

Finalmente, estudiamos las características de las rotaciones infinitesimales, y las utilizamos para describir la aceleración de Coriolis.

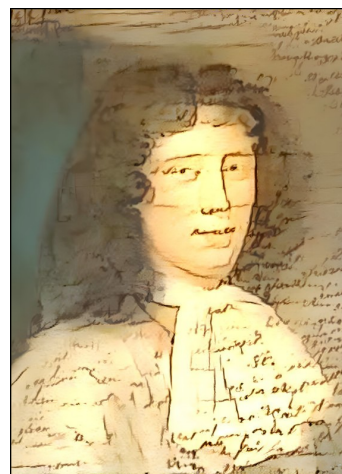
8. Cuerpo rígido II: Cinemática e inercia

8.1 Objetivos

En esta clase definiremos la noción de cuerpo rígido, como un sistema de partículas cuyas distancias relativas están fijas. Dejaremos para más adelante la discusión sobre cuáles son las fuerzas que aseguran tal rigidez, y sobre qué sucede cuando estas fuerzas se relajan.

Comenzaremos con la cinemática del cuerpo rígido, identificando los elementos necesarios para poder más adelante describir su evolución temporal.

Luego analizaremos la manera en la que la inercia se manifiesta en estos sistemas, construyendo otro de los ingredientes indispensables para escribir sus ecuaciones de movimiento.



Christiaan Huygens

8.2 Cinemática del cuerpo rígido

Supongamos que tenemos un sistema de N partículas, El conjunto formado por las tres componentes de cada uno de los vectores posición \vec{r}_n constituye naturalmente un sistema de $3N$ coordenadas generalizadas. A continuación vamos a construir otro sistema de coordenadas generalizadas para describir su movimiento, que nos resultará útil para definir un cuerpo rígido.

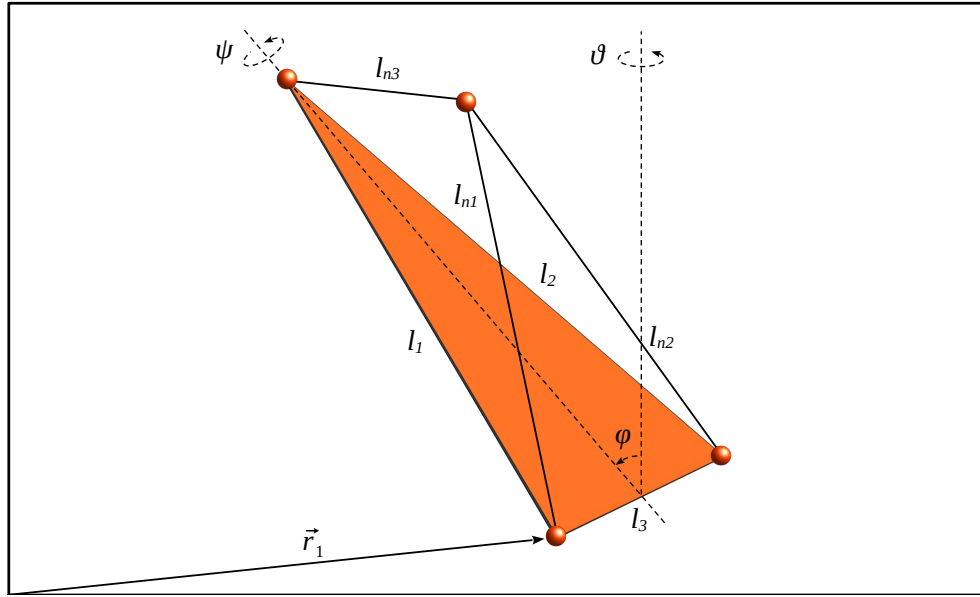


Figura 8.2.1: Coordenadas para la descripción de un cuerpo rígido. Usamos la posición de una de las partículas, la orientación del triángulo formado por tres partículas respecto de una orientación de referencia, y la distancia de cualquier otra partícula a los vértices del triángulo.

Tomamos tres partículas cualesquiera no colineales $n = 1, 2, 3$. Estas partículas forman por supuesto un triángulo. La posición de las tres partículas estará completamente determinada por la forma, la posición y la orientación de este triángulo, ver figura 8.2.1.

- La forma del triángulo está dada por las longitudes de cada uno de sus lados l_1 , l_2 y l_3 . Podemos escribir

$$l_1 = |\vec{r}_1 - \vec{r}_2| \quad l_2 = |\vec{r}_2 - \vec{r}_3| \quad l_3 = |\vec{r}_3 - \vec{r}_1|$$

- La posición del triángulo queda determinada por la posición de uno de sus vértices, digamos por ejemplo \vec{r}_1 .
- La orientación del triángulo se fija en términos de la rotación necesaria para llevar al triángulo a su posición actual comenzando desde una orientación de referencia prefijada (por ejemplo, la orientación de referencia podría ser el triángulo apoyado en el primer cuadrante del plano xy con su lado más largo alineado en la dirección de las x). Esta rotación está definida por tres ángulos, que podemos identificar con sus ángulos de Euler (ϑ , φ , ψ).

La posición de cualquier otra partícula $n > 3$ del sistema estará dada por las distancias l_{n1} , l_{n2} , l_{n3} que la separan de los tres vértices del triángulo.

$$l_{n1} = |\vec{r}_1 - \vec{r}_n| \quad l_{n2} = |\vec{r}_2 - \vec{r}_n| \quad l_{n3} = |\vec{r}_3 - \vec{r}_n|$$

Con esto, nuestro conjunto de coordenadas generalizadas $\{q_i\}$ con $i \in \{1, 2, \dots, 3N\}$ está dado por $\{q_i\} = \{\vec{r}_1, \vartheta, \varphi, \psi, l_1, l_2, l_3, l_{n1}, l_{n2}, l_{n3}\}$ para $n \in \{4, \dots, N\}$. Son tres componentes de la posición

de un vértice, los tres ángulos de Euler que fijan la rotación, y $3 + 3(N - 3) = 3N - 6$ distancias. Comprobamos que es un total de $3N$ coordenadas, que parametrizan el espacio de configuración $\mathcal{C} = \mathbb{R}^{3N}$.

Un *cuerpo rígido* es un sistema de partículas tal que la distancia entre cualquier par de partículas es constante. Por lo tanto, las $3N - 6$ distancias mencionadas más arriba no cambiarán con el tiempo, y podemos dejarlas fijas en nuestra descripción del sistema. Esta truncación puede involucrar alguna sutileza, ya que las coordenadas generalizadas que omitimos también tienen sus ecuaciones de movimiento, y deberíamos asegurarnos que tales ecuaciones se están cumpliendo, es decir que se trate de una truncación consistente. Postpondremos esa discusión para más adelante.

Como conclusión, un cuerpo rígido tiene seis grados de libertad, dados por la posición de una de sus partículas $\vec{r}_1 \in \mathbb{R}^3$ y los tres ángulos de Euler θ, ϕ, ψ que definen la rotación $R \in SO(3)$ que lleva de un sistema de coordenadas inercial externo a un sistema de coordenadas unido al cuerpo. Es decir que su espacio de configuración está dado por $\mathcal{C} = \mathbb{R}^3 \times SO(3)$.

Una partícula cualquiera estará en una posición dada por

$$\vec{r}_n = \vec{r}_1 + \underbrace{(\vec{r}_n - \vec{r}_1)}_{\vec{r}_{1n}} = \vec{r}_1 + \vec{r}_{1n}$$

donde \vec{r}_{1n} es el vector distancia entre la partícula n -ésima y la partícula que usamos para describir la posición del cuerpo rígido. Ahora bien, estas variables cambiarán al transcurrir un instante de tiempo dt , de modo que

$$d\vec{r}_n = d\vec{r}_1 + d\vec{r}_{1n}$$

Dado que por definición las distancias entre cualquier par de partículas de un cuerpo rígido no cambian, entonces $|\vec{r}_{1n}|$ no puede cambiar, por lo que $d\vec{r}_{1n}$ debe corresponder a una rotación. Usando lo que aprendimos en la clase anterior podemos entonces escribir

$$d\vec{r}_n = d\vec{r}_1 + d\theta \, \hat{n} \times \vec{r}_{1n}$$

donde \hat{n} es un eje de rotación que pasa por \vec{r}_1 . Con esto podemos escribir la velocidad de la partícula n -ésima como

$$\dot{\vec{r}}_n = \dot{\vec{r}}_1 + \vec{\omega} \times \vec{r}_{1n}$$

donde hemos definido la velocidad angular en torno a \vec{r}_1 como $\vec{\omega} = \hat{n} d\theta/dt$. En esta expresión, el primer término representa la traslación de una de las partículas del cuerpo rígido respecto de origen del sistema de coordenadas, y el segundo término tiene en cuenta la rotación del cuerpo en torno a esa partícula.

8.3 Momento de inercia

Supongamos que queremos escribir el momento angular del cuerpo rígido. Para esto tenemos que escribir

$$\vec{\ell} = m_n \vec{r}_n \times \dot{\vec{r}}_n$$

donde n es el índice que corre entre todas las partículas del cuerpo, y estamos sumando sobre él. Por el momento, pondremos el origen de nuestro sistema de coordenadas en la partícula 1, de modo que $\vec{r}_1 = 0$. Si se trata o no de un sistema de coordenadas inercial no es importante para lo que discutiremos, por lo que pospondremos esa discusión. Podemos entonces usar la expresión para la velocidad de más arriba con $\dot{\vec{r}}_1 = 0$ y obtener

$$\vec{\ell} = m_n \vec{r}_n \times (\vec{\omega} \times \vec{r}_n) = m_n (\vec{\omega} r_n^2 - \vec{r}_n (\vec{\omega} \cdot \vec{r}_n))$$

Donde en la segunda igualdad usamos la fórmula para un producto vectorial triple $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$. En componentes esta ecuación se puede escribir como

$$\ell_a = m_n (\omega_a r_n^2 - r_a^n \omega_b r_b^n)$$

Donde los índices a, b corren sobre las direcciones cartesianas $\{x, y, z\}$, y en el segundo término hay una suma sobre b . Para poder sacar factor común ω_b insertamos una δ_{ab} en el primer término, y tenemos

$$\ell_a = m_n (\delta_{ab} \omega_b r_n^2 - r_a^n \omega_b r_b^n) = \underbrace{m_n (\delta_{ab} r_n^2 - r_a^n r_b^n)}_{\mathcal{J}_{ab}} \omega_b = \mathcal{J}_{ab} \omega_b$$

Hemos definido aquí las nuevas magnitudes \mathcal{J}_{ab} , que forman una matriz \mathcal{J} llamada *matriz momento de inercia*. Esto nos permite escribir la relación en la forma matricial

$$\vec{\ell} = \mathcal{J} \cdot \vec{\omega}$$

o bien explícitamente en componentes

$$\ell_x = \mathcal{J}_{xx} \omega_x + \mathcal{J}_{xy} \omega_y + \mathcal{J}_{xz} \omega_z$$

$$\ell_y = \mathcal{J}_{yx} \omega_x + \mathcal{J}_{yy} \omega_y + \mathcal{J}_{yz} \omega_z$$

$$\ell_z = \mathcal{J}_{zx} \omega_x + \mathcal{J}_{zy} \omega_y + \mathcal{J}_{zz} \omega_z$$

Nótese que la matriz \mathcal{J} es una matriz simétrica, de acuerdo a su definición. Ahora bien ¿qué es exactamente esta matriz?

Ejercicio:

Calcule las componentes de la matriz anterior para una molécula de metano, representada como un tetraedro con un átomo de hidrógeno en cada vértice y un átomo de carbono en el centro. Elija un sistema de ejes que con su origen el centro del tetraedro y tal que el eje z pase por uno de los vértices ¿Qué sucede si cambia el sistema de coordenadas?

Nota:

La película “2001: una odisea espacial” del director Stanley Kubrick está basada en la novela homónima de Arthur C. Clarke. Una de las escenas icónicas de la película muestra al astronauta David Bowman trotando alrededor del un pasillo circular que rodea completamente la nave *Discovery*. El pasillo gira sobre su eje de modo tal que la fuerza centrífuga provee de un sustituto para la gravedad.

Luego de que la computadora de la nave Hal-9000, con su célebre ojo único de color rojo, se volviera loca y asesinara a su compañero Frank Poole, Bowman desaparece en un portal dimensional que orbita Júpiter. En la novela de Clarke, la narración sucede en órbita de Saturno, pero Kubrick temía no poder retratar satisfactoriamente los anillos, por eso la trasladó al otro gigante gaseoso.

En la secuela de la novela “2010: odisea dos”, la nave rusa *Leonov* llega a Saturno para rescatar los restos de la *Discovery*, pero la encuentra girando descontroladamente alrededor de su centro de masas. El mecanismo que movía el pasillo se atascó, y todo el momento angular del mismo se transfirió a la nave entera.

Con esto, el autor retrata uno de los problemas de construir mecanismos con piezas en rotación relativa en el espacio: si por alguna razón el mecanismo falla, el objeto entero comienza a girar.

8.3.1 Tensor momento de inercia

Aprendimos en los cursos básicos de física que existen magnitudes físicas conocidas como *vectores*, un ejemplo de las cuales es el momento angular $\vec{\ell}$. Estos vectores están definidos con una terna de números $\vec{\ell} = (\ell_x, \ell_y, \ell_z)$ que nos dan sus componentes en un dado sistema de coordenadas. Además, satisfacen una propiedad fundamental que es su regla de transformación frente a rotaciones

$$\vec{\ell}' = R \cdot \vec{\ell}$$

lo que se puede escribir en componentes como

$$\ell_a = R_{ab} \ell_b$$

Cualquier terna de números que se transforma de esa manera frente a las rotaciones se llama *vector contravariante*. Existen también los *vectores covariantes* que transforman con la matriz transpuesta, un ejemplo de ellos es el gradiente $\vec{\nabla} = (\partial_x, \partial_y, \partial_z)$ que cumple la regla de transformación

$$\partial'_a = R_{ba} \partial_b$$

Ejercicio:

Probar la regla de transformación anterior haciendo actuar el gradiente sobre una función arbitraria y utilizando la regla de la cadena.

Uno podría preguntarse entonces ¿cómo se transforman la matriz de 3×3 que llamamos más arriba momento de inercia? Para investigar esta cuestión, aplicamos una rotación al momento angular, obteniendo

$$\vec{\ell}' = R \cdot \vec{\ell} = R \cdot (\mathcal{I} \cdot \vec{\omega})$$

En el miembro derecho podemos insertar la identidad escrita como $I = R^t \cdot R$ y nos queda

$$\vec{\ell}' = R \cdot \vec{\ell} = R \cdot \mathcal{I} \cdot (R^t \cdot R) \cdot \vec{\omega} = \underbrace{(R \cdot \mathcal{I} \cdot R^t)}_{\mathcal{I}'} \cdot \underbrace{R \cdot \vec{\omega}}_{\vec{\omega}'}$$

Con las definiciones de la última igualdad, podemos escribir la relación entre los vectores rotados, en términos de una matriz momento de inercia rotada, según

$$\vec{\ell}' = \mathcal{I}' \cdot \vec{\omega}'$$

Donde hemos entendido que el momento de inercia transforma de acuerdo a la regla

$$\mathcal{I}' = R \cdot \mathcal{I} \cdot R^t$$

lo que en componentes se escribe como

$$\mathcal{I}'_{ab} = R_{ac} \mathcal{I}_{cd} R_{bd} = R_{ac} R_{bd} \mathcal{I}_{cd}$$

Esta regla de transformación es parecida a la de un vector, salvo que hay dos matrices de rotación actuando una sobre cada índice de nuestra matriz. Esto define un nuevo tipo de magnitud, cuyas componentes en un sistema de coordenadas están dadas por una matriz, que llamamos *tensor de dos índices contravariante*. El momento de inercia es entonces un ejemplo de tal tipo de tensores.

Existen tensores de más índices contravariantes, definidos en un sistema de coordenadas por arreglos de números $\mathcal{T}_{abc\dots}$, que se transforman frente a rotaciones según la regla

$$\mathcal{T}'_{abc\dots} = R_{ad} R_{be} R_{cf} \dots \mathcal{T}_{def\dots}$$

Un ejemplo es el *tensor de Levi-Civita* dado en algún sistema de coordenadas por el símbolo $\epsilon_{abc} = \pm 1$ con $\epsilon_{xyz} = 1$ y cambiando de signo por cada permutación de índices. Este tensor tiene la propiedad de ser *invariante*, es decir que sus componentes en cualquier otro sistema de coordenadas son las mismas

$$\epsilon'_{abc} = R_{ad} R_{be} R_{cf} \epsilon_{def} = \text{Det}(R) \epsilon_{abc} = \epsilon_{abc}$$

donde en la segunda igualdad hemos usado la definición del determinante de una matriz, y el hecho de que las rotaciones tienen determinante unidad.

Ejercicio:

Probar la invarianza del tensor de Levi-Civita, y verificar que si bien es invariante frente a rotaciones, no lo es frente a paridad.

Ahora bien, tomemos el producto de las componentes de dos vectores contravariantes cualesquiera \vec{A} y \vec{B} , y construyamos una matriz *producto tensorial* cuyas componentes son el producto de las componentes de los vectores

$$(\vec{A} \otimes \vec{B})_{ab} = A_a B_b$$

Si ahora aplicamos una rotación, tenemos que los elementos de esta matriz se transformarán según

$$A'_a B'_b = R_{ac} A_c R_{bd} B_d = R_{ac} R_{bd} A_c B_d$$

Es decir que el producto tensorial de dos vectores contravariantes forma un tensor de dos índices contravariante. Ese objeto es útil para reescribir la fórmula para el momento de inercia en la forma

$$\mathcal{J} = m_n (r_n^2 I - \vec{r}_n \otimes \vec{r}_n)$$

donde I en el primer término es la matriz identidad, cuyas componentes son $I_{ab} = \delta_{ab}$.

Ahora bien, para entender qué relación tiene este *tensor momento de inercia* con los objetos de los que hablamos en los cursos elementales de física, escribamos la energía cinética de nuestro cuerpo rígido en la forma

$$K = \frac{1}{2} m_n \dot{\vec{r}}_n^2 = \frac{1}{2} m_n \dot{\vec{r}}_n \cdot (\vec{\omega} \times \vec{r}_n) = \frac{1}{2} \underbrace{m_n (\vec{r}_n \times \dot{\vec{r}}_n)}_{\vec{\ell}} \cdot \vec{\omega}$$

donde asumimos que el cuerpo rígido sólo está rotando, y en la última igualdad hicimos una *permutación cíclica* reordenando los factores en el producto mixto, para identificar el momento angular. Si escribimos lo de arriba como un producto de matrices (recordando que los vectores son matrices columna) entonces tenemos

$$K = \frac{1}{2} \vec{\ell} \cdot \vec{\omega} = \frac{1}{2} (\mathcal{J} \cdot \omega)^t \cdot \omega = \frac{1}{2} \vec{\omega}^t \cdot \mathcal{J} \cdot \vec{\omega}$$

Escribiendo la velocidad angular como $\vec{\omega} = \omega \check{n}$, donde \check{n} es la dirección del eje de rotación, nos queda la expresión

$$K = \frac{1}{2} \underbrace{(\check{n}^t \cdot \mathcal{J} \cdot \check{n})}_{\mathcal{J}_{\check{n}}} \omega^2 = \frac{1}{2} \mathcal{J}_{\check{n}} \omega^2$$

Donde hemos definido el *momento de inercia alrededor de un eje \check{n}* según $\mathcal{J}_{\check{n}} = \check{n}^t \cdot \mathcal{J} \cdot \check{n}$. Escribiendo esta expresión en componentes tenemos

$$\mathcal{J}_{\check{n}} = n_a \mathcal{J}_{ab} n_b = n_a m_n (r_n^2 \delta_{ab} - r_{na} r_{nb}) n_b = m_n (r_n^2 - (\vec{r}_n \cdot \check{n})^2)$$

donde en la última igualdad usamos $\delta_{ab} n_a n_b = \check{n} \cdot \check{n} = 1$ y $r_{na} n_a = \vec{r}_n \cdot \check{n}$.

Si ahora identificamos $\vec{r}_n \cdot \check{n}$ como la proyección del vector posición en la dirección del eje de rotación, tenemos un triángulo de lados $|\vec{r}_n|$, $\vec{r}_n \cdot \check{n}$, y r_n^\perp donde esta última variable representa la

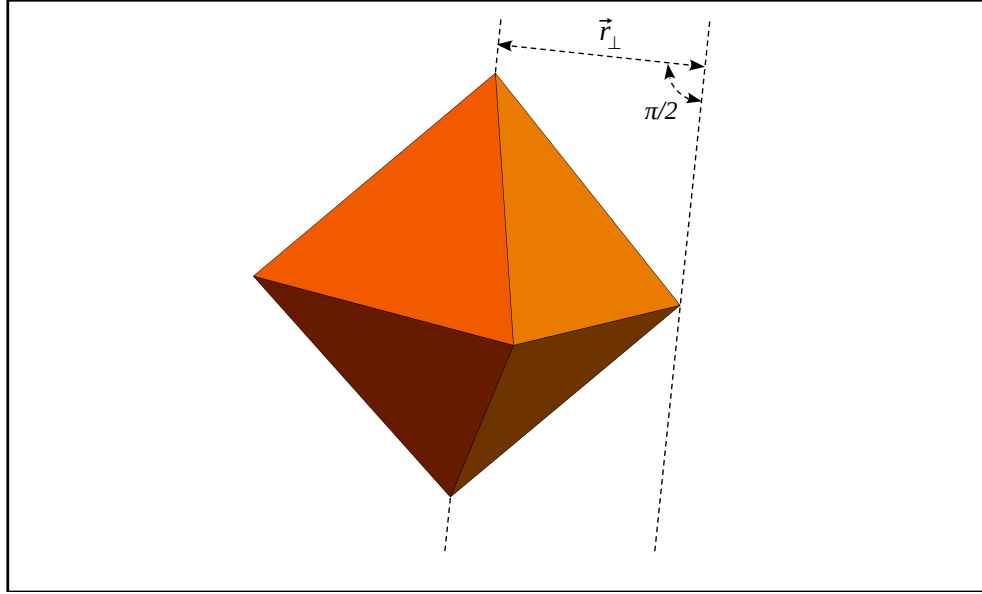


Figura 8.3.1: Esquema del teorema de Huygens-Steiner. El momento de inercia en torno a un eje que pasa por el centro de masas se relaciona con el que corresponde a un eje paralelo cualquiera a través de la distancia perpendicular r_{\perp} .

distancia perpendicular de la partícula n -ésima al eje de rotación. Usando entonces el teorema de Pitágoras podemos escribir

$$\mathcal{I}_{\hat{n}} = m_n (r_n^2 - (\vec{r}_n \cdot \hat{n})^2) = m_n (r_n^{\perp})^2$$

Que no es otra cosa que la definición de momento de inercia alrededor del eje \hat{n} que se aprende en los cursos básicos de física.

La expresión para el tensor momento de inercia en términos de producto tensorial facilita la prueba del teorema de Huygens-Steiner. En efecto, si tenemos

$$\mathcal{I} = m_n (r_n^2 I - r_n \otimes r_n)$$

Usando que $\vec{r}_n = \vec{r}_{\text{cm}} + \vec{r}'_n$ donde \vec{r} es la posición del centro de masas y \vec{r}'_n la posición de la partícula n -ésima en el sistema centro de masas, entonces podemos poner

$$\mathcal{I} = m_n ((\vec{r}_{\text{cm}} + \vec{r}'_n)^2 I - (\vec{r}_{\text{cm}} + \vec{r}'_n) \otimes (\vec{r}_{\text{cm}} + \vec{r}'_n))$$

Como se puede ver usando su definición, el símbolo \otimes funciona de manera distributiva con respecto a la suma, al igual que una multiplicación normal (después de todo no es sino una multiplicación componente a componente). Por lo tanto, si expandimos los productos obtenemos

$$\mathcal{I} = m_n (r_n'^2 I - \vec{r}'_n \otimes \vec{r}'_n) + M (r_{\text{cm}}^2 I - \vec{r}_{\text{cm}} \otimes \vec{r}_{\text{cm}}) + \dots$$

donde $M = \sum_{n=1}^N m_n$ es la masa total, y los términos omitidos se anulan usando que en el sistema centro de masas se verifica que $m_n \vec{r}'_n = 0$. Con esto nos queda

$$\mathcal{I} = \mathcal{I}' + M (r_{\text{cm}}^2 I - \vec{r}_{\text{cm}} \otimes \vec{r}_{\text{cm}})$$

Aquí \mathcal{I}' es el tensor momento de inercia en un sistema de coordenadas con el origen en el centro de masas. Si usamos esta expresión en el momento de inercia alrededor de un eje \hat{n} según lo habíamos definido más arriba $\mathcal{I}_{\hat{n}} = \hat{n}^T \cdot \mathcal{I} \cdot \hat{n}$, obtenemos

$$\mathcal{I}_{\hat{n}} = \mathcal{I}'_{\hat{n}} + Mr_{\perp}^2$$

con $r_{\perp}^2 = r_{\text{cm}}^2 - (r_{\text{cm}} \cdot \hat{n})^2$. Es decir que hemos demostrado el teorema de Huygens-Steiner.

8.3.2 Ejes principales

En los cursos básicos de física aprendimos que los cuerpos rígidos tienen ejes especiales, elegidos de acuerdo a alguna intuición geométrica, que llamamos *ejes principales*. Recordemos que, entre otras propiedades, estos ejes cumplen que alrededor de ellos el momento angular toma una forma muy sencilla

$$\vec{\ell} = \mathcal{I}_{\hat{n}} \vec{\omega}$$

Donde $\mathcal{I}_{\hat{n}}$ es el momento de inercia a lo largo de un eje principal que apunta en la dirección de \hat{n} . Para entender qué son tales ejes principales, empecemos escribiendo la *ecuación de autovalores* en la forma

$$\mathcal{I} \cdot \vec{w} = \mathcal{I} \vec{w}$$

donde en el lado izquierdo tenemos una matriz \mathcal{I} multiplicada por un vector \vec{w} , mientras que en el lado derecho tenemos un número \mathcal{I} multiplicado por el mismo vector. La solución de esta ecuación viene dada por un vector \vec{w} que se llama *autovector* de la matriz \mathcal{I} , y por un escalar \mathcal{I} que se llama *autovalor* de la matriz \mathcal{I} . Por supuesto puede haber varias soluciones $\{\mathcal{I}_a, \vec{w}_a\}$, por lo que es más preciso escribir

$$\mathcal{I} \cdot \vec{w}_a = \mathcal{I}_a \vec{w}_a$$

Nótese que aquí en el miembro derecho no estamos sumando sobre a . Podemos reordenar esta expresión para escribirla en la forma

$$(\mathcal{I} - \mathcal{I}_a I) \cdot \vec{w}_a = 0$$

donde I es la matriz identidad. Esta es una ecuación lineal homogénea para las tres componentes de cada solución \vec{w}_a , que sólo tendrá una solución no trivial si el determinante se anula

$$\text{Det}(\mathcal{I} - \mathcal{I}_a I) = 0$$

Este determinante contiene una suma de términos, cada uno de los cuales es un producto de tres componentes diferentes de la matriz $\mathcal{I} - \mathcal{I}_a I$. Por lo tanto, será un polinomio cúbico en la variable \mathcal{I}_a . La ecuación cúbica resultante tendrá a lo sumo tres soluciones distintas, es decir que el índice a tomará tres valores diferentes.

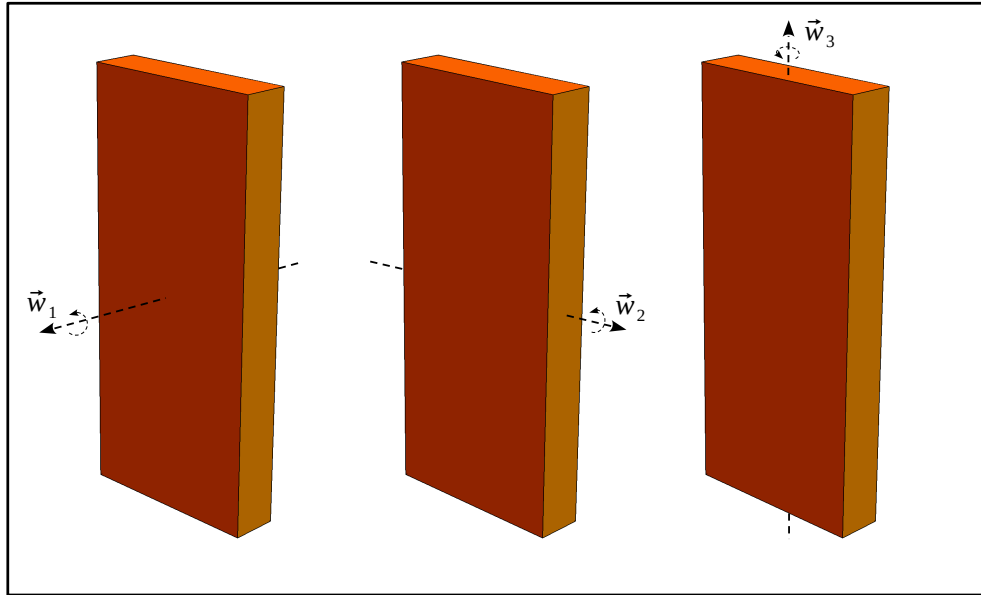


Figura 8.3.2: Ilustración de los ejes principales de un cuerpo rígido. En los cursos básicos de física se los identifica de modo intuitivo. Aquí aprendemos que se trata de los tres autovectores del tensor de inercia.

Vamos a probar ahora que las soluciones \mathcal{J}_a de esta ecuación son números reales. Para esto, multiplicamos escalarmente por el vector conjugado \vec{w}_b^* en la ecuación de autovalores, con lo que la fórmula se escribe

$$(\vec{w}_b^*)^t \cdot \mathcal{J} \cdot \vec{w}_a = \mathcal{J}_a (\vec{w}_b^*)^t \cdot \vec{w}_a$$

En esta expresión, ahora cambiamos a por b y conjugamos el resultado, para obtener

$$(\vec{w}_a)^t \cdot \mathcal{J}^* \cdot \vec{w}_b^* = \mathcal{J}_b^* (\vec{w}_a)^t \cdot \vec{w}_b^*$$

y ahora transponemos

$$(\vec{w}_b^*)^t \cdot \mathcal{J}^\dagger \cdot \vec{w}_a = \mathcal{J}_b^* (\vec{w}_b^*)^t \cdot \vec{w}_a$$

donde hemos usado que la transpuesta del producto es el producto de las transpuestas en orden inverso. Usando el hecho de que \mathcal{J} es una matriz real y simétrica $\mathcal{J}^\dagger = \mathcal{J}$, esta expresión nos queda escrita en la forma

$$(\vec{w}_b^*)^t \cdot \mathcal{J} \cdot \vec{w}_a = \mathcal{J}_b^* (\vec{w}_b^*)^t \cdot \vec{w}_a$$

Restando esta última igualdad de la ecuación con la que comenzamos nuestro cálculo más arriba, tenemos

$$0 = (\mathcal{J}_a - \mathcal{J}_b^*) (\vec{w}_b^*)^t \cdot \vec{w}_a$$

Si elegimos el caso $a = b$ nos queda $(\mathcal{J}_a - \mathcal{J}_a^*) (\vec{w}_a^*)^t \cdot \vec{w}_a = 0$. Nótese que en esta fórmula el miembro derecho es positivo $(\vec{w}_a^*)^t \cdot \vec{w}_a = w_{ax}^* w_{ax} + w_{ay}^* w_{ay} + w_{az}^* w_{az} = |w_{ax}|^2 + |w_{ay}|^2 + |w_{az}|^2 > 0$. Por

lo tanto, si el vector solución \vec{w}_a es no nulo, esta ecuación implica que $\mathcal{I}_a = \mathcal{I}_a^*$. En otras palabras \mathcal{I}_a es real.

Con este resultado, podemos volver a la primera ecuación, para verificar que si \vec{w}_a es solución con autovalor \mathcal{I}_a , entonces \vec{w}_a^* también lo es. Esto implica que la parte real y la parte imaginaria de \vec{w}_a son soluciones reales del problema de autovalores, con autovalor \mathcal{I}_a .

Si en la última ecuación elegimos en cambio $a \neq b$ entonces obtenemos $(\mathcal{I}_a - \mathcal{I}_b)(\vec{w}_b)^t \cdot \vec{w}_a = 0$, de lo que deducimos que

- Si los autovalores son diferentes $\mathcal{I}_a \neq \mathcal{I}_b$ implica $(\vec{w}_b)^t \cdot \vec{w}_a = \vec{w}_b \cdot \vec{w}_a = 0$ por lo tanto los autovectores son ortogonales $\vec{w}_a \perp \vec{w}_b$
- Si los autovalores son iguales $\mathcal{I}_a = \mathcal{I}_b$ y los vectores \vec{w}_a y \vec{w}_b son diferentes, siempre se pueden elegir \vec{w}_a y \vec{w}_b perpendiculares.

Con esto, hemos demostrado que la ecuación de autovalores tiene como solución tres vectores reales y mutuamente perpendiculares. Estos vectores se pueden multiplicar por una constante cualquiera, y seguirán siendo solución. Por lo tanto, podemos normalizarlos para obtener la fórmula general

$$\vec{w}_a \cdot \vec{w}_b = \delta_{ab}$$

Es útil demostrar que estos vectores nos permiten escribir el momento de inercia según la descomposición

$$\mathcal{I} = \mathcal{I}_a(\vec{w}_a \otimes \vec{w}_a)$$

Para probar que esta fórmula es correcta, usamos su expresión en componentes $\mathcal{I}_{cd} = \mathcal{I}_a(\vec{w}_a)_c(\vec{w}_a)_d$ y la insertamos en la ecuación de autovalores $\mathcal{I} \cdot \vec{w}_b = \mathcal{I}_b \vec{w}_b$, obteniendo

$$\mathcal{I}_{cd}(\vec{w}_b)_d = \mathcal{I}_a(\vec{w}_a)_c(\vec{w}_a)_d(\vec{w}_b)_d = \mathcal{I}_a(\vec{w}_a)_c \vec{w}_a \cdot \vec{w}_b = \mathcal{I}_a(\vec{w}_a)_c \delta_{ab} = \mathcal{I}_b(\vec{w}_b)_c$$

donde en la penúltima igualdad utilizamos la ortogonalidad de los vectores \vec{w}_a . Esto demuestra que, en efecto, la descomposición que hemos escrito para \mathcal{I} cumple la ecuación de autovalores con soluciones \mathcal{I}_a . Nótese que esto implica que el momento de inercia alrededor del eje determinado por \vec{w}_a es \mathcal{I}_a , según nuestra definición

$$\mathcal{I}_{\vec{w}_a} = \vec{w}_a^t \cdot \mathcal{I} \cdot \vec{w}_a = \vec{w}_a^t \cdot (\mathcal{I} \cdot \vec{w}_a) = \vec{w}_a^t \cdot (\mathcal{I}_a \vec{w}_a) = \mathcal{I}_a \vec{w}_a^t \cdot \vec{w}_a = \mathcal{I}_a$$

donde en la penúltima igualdad usamos la ecuación de autovalores, y en la última el hecho de que los vectores \vec{w}_a están normalizados.

Ahora vamos a probar la relación de estos autovalores y autovectores con los ejes principales y sus momentos de inercia. Para eso, escribimos el momento angular como $\vec{\ell} = \mathcal{I} \cdot \vec{\omega}$, y suponemos que la velocidad angular apunta en la dirección de alguno de los autovectores $\vec{\omega} = \vec{w}_a \omega$. En ese

caso podemos escribir

$$\vec{\ell} = \mathcal{I} \cdot \vec{\omega} = \mathcal{I} \cdot (\vec{w}_a \omega) = \mathcal{I} \cdot \vec{w}_a \omega$$

y usando en esta expresión la ecuación de autovalores, tenemos

$$\vec{\ell} = \mathcal{I}_a \vec{w}_a \omega = \mathcal{I}_{\vec{w}_a} \vec{\omega}$$

Es decir que los ejes determinados por $\vec{w}_x, \vec{w}_y, \vec{w}_z$ son los ejes principales del cuerpo rígido, en el sentido que le dimos a esa expresión más arriba. Los autovalores $\mathcal{I}_x, \mathcal{I}_y, \mathcal{I}_z$ se llaman *momentos de inercia principales*.

Si elegimos nuestro sistema de coordenadas de manera que sus ejes apunten en las direcciones de nuestros tres vectores \vec{w}_a mutuamente ortogonales, podemos escribir $\vec{\omega} = \omega_x \vec{w}_x + \omega_y \vec{w}_y + \omega_z \vec{w}_z$. Replicando en cada término el cálculo anterior, obtenemos

$$\vec{\ell} = \mathcal{I}_x \omega_x \vec{w}_x + \mathcal{I}_y \omega_y \vec{w}_y + \mathcal{I}_z \omega_z \vec{w}_z$$

Descomponiendo al momento angular según sus componentes en este sistema de ejes $\vec{\ell} = \ell_x \vec{w}_x + \ell_y \vec{w}_y + \ell_z \vec{w}_z$, tenemos que

$$\ell_x = \mathcal{I}_x \omega_x \quad \ell_y = \mathcal{I}_y \omega_y \quad \ell_z = \mathcal{I}_z \omega_z$$

Por otro lado, para la energía cinética de rotación podemos escribir

$$K = \frac{1}{2} \vec{\omega}^t \cdot \mathcal{I} \cdot \vec{\omega} = \frac{1}{2} (\mathcal{I}_x \omega_x^2 + \mathcal{I}_y \omega_y^2 + \mathcal{I}_z \omega_z^2) = \frac{1}{2} \mathcal{I}_x \ell_x^2 + \frac{1}{2} \mathcal{I}_y \ell_y^2 + \frac{1}{2} \mathcal{I}_z \ell_z^2$$

Como vemos, las expresiones en términos de los momentos principales de inercia son mucho más simples y no incluyen matrices, por lo que serán útiles para escribir las ecuaciones de movimiento de un cuerpo rígido.

La expresión para el tensor momento de inercia también se simplifica, ya que podemos usar que en este sistema de coordenadas $\vec{w}_x = (1, 0, 0)$, $\vec{w}_y = (0, 1, 0)$ y $\vec{w}_z = (0, 0, 1)$, para obtener

$$w_x \otimes w_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad w_y \otimes w_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad w_z \otimes w_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

con lo cual

$$\mathcal{I} = \mathcal{I}_a (\vec{w}_a \otimes \vec{w}_a) = \begin{pmatrix} \mathcal{I}_1 & 0 & 0 \\ 0 & \mathcal{I}_2 & 0 \\ 0 & 0 & \mathcal{I}_3 \end{pmatrix}$$

Hemos demostrado que en el sistema de ejes principales, la matriz momento de inercia es diagonal.

8.4 Resumen

En esta clase encontramos una descripción para la cinemática del cuerpo rígido, donde los grados de libertad están dados por la posición del cuerpo, es decir la posición de alguna de sus partículas, y su orientación, dada por una rotación respecto de alguna orientación de referencia. Con esto, fuimos capaces de descomponer cualquier el movimiento del cuerpo en términos de una traslación y una rotación.

Escribiendo el momento angular en función de la velocidad angular asociada a la rotación, pudimos encontrar una expresión para el momento de inercia que nos permitió identificarlo como un tensor de dos índices contravariante. Esto nos permitió definir los ejes principales a través de una ecuación de autovalores.

9. Cuerpo rígido III: Dinámica

9.1 Objetivos

En esta clase escribiremos las ecuaciones que determinan el movimiento del cuerpo rígido. Lo haremos primero según el formalismo newtoniano, para lo cual emplearemos el torque para escribir la derivada temporal del momento angular. Luego utilizaremos el formalismo lagrangiano, lo que nos permitirá hacer uso de coordenadas generalizadas.

Nos concentraremos en dos situaciones de interés: el caso del cuerpo rígido libre sobre el que no actúa ninguna fuerza externa, y el caso del trompo, es decir un cuerpo rígido que tiene un punto que se mantiene quieto mientras se mueve bajo la influencia de la gravedad.



Vladimir Dzhanibekov

9.2 Dinámica del cuerpo rígido

9.2.1 Ecuaciones de Euler

Con los elementos que estudiamos en las clases anteriores, estamos en condiciones de escribir las ecuaciones de movimiento del cuerpo rígido. Como aprendimos en los cursos básicos de física,

estas ecuaciones tienen la forma

$$\frac{d\vec{\ell}}{dt} = \vec{\tau}$$

Donde $\vec{\tau}$ es el torque total que actúa sobre el sistema. En un sistema de coordenadas inercial S , las componentes de esta ecuación se leen

$$\left(\frac{d\vec{\ell}}{dt}\right)_S = (\vec{\tau})_S$$

Vamos a asumir que este sistema inercial se puede colocar en el centro de masas del cuerpo, lo que es posible solo si la suma de fuerzas sobre el mismo se anula. Esto resulta en una simplificación de los cálculos que siguen, que pueden luego generalizarse al caso de un cuerpo acelerado.

Pasando ahora a un sistema de coordenadas no inercial B cuyos ejes están fijos respecto del cuerpo, debemos adicionar a la derivada temporal la contribución que discutimos unas clases atrás, originada en la rotación del sistema de coordenadas. En otras palabras, las componentes del momento angular cumplen

$$\left(\frac{d\vec{\ell}}{dt}\right)_B + (\vec{\omega} \times \vec{L})_B = (\vec{\tau})_B$$

Si recordamos que $\vec{\ell} = \mathcal{I} \cdot \vec{\omega}$ entonces

$$\left(\frac{d(\mathcal{I} \cdot \vec{\omega})}{dt}\right)_B + (\vec{\omega} \times (\mathcal{I} \cdot \vec{\omega}))_B = (\vec{\tau})_B$$

Escribiendo esto explícitamente en componentes, tenemos

$$\mathcal{I}_{ab} \frac{d\omega_b}{dt} + \epsilon_{abc} \omega_b \mathcal{I}_{cd} \omega_d = \tau_a$$

donde como siempre estamos considerando que hay una suma sobre cualquier índice que se repita dos veces. Si agregamos una nueva convención notacional, definiendo que no habrá una suma cuando alguno de los índices del par aparece entre paréntesis, podemos escribir $\mathcal{I}_{ab} = \mathcal{I}_{(b)} \delta_{ab}$ en el sistema de ejes principales del cuerpo. En esta expresión no hay suma sobre el índice b . Con esto la ecuación de movimiento nos queda

$$\mathcal{I}_{(a)} \frac{d\omega_a}{dt} + \epsilon_{abc} \omega_b \mathcal{I}_{(c)} \omega_c = \tau_a$$

hemos hecho explícitamente algunas de las sumas para eliminar las deltas de Kronecker. Separando ahora cada componente

$$\mathcal{I}_x \frac{d\omega_x}{dt} + (\mathcal{I}_z - \mathcal{I}_y) \omega_y \omega_z = \tau_x$$

$$\mathcal{I}_y \frac{d\omega_y}{dt} + (\mathcal{I}_z - \mathcal{I}_x) \omega_x \omega_z = \tau_y$$

$$\mathcal{I}_z \frac{d\omega_z}{dt} + (\mathcal{I}_x - \mathcal{I}_y) \omega_x \omega_y = \tau_z$$

Estas ecuaciones se conocen como *ecuaciones de Euler* y determinan completamente el movimiento de un cuerpo rígido sometido a un torque, en el sistema de coordenadas que está fijo al cuerpo.

Dependiendo de la forma en la que el torque dependa de los ángulos, las ecuaciones de Euler pueden resultar arbitrariamente complicadas de resolver. En la sección que siguen estudiaremos el caso libre, en el cual se puede encontrar una solución explícita.

9.2.2 Cuerpo rígido libre

Sobre un cuerpo rígido libre el torque externo total se anula $\vec{\tau} = 0$. Por lo tanto, las ecuaciones de Euler quedan escritas de una manera particularmente sencilla

$$\mathcal{I}_x \frac{d\omega_x}{dt} + (\mathcal{I}_z - \mathcal{I}_y) \omega_y \omega_z = 0$$

$$\mathcal{I}_y \frac{d\omega_y}{dt} + (\mathcal{I}_z - \mathcal{I}_x) \omega_x \omega_z = 0$$

$$\mathcal{I}_z \frac{d\omega_z}{dt} + (\mathcal{I}_x - \mathcal{I}_y) \omega_x \omega_y = 0$$

Como no hay una dependencia en los ángulos, estas ecuaciones son ahora ecuaciones diferenciales de primer orden para las componentes de la velocidad angular. Se trata de ecuaciones no-lineales, debido a los términos cuadráticos que aparecen en ellas.

Para resolverlas, estudiemos primero sus *puntos fijos*. Es decir, preguntémonos qué tiene que suceder para que todas las derivadas se anulen y una velocidad angular constante resuelva las ecuaciones. Es fácil ver que esto demanda

$$(\mathcal{I}_z - \mathcal{I}_y) \omega_y \omega_z = 0$$

$$(\mathcal{I}_z - \mathcal{I}_x) \omega_x \omega_z = 0$$

$$(\mathcal{I}_x - \mathcal{I}_y) \omega_x \omega_y = 0$$

De aquí podemos ver inmediatamente que

- Si los tres momentos principales de inercia son iguales $\mathcal{I}_x = \mathcal{I}_y = \mathcal{I}_z$, entonces cualquier valor constante de ω_x , ω_y y ω_z satisface estas ecuaciones. Es decir que la solución para la velocidad angular es cualquier vector constante $\vec{\omega} = (\omega_x^0, \omega_y^0, \omega_z^0)$. El cuerpo rígido libre completamente simétrico gira en torno a cualquier eje con velocidad angular constante.
- Si dos de los tres momentos principales de inercia son iguales, digamos $\mathcal{I}_x = \mathcal{I}_y$, entonces la ecuación que contiene la resta no pone restricciones en las componentes de la velocidad angular, en este caso tenemos que ω_x y ω_y son arbitrarias.

Si elegimos $\omega_x = \omega_y = 0$, entonces $\omega_z = \omega_z^0$ es una constante arbitraria y la solución de las ecuaciones es $\vec{\omega} = (0, 0, \omega_z^0)$. El cuerpo rígido libre con un eje de simetría está girando alrededor del mismo.

Si en cambio o bien ω_x o bien ω_y son constantes no nulas, las dos ecuaciones restantes implican que $\omega_z = 0$. La solución en ese caso es $\vec{\omega} = (\omega_x^0, \omega_y^0, 0)$. El cuerpo rígido libre con un eje de simetría está girando alrededor de un eje arbitrario perpendicular al mismo.

- Si los tres momentos de inercia principales son diferentes, entonces las ecuaciones implican que dos de las tres componentes de la velocidad angular deben anularse, digamos $\omega_y = \omega_z = 0$. La solución es entonces $\vec{\omega} = (\omega_x^0, 0, 0)$. Es decir que el cuerpo rígido libre sin simetrías está girando alrededor de alguno de sus ejes principales.

Vamos a concentrarnos en el último caso. Una pregunta natural es qué pasa si perturbamos ese movimiento ligeramente, es decir si hacemos

$$\omega_x = \omega_x^0 + \varepsilon \delta \omega_x \quad \omega_y = \varepsilon \delta \omega_y \quad \omega_z = \varepsilon \delta \omega_z$$

con ε una magnitud muy pequeña. Insertando esto en las ecuaciones de Euler nos queda

$$\begin{aligned} \varepsilon \mathcal{I}_x \frac{d\delta \omega_x}{dt} + \varepsilon^2 (\mathcal{I}_z - \mathcal{I}_y) \delta \omega_y \delta \omega_z &= 0 \\ \varepsilon \mathcal{I}_y \frac{d\delta \omega_y}{dt} + \varepsilon (\mathcal{I}_z - \mathcal{I}_x) (\omega_x^0 + \varepsilon \delta \omega_x) \delta \omega_z &= 0 \\ \varepsilon \mathcal{I}_z \frac{d\delta \omega_z}{dt} + \varepsilon (\mathcal{I}_x - \mathcal{I}_y) (\omega_x^0 + \varepsilon \delta \omega_x) \delta \omega_y &= 0 \end{aligned}$$

Expandiendo estas ecuaciones y descartando los órdenes cuadráticos en ε , nos queda

$$\begin{aligned} \mathcal{I}_x \frac{d\delta \omega_x}{dt} &= 0 \\ \mathcal{I}_y \frac{d\delta \omega_y}{dt} + (\mathcal{I}_z - \mathcal{I}_x) \omega_x^0 \delta \omega_z &= 0 \\ \mathcal{I}_z \frac{d\delta \omega_z}{dt} + (\mathcal{I}_x - \mathcal{I}_y) \omega_x^0 \delta \omega_y &= 0 \end{aligned}$$

Es evidente que una solución de la primera ecuación es $\delta \omega_x = \delta \omega_x^0$ constante. Para resolver las dos ecuaciones que faltan, podemos tomar una derivada temporal adicional en la segunda

$$\mathcal{I}_y \frac{d^2 \delta \omega_y}{dt^2} + (\mathcal{I}_z - \mathcal{I}_x) \omega_x^0 \frac{d\delta \omega_z}{dt} = 0$$

y reemplazar en ella la derivada primera de $\delta \omega_z$ obtenida de la tercera

$$\frac{d\delta \omega_z}{dt} = - \frac{\mathcal{I}_x - \mathcal{I}_y}{\mathcal{I}_z} \omega_x^0 \delta \omega_y$$

con esto obtenemos

$$\frac{d^2 \delta \omega_y}{dt^2} + \underbrace{\frac{(\mathcal{I}_x - \mathcal{I}_z)(\mathcal{I}_x - \mathcal{I}_y)}{\mathcal{I}_y \mathcal{I}_z}}_{\alpha^2} (\omega_x^0)^2 \delta \omega_y = 0$$

Donde hemos supuesto que el factor sobre la llave es positivo. Es decir que la ecuación para $\delta \omega_y$ vendrá dada por

$$\frac{d^2 \delta \omega_y}{dt^2} + \alpha^2 \delta \omega_y = 0$$

Una vez más, como ya vimos en clases previas cuando perturbamos el movimiento del problema de dos cuerpos para obtener el teorema de Bertrand, llegamos a la ecuación de un oscilador armónico. Esta tiene la solución inmediata

$$\delta \omega_y = \delta \omega_y^0 \cos(\alpha(t - t_0))$$

donde $\delta\omega_y^0$ y t_0 son constantes de integración. Con este resultado, la solución para $\delta\omega_z$ se obtiene a partir de la ecuación que escribimos más arriba para su derivada primera, en la forma

$$\delta\omega_z = -\frac{\mathcal{I}_x - \mathcal{I}_y}{\alpha \mathcal{I}_z} \omega_x^0 \delta\omega_y^0 \sin(\alpha(t - t_0)) = \delta\omega_z^0 \sin(\alpha(t - t_0))$$

Esto claramente implica que la dirección de la velocidad angular perturbada rotará describiendo una elipse alrededor del eje original de rotación x . Es decir que si la perturbación comienza siendo pequeña, se conserva pequeña al transcurrir el tiempo. En otras palabras, el sistema es *estable*.

Pero ¿qué pasa si el factor sobre la llave en la ecuación de más arriba es negativo? En ese caso escribimos $\alpha^2 = -|\alpha|^2$ con lo que la solución para $\delta\omega_y$ es ahora

$$\delta\omega_y = \delta\omega_y^+ e^{+|\alpha|t} + \delta\omega_y^- e^{-|\alpha|t}$$

con $\delta\omega_y^\pm$ constantes de integración. Para $\delta\omega_z$ tendremos

$$\delta\omega_z = \delta\omega_z^+ e^{+|\alpha|t} + \delta\omega_z^- e^{-|\alpha|t}$$

con las $\delta\omega_z^\pm$ proporcionales a las constantes $\delta\omega_y^\pm$. Es decir que ahora la perturbación que comienza pequeña crecerá exponencialmente. Esto quiere decir que el eje de rotación del cuerpo rígido se *desestabiliza*, cambiando de dirección.

Ahora bien ¿de qué depende el signo de la magnitud sobre la llave? Claramente su origen está en la combinación

$$(\mathcal{I}_x - \mathcal{I}_z)(\mathcal{I}_x - \mathcal{I}_y)$$

Nótese que esta combinación será positiva tanto cuando \mathcal{I}_x sea el mayor de los momentos de inercia principales, como cuando sea el menor, y será negativa en cambio cuando \mathcal{I}_x sea el momento de inercia intermedio. En otras palabras, la rotación de un cuerpo rígido es estable cuando gira alrededor de sus ejes principales con momento de inercia mayor y menor, pero es inestable cuando lo hace en torno al eje que tiene el momento de inercia intermedio.

Este efecto se conoce como *teorema de la raqueta* porque se hace evidente al arrojar al aire una raqueta de tenis, y también con el nombre de *efecto Dzhanibekov*, por el cosmonauta ruso que lo filmó en ausencia de gravedad en la estación espacial soviética *Salyut-7* en los años 80'.

Ejercicio:

Perturbe el movimiento para el caso del cuerpo rígido libre simétrico del que hablamos más arriba $\mathcal{I}_x = \mathcal{I}_y = \mathcal{I}_z$, y describa sus propiedades de estabilidad.

Haga lo mismo para el caso del cuerpo rígido libre con un eje de simetría $\mathcal{I}_x = \mathcal{I}_y \neq \mathcal{I}_z$, para los dos tipos de movimiento que encontramos en ese caso.

Nota:

La estabilidad de la rotación de un cuerpo rígido en torno a sus ejes principales con momentos de inercia mayor y menor se puede utilizar para estabilizar naves espaciales de una manera que consume muy poco combustible.

Por ejemplo, las sondas *Pioneer 10* y *Pioneer 11*, lanzadas por la NASA en 1972 y 1973 respectivamente, fueron *estabilizadas por rotación*. Cada una de ellas estaba construida con su antena de comunicaciones dispuesta a lo largo de uno de sus ejes principales estables. Poco después del lanzamiento, se apuntó la antena hacia la Tierra y se le imprimió a la nave una velocidad angular a lo largo de tal eje estable. De este modo se logró que la antena permaneciera apuntando hacia la base durante largos períodos, sin tener que disparar los cohetes de control.

Estas sondas fueron extremadamente durables, permaneciendo funcionales durante 30 años. Junto con las *Voyager 1* y *Voyager 2*, son los primeros objetos hechos por el hombre que abandonaron el Sistema Solar. Debido a que no era necesario disparar sus cohetes de control para mantener la comunicación, se pudo recolectar una gran cantidad de datos sobre su trayectoria libre. A fines de la década de 1990, se descubrió en esos datos algo muy extraño: había algún efecto desconocido que las estaba frenando, haciendo que se alejaran del Sol cada vez más lentamente.

Se generó una enorme actividad de investigación científica intentando explicar la *anomalía de las Pioneer*. Se propusieron todo tipo de hipótesis, desde una simple falla en la telemetría o un medio interestelar desconocido, hasta modificaciones de la ley de gravitación universal a largas distancias o la presencia de una quinta fuerza fundamental.

Finalmente, la explicación resultó tan simple como sorprendente: las naves recibían una minúscula cantidad de calor, y generaban una cantidad mucho mayor en su funcionamiento. Este calor se emitía al espacio en forma de radiación infrarroja la cual, debido a la forma de la sonda, no tenía la misma intensidad en todas las direcciones. Esta *emisión anisotrópica* provocaba una fuerza neta en dirección al Sol, frenando de este modo su carrera.

9.2.3 Trompo

Usando lo que aprendimos en las clases anteriores, tenemos todos los elementos para escribir la expresión del lagrangiano de un cuerpo rígido. Comencemos por elegir un sistema de coordenadas en el cual la partícula en \vec{r}_1 está en reposo $\dot{\vec{r}}_1 = 0$. En este sistema, la energía cinética toma la forma

$$K = \frac{1}{2} \vec{\omega}^t \cdot \mathcal{I} \cdot \vec{\omega}$$

Este sistema debe ser un sistema inercial, de modo tal de poder escribir el Lagrangiano según

$$L = \frac{1}{2} \vec{\omega}^t \cdot \mathcal{I} \cdot \vec{\omega} - V(\vartheta, \varphi, \psi)$$

Para darle un uso efectivo a esta expresión, vamos a escribir su energía cinética de acuerdo a su descomposición en los ejes principales

$$K = \frac{1}{2} \mathcal{I}_x \omega_x^2 + \frac{1}{2} \mathcal{I}_y \omega_y^2 + \frac{1}{2} \mathcal{I}_z \omega_z^2$$

Debemos encontrar una forma explícita a las componentes de $\vec{\omega}$ en términos de algún conjunto de coordenadas generalizadas, por ejemplo los ángulos de Euler.

Como se ve en el dibujo, podemos escribir la velocidad angular como una superposición de una rotación alrededor de cada uno de los ejes respecto de los cuales definimos los ángulos de Euler, según

$$\vec{\omega} = \vec{\omega}_\psi + \vec{\omega}_\theta + \vec{\omega}_\phi = \dot{\psi} \check{k}'' + \dot{\vartheta} \check{i}' + \dot{\phi} \check{k}$$

Ahora bien, podemos elegir nuestro sistema de coordenadas de modo tal que uno de los ejes principales del cuerpo apunte en la dirección z'' , mientras que los otros dos apuntarán en las direcciones x'' e y'' . Las componentes de la velocidad angular en este sistema de ejes serán

$$\omega_x = \vec{\omega} \cdot \check{i}'' = \dot{\vartheta} \quad \omega_y = \vec{\omega} \cdot \check{j}'' = \dot{\phi} \sin \vartheta \quad \omega_z = \vec{\omega} \cdot \check{k}'' = \dot{\psi} + \dot{\phi} \cos \vartheta$$

Esto nos permite escribir inmediatamente para la energía cinética

$$K = \frac{1}{2} \mathcal{I}_x \dot{\vartheta}^2 + \frac{1}{2} \mathcal{I}_y \sin^2 \vartheta \dot{\phi}^2 + \frac{1}{2} \mathcal{I}_z (\dot{\psi} + \dot{\phi} \cos \vartheta)^2$$

Un *trompo* es un cuerpo rígido que es simétrico de modo tal $\mathcal{I}_2 = \mathcal{I}_3$, es decir que dos de sus ejes principales son iguales. Más aún, el potencial al que está sometido es puramente gravitatorio, con lo que tenemos que

$$V = mgh = mgl \cos \vartheta$$

Donde $h = l \cos \vartheta$ es la altura del centro de masas, el cual para un trompo homogéneo se encuentra sobre el eje a una distancia l del punto fijo del mismo.

Con lo anterior, el lagrangiano toma la forma

$$L = \frac{1}{2} \mathcal{I}_x (\dot{\vartheta}^2 + \sin^2 \vartheta \dot{\phi}^2) + \frac{1}{2} \mathcal{I}_z (\dot{\psi} + \dot{\phi} \cos \vartheta)^2 - mgl \cos \vartheta$$

La primera observación que se puede hacer aquí es que tanto ϕ como ψ son coordenadas cíclicas, con lo que los momentos generalizados asociados se conservan

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = \mathcal{I}_z (\dot{\psi} + \dot{\phi} \cos \vartheta) \cos \vartheta + \mathcal{I}_x \sin^2 \vartheta \dot{\phi} = p_\psi \cos \vartheta + \mathcal{I}_x \sin^2 \vartheta \dot{\phi} = \text{constante}$$

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = \mathcal{I}_z (\dot{\psi} + \dot{\phi} \cos \vartheta) = \text{constante}$$

Esto nos permite despejar las velocidades generalizadas, en la forma

$$\dot{\phi} = \frac{p_\phi - p_\psi \cos \vartheta}{\mathcal{I}_x \sin^2 \vartheta} \quad \dot{\psi} = \frac{p_\psi}{\mathcal{I}_z} - \frac{p_\phi - p_\psi \cos \vartheta}{\mathcal{I}_x \sin^2 \vartheta} \cos \vartheta$$

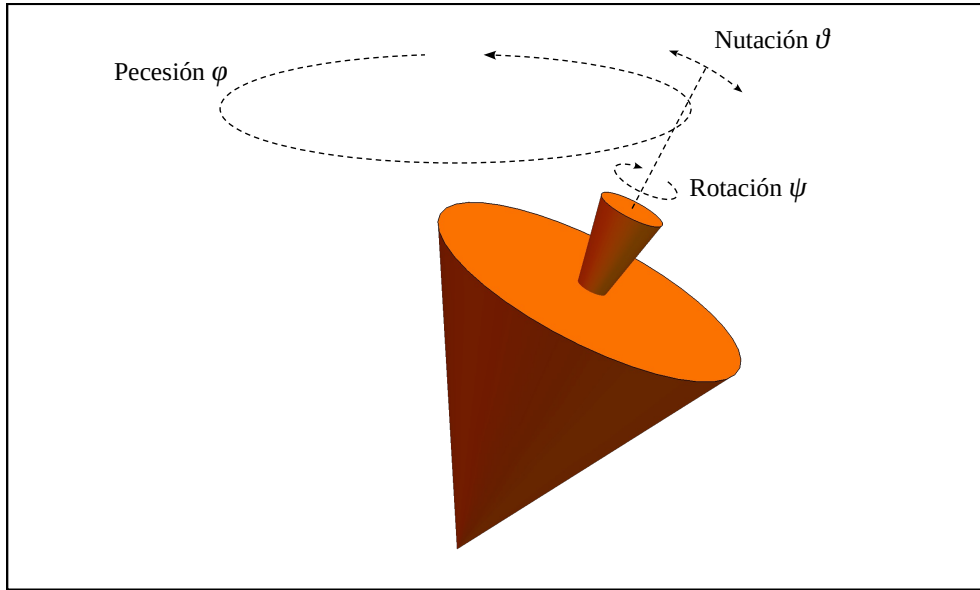


Figura 9.2.1: Movimientos del trompo. La rotación se produce a en torno al eje de simetría, la precesión tiene lugar alrededor del eje vertical, mientras que la nutación tiene en cuenta el balanceo del eje de rotación.

El movimiento en la dirección ψ se llama *rotación* ya que corresponde a la rotación del trompo alrededor de su eje principal. El movimiento en la dirección ϕ indica que el mencionado eje principal está girando alrededor del eje z , y se denomina *precesión*. Estas ecuaciones indican que si conocemos el movimiento de ϑ podemos obtener inmediatamente el de ϕ y ψ mediante una integral temporal. Para esta última coordenada, la ecuación a resolver será la ecuación de Lagrange

$$\mathcal{J}_x \ddot{\vartheta} + \mathcal{J}_z (\dot{\psi} + \dot{\phi} \cos \vartheta) \sin \vartheta \dot{\phi} - \mathcal{J}_x \dot{\phi}^2 \sin \vartheta \cos \vartheta - mgl \sin \vartheta = 0$$

Esta ecuación se simplifica si reemplazamos en ella los valores de $\dot{\psi}$ y $\dot{\phi}$ que se obtienen de los momentos generalizados

$$\mathcal{J}_x \ddot{\vartheta} + p_\psi \sin \vartheta \left(\frac{p_\phi - p_\psi \cos \vartheta}{\mathcal{J}_x \sin^2 \vartheta} \right) - \mathcal{J}_x \left(\frac{p_\phi - p_\psi \cos \vartheta}{\mathcal{J}_x \sin^2 \vartheta} \right)^2 \sin \vartheta \cos \vartheta - mgl \sin \vartheta = 0$$

Esta ecuación se puede reescribir según

$$\mathcal{J}_x \ddot{\vartheta} = - \frac{d}{d\vartheta} \left(\underbrace{\frac{1}{2\mathcal{J}_x} \left(\frac{p_\phi - p_\psi \cos \vartheta}{\sin \vartheta} \right)^2 + mgl \cos \vartheta}_{V_{1d}} \right)$$

Donde vemos inmediatamente que se puede derivar del Lagrangiano truncado

$$L_{1d} = \frac{1}{2} \mathcal{J}_x \dot{\vartheta}^2 - V_{1d}$$

Es decir que, de modo similar a lo que encontramos en el caso del problema de dos cuerpos, terminamos con un sistema unidimensional equivalente, ahora para la coordenada ϑ . Dado que ϑ toma valores a lo largo de un meridiano, el espacio de configuración es un semicírculo $\mathcal{C} = S_1/\mathbb{Z}_2$.

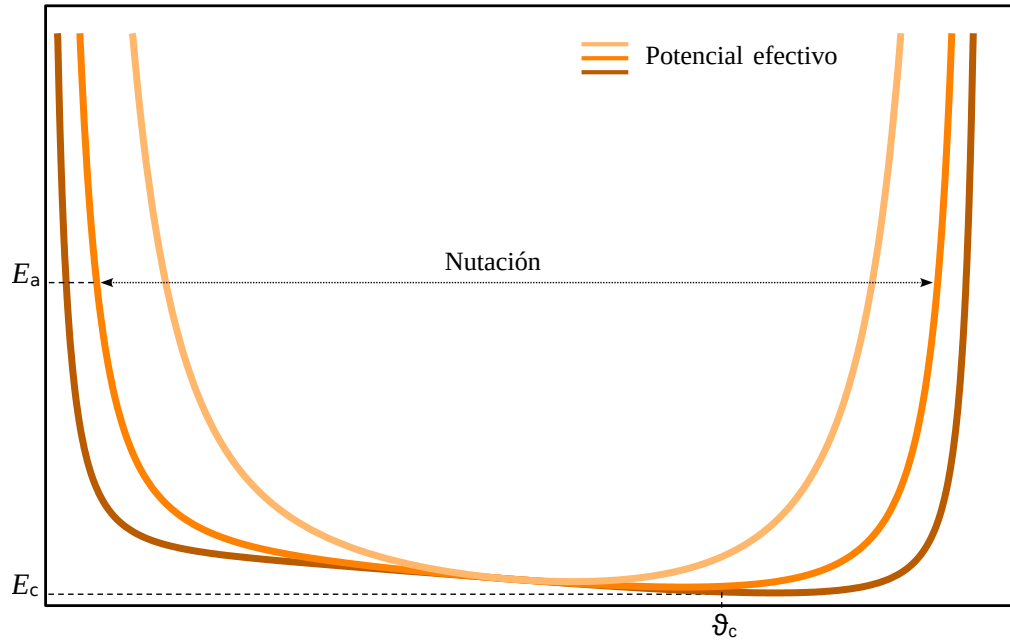


Figura 9.2.2: Potencial efectivo para el trompo: el potencial efectivo como función de ϑ diverge en los valores extremos $\vartheta = 0$ y $\vartheta = \pi$, lo que implica que el movimiento de nutación es siempre acotado. Existe un valor mínimo de la energía en el que no hay nutación y se produce precesión regular.

Le energía correspondiente toma la forma

$$\tilde{E} = \frac{1}{2} \mathcal{J}_x \dot{\vartheta}^2 + V_{1d}(\vartheta)$$

De aquí se puede despejar la velocidad angular según

$$\dot{\vartheta} = \sqrt{\frac{2}{\mathcal{J}_x} (\tilde{E} - V_{1d}(\vartheta))}$$

Lo que se puede integrar, obteniendo la solución completa del problema

$$t = t_0 + \int_{\vartheta_0}^{\vartheta} \frac{d\vartheta}{\sqrt{\frac{2}{\mathcal{J}_x} (\tilde{E} - V_{1d}(\vartheta))}}$$

Reemplazando en las fórmulas de más arriba para las velocidades $\dot{\phi}$ y $\dot{\psi}$ e integrando, tenemos una solución completa del problema.

Sin embargo, de modo similar a lo que pasaba en el problema de dos cuerpos, podemos tener una descripción cualitativa de las características principales del movimiento sin necesidad de hacer las mencionadas integrales, simplemente analizando el potencial efectivo V_{1d} .

Primero notemos que, debido al $\sin \vartheta$ en el denominador, el potencial efectivo V_{1d} diverge para $\vartheta = 0$ y para $\vartheta = \pi$. Es decir que sea cual fuere la energía \tilde{E} , siempre tendremos dos puntos de retorno $0 < \vartheta_2 < \vartheta < \vartheta_1 < \pi$. Esto implica que el eje principal del trompo realizará un movimiento de *nutación* o cabeceo en el ángulo ϑ entre estos dos valores $\vartheta_1 \leq \vartheta \leq \vartheta_2$, a medida que precesa alrededor del eje z . Dada la fórmula de más arriba para la velocidad de precesión $\dot{\phi}$ podemos ver que en ambos extremos de la oscilación esta velocidad tomará los valores

$$\dot{\phi}_2 = \frac{p_\phi - p_\psi \cos \vartheta_2}{\mathcal{J}_x \sin^2 \vartheta_2} \quad \dot{\phi}_1 = \frac{p_\phi - p_\psi \cos \vartheta_1}{\mathcal{J}_x \sin^2 \vartheta_1}$$

Vemos que

- Estos valores de la velocidad de precesión $\dot{\phi}$ tendrán el mismo signo si $p_\phi/p_\psi > \cos \vartheta_2$ o si $p_\phi/p_\psi < \cos \vartheta_1$. Este caso el eje del trompo se estará moviendo alrededor del eje z en la misma dirección en ambos extremos de la oscilación. Estamos en presencia de una *precesión cuasiperiódica*.
- Por otro lado, los valores de $\dot{\phi}$ tendrán signos diferentes si $\cos \vartheta_1 < p_\phi/p_\psi < \cos \vartheta_2$, con lo que el eje del trompo se estará moviendo en direcciones opuestas en cada extremo de la oscilación. Esto se denomina *precesión alternada*.
- Existe un caso límite entre ambos comportamientos cuando $p_\phi/p_\psi = \cos \vartheta_2$, lo que implica que la velocidad de precesión en el extremo superior de la nutación se anula. Este caso se realiza cuando “soltamos” el trompo en un ángulo fijo, de modo que su velocidad angular inicial tenga $\dot{\phi}_0 = 0$. En ese caso tendremos

$$p_\phi = \mathcal{J}_z \dot{\psi}_0 \cos \vartheta_0 \quad p_\psi = \mathcal{J}_z \dot{\psi}_0$$

De donde $p_\phi/p_\psi = \cos \vartheta_0$. Además dado que $\dot{\vartheta}_0 = 0$ es evidente que estamos en uno de los puntos de retorno, con lo que naturalmente podemos identificar $\vartheta_0 = \vartheta_2$.

- Cuando la energía \tilde{E} toma el valor del mínimo del potencial efectivo, entonces no hay movimiento de nutación $\dot{\vartheta} = 0$ y estamos en presencia de una *precesión regular*. El mínimo mencionado ocurre cuando

$$V'_{1d} = \frac{1}{\mathcal{J}_x} \left(\frac{p_\phi - p_\psi \cos \vartheta}{\sin \vartheta} \right) \left(p_\psi + \left(\frac{p_\phi - p_\psi \cos \vartheta}{\sin^2 \vartheta} \right) \cos \vartheta \right) - mgl \sin \vartheta = 0$$

lo que se puede reescribir en términos de las velocidades que despejamos más arriba como $\dot{\phi} (p_\psi + \dot{\phi} \cos \vartheta) = \mathcal{J}_x mgl$, de donde se puede despejar la velocidad $\dot{\phi}$ del movimiento de precesión regular.

Ejercicio:

Usando un software que permita dibujar curvas como Mathematica o GeoGebra, estudie la forma del potencial como función de $\cos \theta$ para valores fijos de p_ψ , p_ϕ y el parámetro $\mathcal{J}_z m l$.

¿Qué sucede con el potencial cuando el trompo gira muy rápidamente? ¿Qué implica esto para el movimiento resultante?

¿Qué sucede cuando el roce frena el movimiento de rotación y el trompo gira demasiado lentamente? ¿Qué implica esto para el protagonista de la película “Inception”?

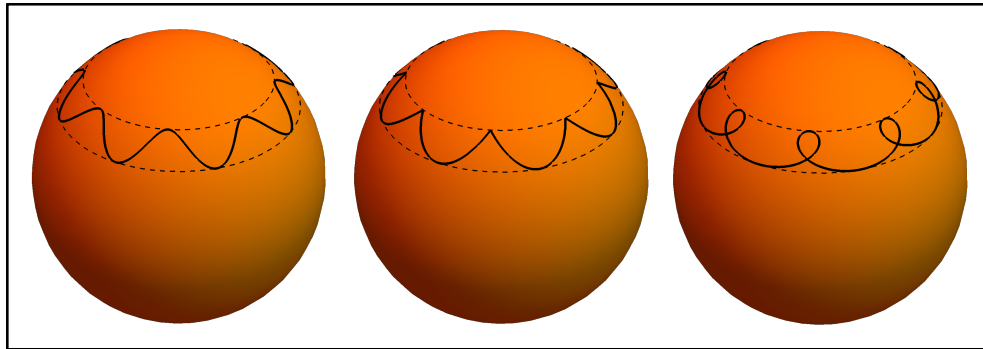


Figura 9.2.3: *Movimiento de precesión. Dependiendo de los valores de los momentos p_ϕ y p_ψ podemos tener precesión cuasiperiódica (izquierda) precesión extrema (centro) y precesión alternada (derecha).*

9.3 Resumen

En esta clase obtuvimos las ecuaciones que rigen el movimiento de un cuerpo rígido, y encontramos sus soluciones para dos casos particulares: el cuerpo rígido libre y el trompo.

En el caso de cuerpo rígido libre asimétrico, aprendimos que la rotación será estable alrededor de los ejes principales con momento de inercia mayor y menor, e inestable alrededor del eje con momento de inercia intermedio.

En el caso de un trompo, fuimos capaces de aislar y describir las tres componentes de su movimiento, a saber la rotación, la precesión, y la nutación.



Formulación lagrangiana

Clase 10 Transformaciones de Legendre 133

- 10.1 Objetivos
- 10.2 Transformación de Legendre
- 10.3 Truncaciones
- 10.4 Resumen

Clase 11 Vínculos 147

- 11.1 Objetivos
- 11.2 Origen físico de los vínculos
- 11.3 Resumen

Clase 12 Pequeñas oscilaciones 169

- 12.1 Objetivos
- 12.2 Oscilaciones en sistemas autónomos
- 12.3 Generalización
- 12.4 Resumen

Clase 13 Principio de acción estacionaria 183

- 13.1 Objetivos
- 13.2 Funcionales y cálculo variacional
- 13.3 Principio de Hamilton de acción estacionaria
- 13.4 Resumen

Clase 14 Simetrías 205

- 14.1 Objetivos
- 14.2 Simetrías
- 14.3 Resumen

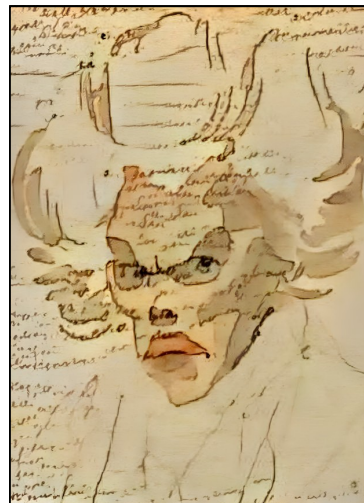
10. Transformaciones de Legendre

10.1 Objetivos

En las clases previas, definimos los momentos generalizados como las derivadas del lagrangiano respecto de las velocidades generalizadas. Estos son naturalmente funciones de las coordenadas y las velocidades.

En varios casos, invertimos estas funciones para escribir algunas de las velocidades generalizadas en función de las coordenadas y los impulsos. Al reemplazar en las ecuaciones de movimiento, esto frecuentemente simplificó su solución. Utilizamos tal método al describir el movimiento del sistema de dos cuerpos y del trompo.

En la presente clase, estudiaremos la forma de generalizar esta construcción.



Adrien-Marie Legendre

10.2 Transformación de Legendre

Comenzamos la discusión presentando en esta sección una técnica matemática general, que aplicaremos en las secciones subsiguientes a la descripción del movimiento de un sistema mecánico.

Supongamos una cualquiera curva en el plano. Como se puede ver en la figura 10.2.1, la información sobre forma de la curva está completamente contenida en su *haz de tangentes*. En efecto,

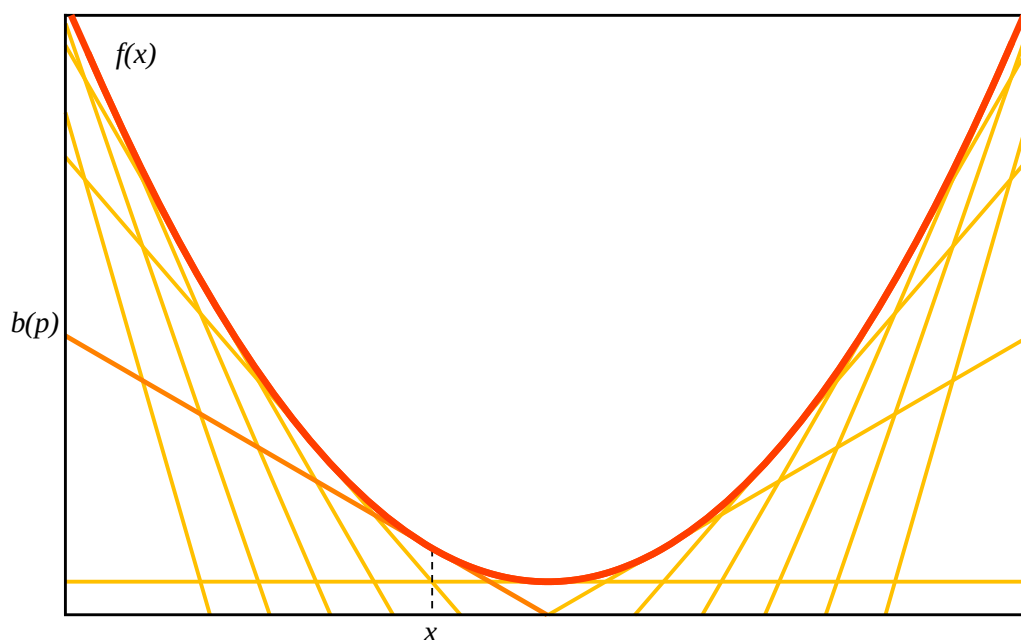


Figura 10.2.1: Ilustración de una función $f(x)$ y su haz de rectas tangentes. Conociendo la ordenada al origen $b(p)$ que corresponde a la recta tangente con pendiente p es posible reconstruir completamente la función.

si conocemos la ordenada al origen de cada recta tangente como una función de su pendiente, podemos trazar cada recta del haz y vislumbrar en el gráfico resultante una curva envolvente, que determina completamente el perfil de nuestra curva original.

Para encontrar este mapeo de manera explícita, supongamos que la curva está definida en el plano $\{x, z\}$ por la fórmula

$$z = f(x)$$

Una tangente cualquiera estará entonces descrita por una recta, que satisface la ecuación

$$z = px - b$$

Donde p es la pendiente y $-b$ es la ordenada al origen. Si asumimos que se trata de la recta particular que toca la curva en el punto donde la variable independiente toma el valor x , sabemos que debe satisfacer en dicho punto las siguientes condiciones de tangencia

$$px - b = f(x)$$

$$p = f'(x)$$

Asumiendo las condiciones de regularidad necesarias sobre la derivada $f'(x)$, podemos invertir localmente la segunda ecuación de manera de obtener el valor de x como una función de la pendiente, en la forma

$$x = f'^{-1}(p) \equiv v(p)$$

Donde hemos supuesto que $f'(x)$ es invertible en el punto x , y llamamos $v(p)$ a su función inversa en el valor p de la pendiente. Por otro lado, es evidente de la primera ecuación que la ordenada al origen se puede despejar como una función de x , en la forma $b(x) = px - f(x)$. Sin pérdida de generalidad podemos insertar la inversa $x = v(p)$ para escribirla como una función de la pendiente p según

$$b(p) = pv(p) - f(v(p))$$

Esta fórmula describe completamente el haz de tangentes, ya que provee una ordenada al origen para cada valor de la pendiente. Se denomina la *transformada de Legendre* de la función $f(x)$.

Dada cualquier función $f(x)$ de una variable x , su transformada de Legendre $b(p)$ es una función de la pendiente p , que se calcula usando la fórmula de arriba, donde $v(p)$ es la función inversa local de $f'(x)$. Como veremos en lo que sigue, la transformación de Legendre se puede generalizar al caso de varias variables, donde podemos transformar respecto de todas las variables o sólo de algunas de ellas.

Ejemplo: derivada de la transformada de Legendre

Cuando calculamos la derivada de la transformación de Legendre respecto de su variable, obtenemos

$$b'(p) = v(p) + pv'(p) - f'(v(p))v'(p) = f'^{-1}(p)$$

donde usamos que $v(p) = f'^{-1}(p)$ para cancelar los dos últimos términos.

Esto implica que la derivada de la transformada de Legendre de una función cualquiera está dada por la función inversa de la derivada de la misma.

Ejemplo: transformada de Legendre inversa

Supongamos ahora que dibujamos la curva en el plano (p, h) dada por la función $b(p)$ en la forma

$$h = b(p)$$

Una recta cualquiera en este plano se puede describir con

$$h = xp - g$$

donde x es ahora la pendiente y $-g$ es la ordenada al origen. Queremos encontrar $g(x)$ que es la transformada de Legendre de $b(p)$. En el valor de la variable independiente p donde la curva entra en contacto con la recta, se deben cumplir las condiciones de tangencia

$$xp - g = b(p)$$

$$x = b'(p)$$

Asumiendo condiciones suficientes de regularidad sobre $b'(p)$, podemos invertir localmente la segunda ecuación para obtener p en la forma

$$p = b'^{-1}(x) \equiv w(x)$$

donde hemos definido la función $w(x)$ como la inversa de $b'(p)$. Esto nos permite escribir la transformada de Legendre de $b(p)$ como

$$g(x) = xw(x) - b(w(x))$$

Usando la identidad que escribimos más arriba para la derivada $b'(p)$ vemos que se cumple que $w(x) = f'(x)$ con lo cual

$$g(x) = f'(x)x - b = f(x)$$

Donde hemos utilizado en la segunda igualdad las condiciones de tangencia en el punto x que escribimos al principio de la discusión. Es decir que una segunda transformada de Legendre nos devuelve la función original $f(x)$.

La transformación de Legendre se puede generalizar al caso en el que la función depende de varias variables. Por ejemplo, supongamos que tenemos una función de dos variables que describe una superficie en el espacio $\{x, y, z\}$ definida como

$$z = f(x, y)$$

Esta función puede reemplazarse por su haz de planos tangentes, el cual contiene la misma información. Tales planos están descriptos por la ecuación

$$z = p_x x + p_y y - b$$

donde p_x, p_y denotan las inclinaciones del plano en las direcciones de cada uno de los ejes, y $-b$ es la ordenada al origen. En el punto (x, y) donde un dado plano entra en contacto con la superficie, se cumplen las condiciones de tangencia

$$f(x, y) = p_x x + p_y y - b$$

$$\partial_x f(x, y) = p_x \quad \partial_y f(x, y) = p_y$$

Supongamos que las derivadas parciales cumplen las condiciones de regularidad necesarias tales que podemos invertir las dos últimas ecuaciones, para obtener las coordenadas del punto de contacto en la forma

$$x = v_x(p_x, p_y) \quad y = v_y(p_x, p_y)$$

Esto nos permite entonces reemplazar en la primera ecuación, y despejar de allí para obtener la ordenada al origen b en términos de la fórmula

$$b(p_x, p_y) = p_x v_x(p_x, p_y) + p_y v_y(p_x, p_y) - f(v_x(p_x, p_y), v_y(p_x, p_y))$$

Hemos obtenido la transformada de Legendre en dos variables de la función $f(x, y)$. La generalización para más variables es completamente directa.

Ejercicio:

Pruebe que las derivadas de la transformada de Legendre de arriba se pueden calcular en la forma

$$\partial_{p_x} b = v_x(p_x, p_y) \quad \partial_{p_y} b = v_y(p_x, p_y)$$

donde al igual que en el caso previo hemos cancelado términos usando la definición de $v_x(p_x, p_y)$ y $v_y(p_x, p_y)$.

Una vez más reobtenemos las funciones que nos permitieron invertir las derivadas de la función original.

Ejercicio:

Calcule la transformada de Legendre de un polinomio cuadrático que depende de dos variables $\{x, y\}$, identifique las funciones $v_x(p_x, p_y)$ y $v_y(p_x, p_y)$, y verifique la propiedad que cumplen las derivadas según el ejercicio anterior.

En el caso de una función que depende de varias variables, podemos también definir una transformada de Legendre parcial. Esto significa que tomaremos algunas de las coordenadas como parámetros. Por ejemplo, para el caso de una función de dos variables, podemos considerar que la ecuación

$$z = f(x, y)$$

define un conjunto de curvas en el plano (x, z) parametrizadas por el parámetro y . Para cada valor del parámetro y , la curva correspondiente puede ser reemplazada por su haz de tangentes. Por supuesto, la pendiente y ordenada al origen de cada recta del mencionado haz dependerán del valor del parámetro. Podemos entonces escribir

$$y = p_x(y)x - b(y)$$

Siguiendo pasos análogos a los de más arriba, encontramos su transformada de Legendre, escribiendo en el punto de tangencia $x(y)$

$$f(x(y), y) = p_x(y)x(y) - b(y)$$

$$\partial_x f(x(y), y) = b(y)$$

Invertiendo la segunda ecuación

$$x(y) = v_x(p_x, y)$$

donde ahora la función inversa tendrá también una dependencia en el parámetro y . Reemplazando en la primera ecuación obtenemos

$$b(p_x, y) = p_x v_x(p_x, y) - f(v_x(p_x, y), y)$$

Lo que define la transformada parcial de Legendre de la función $f(x, y)$ respecto de la variable x . Esto se puede generalizar sin dificultad para un número arbitrario de variables.

Ejemplo: derivadas de la transformada parcial

La transformada parcial de Legendre tiene la propiedad de que sus derivadas satisfacen

$$\partial_{p_x} b = v_x(p_x, y)$$

$$\partial_y b = p_x \partial_y v_x - \partial_x f \partial_y v_x - \partial_y f = -\partial_y f$$

En la igualdad en la segunda línea usamos la condición de tangencia $p_x = \partial_x f$ para eliminar los dos primeros términos. Este resultado es particularmente interesante, porque la derivada de la izquierda corresponde a la derivada parcial de una función que depende de p_x e y , mientras que la derecha actúa sobre una función de x e y . Es decir que mientras una se toma manteniendo p_x constante, la otra se calcula con x constante.

Esta observación resultará de enorme utilidad en lo que sigue.

Nota

Durante más de dos siglos, hasta 2005, la imagen de Adrien-Marie Legendre fue confundida con la de un político francés llamado Louis Legendre, que apareció sin el nombre y junto con algunos matemáticos contemporáneos tales como Lagrange en un libro de la época.

El único retrato conocido de Adrien-Marie Legendre es la acuarela en la que se basa la imagen con la que abrimos este capítulo, que se descubrió en 2008 en un libro de caricaturas de los miembros del *Institut de France*.

10.2.1 Hamiltoniano

Veamos ahora cómo podemos aplicar la idea de transformada de Legendre a la descripción del movimiento.

Supongamos que tenemos un sistema mecánico de N partículas, descrito en términos de coordenadas generalizadas $\{q_i\}$ donde $i \in \{1, \dots, 3N\}$. Su lagrangiano será una función de las coordenadas y las velocidades generalizadas $L(\dot{q}_i, q_i, t)$, que nos permite escribir las ecuaciones de Lagrange en la forma

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

Ahora bien, si recordamos la definición de los momentos generalizados

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

vemos los mismos son naturalmente funciones de las coordenadas y velocidades generalizadas.

Si asumimos que podemos despejar las velocidades generalizadas $\{\dot{q}_i\}$, tenemos que

$$\dot{q}_i = v_i(p_j, q_j, t) \quad \text{donde } i, j \in \{1, \dots, 3N\},$$

Esto nos permite definir la transformada de Legendre del lagrangiano con respecto a las velocidades generalizadas $\{\dot{q}_i\}$ en la forma

$$H(p_i, q_i, t) = p_i v_i(p_j, q_j, t) - L(v_i(p_j, q_j, t), q_i, t)$$

donde en el primer término estamos sumando sobre el índice i . La función $H(p_i, q_i, t)$ se denomina *hamiltoniano* del problema, y es naturalmente una función de las coordenadas $\{q_i\}$ y de los *impulsos* $\{p_i\}$. Como veremos, toda la información dinámica del problema, que estaba contenida en el lagrangiano, ahora se puede reobtener a partir del hamiltoniano. En el contexto hamiltoniano, el conjunto de todos los pares de coordenadas e impulsos $\{(p_i, q_i)\}$ se denomina *variables canónicas*.

En la transformación de Legendre que acabamos de realizar, las coordenadas generalizadas $\{q_i\}$ cumplen el rol de parámetros. De acuerdo a lo que aprendimos en la sección anterior, esto implica que se cumple la siguiente propiedad importante

$$\frac{\partial H}{\partial q_i} = \frac{\partial v_j}{\partial \dot{q}_i} p_j - \frac{\partial L}{\partial \dot{q}_j} \frac{\partial v_j}{\partial q_i} - \frac{\partial L}{\partial q_i} = - \frac{\partial L}{\partial q_i}$$

dado que los dos primeros términos se cancelan. En estas expresiones, las derivadas parciales del lado derecho se calculan manteniendo las velocidades $\{\dot{q}_i\}$ constantes, ya que son derivadas de $L(\dot{q}_i, q_i, t)$ que es una función de las coordenadas y velocidades generalizadas $\{\dot{q}_i, q_i\}$. Por otra parte, las derivadas del lado izquierdo se toman manteniendo los impulsos $\{p_i\}$ constantes, ya que actúan sobre $H(p_i, q_i, t)$ que es una función de las variables canónicas $\{(p_i, q_i)\}$.

Las identidades de arriba nos permiten reescribir las ecuaciones de Lagrange para las coordenadas $\{q_i\}$ en la forma

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = \dot{p}_i - \frac{\partial L}{\partial q_i} = \dot{p}_i + \frac{\partial H}{\partial q_i} = 0$$

En la última expresión de la derecha, las velocidades $\{\dot{q}_i\}$ han desaparecido completamente, ya que el hamiltoniano no depende de ellas. Esto nos permite reinterpretarla como una ecuación diferencial directamente para los impulsos $\{p_i\}$. Sin embargo, en esta ecuación la función $H(p_i, q_i, t)$ depende de las coordenadas $\{q_i\}$, por lo que necesitamos también alguna manera de obtener la dependencia temporal de estas últimas. Para eso utilizamos la relación

$$\dot{q}_i = v_i(p_j, q_j, t) = \frac{\partial H}{\partial p_i}$$

Hemos obtenido entonces un conjunto completo de ecuaciones para describir el movimiento del sistema en términos del hamiltoniano $H(p_i, q_i, t)$ en lugar del lagrangiano $L(q_i, \dot{q}_i, t)$. Estas ecuaciones se leen

$$\dot{p}_i = - \frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

y se conocen como *ecuaciones de Hamilton*.

Es decir que hemos remplazado un conjunto de $3N$ ecuaciones diferenciales de segundo orden para las $3N$ variables $\{q_i\}$ dadas por las ecuaciones de Lagrange, por un conjunto de $6N$ ecuaciones

de primer orden para las $3N$ coordenadas y los $3N$ impulsos $\{(q_i, p_i)\}$ dadas por las ecuaciones de Hamilton. Por cada una de las ecuaciones necesitaremos una condición inicial, por lo que nuestro espacio de estados sigue siendo $\mathcal{E} = \mathbb{R}^{3N}$. En este sentido, es correcto decir que así como las ecuaciones de Lagrange describen el movimiento del sistema en el espacio de configuración $\mathcal{C} = \mathbb{R}^{3N}$, trazando una curva dada por las funciones $\{q_i(t)\}$ a medida que transcurre el tiempo, las ecuaciones de Hamilton lo describen en el espacio de estados \mathcal{E} en términos de la curva dada por $\{(q_i(t), p_i(t))\}$.

Ejemplo: relación con la energía

La fórmula del Hamiltoniano que hemos escrito más arriba resulta muy similar a la que escribimos en el capítulo 2 para la energía mecánica.

$$E(\dot{q}_i, q_i) = \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L(\dot{q}_i, q_i)$$

Sin embargo, en la fórmula anterior la energía es una función de las coordenadas y velocidades generalizadas. Para obtener su relación con el hamiltoniano, tenemos que usar la funciones $v_i(p_j, q_j, t)$ que nos dan las velocidades generalizadas como función de los impulsos, para escribir

$$H(p_i, q_i, t) = E(v_i(p_j, q_j, t), q_i, t)$$

Es decir que el hamiltoniano es en efecto la energía, pero escrita como función de los impulsos generalizados.

Por supuesto la construcción funciona en sentido contrario, usando el hecho de que una segunda transformación de Legendre nos devuelve la función original. En efecto, dado un Hamiltoniano $H(p_i, q_i, t)$ para un sistema mecánico, podemos definir la velocidad generalizada usando las ecuaciones de Hamilton según

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$

de donde podemos invertir para obtener

$$p_i = w_i(\dot{q}_i, q_i, t)$$

Con esto podemos recuperar el lagrangiano

$$L(\dot{q}_i, q_i, t) = w_i(\dot{q}_j, q_j, t) \dot{q}_i - H(w(\dot{q}_j, q_j, t), q_i, t)$$

del cual pueden entonces obtenerse las ecuaciones de Lagrange.

Ejercicio:

Demuestre que usando las ecuaciones de Hamilton y la construcción anterior para el lagrangiano, se recuperan las ecuaciones de Lagrange.

Ejemplo: partícula en una dimensión

Dado el lagrangiano para una partícula de masa m que se mueve en una dimensión bajo la acción de un potencial

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

Obtenemos el momento generalizado según la definición

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

esto se puede invertir para obtener la velocidad como función de los impulsos, en la forma

$$\dot{x} = \frac{p}{m} \equiv v(p)$$

Con esta fórmula, podemos computar la transformada de Legendre del lagrangiano para obtener el hamiltoniano, según

$$H(p, x, t) = p v(p) - L(v(p), x, t) = \frac{p^2}{2m} + V(x)$$

El cual resulta ser la energía mecánica escrita en términos del impulso, como demostramos más arriba para el caso general.

Las ecuaciones de Hamilton de este hamiltoniano se obtienen tomando las derivadas correspondientes, que resultan en

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m} \qquad \dot{p} = -\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x}$$

Resolviendo la primera ecuación para obtener p y reemplazando en la segunda, se recupera la segunda ley de Newton

$$m\ddot{x} = -\frac{\partial V}{\partial x}$$

con lo que vemos que la descripción es completamente equivalente.

En el estudio de un sistema mecánico cualquiera, en ocasiones puede resultar ventajosa su descripción en términos de ecuaciones de Lagrange, y en otras situaciones puede ser conveniente el uso de ecuaciones de Hamilton. Nos ocuparemos de describir las ventajas y desventajas de ambos enfoques en las clases por venir. Existe incluso un enfoque mixto, donde algunas de las coordenadas satisfacen ecuaciones de Lagrange, mientras que otras cumplen ecuaciones de Hamilton, como describiremos en la siguiente sección.

10.2.2 Ruthiano

Vimos más arriba que en el caso de una función de varias variables, podíamos definir una transformada de Legendre parcial respecto de un subconjunto de las mismas. Resulta natural

entonces preguntarse si, así como al transformar el lagrangiano respecto de todas las velocidades generalizadas obtenemos una descripción equivalente del movimiento, podríamos hacer lo mismo en términos de una transformada parcial.

Supongamos que tenemos un sistema mecánico de N partículas, descrito en términos de coordenadas generalizadas $\{q_i\}$ donde $i \in \{1, \dots, 3N\}$. Podemos separar las coordenadas en dos subconjuntos, el primero $\{q_k\}$ con $k \in \{1, \dots, D\}$ y el segundo $\{q_l\}$ con $l \in \{D+1, \dots, 3N\}$. Escribimos entonces los momentos generalizados del segundo subconjunto de acuerdo con su definición

$$p_l = \frac{\partial L}{\partial \dot{q}_l} \quad \text{donde } l \in \{D+1, \dots, 3N\}$$

Asumiendo que podemos despejar las velocidades generalizadas del subconjunto $\{\dot{q}_l\}$, tenemos que

$$\dot{q}_l = v_l(\dot{q}_k, p_l, q_i, t) \quad \text{donde } k \in \{1, \dots, D\}, l \in \{D+1, \dots, 3N\}$$

Con lo que podemos escribir la transformada parcial de Legendre del lagrangiano con respecto a un subconjunto particular de sus variables dado por las velocidades generalizadas $\{\dot{q}_l\}$ según

$$R(\dot{q}_k, p_l, q_i, t) = v_l(\dot{q}_k, p_l, q_i, t)p_l - L(\dot{q}_k, v_l(\dot{q}_k, p_l, q_i, t), q_i, t)$$

donde estamos sumando sobre el índice l . La transformada $R(\dot{q}_k, p_l, q_i)$ se denomina *ruthiano* del problema. En ella hemos reemplazado la dependencia del lagrangiano en el subconjunto $\{\dot{q}_l\}$ de las velocidades generalizadas, por la dependencia del ruthiano en los correspondientes impulsos $\{p_l\}$.

Repetiendo pasos análogos a los de la sección anterior, demostramos que las variables canónicas $\{(q_l, p_l)\}$ con $l \in \{D+1, \dots, 3N\}$ satisfacen ecuaciones de Hamilton

$$\dot{p}_l = -\frac{\partial R}{\partial q_l} \quad \dot{q}_l = \frac{\partial R}{\partial p_l} \quad \text{donde } l \in \{D+1, \dots, 3N\}$$

En cuanto a las coordenadas restantes $\{q_k\}$ con $k \in \{1, \dots, D\}$, estas satisfacen las ecuaciones de lagrange

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = - \left(\frac{d}{dt} \left(\frac{\partial R}{\partial \dot{q}_k} \right) - \frac{\partial R}{\partial q_k} \right) = 0 \quad \text{donde } k \in \{1, \dots, D\}$$

donde utilizamos el hecho de que en la transformada de Legendre que define el ruthiano, las velocidades generalizadas $\{\dot{q}_k\}$ y la totalidad de las coordenadas generalizadas $\{q_i\}$ cumplen el rol de parámetros, y que por lo tanto se cumple que

$$\frac{\partial R}{\partial q_k} = -\frac{\partial L}{\partial q_k} \quad \frac{\partial R}{\partial \dot{q}_k} = -\frac{\partial L}{\partial \dot{q}_k}$$

En estas fórmulas, las derivadas parciales del lado derecho se calculan manteniendo las velocidades $\{\dot{q}_l\}$ constantes, mientras que las derivadas del lado izquierdo se toman manteniendo los impulsos $\{p_l\}$ constantes.

Es decir que en términos del routhiano, un subconjunto de las coordenadas generalizadas $\{q_k\}$ con $k \in \{1, \dots, D\}$ cumplen ecuaciones de Lagrange, mientras las coordenadas restantes se acomodan en pares canónicos con los correspondientes impulsos $\{(q_l, p_l)\}$ con $l \in \{D+1, \dots, 3N\}$ y satisfacen ecuaciones de Hamilton.

10.3 Truncaciones

En el caso particular en el que las coordenadas $\{q_l\}$ con $l \in \{D+1, \dots, 3N\}$ son cíclicas, el correspondiente routhiano resulta de utilidad para describir el movimiento del sistema en un espacio de configuración reducido. En efecto, la primera ecuación de Hamilton para $R(\dot{q}_k, p_l, q_k, t)$ toma la forma

$$\dot{p}_l = \frac{\partial R}{\partial q_l} = 0$$

con lo que los impulsos $\{p_l\}$ se conservan. Por supuesto, esto ya era explícito en las ecuaciones de lagrange de $L(q_i, \dot{q}_i)$. La ventaja del nuevo formalismo es que nos permite reemplazar los valores constantes de los $\{p_l\}$ directamente en el routhiano, algo que no podríamos hacer en el lagrangiano original.

Las ecuaciones de movimiento para las coordenadas restantes $\{q_k\}$ con $k \in \{1, \dots, D\}$ serán las ecuaciones de Lagrange que se obtienen a partir de $R(\dot{q}_k, p_l, q_k, t)$. Esto implica que podemos definir un lagrangiano truncado en la forma

$$L_{\text{trunc}}(q_k, \dot{q}_k, t) = -R(q_k, \dot{q}_k, p_l, t)$$

Este lagrangiano describe un sistema en dimensión $D < 3N$, con lo que hemos en efecto reducido la dimensionalidad del espacio de configuración \mathcal{C} .

Nótese que la truncación se puede llevar a cabo incluso cuando las coordenadas $\{q_l\}$ no son cíclicas, siempre que exista una solución donde tanto ellas como los $\{p_l\}$ sean constantes. Otra observación es que podemos sumar a L_{trunc} una función arbitraria de los $\{p_l\}$ constantes, lo que no afectará las ecuaciones de Lagrange.

Ejemplo: problema equivalente de una partícula en el problema de dos cuerpos

En el caso del problema de dos cuerpos, comenzamos con un sistema de dos partículas con un espacio de configuración de seis dimensiones $\mathcal{C} = \mathbb{R}^6$, cuyo lagrangiano se puede escribir, en coordenadas esféricas, en la forma

$$L(\vec{r}_{\text{cm}}, \dot{\vec{r}}, \vec{r}) = \frac{1}{2} m_{\text{tot}} \dot{\vec{r}}_{\text{cm}}^2 + \frac{1}{2} \mu (\dot{r}^2 + r^2 (\dot{\phi}^2 + \dot{\theta}^2 \sin^2 \phi)) - V(r)$$

Lo que nos permite identificar el momento lineal del centro de masas como

$$\vec{p}_{\text{cm}} = m_{\text{tot}} \dot{\vec{r}}_{\text{cm}}$$

Esto se puede resolver para la velocidad del centro de masas

$$\dot{\vec{r}}_{\text{cm}} = \frac{\vec{p}_{\text{cm}}}{m_{\text{tot}}} \equiv \vec{v}(\vec{p}_{\text{cm}})$$

Con esto, el ruthiano se lee

$$R(\dot{\vec{r}}, \vec{p}_{\text{cm}}, \vec{r}) = \vec{p}_{\text{cm}} \cdot \vec{v}(\vec{p}_{\text{cm}}) - L(\vec{v}(\vec{p}_{\text{cm}}), \dot{\vec{r}}, \vec{r})$$

Explícitamente

$$R(\dot{\vec{r}}, \vec{p}_{\text{cm}}, \vec{r}) = \frac{\vec{p}_{\text{cm}}^2}{2m_{\text{tot}}} - \frac{1}{2}\mu (\dot{r}^2 + r^2 (\dot{\phi}^2 + \dot{\theta}^2 \sin^2 \phi)) + V(r)$$

Dado que \vec{p}_{cm} se conserva, podemos reemplazar su valor constante en el ruthiano y definir el lagrangiano truncado

$$L_{1\text{part}}(\dot{\vec{r}}, \vec{r}) = -R(\dot{\vec{r}}, \vec{p}_{\text{cm}}, \vec{r}) + \frac{\vec{p}_{\text{cm}}^2}{2m_{\text{tot}}} = \frac{1}{2}\mu (\dot{r}^2 + r^2 (\dot{\phi}^2 + \dot{\theta}^2 \sin^2 \phi)) - V(r)$$

donde sumamos $\vec{p}_{\text{cm}}^2/2m_{\text{tot}}$ para eliminar una constante aditiva que no afecta las ecuaciones de movimiento para \vec{r} . Este problema tiene un espacio de configuración descrito por \vec{r} , es decir que $\mathcal{C} = \mathbb{R}^3$. Nótese que si hubiéramos reemplazado $\dot{\vec{r}}_{\text{cm}}$ directamente en L hubiéramos obtenido un término adicional constante.

Ejemplo: problema equivalente en el plano para el problema de dos cuerpos

Si ahora estudiamos el sistema de una partícula que acabamos de obtener vemos que el momento generalizado asociado a ϕ toma la forma

$$p_{\phi} = \mu r^2 \dot{\phi}$$

lo que nos permite despejar

$$\dot{\phi} = \frac{p_{\phi}}{\mu r^2} \equiv v_{\phi}(p_{\phi}, r)$$

Con esto podemos encontrar el ruthiano correspondiente como

$$R(\dot{r}, \dot{\theta}, p_{\phi}, r, \phi) = p_{\phi} v_{\phi}(p_{\phi}, r) - L_{1\text{part}}(\dot{r}, \dot{\theta}, v_{\phi}(r, p_{\phi}), r, \phi)$$

es decir

$$R(\dot{r}, \dot{\theta}, p_{\phi}, r, \phi) = \frac{p_{\phi}^2}{2\mu r^2} - \frac{1}{2}\mu (\dot{r}^2 + r^2 \dot{\theta}^2 \sin^2 \phi) + V(r)$$

las correspondientes ecuaciones de Hamilton para p_{ϕ}, ϕ se pueden escribir

$$\dot{p}_{\phi} = \mu r^2 \dot{\theta}^2 \sin \phi \cos \phi$$

$$\dot{\phi} = \frac{p_{\phi}}{\mu r^2}$$

donde es evidente que existe una solución con $\phi = \pi/2$ y $p_{\phi} = 0$ constantes. Reemplazando estos valores en el ruthiano obtenemos el lagrangiano truncado

$$L_{2d}(\dot{r}, \dot{\theta}, r) = -R(\dot{r}, \dot{\theta}, 0, r, \pi/2) = \frac{1}{2}\mu (\dot{r}^2 + r^2 \dot{\theta}^2) - V(r)$$

El espacio de configuración para este problema es $\mathcal{C} = \mathbb{R}^2$, con lo que de nuevo hemos reducido la dimensión del problema.

Ejemplo: problema unidimensional equivalente en el problema de dos cuerpos

Dado el lagrangiano efectivo L_{2d} podemos obtener el momento generalizado asociado a la coordenada θ según

$$p_\theta = \mu r^2 \dot{\theta}$$

De aquí obtuvimos una expresión para la velocidad angular en la forma

$$\dot{\theta} = \frac{p_\theta}{\mu r^2} \equiv v(\dot{\theta}, p_\theta, r)$$

Con esto el ruthiano se lee

$$R(\dot{r}, p_\theta, r) = p_\theta v(\dot{\theta}, p_\theta, r) - L_{2d}(\dot{r}, v(\dot{\theta}, p_\theta, r), r)$$

es decir

$$R(\dot{r}, p_\theta, r) = -\frac{1}{2}\mu\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + V(r)$$

La ecuación de Hamilton para p_θ implica que toma un valor constante $p_\theta = \ell$, con lo que podemos reemplazarlo en el ruthiano para construir el lagrangiano truncado

$$L_{1d}(\dot{r}, r) = -R(\dot{r}, \ell, r) = \frac{1}{2}\mu\dot{r}^2 - \underbrace{\left(\frac{\ell^2}{2\mu r^2} + V(r)\right)}_{V_{1d}} = \frac{1}{2}\mu\dot{r}^2 - V_{1d}$$

Que corresponde a un sistema de una partícula en una dimensión, descrito en términos de un espacio unidimensional $\mathcal{C} = \mathbb{R}$. Nótese que este lagrangiano no coincide con el que se obtendría de reemplazar $\dot{\theta}$ directamente en L_{2d} .

Ejemplo: sistema unidimensional equivalente para el trompo

Para el caso del trompo, escribimos el lagrangiano en la forma

$$L(\vartheta, \dot{\vartheta}, \dot{\varphi}, \dot{\psi}, \vartheta) = \frac{1}{2}\mathcal{J}_x(\dot{\vartheta}^2 + \sin^2 \vartheta \dot{\varphi}^2) + \frac{1}{2}\mathcal{J}_z(\dot{\psi} + \dot{\varphi} \cos \vartheta)^2 - mgl \cos \vartheta$$

Los momentos generalizados asociados a las variables φ y ψ se pueden escribir según

$$p_\varphi = \mathcal{J}_z(\dot{\psi} + \dot{\varphi} \cos \vartheta) \cos \vartheta + \mathcal{J}_x \sin^2 \vartheta \dot{\varphi}$$

$$p_\psi = \mathcal{J}_z(\dot{\psi} + \dot{\varphi} \cos \vartheta)$$

se conservan. Los usamos para obtener las velocidades generalizadas, según

$$\dot{\varphi} = \frac{p_\varphi - p_\psi \cos \vartheta}{\mathcal{J}_x \sin^2 \vartheta} \equiv v_\varphi(\vartheta, p_\varphi, p_\psi)$$

$$\dot{\psi} = \frac{p_\psi}{\mathcal{J}_z} - \frac{p_\varphi - p_\psi \cos \vartheta}{\mathcal{J}_x \sin^2 \vartheta} \cos \vartheta \equiv v_\psi(\vartheta, p_\varphi, p_\psi)$$

El correspondiente ruthiano tomará entonces la forma

$$R(\dot{\vartheta}, p_\varphi, p_\psi) = p_\varphi v_\varphi(\vartheta, p_\varphi, p_\psi) + p_\psi v_\psi(\vartheta, p_\varphi, p_\psi) - L(\dot{\vartheta}, v_\varphi(\vartheta, p_\varphi, p_\psi), v_\psi(\vartheta, p_\varphi, p_\psi), \vartheta)$$

o en otras palabras

$$R(\dot{\vartheta}, p_\varphi, p_\psi) = - \left(\frac{1}{2} \mathcal{J}_x \dot{\vartheta}^2 - \left(\frac{1}{2 \mathcal{J}_x} \left(\frac{p_\varphi - p_\psi \cos \vartheta}{\sin \vartheta} \right)^2 + mgl \cos \vartheta \right) \right)$$

Dado que los impulsos p_φ, p_ψ se conservan, según se deduce de las correspondientes ecuaciones de Hamilton, podemos reemplazar sus valores constantes para definir el lagrangiano truncado unidimensional según $L_{1d} = -R$, obteniendo

$$L_{1d}(\dot{\vartheta}, \vartheta) = \frac{1}{2} \mathcal{J}_x \dot{\vartheta}^2 - \underbrace{\left(\frac{1}{2 \mathcal{J}_x} \left(\frac{p_\varphi - p_\psi \cos \vartheta}{\sin \vartheta} \right)^2 + mgl \cos \vartheta \right)}_{V_{1d}} = \frac{1}{2} \mathcal{J}_x \dot{\vartheta}^2 - V_{1d}$$

Con esto, reducimos el espacio de configuración para quedarnos con $\mathcal{C} = S_1/\mathbb{Z}_2$. Es importante mencionar que este lagrangiano no corresponde al que se obtendría de reemplazar ϕ y ψ directamente en L .

En consecuencia, en todos los casos que hemos explorado previamente, cada vez que pudimos escribir un lagrangiano truncado que redujera el número de grados de libertad, estábamos en realidad obteniendo un ruthiano respecto de los impulsos conservados. Es importante señalar que el reemplazar los valores constantes de los impulsos en el correspondiente lagrangiano no lleva al resultado correcto. Esto sucede porque las ecuaciones de Lagrange se obtienen tomando derivadas parciales en las que son las velocidades y no los impulsos las que se mantienen constantes.

10.4 Resumen

En esta clase aprendimos que dado un lagrangiano cualquiera, podemos construir un hamiltoniano como su transformada de Legendre respecto de las velocidades generalizadas. Tal función también permite describir el movimiento, en términos de ecuaciones de Hamilton para las correspondientes coordenadas e impulsos generalizados.

En nuestra construcción, asumimos que siempre era posible despejar las velocidades generalizadas en términos de las variables canónicas. Posponemos para la última clase la discusión de qué sucede cuando tal condición no se cumple.

Vimos también que la transformada de Legendre puede realizarse respecto de un subconjunto particular de las velocidades generalizadas, para obtener un ruthiano. Las ecuaciones de movimiento se obtienen entonces como ecuaciones de Hamilton para el subconjunto correspondiente de coordenadas e impulsos generalizados, y ecuaciones de lagrange para las coordenadas generalizadas restantes.

Cuando existen soluciones con impulsos constantes, lo que sucede por ejemplo cuando las coordenadas correspondientes son cíclicas, podemos reemplazar directamente tales valores constantes en el ruthiano y obtener un lagrangiano truncado que describe el movimiento del resto de las coordenadas generalizadas.

11. Vínculos

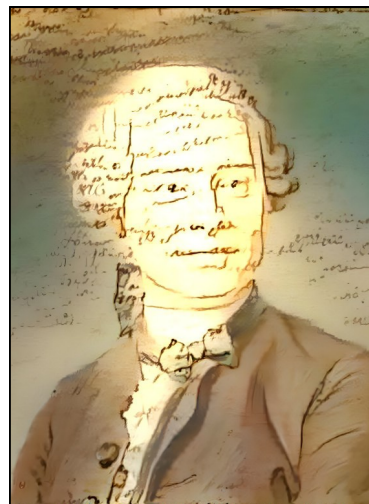
11.1 Objetivos

Comenzamos este curso con un sistema de N partículas cuyas posiciones estaban determinadas por vectores \vec{r}_n . Como consecuencia, nuestro sistema físico tenía $3N$ grados de libertad, cada uno de los cuales tomaba valores en la recta real. Es decir que su espacio de configuración estaba dado por $\mathcal{C} = \mathbb{R}^{3N}$. Al describir el movimiento en coordenadas generalizadas, simplemente cambiamos de coordenadas en el mismo espacio $\mathcal{C} = \mathbb{R}^{3N}$.

Sin embargo, cuando estudiamos más adelante el cuerpo rígido, lo definimos como un sistema de partículas cuyas distancias relativas están fijas. Para esto, usamos como coordenadas generalizadas la posición de una de tales partículas $\vec{r}_1 \in \mathbb{R}^3$, la orientación del cuerpo respecto de un sistema de ejes de referencia $\{\phi, \theta, \psi\} \in SO(3)$, y las distancias entre las diferentes partículas que lo componen $\{l_1, l_2, l_3, l_{n1}, l_{n2}, l_{n3}\}$. Acto seguido, impusimos la condición de que el valor de tales distancias permaneciera fijo. Es decir que redujimos el espacio de configuración del sistema a $\mathcal{C} = \mathbb{R}^3 \times SO(3)$, que es una variedad no trivial dentro del espacio original \mathbb{R}^{3N} .

Ahora bien ¿es correcto hacer esta reducción? ¿Cómo sabemos que al dejar fijas las distancias se cumplen (siquiera de forma aproximada) las ecuaciones de movimiento que obtendríamos del lagrangiano tomando las derivadas correspondientes?

A continuación vamos a estudiar las condiciones bajo las cuales tales *vínculos* tienen sentido físico. Motivaremos la intuición física en base a ejemplos, y luego la generalizaremos.



Jean-Baptiste le Rond d'Alembert

11.2 Origen físico de los vínculos

Un *vínculo* aparece cuando hay una función de las coordenadas y las velocidades generalizadas que se debe mantener constante durante el movimiento. Por ejemplo, la longitud de la soga de un péndulo o el ángulo de un plano inclinado. En lo que sigue, veremos cuál es el origen microscópico de tales restricciones, y cómo se incorporan en el formalismo lagrangiano.

11.2.1 Vínculos holonómicos y coordenadas adaptadas

Para entender qué es un vínculo, es conveniente ilustrarlo con una serie de ejemplos, que luego extenderemos al caso general. Comenzaremos en una dimensión, y luego analizaremos ejemplos bidimensionales en coordenadas cartesianas y polares.

Ejemplo: partícula en una dimensión

Comencemos con un ejemplo sencillo: sea un sistema de una partícula en una dimensión, sometida a una fuerza conservativa. El lagrangiano toma la forma

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

La energía de este sistema vendrá dada por

$$E = \frac{1}{2}m\dot{x}^2 + V(x)$$

De aquí se puede despejar la velocidad, según

$$\dot{x} = \sqrt{\frac{2(E - V(x))}{m}}$$

donde vemos que será real en las regiones en las cuales $V(x) < E$. Esto es la forma matemática de la intuición sencilla de que el sistema se moverá en la región en la que su energía cinética sea mayor o igual a cero.

Ahora bien, supongamos que el potencial tiene la forma de *pared infinita*, es decir

$$V(x) = \begin{cases} 0 & x \leq x_+ \\ \infty & x > x_+ \end{cases}$$

En este caso es evidente que el sistema se moverá libremente pero confinado a la región $x \leq x_+$ a lo largo de toda su evolución temporal. Esto es lo que se conoce como un *vínculo* (o *ligadura*) *unilateral*.

Sin embargo, un potencial de este tipo puede parecer muy poco realista. Para acercarlo un poco más a algo realizable, supongamos que tenemos un potencial $V(x)$ que tiene una asíntota vertical, es decir

$$V(x) \xrightarrow{x \rightarrow x_+} \infty$$

Ahora el potencial no es algo tan rudo como una pared infinita, pero de nuevo el sistema se moverá en la región $x \leq x_+$. Si el potencial es pequeño cuando $x \ll x_+$ y

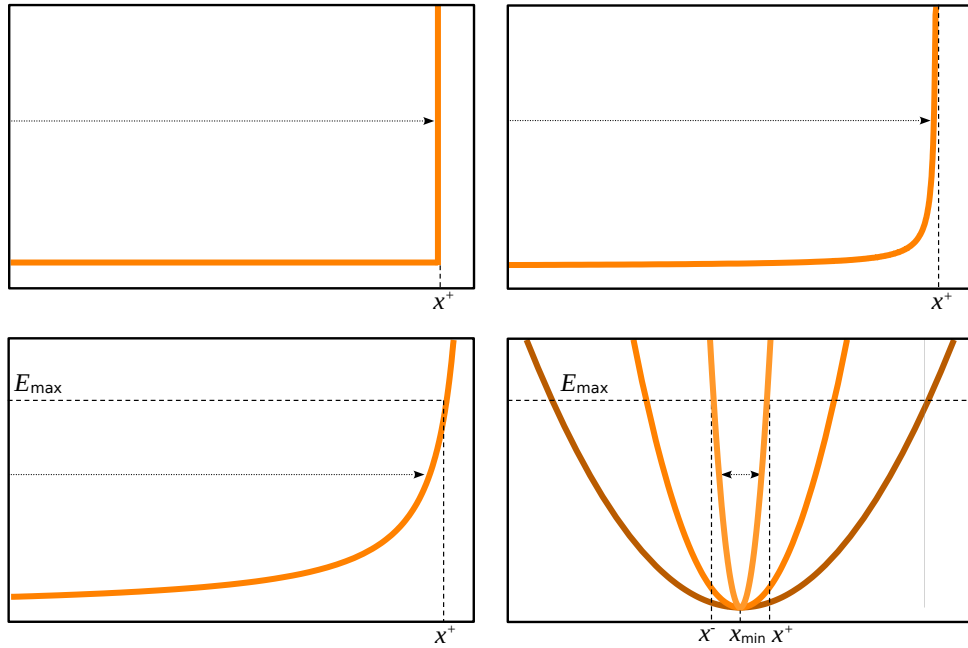


Figura 11.2.1: *Vínculos en una dimensión. Un potencial con una pared no permite que la partícula pase de x_+ , y lo mismo es cierto para un potencial con una asíntota (arriba). En un caso realista, la energía está acotada lo que implica que la partícula no puede pasar de x_+ (abajo a la izquierda). Si el potencial tiene un mínimo la partícula se moverá en la región $x_- < x < x_+$, y si el mínimo es muy marcado estará confinada a $x \approx x_{\min}$.*

su crecimiento es muy rápido cuando nos acercamos a $x \sim x_+$, podemos decir que el sistema se moverá libremente confinado a la región $x \leq x_+$, realizando entonces un vínculo unilateral.

En una situación real no es necesario que haya realmente una asíntota. En efecto, dado que cualquier sistema tiene una energía máxima finita E_{\max} , su movimiento estará limitado a la región del espacio donde se satisface $V(x) < E < E_{\max}$. Si suponemos que el potencial crece con x (si decrece, simplemente podemos cambiar x por $-x$ en la discusión que sigue) entonces habrá un punto x_+ donde se cumple que

$$V(x_+) = E_{\max}$$

Por lo que es correcto decir que la evolución temporal del sistema cumplirá $x \leq x_+$. Si el crecimiento es muy rápido al acercarnos a x_+ y el potencial es pequeño lejos de ese punto, entonces esto realiza un vínculo unilateral en un sistema real.

Ahora bien, si se cumple que a la izquierda de x_+ el potencial tiene un mínimo, digamos en $x = x_{\min}$, y al continuar hacia la izquierda en la dirección $x < x_{\min}$ el potencial también crece muy rápidamente, entonces existirá también un segundo punto x_- tal que

$$V(x_-) = E_{\max}$$

lo que por su parte implica el vínculo unilateral $x \geq x_-$. Por lo tanto es correcto decir que el sistema se moverá alrededor de x_{\min} entre dos extremos $x_- < x < x_+$.

Si el potencial crece muy rápidamente cuando x se aleja de x_{\min} en ambas direcciones, esto que implica que la energía alcanzará el valor E_{\max} en valores de x que son muy cercanos a x_{\min} tanto a su derecha como a su izquierda. En otras palabras, tendremos que $x_+ \approx x_-$, por lo que podemos decir que el sistema está *confinado* al punto

$$x = x_{\min}$$

Esto es lo que se conoce como un *vínculo* (o *ligadura*) *bilateral*.

Ejemplo: partícula en una dimensión, forma alternativa

Si el potencial es diferenciable en el mínimo, hay una manera alternativa de entender el ejemplo anterior. Podemos escribir un desarrollo en serie de Taylor alrededor del mínimo según

$$V(x) \approx V(x_{\min}) + \frac{1}{2}V''(x_{\min})(x - x_{\min})^2 + \mathcal{O}(x - x_{\min})^3$$

Por lo tanto, vemos que cerca del mínimo el sistema se comporta como un oscilador armónico con constante restauradora $k = V''(x_{\min})$. La solución para el movimiento de dicho oscilador armónico es la que conocemos de los cursos básicos

$$x = x_{\min} + A \cos \left(\sqrt{\frac{V''(x_{\min})}{m}} (t - t_0) \right)$$

siendo A es la amplitud de la oscilación. Podemos obtener entonces la energía del sistema usando la fórmula del oscilador armónico $E = kA^2/2$, es decir

$$E = \frac{1}{2}V''(x_{\min})A^2 \leq E_{\max}$$

Donde hemos escrito explícitamente la cota máxima para la energía. Invirtiendo la expresión anterior podemos obtener una cota para la amplitud de oscilación, con la forma

$$A \leq \sqrt{\frac{2E_{\max}}{V''(x_{\min})}}$$

Cualquier sistema real tiene una cantidad máxima de energía E_{\max} que es finita. Por lo tanto, la amplitud de su movimiento será más pequeña cuanto mayor sea la constante restauradora $V''(x_{\min})$. Para una constante $V''(x_{\min})$ lo bastante grande, es una buena aproximación afirmar que la amplitud es nula, es decir que el sistema no oscila y está fijo en su punto de equilibrio $x = x_{\min}$. Esta es otra manera de comprender cómo aparece un vínculo bilateral.

Ejemplo: partícula en dos dimensiones, con potencial unidimensional

Un razonamiento completamente análogo funciona cuando el sistema es bidimensional, pero el potencial depende de una sola de las coordenadas.

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x)$$

La energía de este sistema vendrá dada por

$$E = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + V(x)$$

En cualquier situación real, el sistema tiene una energía máxima finita E_{\max} , es decir que está confinado a moverse en la región del espacio \mathbb{R}^2 donde se satisface $V(x) < E < E_{\max}$. Si el potencial es creciente en ambas direcciones $x \rightarrow \pm\infty$, esto implica la existencia de los vínculos unilaterales

$$x_- < x < x_+$$

A medida que el potencial se haga más empinado, creciendo más rápidamente cuando nos alejamos hacia la izquierda o hacia la derecha del mínimo, el sistema estará cada vez más limitado en su movimiento en la dirección x , obteniéndose en el límite el vínculo bilateral

$$x = x_{\min}$$

En otras palabras, el sistema reduce su dimensión, estando ahora confinado a moverse solamente en y .

Veamos ahora si podemos encontrar un lagrangiano efectivo para describir el movimiento en la variable y . Para esto escribimos las ecuaciones de movimiento, en la forma

$$m\ddot{x} = -\partial_x V(x)$$

$$m\ddot{y} = 0$$

La condición $x = x_{\min}$ resuelve la primera, mientras que la segunda se puede obtener del lagrangiano efectivo

$$L_{\text{eff}} = \frac{1}{2}m\dot{y}^2$$

Este lagrangiano es, a menos de una constante que no contribuye a las ecuaciones de movimiento, el que se hubiera obtenido de truncar el lagrangiano inicial reemplazando directamente en él la condición de vínculo $x = x_{\min}$.

Este resultado puede parecer trivial, ya que la dimensión extra no juega ningún rol en la discusión y estamos en realidad tratando con un problema unidimensional. Veremos a continuación un caso más general genuinamente bidimensional.

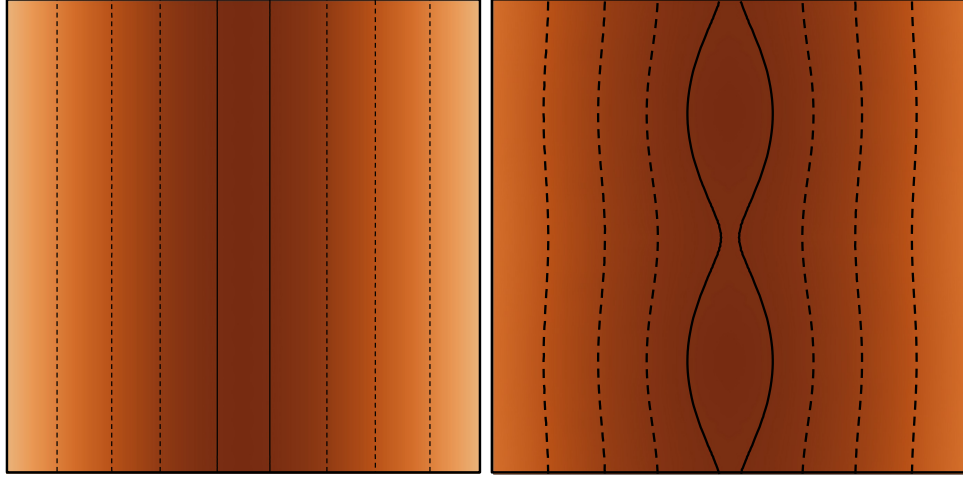


Figura 11.2.2: Vínculos en dos dimensiones. Vemos las curvas de nivel de un potencial que depende de una sola coordenada cartesiana (izquierda), y de uno que depende de ambas coordenadas cartesianas, pero más fuertemente de una que de la otra (derecha). Si el pozo de potencial es lo bastante marcado, el movimiento se producirá entre las líneas continuas.

Ejemplo: partícula dos dimensiones, caso más general

La construcción del ejemplo anterior se puede generalizar al caso en el que el potencial depende de las dos variables, pero mucho más fuertemente de una (digamos x) que de la otra. Comenzamos con el lagrangiano

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y)$$

Asumiendo que el potencial crece en ambas direcciones de la variable x , la condición de energía acotada $V(x, y) < E < E_{\max}$ puede ser resuelta de la siguiente forma

$$x_-(y) < x < x_+(y)$$

donde ahora $x_{\pm}(y)$ son funciones de la variable y que corresponden a las curvas de nivel $V(x_{\pm}(y), y) = E_{\max}$. La condición de que el potencial dependa mucho más fuertemente de x que de y se traduce en que las funciones $x_{\pm}(y)$ son muy suaves, cambiando perceptiblemente en una distancia Δy que es mucho mayor que $\Delta x(y) = x_+(y) - x_-(y)$.

De nuevo, los vínculos a izquierda y derecha se irán acercando a medida que el potencial se hace más empinado en la dirección x , resultando finalmente en el confinamiento del sistema, que sólo podrá moverse en una dimensión, a lo largo del vínculo bilateral

$$x = x_{\min}$$

Escribiendo las ecuaciones de movimiento que se obtienen de nuestro lagrangiano

$$m\ddot{x} = -\partial_x V(x, y)$$

$$m\ddot{y} = -\partial_y V(x, y)$$

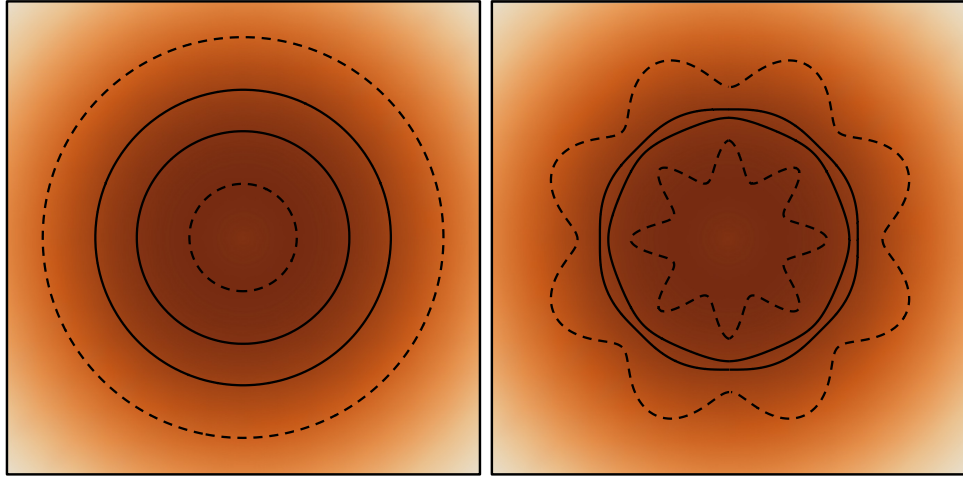


Figura 11.2.3: *Vínculos en dos dimensiones. Vemos las curvas de nivel de un potencial que depende del radio polar (izquierda), y de uno que depende también del ángulo, pero más débilmente que del radio (derecha). Si el pozo de potencial es lo bastante marcado, el movimiento se producirá entre las líneas continuas.*

Podemos reemplazar en ellas el vínculo, para obtener

$$0 = -\partial_x V(x_{\min}, y)$$

$$m\ddot{y} = -\partial_y V(x_{\min}, y)$$

Vemos que la primera es la condición de mínimo en la dirección x y se satisface automáticamente, mientras que la segunda se puede obtener a partir del lagrangiano

$$L_{\text{eff}} = \frac{1}{2}m\dot{y}^2 - V(x_{\min}, y)$$

Nuevamente, este lagrangiano coincide con lo que se obtiene de reemplazar directamente en el lagrangiano original la condición de vínculo.

Nótese que en los cálculos anteriores hemos supuesto que las dos curvas $x_{\pm}(y)$ se juntan en una recta $x = x_{\min}$ cuando el potencial se hace más empinado. Este no tiene por qué ser el caso, podríamos tener una condición de vínculo que también está dada por una curva $x = x_{\min}(y)$. Para entenderlo mejor podemos cambiar coordenadas, como haremos en el ejemplo que sigue.

Ejemplo: partícula en dos dimensiones en coordenadas polares

Supongamos una partícula en el plano descrita en coordenadas polares, con un potencial que depende solo del radio

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

Si el potencial tiene un mínimo en $r = r_{\min}$, el hecho de que siempre habrá una cantidad finita de energía $E < E_{\max}$ hace que el sistema se mueva en una región acotada por

el vínculo unilateral $r_- < r < r_+$. Cuando el potencial sea muy empinado, el sistema deberá satisfacer el vínculo bilateral $r = r_{\min}$. Esta restricción establece una curva en el plano, a diferencia de las rectas de los ejemplos anteriores.

Escribiendo las ecuaciones de movimiento

$$m\ddot{r} = -\partial_r \left(V(r) - \frac{1}{2}mr^2\dot{\theta}^2 \right)$$

$$\frac{d}{dt} (mr^2\dot{\theta}) = 0$$

podemos evaluarlas en un valor constante del radio $r = r_{\text{sol}}$ para verificar si son consistentes con una condición de vínculo

$$0 = -\partial_r \left(V(r) - \frac{1}{2}mr^2\dot{\theta}^2 \right) \Big|_{r=r_{\text{sol}}}$$

$$\frac{d}{dt} (mr_{\text{sol}}^2\dot{\theta}) = 0$$

La primera ecuación es una condición de mínimo que determina r_{sol} , pero modificada con un término rotacional que contiene $\dot{\theta}^2$. Para ver los efectos de este término, desarrollamos el miembro derecho alrededor del mínimo de $V(r)$, que habíamos llamado r_{\min} , obteniendo

$$0 = (m\dot{\theta}^2 - V''(r_{\min}))(r_{\text{sol}} - r_{\min}) + mr_{\min}\dot{\theta}^2$$

En esta ecuación es importante notar que, dado que la energía está acotada, también lo estará su componente rotacional, proporcional a $\dot{\theta}^2$. Esto significa que para un potencial lo bastante empinado, el término que contiene la derivada segunda $V''(r_{\min})$ domina la ecuación

$$0 = -V''(r_{\min})(r_{\text{sol}} - r_{\min})$$

Lo que se resuelve con $r_{\text{sol}} = r_{\min}$. En otras palabras, el sistema está confinado en el mínimo del potencial, y el término rotacional no altera este hecho.

La segunda ecuación de movimiento se puede obtener del lagrangiano para una partícula que se mueve en un círculo

$$L_{\text{eff}} = \frac{1}{2}mr_{\min}^2\dot{\theta}^2$$

Nuevamente, este lagrangiano se obtiene de reemplazar directamente la condición $r = r_{\min}$ en el lagrangiano original.

Como en el ejemplo anterior, podría existir una dependencia explícita del potencial en la variable adicional, en este caso θ , siempre que sea mucho más suave que la dependencia en la variable vinculada, en este caso r . Si así fuera, siguiendo pasos completamente análogos, obtendríamos un lagrangiano efectivo

$$L_{\text{eff}} = \frac{1}{2}mr_{\min}^2\dot{\theta}^2 - V(r_{\min}, \theta)$$

Que describe un problema unidimensional.

Ejercicio:

Estudie el caso de un cuerpo que desciende por un plano inclinado. El lagrangiano será

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy - V(y \cos \theta - x \sin \theta)$$

Además del término gravitacional, el potencial tiene un término adicional que es el que provee la condición de vínculo, creciendo muy rápidamente en la dirección transversal al plano ¿Cuál es la fuerza que se deriva de este potencial? ¿Se trata de un vínculo unilateral o bilateral?

Escriba un lagrangiano efectivo para describir el sistema unidimensional.

Ejercicio:

Estudie el caso de un péndulo plano. El lagrangiano se puede escribir en la forma

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy - V(x^2 + y^2)$$

En este sistema, la longitud de la cuerda se mantiene invariante, lo que significa que además del término gravitacional el potencial tiene un término con un crecimiento muy marcado en la dirección en la que se estira la cuerda ¿cómo se llama la fuerza que se obtiene a partir de tal potencial?

Reemplace la condición de vínculo en el lagrangiano, para obtener un lagrangiano efectivo capaz de describir el sistema truncado.

La lección de los ejemplos anteriores es que, cuando el potencial presenta un mínimo muy marcado a lo largo de alguna coordenada, podemos reemplazar directamente en el lagrangiano el valor de la misma en el mínimo. El lagrangiano efectivo resultante nos proporcionará ecuaciones de movimiento para las coordenadas que restan, que son las mismas que las que hubiéramos obtenido del lagrangiano original.

Esta idea se puede extender a un caso general. Escribamos el lagrangiano para un sistema de N partículas en términos de las $3N$ coordenadas cartesianas \vec{r}_n que parametrizan el espacio de configuración \mathbb{R}^{3N} . Éste toma la forma

$$L = K(\dot{\vec{r}}_n^2) - V(\vec{r}_n)$$

La energía del sistema está acotada, y por lo tanto se cumple que $V(\vec{r}_n) < E < E_{\max}$. Al igual que en los ejemplos, esto establece una región dentro de \mathbb{R}^{3N} en la cual se puede mover el sistema, es decir un vínculo unilateral.

Supongamos ahora que, dentro de la región determinada por el vínculo unilateral, el potencial crece muy rápidamente al alejarnos de alguna hipersuperficie de dimensión D , sobre la cual tiene además una variación suave. Entonces podemos limitarnos a describir el movimiento exclusivamente a

lo largo de esa hipersuperficie, dado que prácticamente no habrá movimiento en las direcciones transversales. Como cualquier hipersuperficie D -dimensional embebida en \mathbb{R}^{3N} , tal *hipersuperficie de vínculo* estará descrita por un conjunto de restricciones de la forma

$$f_l(\vec{r}_n) = c_l \text{ constante, con } l \in \{D+1, \dots, 3N\}$$

Este tipo de restricciones se conocen como vínculos *holonómicos*. Notemos que, en presencia de estos vínculos holonómicos, el espacio de configuración original \mathbb{R}^{3N} se redujo a una hipersuperficie $\mathcal{C} \subset \mathbb{R}^{3N}$ de dimensión $D \leq 3N$. La misma puede tener una forma y topología arbitrarias, por lo que se suele hablar de una *variedad* de configuración \mathcal{C} .

Ahora bien, sabemos que podemos describir el sistema en términos de un conjunto cualquiera de coordenadas generalizadas $\{q_i\}$ con $i \in \{1, \dots, 3N\}$, las cuales podemos elegir a nuestra conveniencia. En términos de ellas, el lagrangiano se escribe

$$L = K(\dot{q}_i, q_i) - V(q_i)$$

Una elección particularmente útil de las coordenadas generalizadas es la que se conoce como *coordenadas adaptadas*. Se define según la condición

$$q_k = q_k(\vec{r}_n) \text{ arbitrario para } k \in \{1, \dots, D\}$$

$$q_l = f_l(\vec{r}_n) \text{ para } l \in \{D+1, \dots, 3N\}$$

Se dice que estas coordenadas *resuelven los vínculos*. En efecto, al movernos a lo largo de las coordenadas q_k dejando las demás fijas el potencial varía suavemente, mientras que a lo largo de q_l existe un mínimo marcado en los valores $q_l = c_l$ y la dependencia del potencial al alejarnos es mucho más fuerte. Esto implica que podemos dejar fijos los valores de tales coordenadas en el mínimo de potencial $q_l = c_l$. De manera análoga a los ejemplos de más arriba, el lagrangiano efectivo definido según

$$L_{\text{eff}} = K(\dot{q}_k, \dot{q}_l = 0, q_k, q_l = c_l) - V(q_k, q_l = c_l)$$

dará una descripción del movimiento consistente con las ecuaciones del lagrangiano original.

Nota:

Este procedimiento es exactamente el que seguimos, sin hacerlo explícito, cuando hablamos del cuerpo rígido.

En efecto, las coordenadas a lo largo de las cuales el potencial presenta mínimos marcados (y que por lo tanto podemos eliminar de nuestra descripción fijándolas a un valor constante) son las distancias entre las partículas $\{q_l\} = \{l_1, l_2, l_3, l_{1n}, l_{2n}, l_{3n}\}$. Esto constituye una realización física de la idea de *rigidez*. Por otro lado, aquéllas a lo largo de las cuales el potencial presenta una dependencia suave resultan ser los grados de libertad del cuerpo rígido $\{q_i\} = \{\vec{r}_1, \phi, \theta, \psi\}$.

El espacio de configuración resultante de fijar las distancias está dado por la variedad $\mathcal{C} = \mathbb{R}^3 \times SO(3) \subset \mathbb{R}^{3N}$ y tiene dimensión $D = 6$.

Los vínculos holonómicos descritos hasta aquí se conocen como vínculos *esclerónomos*, ya que las restricciones $f_i(\vec{r}_n)$ no dependen del tiempo. Sin embargo y como clarificarán los siguientes ejemplos, el tratamiento se puede extender de manera más o menos inmediata al caso donde existe dependencia temporal.

Ejemplo: partícula en una dimensión, con fuerza monogénica

Sea una partícula en una dimensión sometida a una fuerza monogénica, es decir una fuerza que se deriva de un potencial dependiente del tiempo. El lagrangiano se escribe

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

La ecuación de movimiento correspondiente será

$$m\ddot{x} = -\partial_x V(x, t)$$

Nos preguntamos si, al igual que sucedía cuando la fuerza era conservativa, existirá una solución a las ecuaciones de movimiento donde la partícula esté confinada al punto $x = x_{\min}$ donde el potencial tiene su mínimo. En ese caso debería cumplirse

$$0 = -\partial_x V(x, t)|_{x=x_{\min}}$$

Dado que el potencial depende del tiempo, al resolver esta condición de mínimo obtenemos una función $x_{\min}(t)$. Esto implica que en principio $\ddot{x}_{\min}(t) \neq 0$, por lo que la condición de que la partícula esté confinada al mínimo de potencial no satisface la ecuación de movimiento.

Sin embargo, aún nos falta imponer la condición de que el mínimo sea muy marcado en todo instante durante el movimiento. Para hacer esto, usamos la regla de la función implícita, escribiendo las derivadas temporales como

$$\dot{x}_{\min}(t) = -\frac{\partial_t \partial_x V(x, t)}{\partial_x^2 V(x, t)} \Big|_{x=x_{\min}(t)} \Rightarrow \ddot{x}_{\min}(t) = \frac{1}{\partial_x^2 V}(\dots)$$

Donde vemos que cuando la derivada segunda en el denominador es lo bastante grande, como sucede para un potencial muy empinado, se cumple que $\ddot{x}_{\min}(t) \approx 0$. Es decir que $x = x_{\min}(t)$ resuelve las ecuaciones de movimiento en este límite.

Esto demuestra que en el caso de una fuerza monogénica, si el potencial tiene un mínimo muy marcado durante toda la duración del movimiento, entonces el sistema está confinado al mínimo del potencial.

Ejercicio: partícula en dos dimensiones, con potencial unidimensional monogénico

Extienda el ejemplo anterior al caso de una partícula en dos dimensiones, que se mueve bajo la influencia de una fuerza monogénica con potencial unidimensional. El lagrangiano será

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, t)$$

Usando un razonamiento análogo al del ejemplo previo, demuestre que el movimiento se confina al mínimo del potencial $x = x_{\min}(t)$. Pruebe que en ese caso se puede reemplazar el lagrangiano por la forma efectiva

$$L_{\text{eff}} = \frac{1}{2}m\dot{y}^2$$

la cual se obtiene al reemplazar el vínculo directamente en el lagrangiano original.

Ejercicio: partícula en dos dimensiones, caso más general monogénico

Generalice el resultado del ejercicio anterior al caso con lagrangiano

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y, t)$$

Asumiendo que durante toda la duración del movimiento la dependencia del potencial en la variable y es mucho más suave que la dependencia en x , y que el potencial tiene un mínimo marcado en esta última dirección en $x = x_{\min}(t)$, pruebe que las ecuaciones de movimiento se pueden obtener de la forma efectiva

$$L_{\text{eff}} = \frac{1}{2}m\dot{y}^2 - V(x_{\min}(t), y, t)$$

El ejemplo y los ejercicios anteriores deberían dejar en claro que el razonamiento que hicimos más arriba para el caso de fuerzas conservativas se puede generalizar inmediatamente para fuerzas monogénicas. Comenzamos con un lagrangiano

$$L = K(\dot{\vec{r}}_n^2) - V(\vec{r}_n, t)$$

Si durante toda la duración del movimiento el potencial varía suavemente sobre una hipersuperficie de dimensión D , mientras que crece muy rápidamente al alejarnos de ella, el movimiento se confina a tal hipersuperficie. Podemos repetir el razonamiento que hicimos antes, teniendo en cuenta que en este caso las condiciones de vínculo holonómico

$$f_l(\vec{r}_n, t) = c_l \text{ constante, con } l \in \{D+1, \dots, 3N\}$$

serán dependientes del tiempo. Este tipo de vínculos holonómicos con dependencia temporal se denominan vínculos *reónomos*.

Nuevamente podemos elegir coordenadas adaptadas, con la única diferencia de que el cambio de variables será ahora dependiente del tiempo. La condición de mínimo $q_l = c_l$ se puede reemplazar entonces en el lagrangiano resultante, obteniendo un lagrangiano efectivo

$$L_{\text{eff}} = K(\dot{q}_k, \dot{q}_l = 0, q_k, q_l = c_l) - V(q_k, q_l = c_l, t)$$

que resultará en las mismas ecuaciones de movimiento.

11.2.2 Vínculos no holonómicos

Además de los vínculos holonómicos que discutimos en la sección anterior, existe un tipo diferente de vínculos que no se pueden expresar como un conjunto de restricciones (tal vez dependientes del tiempo) en las coordenadas \vec{r}_n de las partículas que forman el sistema. Para entender de qué estamos hablando, veamos primero un ejemplo de este tipo de vínculos.

Ejemplo: efecto Hall

El efecto Hall es un estado de la materia que aparece cuando ponemos partículas cargadas en un campo magnético intenso.

Para tener una imagen sencilla de este problema, supongamos que tenemos una partícula que se mueve en tres dimensiones bajo la influencia de un campo magnético uniforme en la dirección z . El potencial vector que genera dicho campo estará dado por

$$\vec{A} = cBx\hat{y}$$

De modo tal que $(\vec{\nabla} \times \vec{A})/c = B\hat{k}$. El lagrangiano para una partícula de carga e tendrá la forma que discutimos en la sección 2.2.3, es decir

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e(\Phi(x, y, z) - Bx\dot{y})$$

donde Φ es el potencial eléctrico. Las ecuaciones de movimiento resultantes se leen

$$m\ddot{x} = e(E_x - B\dot{y}) \quad m\ddot{y} = e(E_y - B\dot{x}) \quad m\ddot{z} = eE_z$$

Donde $E_a = -\partial_a\Phi$ es el campo eléctrico. Ahora bien, si suponemos que el campo magnético es muy intenso, de modo tal que durante todo el movimiento se cumple que $\dot{x}/\dot{y} \ll eB/m$ y $\dot{y}/\dot{x} \ll eB/m$, estas ecuaciones se pueden reemplazar por

$$0 = e(E_x - B\dot{y}) \quad 0 = e(E_y + B\dot{x}) \quad m\ddot{z} = eE_z$$

Dado que las dos primeras ecuaciones no contienen derivadas segundas, se pueden considerar como vínculos que incluyen a las velocidades. Este es el primer ejemplo de un vínculo no holonómico. En este caso particular, los vínculos implican que la carga se puede mover solamente donde hay un campo eléctrico no nulo, lo que normalmente sucede en los bordes de una muestra material.

Podríamos intentar resolver los vínculos y reemplazarlos en el lagrangiano, como hicimos más arriba para el caso de los vínculos holonómicos. Eso nos daría el lagrangiano efectivo

$$L_{\text{eff}} = \frac{1}{2}m\dot{z}^2 + \frac{mc^2}{2B^2}(E_x^2 + E_y^2) - e(\Phi(x, y, z) - E_x x)$$

Sin embargo, los nuevos términos con las componentes del campo eléctrico actúan como un término extra en el potencial, resultando en que la ecuación de movimiento para z no corresponde a la que se obtiene del lagrangiano original. Por lo tanto, el método de evaluar el lagrangiano en los vínculos no funciona en el caso de vínculos no holonómicos.

Ejemplo: rodadura en una dimensión

Supongamos que tenemos una rueda de radio R en que se mueve en una dimensión. El estado de la rueda está caracterizado por su posición en la recta x y por el ángulo ϕ que ha rotado respecto de la dirección vertical en la dirección contraria a las agujas del reloj. Estas variables satisfacen el vínculo de rodadura sin deslizamiento, que es una restricción diferencial de la forma

$$dx = -R d\phi$$

Dividiendo por dt esto se puede reescribir en términos de las velocidades

$$\dot{x} + R\dot{\phi} = 0$$

lo que claramente constituye un vínculo no-holonómico.

En este caso particular unidimensional, el vínculo no holonómico se puede integrar, obteniendo la expresión

$$x - x_0 + R(\phi - \phi_0) = 0$$

que tiene todo el aspecto de un vínculo holonómico, más allá de que incluye constantes de integración. Sin embargo, esto no se puede extender al caso de una rueda que gira sobre una superficie bidimensional, como veremos en el ejemplo que sigue.

Ejemplo: rodadura en dos dimensiones

En el caso de una rueda de radio R que describe una curva sobre un plano, el vínculo no holonómico no se puede integrar. En efecto, la distancia recorrida por la rueda debe satisfacer la restricción diferencial

$$dl = -R d\phi$$

lo que se puede reescribir como

$$\sqrt{dx^2 + dy^2} = -R d\phi$$

En términos de las velocidades esto toma la forma

$$\sqrt{\dot{x}^2 + \dot{y}^2} + R\dot{\phi} = 0$$

Este vínculo no se puede integrar para obtener una relación entre las coordenadas, por lo que es inherentemente no holonómico.

Generalizando los ejemplos, cuando el sistema está sometido a vínculos no holonómicos habrá restricciones sobre su movimiento que involucran a las velocidades. Estas se pueden escribir en la forma

$$f_l(\vec{r}_n, \dot{\vec{r}}_n, t) = \text{constante, con } l \in \{D+1, \dots, 3N\}$$

En algunos casos particulares, estas condiciones de vínculo se pueden integrar y los vínculos se transforman en vínculos de tipo holonómico como los estudiados anteriormente. Pero en el caso general, al ser dependientes de las velocidades, estos vínculos no definen una hipersuperficie de dimensión D en el espacio de configuración original \mathbb{R}^{3N} . Por lo tanto, no se pueden resolver usando coordenadas adaptadas, es decir que no pueden incorporarse en el Lagrangiano de manera sencilla.

11.2.3 Multiplicadores de Lagrange

Hemos visto que los vínculos holonómicos, tanto los esclerónomos como los reónomos, se pueden tratar utilizando coordenadas adaptadas, para las cuales los vínculos se reducen a fijar el valor de algunas de ellas. En este caso, las ecuaciones de movimiento se pueden obtener a partir de un lagrangiano efectivo que se obtiene de aplicar el vínculo directamente en el lagrangiano del sistema. Sin embargo, vimos también que existen vínculos más generales conocidos como no holonómicos, para los cuales el método de las coordenadas adaptadas no funciona. En esta sección reexaminaremos el problema de los vínculos desde una óptica más amplia, con el objeto de obtener un tratamiento más general.

Ejemplo: péndulo esférico

La segunda ley de Newton para un péndulo en tres dimensiones se escribe

$$m\ddot{x} = T_x \quad m\ddot{y} = T_y \quad m\ddot{z} = T_z - mg$$

donde hemos llamado $\vec{T} = (T_x, T_y, T_z)$ a la tensión de la soga de la que cuelga el péndulo. No conocemos la forma explícita de \vec{T} como función de las coordenadas, solo sabemos que toma los valores que deba tomar para que se cumpla la restricción

$$\sqrt{x^2 + y^2 + z^2} = l$$

donde hemos supuesto que la cuerda es inextensible y tiene longitud l . Esto implica en particular que se trata de una fuerza que apunta en la dirección de la cuerda, por lo que podemos escribir

$$\vec{T} = \lambda \vec{r} = \lambda \vec{\nabla}(\sqrt{x^2 + y^2 + z^2} - l)$$

donde λ es una cantidad desconocida, y en la segunda igualdad notamos que el versor \vec{r} se puede obtener como el gradiente de la condición de vínculo. Con esto tenemos que las ecuaciones a resolver son

$$\begin{aligned} m\ddot{x} &= \lambda \partial_x(\sqrt{x^2 + y^2 + z^2} - l) & m\ddot{y} &= \lambda \partial_y(\sqrt{x^2 + y^2 + z^2} - l) \\ m\ddot{z} &= \lambda \partial_z(\sqrt{x^2 + y^2 + z^2} - l) - mg & 0 &= \sqrt{x^2 + y^2 + z^2} - l \end{aligned}$$

Se trata de cuatro ecuaciones para determinar cuatro funciones desconocidas del tiempo, a saber las coordenadas $\{x, y, z\}$ y la nueva magnitud λ . Es interesante remarcar

que estas ecuaciones se pueden obtener del lagrangiano

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz + \underbrace{\lambda(\sqrt{x^2 + y^2 + z^2} - l)}_{-V_{\vec{T}}}$$

Donde las ecuaciones de Lagrange para $\{x, y, z\}$ resultan en las ecuaciones de movimiento, mientras que la ecuación de Lagrange para λ nos devuelve el vínculo. En este lagrangiano, los dos primeros términos corresponderían al problema en ausencia de vínculos, mientras que el término adicional, proporcional a la nueva variable λ , se puede asociar a un potencial $V_{\vec{T}}$ para la tensión \vec{T} .

¿Cuáles fueron los pasos cruciales que nos permitieron llegar al lagrangiano? El primero fue identificar la dirección de la tensión, lo que redujo el número de magnitudes independientes a determinar con nuestras ecuaciones. Y el segundo fue notar que dicha dirección coincide con la del gradiente de la condición de vínculo, lo que nos permitió darle una forma concreta a nuestro lagrangiano.

Por supuesto que en este caso el vínculo es holonómico y esclerónomo, por lo que podríamos haber usado el método de las coordenadas adaptadas, pasando a coordenadas esféricas $\{r, \theta, \phi\}$ y luego fijando $r = l$ directamente en el lagrangiano del problema. Esto se puede hacer incluso sin conocer la forma explícita de $V_{\vec{T}}$, ya que al evaluarlo en el vínculo sabemos que tomará su valor mínimo, y por lo tanto lo podemos reemplazar por una constante.

Ejercicio: péndulo forzado

Repita el análisis anterior para el caso de un péndulo colgado de un soporte móvil que está a una altura $z_{\text{sop}}(t)$. En este caso la restricción de que la soga es inextensible se ve modificada según

$$\sqrt{x^2 + y^2 + (z - z_{\text{sop}}(t))^2} = l$$

Convénzase de que también en este caso la tensión es proporcional al gradiente de la condición de vínculo

$$\vec{T} = \lambda \vec{\nabla}(\sqrt{x^2 + y^2 + (z - z_{\text{sop}}(t))^2} - l)$$

Demuestre que al reemplazar esta condición en la segunda ley de Newton se obtienen tantas ecuaciones como magnitudes desconocidas, que se pueden derivar del lagrangiano

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz + \lambda(\sqrt{x^2 + y^2 + (z - z_{\text{sop}}(t))^2} - l)$$

Nótese que seguimos en presencia de un vínculo holonómico, en este caso reónomo, para el cual podríamos haber llevado adelante la técnica de las coordenadas adaptadas de las secciones previas, incluso sin conocer explícitamente la forma del potencial que da origen a la tensión.

Ejemplo: planos inclinados y poleas

Supongamos que tenemos un bloque de masa m_B que desliza sin fricción a lo largo de un plano inclinado. El bloque está atado a una cuerda que pasa por una polea sin masa en la parte superior del plano y de la cual cuelga una pesa de masa m_P . Poniendo el origen de coordenadas en la polea y el eje y en la dirección vertical, tenemos que la segunda ley de Newton para el bloque y la pesa toma la forma

$$m_B \ddot{x}_B = N_x - T_x \quad m_B \ddot{y}_B = N_y + T_y - m_B g \quad m_P \ddot{y}_P = T - m_P g$$

Donde $\vec{T} = (-T_x, T_y)$ es la tensión de la soga y $\vec{N} = (N_x, N_y)$ es la normal del plano. La condición de que el bloque se mueva sobre el plano que forma un ángulo θ con la horizontal se puede escribir como $\tan \theta = y_B/x_B$. Esto toma la forma de un vínculo

$$x_B \sin \theta - y_B \cos \theta = 0$$

Por otro lado, la condición de soga inextensible se transforma en el vínculo

$$\sqrt{x_B^2 + y_B^2} - y_P = l$$

Para la fuerza normal, tenemos que

$$\vec{N} = \lambda_N (\sin \theta, -\cos \theta) = \lambda_N \vec{\nabla}_B (x_B \sin \theta - y_B \cos \theta)$$

donde λ_N es una nueva magnitud desconocida, que debe ser negativa para que la fuerza tenga el sentido correcto. En la segunda igualdad hemos notado que, una vez más, la fuerza tiene la dirección del gradiente de la condición de vínculo.

En cuanto a la tensión, podemos escribir

$$\vec{T} = \lambda_T (\cos \theta, \sin \theta) = \lambda_T \vec{\nabla}_B (\sqrt{x_B^2 + y_B^2} - y_P - l)$$

donde en la segunda igualdad usamos que $\cos \theta = x_B/\sqrt{x_B^2 + y_B^2}$ y $\sin \theta = y_B/\sqrt{x_B^2 + y_B^2}$, y la nueva magnitud λ_T debe ser negativa para que el sentido de la fuerza sea el correcto. Nótese que también se cumple

$$T = -\lambda_T = \lambda_T \partial_{y_P} (\sqrt{x_B^2 + y_B^2} - y_P - l)$$

Una vez más, las fuerzas que entran en la segunda ley de Newton con el objeto de garantizar los vínculos resultan ser proporcionales a los gradientes de los mismos.

Con todo lo anterior, podemos escribir el conjunto de ecuaciones en la forma

$$\begin{aligned} m_B \ddot{x}_B &= \lambda_N \partial_{x_B} (x_B \sin \theta - y_B \cos \theta) + \lambda_T \partial_{x_B} (\sqrt{x_B^2 + y_B^2} - y_P - l) \\ m_B \ddot{y}_B &= \lambda_N \partial_{y_B} (x_B \sin \theta - y_B \cos \theta) + \lambda_T \partial_{y_B} (\sqrt{x_B^2 + y_B^2} - y_P - l) - m_B g \\ m_P \ddot{y}_P &= \lambda_T \partial_{y_P} (\sqrt{x_B^2 + y_B^2} - y_P - l) - m_P g \\ x_B \sin \theta - y_B \cos \theta &= 0 \quad \sqrt{x_B^2 + y_B^2} - y_P = l \end{aligned}$$

donde hay cinco ecuaciones para las cinco magnitudes desconocidas dadas por las coordenadas del bloque $\{x_B, y_B\}$, de la polea y_P y las nuevas magnitudes $\{\lambda_T, \lambda_B\}$. Estas ecuaciones se pueden obtener del lagrangiano

$$L = \frac{1}{2}m_P \dot{y}_P^2 + \frac{1}{2}m_B(\dot{x}_B^2 + \dot{y}_B^2) - m_P g y_P - m_B g y_B + \\ + \lambda_T(\sqrt{x_B^2 + y_B^2} - y_P - l) + \lambda_N(x_B \sin \theta - y_B \cos \theta)$$

Nótese que esto corresponde al lagrangiano que tendríamos en ausencia de vínculo, más los términos proporcionales a los vínculos, en los que introdujimos una nueva variable por cada vínculo. Las ecuaciones de Lagrange para $\{x_B, y_B\}$ dan las ecuaciones de movimiento del bloque, aquéllas para y_P dan las de la polea, mientras que las que corresponden a λ_T y λ_N devuelven los vínculos.

Ejemplo: rodadura en una dimensión

Para intentar generalizar lo anterior al caso de vínculos no holonómicos, comencemos con el ejemplo de una rueda que corre a lo largo de un eje. Podemos escribir sus ecuaciones de movimiento en la forma

$$m\ddot{x} = F_x - V'(x)$$

$$I\ddot{\phi} = \tau$$

donde F_x es la fuerza de roce estático entre la rueda y la superficie, τ es torque que actúa sobre la rueda, siendo I su momento de inercia, y hemos supuesto que hay una fuerza externa cuyo potencial es $V(x)$ que actúa sobre el centro de la rueda. El vínculo de rodadura establece que

$$\dot{x} + R\dot{\phi} = 0$$

o en su forma integrada

$$x - x_0 + R(\phi - \phi_0) = 0$$

No sabemos nada de F_x , más allá de que debe tomar los valores necesarios para garantizar el vínculo. Siguiendo los razonamientos anteriores, proponemos el lagrangiano

$$L = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\phi}^2 - V(x) + \tilde{\lambda}(x - x_0 + R(\phi - \phi_0))$$

Donde agregamos al lagrangiano libre un término proporcional a la condición de vínculo, con una nueva variable $\tilde{\lambda}$. La ecuación de Lagrange de esta última nos devuelve el vínculo, mientras que las ecuaciones para x y ϕ implican que

$$F_x = \tilde{\lambda} \partial_x (x - x_0 + R(\phi - \phi_0)) = \tilde{\lambda}$$

$$\tau = \tilde{\lambda} \partial_\phi (x - x_0 + R(\phi - \phi_0)) = \tilde{\lambda} R = F_x R$$

Este lagrangiano provee las ecuaciones de movimiento correctas. Sin embargo, está escrito en términos de la forma integrada del vínculo, lo que no nos permite generalizar para vínculos no integrables. Para resolver este problema, podemos redefinir $\tilde{\lambda} = -\dot{\lambda}$ en el lagrangiano,

$$L = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\phi}^2 - V(x) - \tilde{\lambda}(x - x_0 + R(\phi - \phi_0))$$

lo que no altera la forma de F_x , pero ahora la ecuación de lagrange de λ devuelve el vínculo en su forma no integrada.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\lambda}} \right) = -(\dot{x} + R\dot{\phi}) = 0$$

Es fácil ver que las mismas ecuaciones de movimiento se obtienen del lagrangiano en el que incluimos el vínculo multiplicando por una función desconocida λ , al igual que hacíamos para el caso holonómico

$$L = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\phi}^2 - V(x) + \lambda(\dot{x} + R\dot{\phi})$$

que difiere del anterior en una derivada total. En efecto, calculando las ecuaciones de Lagrange de este lagrangiano obtenemos las expresiones

$$F_x = -\frac{d}{dt} \left(\lambda \frac{\partial}{\partial \dot{x}} (\dot{x} + R\dot{\phi}) \right) = -\dot{\lambda}$$

$$\tau = -\frac{d}{dt} \left(\lambda \frac{\partial}{\partial \dot{\phi}} (\dot{x} + R\dot{\phi}) \right) = -\dot{\lambda}R = F_x R$$

Es decir que la fuerza de vínculo se obtiene a partir de un potencial dependiente de las velocidades, que es proporcional a la condición de vínculo. Tenemos entonces tres ecuaciones (las dos ecuaciones de movimiento y el vínculo) para tres variables (las coordenadas $\{x, \phi\}$ y la fuerza F_x).

De los ejemplos anteriores debemos aprender las siguientes propiedades generales

- Cada vez que hay vínculos, aparecen nuevas fuerzas cuya función es garantizar que los vínculos se cumplan. En los problemas de los cursos básicos estas fuerzas corresponden a tensiones, normales, fuerzas de roce estático, etc. Hay una de estas *fuerzas de vínculo* por cada uno de los vínculos.
- Estas fuerzas se obtienen de un potencial que es proporcional a cada vínculo, con un factor de proporcionalidad que es una magnitud desconocida. En el caso de vínculos no holonómicos, éste funciona como un potencial dependiente de las velocidades.
- Con el mencionado potencial se puede construir un lagrangiano. Las ecuaciones de Lagrange para los factores de proporcionalidad resultan en los vínculos.

A la luz de los ejemplos anteriores, analicemos el caso general. Supongamos que tenemos un conjunto de N partículas con posiciones \vec{r}_n , que satisfacen los vínculos

$$f_l(\vec{r}_n, \dot{\vec{r}}_n, t) = \text{constante}, \text{ con } l \in \{D+1, \dots, 3N\}$$

La segunda ley de Newton para este sistema se escribe como

$$\frac{d\vec{p}_n}{dt} = \vec{F}_n + \vec{F}_n^{\text{vinc}}$$

donde hemos separado explícitamente la fuerza de vínculo. De acuerdo a lo que vimos en los ejemplos anteriores, esta fuerza se puede escribir en la forma

$$F_n^{\text{vinc}} = \lambda_l \frac{\partial f_l}{\partial \vec{r}_n} - \frac{d}{dt} \left(\lambda_l \frac{\partial f_l}{\partial \dot{\vec{r}}_n} \right)$$

Donde hemos insertado nuevas magnitudes desconocidas λ_l , una por cada uno de los vínculos, que se denominan *multiplicadores de Lagrange*.

Si suponemos que el resto de la fuerza \vec{F} se obtiene a partir de un potencial $V(\vec{r}_n, \dot{\vec{r}}_n, t)$, entonces podemos escribir un lagrangiano en la forma

$$L = K(\dot{\vec{r}}_n^2) - V(\vec{r}_n, \dot{\vec{r}}_n, t) + \lambda_l f_l(\vec{r}_n, \dot{\vec{r}}_n, t)$$

Las ecuaciones de Lagrange que se obtienen de este lagrangiano para las variables \vec{r}_n nos devuelven la segunda ley de Newton para cada una de las partículas del problema. Por otro lado, aquéllas que se obtienen a partir de los multiplicadores de Lagrange resultan en los vínculos.

El conteo de variables también funciona: cada vínculo suprime un grado de libertad de los $3N$ que tiene originalmente el sistema, y cada multiplicador de Lagrange lo restituye. Tenemos $3N$ ecuaciones de movimiento y $3N - D$ vínculos, para $3N$ coordenadas cartesianas y $3N - D$ multiplicadores de Lagrange.

Por supuesto que podríamos escribir el lagrangiano de arriba en coordenadas adaptadas para eliminar los vínculos holonómicos, quedándonos solamente con multiplicadores para los no holonómicos. Esto pone de manifiesto que el lagrangiano está definido sobre un espacio de configuración \mathcal{C} que ya no es \mathbb{R}^{3N} sino una variedad arbitraria.

Ejercicio:

Escriba el lagrangiano para una partícula cargada sometida a un campo magnético en la dirección x y un campo eléctrico en la dirección z , suponiendo que la partícula está unida al origen por medio de una varilla rígida.

Resuelva primero el problema utilizando multiplicadores de Lagrange. Escriba las ecuaciones de movimiento e identifique la fuerza que realiza la varilla.

Luego resuelva el problema utilizando el método de coordenadas adaptadas, es decir usando coordenadas esféricas y fijando el valor del radio a la longitud de la varilla. Obtenga las ecuaciones de movimiento.

¿Cuál es la relación entre las ecuaciones de movimiento que se obtienen en cada caso?

Podríamos preguntarnos qué característica particular tienen las fuerzas de vínculo que nos permite escribirlas en la forma de arriba. Podríamos recorrer nuevamente a los ejemplos, y notaríamos que en todos los casos tales fuerzas no realizan trabajo para un desplazamiento muy rápido. Para ahorrar tiempo, vamos a probarlo para el caso general, usando la definición de fuerza de vínculo que dimos más arriba y que se aplica a todos los ejemplos que analizamos. Para esto escribimos

$$dW = \vec{F}_n^{\text{vinc}} \cdot d\vec{r}_n$$

vemos que

$$dW = \left(\lambda_l \frac{\partial f_l}{\partial \vec{r}_n} - \frac{d}{dt} \left(\lambda_l \frac{\partial f_l}{\partial \dot{\vec{r}}_n} \right) \right) \cdot d\vec{r}_n$$

Integrando por partes en el segundo término, tenemos

$$dW = \lambda_l \left(\frac{\partial f_l}{\partial \vec{r}_n} \cdot d\vec{r}_n + \frac{\partial f_l}{\partial \dot{\vec{r}}_n} \cdot d\dot{\vec{r}}_n \right) - \frac{d}{dt} \left(\lambda_l \frac{\partial f_l}{\partial \dot{\vec{r}}_n} \cdot d\vec{r}_n \right)$$

Si el desplazamiento es instantáneo podemos descartar el último término, resultando en

$$dW = \lambda_l \left(\frac{\partial f_l}{\partial \vec{r}_n} \cdot d\vec{r}_n + \frac{\partial f_l}{\partial \dot{\vec{r}}_n} \cdot d\dot{\vec{r}}_n \right) = \lambda_l df_l = 0$$

donde en la última igualdad hemos usado el hecho de que los vínculos se satisfacen durante el desplazamiento, por lo que $df_l = 0$.

Esto se conoce como *principio de D'Alembert de los trabajos virtuales* y dice que las fuerzas de vínculo no realizan trabajo en un desplazamiento instantáneo.

11.3 Resumen

En esta clase estudiamos el origen físico de los vínculos, y la forma de incorporarlos al tratamiento lagrangiano.

Vimos que cuando hay vínculos holonómicos reónomos o esclerónomos, podemos elegir coordenadas adaptadas, que resuelven los vínculos simplemente fijando el valor de algunas de ellas. Vimos también que cuando hay vínculos no holonómicos, el método de las coordenadas adaptadas no funciona.

Aprendimos la técnica de los multiplicadores de Lagrange, que permite escribir un lagrangiano para un sistema vinculado, sean los vínculos holonómicos o no. Esta construcción tiene la ventaja de permitirnos identificar las fuerzas de vínculo.

Junto con el análisis de las truncaciones que hicimos la clase previa, los presentes resultados demuestran que los sistemas mecánicos pueden moverse sobre una variedad de configuración \mathcal{C} cualquiera recorrida por las coordenadas generalizadas $\{q_i\}$. Con esto, trascendemos nuestra hipótesis inicial de un sistema de N partículas que tenía $\mathcal{C} = \mathbb{R}^{3N}$, construyendo así la primera generalización de los sistemas mecánicos que veremos en este curso.

12. Pequeñas oscilaciones

12.1 Objetivos

Aprendimos la clase pasada que cuando existe en el potencial un pozo profundo y marcado, entonces se puede simplificar la descripción del sistema mediante una condición de vínculo. El sistema se mueve en las direcciones en las que el potencial varía suavemente, confinado dentro de la hipersuperficie determinada por el fondo del pozo.

En esta clase vamos a explorar qué pasa cuando la condición de vínculo se relaja un poco, es decir cuando permitimos que el sistema realice pequeñas oscilaciones que lo alejan ligeramente del fondo del pozo de potencial.

Una vez más, analizaremos primero algunos casos muy simples, para ganar algo de intuición que nos permita tratar el caso general.



Frédéric Chopin

12.2 Oscilaciones en sistemas autónomos

Un sistema *autónomo* es uno donde el lagrangiano no depende explícitamente del tiempo. En particular esto implica que todas las fuerzas que actúan sobre las partículas que lo componen son conservativas, y que todos los vínculos que se impongan sobre el sistema son esclerónomos.

Esta caracterización involucra una clase muy general de sistemas físicos. En esta sección, vamos a estudiar cómo se describen sus pequeñas oscilaciones. Comenzaremos con ejemplos en una y dos dimensiones, y luego estudiaremos el caso general.

Ejemplo: partícula no relativista con fuerza conservativa en una dimensión

Supongamos que tenemos una partícula que se mueve en una dimensión bajo la influencia de una fuerza conservativa a velocidades no relativistas. Su lagrangiano tendrá la forma

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

La ecuación de movimiento resultantes será, por supuesto

$$m\ddot{x} = -V'(x)$$

Puede resultar más o menos complejo resolver esta ecuación para un potencial cualquiera $V(x)$. Sin embargo, si el potencial tiene un punto estacionario en x_{\min} , es decir que $V'(x_{\min}) = 0$, entonces una solución trivial es la constante $x = x_{\min}$.

Vimos en la clase anterior que cuando el potencial crece muy rápidamente al alejarnos del mínimo en ambas direcciones, entonces la solución $x = x_{\min}$ es una buena aproximación de cualquier solución, y el sistema está vinculado. Nos preguntamos ahora qué pasa si el mínimo del potencial no es tan marcado, lo que permitiría a la partícula alejarse ligeramente de $x = x_{\min}$.

Queremos seguir el movimiento expandiendo en potencias del desplazamiento del sistema respecto de x_{\min} , por lo que escribimos $x = x_{\min} + \varepsilon \delta x$ con ε lo suficientemente pequeño. En ese caso, el Lagrangiano toma la forma

$$L = \frac{1}{2}m(\dot{x}_{\min} + \varepsilon \delta \dot{x})^2 - V(x_{\min} + \varepsilon \delta x)$$

Expandiendo en potencias de ε tenemos

$$L \approx -V(x_{\min}) - \varepsilon V'(x_{\min}) \delta x + \varepsilon^2 \left(\frac{1}{2}m \delta \dot{x}^2 - \frac{1}{2}V''(x_{\min}) \delta x^2 \right) + \mathcal{O}(\varepsilon)^3$$

El término $V(x_{\min})$ es una constante que no afectará las ecuaciones de movimiento, ya que las mismas se obtienen de tomar derivadas respecto de δx y $\delta \dot{x}$. Por otro lado, el término lineal $\varepsilon V'(x_{\min}) \delta x$ se anula porque x_{\min} es un punto estacionario del potencial. Nos queda entonces la expresión

$$L \approx \varepsilon^2 \left(\frac{1}{2}m \delta \dot{x}^2 - \frac{1}{2}k \delta x^2 \right) + \mathcal{O}(\varepsilon)^3$$

Donde $k = V''(x_{\min})$. Al orden más bajo en ε este es el lagrangiano de un oscilador armónico, como puede verse obteniendo sus ecuaciones de movimiento

$$m \delta \ddot{x} + k \delta x = 0$$

Nótese que hubiéramos obtenido la misma ecuación si en lugar de expandir en potencias de ε en el lagrangiano, lo hacíamos en la ecuación de movimiento para x .

Esta ecuación es lineal y tiene coeficientes reales, lo que nos permite recurrir a la técnica de *complexificación* para obtener una solución, como se describe a continuación. Para comenzar, duplicamos la ecuación, inventando una segunda variable $\delta\tilde{x}$ que también la satisface, con lo que tenemos

$$m\delta\ddot{x} + k\delta x = 0 \qquad m\delta\ddot{\tilde{x}} + k\delta\tilde{x} = 0$$

Ahora definimos la variable compleja $\delta\zeta = \delta x + i\delta\tilde{x}$. Es fácil ver que, multiplicando por i la segunda ecuación y sumándola a la primera, se obtiene una ecuación para $\delta\zeta$ con la forma

$$m\delta\ddot{\zeta} + k\delta\zeta = 0$$

Si ahora escribimos una solución tentativa o *Ansatz* de la forma $\delta\zeta = c e^{-i\omega t}$ con c y ω constantes, podemos insertarlo en la ecuación para obtener

$$(m - \omega^2 k)c = 0$$

Esto se satisface siempre que $\omega = \pm\sqrt{k/m}$, para cualquier valor de c , lo que nos permite elegir convencionalmente $c = 1/\sqrt{m}$. Hemos obtenido entonces dos soluciones, una para cada elección del signo de ω . Como se trata de una ecuación lineal, podemos escribir la solución general como una combinación lineal de ambas, en la forma

$$\delta\zeta = \frac{1}{\sqrt{m}} (A^+ e^{i\omega t} + A^- e^{-i\omega t})$$

con A^\pm constantes arbitrarias. De aquí obtenemos δx tomando la parte real

$$\delta x = \frac{1}{\sqrt{m}} \Re (A^+ e^{i\omega t} + A^- e^{-i\omega t})$$

Nótese que ω^2 viene dado por $k/m = V''(x_{\min})/m$. Por lo tanto, si el punto estacionario x_{\min} del potencial es un mínimo, ω será real y las exponenciales en la expresión anterior serán sumas de senos y cosenos, por lo que al tomar la parte real obtenemos

$$\delta x = A \cos(\omega t + \varphi)$$

Donde A y φ se pueden escribir fácilmente en términos de A^\pm y m .

Si en cambio x_{\min} es un máximo del potencial, todos los cálculos que hicimos para obtener la solución se mantienen sin cambios, sólo que al final obtendremos un valor de ω que será imaginario. En ese caso la solución general queda escrita como

$$\delta x = \frac{1}{\sqrt{m}} \Re (A^+ e^{-|\omega|t} + A^- e^{|\omega|t})$$

Aquí vemos que uno de los términos tiende a cero exponencialmente, mientras que el otro diverge. Esto quiere decir que cualquier perturbación δx alrededor del máximo x_{\min}

saca al sistema del punto estacionario y lo hace caer por el potencial. Por supuesto cuando esto sucede, la solución que obtuvimos sólo es válida durante un pequeño intervalo de tiempo, ya que en su deducción asumimos que δx era pequeño.

En el análisis de este problema sencillo, aislamos los pasos cruciales que nos permitirán generalizar en lo que sigue. Primero linealizamos, es decir escribir las ecuaciones a primer orden en una perturbación alrededor de la solución de equilibrio. Luego complexificamos, es decir transformamos nuestra variable dependiente en una cantidad compleja. Finalmente propusimos una solución tentativa o Ansatz conteniendo exponenciales complejas, lo que simplificó el cálculo de las derivadas. Repetiremos estos pasos en un caso bidimensional en el ejemplo siguiente.

Ejemplo: partícula no relativista con fuerza conservativa en dos dimensiones

Estudiemos ahora el caso de una partícula que se mueve en dos dimensiones sometida a la acción de una fuerza conservativa a velocidades no relativistas. En coordenadas cartesianas, este problema toma la forma

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y)$$

De nuevo, supongamos que $(x_{\text{sol}}, y_{\text{sol}})$ es un punto estacionario del potencial, que cumple que sus derivadas se anulan allí

$$\frac{\partial V}{\partial x}(x_{\text{sol}}, y_{\text{sol}}) = \frac{\partial V}{\partial y}(x_{\text{sol}}, y_{\text{sol}}) = 0$$

Escribiendo una pequeña perturbación alrededor del punto estacionario como $x = x_{\text{sol}} + \varepsilon \delta x$, $y = y_{\text{sol}} + \varepsilon \delta y$ y desarrollando a segundo orden en el parámetro pequeño ε , nos queda

$$L = -V \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} - \varepsilon \left(\frac{\partial V}{\partial x} \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} \delta x + \frac{\partial V}{\partial y} \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} \delta y \right) + \\ + \frac{1}{2} \varepsilon^2 \left(m(\delta \dot{x}^2 + \delta \dot{y}^2) - \frac{\partial^2 V}{\partial x^2} \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} \delta x^2 - \frac{\partial^2 V}{\partial y^2} \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} \delta y^2 - 2 \frac{\partial^2 V}{\partial x \partial y} \Big|_{\substack{x=x_{\text{sol}} \\ y=y_{\text{sol}}}} \delta x \delta y \right) + \mathcal{O}(\varepsilon)^3$$

En la primera línea, el primer término es una constante que no depende de δx ni de δy , ni de sus derivadas, y que por lo tanto no afectará las ecuaciones de movimiento. En cuanto a los siguientes dos términos, se anulan por la condición de punto estacionario. Eso nos deja, al orden más bajo en ε , con el lagrangiano cuadrático

$$L = \varepsilon^2 \left(\frac{1}{2}m(\delta \dot{x}^2 + \delta \dot{y}^2) - \frac{1}{2}k_{xx} \delta x^2 - \frac{1}{2}k_{yy} \delta y^2 - k_{xy} \delta x \delta y \right)$$

donde hemos definido las magnitudes

$$k_{xx} = \frac{\partial^2 V}{\partial x^2}(x_{\text{sol}}, y_{\text{sol}}) \quad k_{yy} = \frac{\partial^2 V}{\partial y^2}(x_{\text{sol}}, y_{\text{sol}}) \quad k_{xy} = \frac{\partial^2 V}{\partial x \partial y}(x_{\text{sol}}, y_{\text{sol}})$$

De este lagrangiano se obtienen las ecuaciones de movimiento

$$m \delta \ddot{x} + k_{xx} \delta x + k_{xy} \delta y = 0 \quad m \delta \ddot{y} + k_{yy} \delta y + k_{xy} \delta x = 0$$

Nótese que estas ecuaciones tienen la forma de dos osciladores armónicos acoplados. Podemos recurrir nuevamente a la técnica de complexificación, para lo cual tenemos que inventar dos nuevas variables dependientes $\delta\tilde{x}$ y $\delta\tilde{y}$, para crear las funciones complejas $\delta\zeta = \delta x + i\delta\tilde{x}$ y $\delta\xi = \delta y + i\delta\tilde{y}$. Las ecuaciones se escriben entonces

$$m\delta\ddot{\zeta} + k_{xx}\delta\zeta + k_{xy}\delta\xi = 0 \quad m\delta\ddot{\xi} + k_{yy}\delta\xi + k_{xy}\delta\zeta = 0$$

Escribimos un Ansatz de la forma $\delta\zeta = c_x e^{i\omega t}$ y $\delta\xi = c_y e^{i\omega t}$. Reemplazando en las ecuaciones obtenemos

$$(k_{xx} - m\omega^2)c_x + k_{xy}c_y = 0 \quad (k_{yy} - m\omega^2)c_y + k_{xy}c_x = 0$$

El caso más sencillo es cuando $k_{xy} = 0$. Entonces, las ecuaciones se *desacoplan*, en el sentido de que la ecuación para c_x no contiene c_y y viceversa. En ese caso, de la primera ecuación podemos despejar la frecuencia obteniendo que $\delta\zeta$ es un oscilador armónico con frecuencia $\omega_1 = \pm\sqrt{k_{xx}/m}$, y haciendo lo mismo en la segunda vemos que $\delta\xi$ es otro oscilador con frecuencia $\omega_2 = \pm\sqrt{k_{yy}/m}$. Además, se ve en las ecuaciones que podemos poner $c_x = c_y = 1/\sqrt{m}$. La solución para $k_{xy} = 0$ se escribe entonces en la forma

$$\delta x = \frac{1}{\sqrt{m}} \Re (A_x^+ e^{i\omega_1 t} + A_x^- e^{-i\omega_1 t}) \quad \delta y = \frac{1}{\sqrt{m}} \Re (A_y^+ e^{i\omega_2 t} + A_y^- e^{-i\omega_2 t})$$

Donde hemos construido para cada una de las direcciones espaciales una combinación lineal de las dos soluciones obtenidas, y hemos tomado la parte real. Aquí vemos que ambas perturbaciones oscilarán cuando el potencial tenga un mínimo en el punto estacionario, ya que en ese caso tanto k_{xx} como k_{yy} serán positivas. Por otro lado, si el potencial tuviera un máximo, ambas perturbaciones serían inestables, creciendo exponencialmente. En el caso de que el punto estacionario sea un punto de ensilladura, entonces una de las dos perturbaciones oscilará mientras que la otra será inestable y crecerá exponencialmente.

Volviendo ahora al caso general en el que $k_{xy} \neq 0$, debemos resolver el sistema acoplado

$$(k_{xx} - m\omega^2)c_x + k_{xy}c_y = 0 \quad (k_{yy} - m\omega^2)c_y + k_{xy}c_x = 0$$

Multiplicando la primera ecuación por $(k_{yy} - m\omega^2)$ y la segunda por k_{xy} y restando, obtenemos

$$((k_{yy} - m\omega^2)(k_{xx} - m\omega^2) - k_{xy}^2)c_x = 0$$

Para que se anule el prefactor tiene que cumplirse que

$$m^2(\omega^2)^2 - m(k_{xx} + k_{yy})\omega^2 + k_{xx}k_{yy} - k_{xy}^2 = 0$$

Esta es una ecuación cuadrática para ω^2 , que resulta en las dos soluciones $\omega_{1,2}^2$

$$\omega_{1,2}^2 = \frac{1}{2m} \left(k_{xx} + k_{yy} \pm \sqrt{(k_{xx} - k_{yy})^2 + 4k_{xy}^2} \right)$$

Por lo tanto, hay dos frecuencias posibles para este sistema dependiendo de la elección de signo en esta solución. Vemos en la fórmula que ambas soluciones satisfacen $\omega_{1,2}^2 \in \mathbb{R}$. Estas se denominan las *frecuencias normales* del sistema. Se puede comprobar que si ponemos $k_{xy} = 0$ recuperamos las frecuencias que obtuvimos más arriba.

Todavía nos falta resolver para c_x y c_y . Primero notemos que las ecuaciones determinan esas magnitudes a menos de una constante multiplicativa. En efecto, si tomamos cualquier solución c_x, c_y y multiplicamos ambas magnitudes por el mismo número, obtenemos otra solución. Esto nos permite elegir $m(c_x^2 + c_y^2) = 1$ o en otras palabras $c_x = (1/\sqrt{m}) \cos \alpha$, $c_y = (1/\sqrt{m}) \sin \alpha$ para algún ángulo α . Con esto, las ecuaciones toman la forma

$$(k_{xx} - m\omega^2) \cos \alpha = -k_{xy} \sin \alpha \quad (k_{yy} - m\omega^2) \sin \alpha = -k_{xy} \cos \alpha$$

Reemplazando cualquiera de nuestras soluciones para ω^2 y dividiendo ambas ecuaciones se obtiene

$$\tan^2 \alpha_{1,2} = \frac{k_{xx} - m\omega_{1,2}^2}{k_{yy} - m\omega_{1,2}^2}$$

donde se ve que hay dos soluciones $\alpha_{1,2}$, correspondientes a las frecuencias $\omega_{1,2}^2$. Luego al tomar la raíz cuadrada, cada una de esas frecuencias aparecerá con ambos signos. Con esto, nuestra solución general para el problema se puede escribir como

$$\delta x = \frac{1}{\sqrt{m}} \Re \left(\cos \alpha_1 (A_1^+ e^{i\omega_1 t} + A_1^- e^{-i\omega_1 t}) + \cos \alpha_2 (A_2^+ e^{i\omega_2 t} + A_2^- e^{-i\omega_2 t}) \right)$$

$$\delta y = \frac{1}{\sqrt{m}} \Re \left(\sin \alpha_1 (A_1^+ e^{i\omega_1 t} + A_1^- e^{-i\omega_1 t}) + \sin \alpha_2 (A_2^+ e^{i\omega_2 t} + A_2^- e^{-i\omega_2 t}) \right)$$

donde $A_{1,2}^\pm$ son constantes arbitrarias que parametrizan una combinación lineal de las cuatro soluciones linealmente independientes, y hemos tomado la parte real. Nótese que ahora la estabilidad o inestabilidad del sistema estará determinada por los valores reales o imaginarios de ambas frecuencias $\omega_{1,2}$.

Ahora bien, las soluciones que hemos obtenido sugieren el cambio de variables

$$\delta x = \frac{1}{\sqrt{m}} (\cos \alpha_1 \delta u_1 + \cos \alpha_2 \delta u_2)$$

$$\delta y = \frac{1}{\sqrt{m}} (\sin \alpha_1 \delta u_1 + \sin \alpha_2 \delta u_2)$$

Podríamos haber comenzado nuestros cálculos haciendo este cambio de variables directamente en el lagrangiano, que tomaría entonces la forma

$$L = \varepsilon^2 \left(\frac{1}{2} (\delta \dot{u}_1^2 + \delta \dot{u}_2^2) + \frac{1}{2} (\omega_1^2 \delta u_1^2 + \omega_2^2 \delta u_2^2) \right)$$

donde hemos utilizado la relación entre $\omega_{1,2}$ y $\tan \alpha_{1,2}$ que escribimos más arriba. Esto significa que δu_1 y δu_2 son dos osciladores armónicos desacoplados, que oscilan con las dos frecuencias que hemos obtenido. Se denominan los *modos normales* del

sistema, y corresponden a rotar δx y δy hacia las direcciones de cambio máximo y de cambio mínimo del potencial.

Todo el cálculo de esta sección se puede simplificar si lo escribimos en términos de matrices. Para eso, ponemos

$$\vec{\delta x} = (\delta x, \delta y)$$

$$k = \begin{pmatrix} k_{xx} & k_{xy} \\ k_{xy} & k_{yy} \end{pmatrix}$$

siendo k la matriz hessiana del potencial, es decir la matriz cuyos elementos son las derivadas segundas de V , evaluada en la solución. Con esto, el lagrangiano se escribe

$$L = \varepsilon^2 \left(\frac{1}{2} m \vec{\delta x}^2 - \frac{1}{2} \vec{\delta x}^t \cdot k \cdot \vec{\delta x} \right)$$

y las ecuaciones de movimiento

$$m \ddot{\vec{\delta x}} + k \cdot \vec{\delta x} = 0$$

Lo que se complexifica de manera inmediata

$$m \ddot{\vec{\delta \zeta}} + k \cdot \vec{\delta \zeta} = 0$$

usando el Ansatz $\vec{\delta \zeta} = \vec{c} e^{-i\omega t}$ tenemos

$$(k - m\omega^2 I) \cdot \vec{c} = 0$$

Esto es una *ecuación de autovalores* donde $m\omega^2$ sería un autovalor de la matriz k . Nótese que esta ecuación determina \vec{c} a menos de una constante multiplicativa, por lo que podemos elegir $m\vec{c}^t \cdot \vec{c} = 1$ de donde $\vec{c} = (1/\sqrt{m})(\cos \alpha, \sin \alpha)$ para algún ángulo α . Para que esta ecuación tenga una solución no trivial, se tiene que cumplir que

$$\text{Det}(k - m\omega^2 I) = 0$$

lo que no es más que la ecuación cuadrática en ω^2 que obtuvimos más arriba, de donde despejamos las dos soluciones $\omega_{1,2}$. Esto implica que habrá en principio dos soluciones diferentes para \vec{c} que podemos llamar $\vec{c}_{1,2}$. Estas soluciones cumplen

$$k \cdot \vec{c}_1 = m\omega_1^2 \vec{c}_1 \quad k \cdot \vec{c}_2 = m\omega_2^2 \vec{c}_2$$

Multiplicando cada ecuación por el vector correspondiente \vec{c}_1 o \vec{c}_2 , obtenemos

$$\vec{c}_{1,2}^t \cdot k \cdot \vec{c}_{1,2} = \omega_{1,2}^2$$

de donde vemos que el signo de $\omega_{1,2}^2$, y consecuentemente la estabilidad del sistema, dependerán del signo de $\vec{c}_{1,2}^t \cdot k \cdot \vec{c}_{1,2}$. Si miramos con atención el desarrollo de Taylor del potencial que escribimos más arriba, podemos comprobar que esta última expresión

corresponde al cambio en el potencial cuando nos alejamos del punto estacionario una distancia $\vec{c}_{1,2}$. Si este cambio es positivo, es decir si estamos en un mínimo de potencial, entonces ambas frecuencias son reales. Por otro lado si el cambio es negativo, es decir si estamos en un máximo del potencial, ambas serán imaginarias y aparecerá una inestabilidad. Puede darse también que estemos en un punto silla, y que el cambio sea positivo para el vector \vec{c}_1 y negativo para \vec{c}_2 , implicando que habrá estabilidad en una dirección e inestabilidad en la otra.

Si ahora multiplicamos la primera ecuación de más arriba por \vec{c}_2 y la segunda por \vec{c}_1 y las restamos tenemos la relación

$$m(\omega_1^2 - \omega_2^2)(\vec{c}_2^t \cdot \vec{c}_1) = 0$$

Donde hemos usado que $k^t = k$. Esto implica que si $\omega_1 \neq \omega_2$ entonces \vec{c}_1 es perpendicular a \vec{c}_2 .

La solución general puede entonces escribirse

$$\vec{\delta x} = \Re \left((A_1^+ e^{i\omega_1 t} + A_1^- e^{-i\omega_1 t}) \vec{c}_1 + (A_2^+ e^{i\omega_2 t} + A_2^- e^{-i\omega_2 t}) \vec{c}_2 \right)$$

con lo que los modos normales nos quedan escritos según

$$\vec{\delta x} = \delta u_1 \vec{c}_1 + \delta u_2 \vec{c}_2$$

Una vez más, podemos reemplazar en el Lagrangiano

$$L = \varepsilon^2 \left(\frac{1}{2} (\dot{\delta u}_1^2 - \omega_1^2 \delta u_1^2) + \frac{1}{2} (\dot{\delta u}_2^2 - \omega_2^2 \delta u_2^2) \right)$$

Donde hemos obtenido dos osciladores armónicos independientes, cada uno oscilando con una de las frecuencias normales $\omega_{1,2}$ en las direcciones normales \vec{c}_1, \vec{c}_2 .

Ahora vamos a generalizar lo que vimos en los ejemplos precedentes en dos direcciones: por un lado, vamos a agregar un número arbitrario de coordenadas, y por otro vamos a permitir un lagrangiano más general, con la forma

$$L(\dot{q}_i, q_i) = \frac{1}{2} K_{ij}(q_k) \dot{q}_i \dot{q}_j - V(q_i)$$

donde la matriz $K_{ij}(q_k)$ es en principio una función sólo de las coordenadas. Este lagrangiano corresponde a un sistema de partículas no relativistas con fuerzas conservativas, sometido a vínculos holonómicos que ya han sido resueltos mediante coordenadas adaptadas.

Las ecuaciones de movimiento se escriben como

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = \frac{d}{dt} (K_{ij}(q_k) \dot{q}_j) - \frac{\partial V(q_k)}{\partial q_i} - \frac{1}{2} \frac{\partial K_{jk}}{\partial q_i} \dot{q}_j \dot{q}_k = 0$$

Estas ecuaciones tienen una solución estática $q_i = q_i^{\text{sol}}$ si el potencial tiene un punto estacionario, es decir si $\partial V / \partial q_i|_{q_k = q_k^{\text{sol}}} = 0$.

Ahora perturbamos el sistema sacándolo del punto estacionario con una pequeña variación de las coordenadas $q_k = q_k^{\text{sol}} + \varepsilon \delta q_k$. En la energía cinética esto nos permite escribir

$$K(\delta \dot{q}_k, q_k^{\text{sol}} + \varepsilon \delta q_k) = \varepsilon^2 \frac{1}{2} K_{ij}(q_k^{\text{sol}} + \varepsilon \delta q_k) \delta \dot{q}_i \delta \dot{q}_j = \varepsilon^2 \frac{1}{2} \underbrace{K_{ij}(q_k^{\text{sol}})}_{m_{ij}} \delta \dot{q}_i \delta \dot{q}_j + \mathcal{O}(\varepsilon)^3$$

Donde la *matriz de masas* se definió como $m_{ij} = K_{ij}(q_k^{\text{sol}})$. Por otro lado en el potencial

$$V(q_i^{\text{sol}} + \varepsilon \delta q_i) = V(q_k^{\text{sol}}) + \varepsilon \frac{\partial V(q_k^{\text{sol}})}{\partial q_i} \delta q_i + \varepsilon^2 \frac{1}{2} \underbrace{\frac{\partial^2 V(q_k^{\text{sol}})}{\partial q_i \partial q_j}}_{k_{ij}} \delta q_i \delta q_j + \mathcal{O}(\varepsilon)^3$$

El término lineal se anula debido a la condición de punto estacionario, y el término constante no contribuye a las ecuaciones de movimiento por lo que lo podemos omitir. La *matriz de restauración* es $k_{ij} = \partial^2 V / \partial q_i \partial q_j|_{q_k=q_k^{\text{sol}}}$. Con esto el lagrangiano al menor orden en ε se lee

$$L = \varepsilon^2 \left(\frac{1}{2} m_{ij} \delta \dot{q}_i \delta \dot{q}_j - \frac{1}{2} k_{ij} \delta q_i \delta q_j \right)$$

De donde se obtienen inmediatamente las ecuaciones de movimiento

$$m_{ij} \ddot{\delta q}_j + k_{ij} \delta q_j = 0$$

Complexificando las variables $\delta \zeta_j = \delta q_j + i \delta \tilde{q}_j$ podemos proponer el Ansatz $\delta \zeta_i = c_i e^{-i\omega t}$ e insertarlo en la ecuación. Nos queda

$$(k_{ij} - \omega^2 m_{ij}) c_j = 0$$

o bien, en notación matricial

$$(k - \omega^2 m) \cdot \vec{c} = 0$$

Esto es una ecuación de autovalores, la cual tiene una solución no trivial siempre que el determinante se anule

$$\text{Det}(k - \omega^2 m) = 0$$

Si tenemos D coordenadas generalizadas $\{q_i\}$, entonces la matriz entre paréntesis tiene $D \times D$ componentes. Esto implica que el determinante será un polinomio de grado D en los elementos de la matriz, y por lo tanto es un polinomio de ese grado en la variable ω^2 . Las raíces de este polinomio nos darán los posibles valores de ω^2 , que serían en principio D valores diferentes ω_i^2 con $i \in \{1 \dots D\}$. Para cada uno de estos valores, deberíamos obtener el autovector correspondiente resolviendo la ecuación

$$(k - \omega_i^2 m) \cdot \vec{c}_i = 0$$

En principio hay D vectores diferentes \vec{c}_i con $i \in \{1 \dots D\}$. Nótese que cada uno de ellos tiene D componentes c_{ij} con $j \in \{1 \dots D\}$. Se puede probar que estos vectores son reales, y lo asumiremos en lo que sigue. Si multiplicamos por \vec{c}_j^t la ecuación para \vec{c}_i tenemos

$$\vec{c}_j^t \cdot k \cdot \vec{c}_i = \omega_i^2 \vec{c}_j^t \cdot m \cdot \vec{c}_i$$

Si en cambio multiplicamos por \vec{c}_i^t la ecuación para \vec{c}_j nos queda

$$\vec{c}_i^t \cdot k \cdot \vec{c}_j = \omega_j^2 \vec{c}_i^t \cdot m \cdot \vec{c}_j$$

Transponiendo la segunda ecuación y restándola de la primera

$$(\omega_j^2 - \omega_i^2) (\vec{c}_i^t \cdot m \cdot \vec{c}_j) = 0$$

Con lo que tenemos que cuando $i \neq j$ se cumple que $\vec{c}_i^t \cdot m \cdot \vec{c}_j = 0$. Esta ecuación no nos dice nada del caso $i = j$, pero dado que los vectores \vec{c}_i están determinados a menos de una constante multiplicativa, en ese caso podemos definir $\vec{c}_i^t \cdot m \cdot \vec{c}_i = 1$. Podemos resumir estos resultados en la ecuación

$$\vec{c}_i^t \cdot m \cdot \vec{c}_j = \delta_{ij}$$

Esto a su vez implica, usando una de las expresiones que escribimos más arriba, que

$$\vec{c}_i^t \cdot k \cdot \vec{c}_j = \omega_j^2 \delta_{ij}$$

Poniendo $i = j$ podemos despejar

$$\omega_i^2 = \frac{\vec{c}_i^t \cdot k \cdot \vec{c}_i}{\vec{c}_i^t \cdot m \cdot \vec{c}_i}$$

Con lo que vemos que el signo de ω^2 está dado por el signo de $\vec{c}_i^t \cdot k \cdot \vec{c}_i$ y de $\vec{c}_i^t \cdot m \cdot \vec{c}_i$. La primera expresión representa la variación del potencial cuando nos apartamos del punto estacionario una cantidad $\delta q_j = c_{ij}$. Por lo tanto será positivo si el punto estacionario es un mínimo, y negativo en caso contrario. En este último caso decimos que tenemos una inestabilidad taquiónica. La segunda expresión representa la energía cinética de una perturbación con velocidad $\delta \dot{q}_j = c_{ij}$, y por lo tanto debe ser positiva para cualquier sistema estable.

La solución general del problema puede entonces escribirse como

$$\delta q_j = \Re \left((A_i^+ e^{+i\omega_i t} + A_i^- e^{-i\omega_i t}) c_{ij} \right)$$

con lo que podemos definir los modos normales según $\delta q_j = c_{ij} \delta u_i$, o bien en términos matriciales

$$\vec{\delta q} = \vec{c}_i \delta u_i$$

Reemplazando en el lagrangiano

$$L = \varepsilon^2 \left(\frac{1}{2} \dot{\vec{\delta q}}^t \cdot m \cdot \dot{\vec{\delta q}} - \frac{1}{2} \vec{\delta q}^t \cdot k \cdot \vec{\delta q} \right) = \varepsilon^2 \left(\frac{1}{2} \delta \dot{u}_i \vec{c}_i^t \cdot m \cdot \vec{c}_j \delta \dot{u}_j - \frac{1}{2} \delta u_i \vec{c}_i^t \cdot k \cdot \vec{c}_j \delta u_j \right)$$

En el primer término podemos usar la relación de ortogonalidad que habíamos obtenido antes $\vec{c}_i^t \cdot m \cdot \vec{c}_j = \delta_{ij}$ mientras que en el segundo término usamos la ecuación de autovalores $\vec{c}_i^t \cdot k \cdot \vec{c}_j = \omega_j^2 \delta_{ij}$. Esto nos deja con

$$L = \varepsilon^2 \sum_i \frac{1}{2} \left(\delta \dot{u}_i^2 - \omega_i^2 \delta u_i^2 \right)$$

Donde por claridad escribimos explícitamente la suma en i . El sistema ha sido descompuesto entonces en un conjunto de D osciladores armónicos desacoplados, oscilando con las frecuencias normales. Estas frecuencias se conocen como el *espectro* del sistema.

Nota:

Para tener una interpretación del resultado que hemos obtenido, imaginemos un sistema mecánico cualquiera que está en reposo en alguna de sus posiciones de equilibrio.

Si le aplicamos un estímulo que le entrega una cierta cantidad de energía, sus modos normales comenzarán a oscilar, cada uno en su frecuencia característica y con una cierta amplitud que dependerá del estímulo y de las particularidades del sistema.

Para dar una forma concreta a esta discusión, consideremos el sistema compuesto por cuerdas, parches, lengüetas, cajas de resonancia y una cierta cantidad de aire contenido dentro de las mismas. Es decir, un instrumento musical cualquiera.

Se trata de un número gigantesco de partículas, si quisiéramos tener en cuenta cada uno de los átomos que constituyen el instrumento. Y sin embargo, el análisis que acabamos de realizar se aplica, y por lo tanto sabemos que sus pequeñas oscilaciones deben comportarse como predijimos.

Esto nos permite interpretar que los modos normales corresponden a las diferentes maneras en las que el instrumento puede vibrar, las frecuencias de su espectro determinan los tonos en los que puede sonar, y la amplitud de cada modo establece el timbre característico que distingue una guitarra de un piano o de un tambor.

En algún sentido, nuestro resultado implica que *todas las cosas están llenas de música*.

Sin embargo, no cualquier sistema mecánico “suena bien”. Esto se debe a nuestra propensión psicológica a considerar agradables ciertas combinaciones *armónicas* de frecuencias, tales que los cocientes entre frecuencias sucesivas correspondan a números enteros pequeños. El arte del luthier consiste entonces en construir sistemas mecánicos que tengan el espectro correcto, de modo de satisfacer tal condición.

Nota:

Una pregunta interesante es cuánta información se puede obtener de un dado sistema mecánico a partir de su espectro. En palabras del matemático Mark Kac “¿Se puede escuchar la forma de un tambor?”.

Esta pregunta, planteada en 1966, fue respondida recién en 1994 cuando Carolyn Gordon, David Webb, and Scott Wolpert encontraron tambores *isoespectrales* que suenan igual a pesar de tener formas diferentes.

12.3 Generalización

En esta sección delinearemos la construcción de las pequeñas oscilaciones de un sistema con un lagrangiano general. Esto incluye los sistemas no autónomos en los que el lagrangiano depende del tiempo, que se obtienen al considerar fuerzas monogénicas o potenciales dependientes de las velocidades, o bien sistemas vínculos reónomos. Esto también incluye sistemas en los que la energía cinética no es cuadrática en las velocidades.

Ejemplo: caso unidimensional general

Para un sistema unidimensional general que depende de la coordenada generalizada q , el lagrangiano se escribe

$$L = L(\dot{q}, q, t)$$

Vamos a suponer que conocemos una solución $q = q_{\text{sol}}(t)$, que cumple las ecuaciones de movimiento de este problema

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \Big|_{q=q_{\text{sol}}} \right) - \frac{\partial L}{\partial q} \Big|_{q=q_{\text{sol}}} = 0$$

Ahora bien, como queremos expandir alrededor de esta solución, escribimos $q = q_{\text{sol}} + \varepsilon \delta q$ y reemplazamos en L para obtener

$$\begin{aligned} L = L|_{q=q_{\text{sol}}} + \varepsilon \left(\frac{\partial L}{\partial \dot{q}} \Big|_{q=q_{\text{sol}}} \delta \dot{q} + \frac{\partial L}{\partial q} \Big|_{q=q_{\text{sol}}} \delta q \right) + \\ + \varepsilon^2 \frac{1}{2} \left(\frac{\partial^2 L}{\partial \dot{q}^2} \Big|_{q=q_{\text{sol}}} \delta \dot{q}^2 + \frac{\partial^2 L}{\partial q^2} \Big|_{q=q_{\text{sol}}} \delta q^2 + 2 \frac{\partial^2 L}{\partial \dot{q} \partial q} \Big|_{q=q_{\text{sol}}} \delta q \delta \dot{q} \right) + \mathcal{O}(\varepsilon)^3 \end{aligned}$$

En esta expresión, el primer término no depende de las variables δq y $\delta \dot{q}$ respecto de las cuales vamos a derivar para obtener las ecuaciones de movimiento, por lo que no influirá en ellas y puede ser omitido. Los términos lineales tampoco influirán en las ecuaciones de movimiento para δq , porque contribuyen a la ecuación de Lagrange para δq con la ecuación para q evaluada en $q = q_{\text{sol}}$, que sabemos que se anula. Esto nos deja con el lagrangiano al orden más bajo

$$L = \varepsilon^2 \frac{1}{2} \left(\frac{\partial^2 L}{\partial \dot{q}^2} \Big|_{q=q_{\text{sol}}} \delta \dot{q}^2 + \frac{\partial^2 L}{\partial q^2} \Big|_{q=q_{\text{sol}}} \delta q^2 + 2 \frac{\partial^2 L}{\partial \dot{q} \partial q} \Big|_{q=q_{\text{sol}}} \delta q \delta \dot{q} \right)$$

Este lagrangiano resulta en las ecuaciones de movimiento

$$\underbrace{\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{q}^2} \Big|_{q=q_{\text{sol}}} \delta \dot{q} \right)}_{m(t)} + \underbrace{\left(\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{q} \partial q} \Big|_{q=q_{\text{sol}}} \right) - \frac{\partial^2 L}{\partial q^2} \Big|_{q=q_{\text{sol}}} \right)}_{k(t)} \delta q = 0$$

En otras palabras, terminamos con una ecuación diferencial lineal de segundo orden escrita en la forma general *de Sturm-Liouville*

$$\frac{d}{dt} \left(m(t) \frac{d\delta q}{dt} \right) + k(t) \delta q = 0$$

Esta ecuación se estudia en los cursos de ecuaciones diferenciales. Aquí solo diremos que tiene en general dos soluciones linealmente independientes $f^{\pm}(t)$, y que cualquier otra solución se puede escribir como una combinación lineal de la forma

$$\delta q(t) = A^+ f^+(t) + A^- f^-(t)$$

Nótese que aquí \pm es solo una notación para distinguir dos soluciones, y no se refiere al signo de ninguna frecuencia, ya que las $f^{\pm}(t)$ no son en general funciones trigonométricas ni exponenciales.

Aplicando el resultado al caso particular en el que ni m ni k dependen del tiempo, vemos que la ecuación recupera la forma de un oscilador armónico

$$m \delta \ddot{q} + k \delta q = 0$$

La diferencia con el caso de una partícula que se mueve en una dimensión bajo los efectos de una fuerza conservativa, es que ahora m y k se obtienen a partir del lagrangiano según las fórmulas de arriba. Por supuesto, la solución estará como antes dada por la expresión

$$\delta q(t) = \frac{1}{\sqrt{m}} \Re (A^+ e^{i\omega t} + A^- e^{-i\omega t})$$

Y será oscilatoria siempre que m y k sean positivos. Nótese que ahora hay dos posibles fuentes de inestabilidad, que podrían hacer que $\omega = \sqrt{k/m}$ fuera imaginario: podríamos tener k negativo, lo que implicaría un máximo de potencial, o podríamos tener m negativo, que implicaría que la inestabilidad proviene de la energía cinética. La primera inestabilidad se llama *taquiónica* o *por taquiones* y la segunda se denomina *por fantasmas*.

Ahora vamos a generalizar lo que hemos encontrado en el ejemplo al caso de un lagrangiano sin restricciones que depende de varias coordenadas y velocidades generalizadas, con la forma $L(\dot{q}_i, q_i, t)$. Si tenemos una solución $q_k = q_k^{\text{sol}}$ que cumple las ecuaciones de movimiento

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \Big|_{q=q_i^{\text{sol}}} \right) - \frac{\partial L}{\partial q} \Big|_{q=q_i^{\text{sol}}} = 0$$

Escribiendo $q = q_i^{\text{sol}} + \varepsilon \delta q_i$ y expandiendo el lagrangiano a orden cuadrático en ε alrededor de la solución, tenemos al orden más bajo

$$L = \varepsilon^2 \frac{1}{2} \left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \Big|_{q_k=q_k^{\text{sol}}} \delta \dot{q}_i \delta \dot{q}_j + \frac{\partial^2 L}{\partial q_i \partial q_j} \delta q_i \delta q_j \Big|_{q_k=q_k^{\text{sol}}} + 2 \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j} \Big|_{q_k=q_k^{\text{sol}}} \delta \dot{q}_i \delta q_j \right)$$

Donde hemos omitido los términos constantes y lineales porque no contribuyen a las ecuaciones de movimiento. Estas ecuaciones toman la forma

$$\underbrace{\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \Big|_{q_k=q_k^{\text{sol}}} \delta \dot{q}_j \right)}_{m_{ij}(t)} + \underbrace{\left(\frac{d}{dt} \left(\frac{\partial^2 L}{\partial q_j \partial \dot{q}_i} \right) - \frac{\partial^2 L}{\partial q_i \partial q_j} \right) \Big|_{q_k=q_k^{\text{sol}}}}_{k_{ij}(t)} \delta q_j = 0$$

o en otras palabras

$$\frac{d}{dt} \left(m_{ij}(t) \frac{d \delta q_j}{dt} \right) + k_{ij}(t) \delta q_j = 0$$

Esto es un conjunto de ecuaciones diferenciales lineales de segundo orden acopladas, en la forma de Sturm-Liouville. En general no es fácil desacoplar esos sistemas para encontrar una solución, por lo que se estudian con técnicas numéricas.

En el caso particular en que ni k_{ij} ni m_{ij} dependan del tiempo, recuperamos la ecuación con la que trabajamos en la sección anterior

$$m_{ij} \delta \ddot{q}_j + k_{ij} \delta q_j = 0$$

Un detalle a tener en cuenta es que ahora la matriz m no tiene por qué ser definida positiva. Por lo tanto, en la expresión

$$\omega_i^2 = \frac{\vec{c}_i^t \cdot k \cdot \vec{c}_i}{\vec{c}_i^t \cdot m \cdot \vec{c}_i}$$

pueden ahora aparecer, además de las inestabilidades taquiónicas originadas en el signo negativo del numerador, otras inestabilidades por fantasmas originadas en un signo negativo en el denominador.

12.4 Resumen

En esta clase hemos aprendido a *relajar los vínculos*, resolviendo el problema de las pequeñas oscilaciones de un sistema físico en torno a una posición de equilibrio. Hemos identificado a los sistemas estables e inestables, y clasificamos esas inestabilidades como taquiones o fantasmas. También identificamos el espectro y los modos normales de un sistema, como la solución a un problema de autovalores y autovectores.

Dos generalizaciones que no hemos explorado por razones de espacio, pero que se pueden obtener con los elementos aprendidos aquí, son las siguientes

1. Desarrollo en torno a un punto que no es una solución:

Por ejemplo, podríamos querer obtener las pequeñas oscilaciones alrededor de un mínimo de potencial cuando el potencial depende del tiempo. En este caso, el mínimo $q_j = q_j^{\min}(t)$ no es estático y por lo tanto tampoco es una solución. Esto implica que los términos lineales no desaparecen del lagrangiano para las perturbaciones, resultando en inhomogeneidades en las ecuaciones de movimiento. Es decir que tenemos un sistema de *osciladores forzados*. Es fácil probar que el fenómeno de *resonancia* se producirá cuando cualquiera de las frecuencias normales del sistema coincida con alguna frecuencia de la fuerza forzadora.

2. Desarrollo en órdenes mayores de la perturbación:

Podríamos también querer perturbar más allá del orden lineal, escribiendo $q_i = q_i^{\text{sol}} + \varepsilon \delta q_i + \varepsilon^2 \delta^{(2)} q_i$, desarrollando las ecuaciones de movimiento a orden cuadrático en ε . Esto es exactamente lo que hicimos cuando estudiamos el teorema de Bertrand en la clase sobre el problema de dos cuerpos. Es fácil ver que se obtienen también osciladores forzados, y que en este caso el fenómeno de resonancia podría generar *inestabilidades no lineales*.

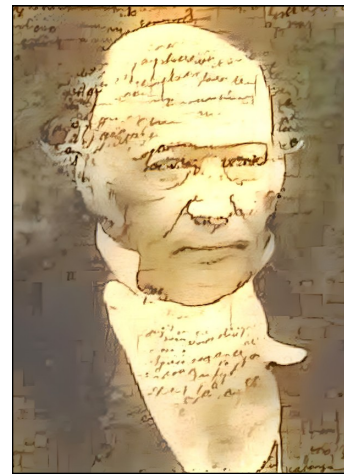
13. Principio de acción estacionaria

13.1 Objetivos

En esta clase estudiaremos un principio general que nos permitirá obtener las ecuaciones de Lagrange a partir de una sola magnitud física denominada *acción*.

Tal principio nos guiará para generalizar la Mecánica más allá de los sistemas de partículas relativistas y no relativistas con o sin vínculos que venimos estudiando hasta ahora. Veremos que hay una amplia variedad de sistemas físicos cuya estructura granular o de partículas no resulta distinguible, y que sin embargo se pueden considerar sistemas mecánicos ya que obedecen las leyes generales de la Mecánica Analítica.

Por otro lado, seremos en esta clase capaces de vislumbrar los límites de la Mecánica Analítica. Veremos que debe estar necesariamente inscrita dentro de una teoría mayor que la contenga, imaginando así la necesidad de una Mecánica Cuántica.



William Rowan Hamilton

13.2 Funcionales y cálculo variacional

En esta sección vamos a introducir el concepto matemático de *funcional*, que es una generalización de la idea de función. Además, vamos a estudiar los fundamentos del *cálculo variacional* que es el área de las matemáticas que amplía el concepto de derivada a estos nuevos objetos. Nos

motivaremos principalmente en el ejemplo sencillo de la longitud de arco de una curva que une dos puntos en el plano euclidiano.

Comencemos con la definición: a los fines de este curso, una *funcional* F es cualquier aplicación que tome una función f y nos devuelva un número. Las funcionales son objetos diferentes de las funciones que conocemos de los cursos de Análisis. En efecto, una función f nos devuelve un número $f(x)$ para cada valor de la variable x , mientras que una funcional F nos devuelve un número $F[f]$ para cada función f . Nótese que para las funcionales usamos la notación de corchete $F[f]$, que se distingue de la notación de paréntesis $f(x)$ comúnmente usada para las funciones.

Ejemplos: algunas funcionales sencillas

Una funcional que usamos todo el tiempo sin llamarla de ese modo es la *evaluación* E_{x_0} en un punto x_0 , definida como

$$E_{x_0}[f] = f(x_0)$$

Para cada función f esta funcional nos devuelve un número $f(x_0)$, que es el valor de la función evaluada en x_0 .

Una funcional un poco más compleja se puede definir a partir de una función auxiliar $g(x, y)$ y un punto x_0 de acuerdo a

$$F[f] = g(x_0, f(x_0))$$

Nuevamente obtenemos un número para cada función f , que es el valor de la función auxiliar g sobre la curva descrita en el plano por $y = f(x)$, en el punto x_0 . Si tenemos varios puntos x_n , otra funcional que podemos construir con la misma función auxiliar esta dada por

$$F[f] = \sum_n g(x_n, f(x_n))$$

que corresponde a sumar la funcional previa sobre varios puntos del eje x .

Podemos tener funcionales que dependan de varias funciones, o de su valor en varios puntos. Por ejemplo, dada una función auxiliar $g(x_n, f_i, h_i)$ donde hay D variables f_i y otras D variables h_i además de las variables x_n , podemos definir

$$F[f_i] = \sum_n f(x_n, f_i(x_n), f_i(x_{n+1}))$$

Otra funcional que usamos continuamente desde los cursos básicos sin identificarla como tal es la integral definida

$$F[f] = \int_{x_0}^{x_1} f(x) dx$$

Con ayuda de una función auxiliar $g(x, y)$, esta funcional se puede generalizar a

$$F[f] = \int_{x_0}^{x_1} g(x, f(x)) dx$$

Para otro ejemplo similar de funcional, podemos hacer uso de una función auxiliar $g(x, y, z)$ para escribir

$$F[f] = \int_{x_0}^{x_1} f(x, f(x), f'(x)) dx$$

Se hace evidente a partir de estos ejemplos que podemos definir funcionales arbitrariamente complicadas.

Ejemplo: la distancia entre dos puntos en el plano

Consideremos una curva arbitraria en el plano que una los puntos (x_0, y_0) y (x_1, y_1) , como la que se puede ver en la figura 13.2.1. Estamos interesados en medir la longitud de esa curva.

En el triángulo que se ve en el dibujo, podemos relacionar la longitud de la hipotenusa con la de sus catetos, usando el teorema de Pitágoras, según la fórmula

$$\Delta h^2 = \Delta x^2 + \Delta y^2$$

En el límite en el que Δx y Δy son muy pequeños, podemos usar la forma infinitesimal

$$dh^2 = dx^2 + dy^2 = dl^2$$

donde en la segunda igualdad aprovechamos el hecho de que la longitud de la hipotenusa es indistinguible de la longitud del fragmento de arco de la curva en el límite infinitesimal. Si describimos la curva como una función $y = f(x)$, entonces de la fórmula anterior deducimos que

$$dl^2 = (1 + f'^2) dx^2$$

Esto que nos permite escribir la longitud total simplemente integrando la longitud infinitesimal de arco a lo largo de toda la curva,

$$L[y] = \int_{x_0}^{x_1} \sqrt{1 + f'^2(x)} dx$$

donde integramos entre el valor inicial x_0 y el valor final x_1 de la variable independiente.

La fórmula de arriba nos devuelve un número $L[f]$ para cada función f , es decir que es una funcional en el sentido que hemos definido. Su dominio son todas las funciones derivables en el intervalo (x_0, x_1) tal que la expresión $\sqrt{1 + f'^2}$ tenga una integral finita.

Dependiendo de la aplicación, podemos querer restringir el mencionado dominio. Por ejemplo, para comparar longitudes de diferentes curvas que unen los mismos puntos, como haremos en el ejemplo siguiente, nos interesan funciones que cumplan $f(x_0) = y_0$ y $f(x_1) = y_1$.

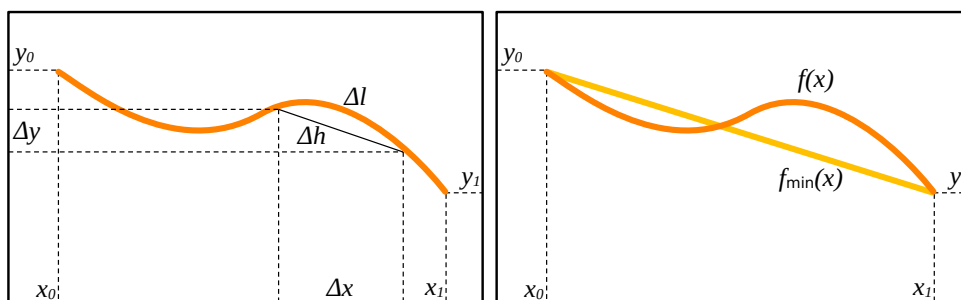


Figura 13.2.1: Longitud de arco de una curva. La fórmula para el teorema de Pitágoras aplicada al triángulo de la figura en su límite infinitesimal nos permite escribir una forma integral para la longitud de arco de una curva. La recta es la distancia más corta entre dos puntos, para lo que se deben comparar curvas que unan los puntos (x_0, y_0) y (x_1, y_1) .

En este curso vamos a estar interesados exclusivamente en funcionales de la forma general

$$F[f_i] = \int_{x_0}^{x_1} g(f'_i(x), f_i(x), x) dx$$

Donde la función auxiliar $g(f'_i, f_i, x)$ estará evaluada en las funciones f_i y sus derivadas f'_i , además de la variable x .

Un punto importante a considerar es cuál será el *dominio* de estas funcionales, es decir sobre qué tipo de funciones actuarán. Es evidente a partir de su definición que las funcionales del tipo arriba mencionado sólo pueden tomar valores en funciones f_i que tengan una derivada f'_i y para las cuales la magnitud $g(f'_i, f_i, x)$ tenga una integral finita entre los puntos x_1 y x_2 . Esto restringe el conjunto de todas las funciones posibles a un subconjunto. En lo que sigue, vamos a limitar aún más las funciones sobre las que actuarán nuestras funcionales fijando sus valores en los puntos inicial y final: nos vamos a limitar a funciones que cumplan que $f_i(x_0) = y_0^0$ y $f(x_1) = y_1^1$.

Supongamos ahora que estamos interesados en buscar, dentro de tal dominio, las funciones que maximizan o minimizan el valor de una funcional del tipo arriba definido. Es evidente que las técnicas que conocemos para encontrar los máximos o mínimos de una función no son inmediatamente aplicables aquí. En efecto, no podemos tomar la derivada respecto de la variable porque la variable misma es una función. Debemos entonces encontrar una generalización de la idea de derivada que se aplique a las funcionales del tipo propuesto.

Ejemplo: distancia más corta entre dos puntos en el plano

Conocemos la propiedad de la geometría euclidiana que dice que *la distancia más corta entre dos puntos es una recta*. Sin embargo ¿podemos probar esta afirmación?

Supongamos que $y = f_{\min}(x)$ es la curva de menor longitud que une (x_0, y_0) con (x_1, y_1) .

Entonces cualquier curva *perturbada*, es decir cualquiera que podamos describir como $y = f_{\min}(x) + \varepsilon \delta f(x)$ con ε pequeño, tendrá necesariamente una longitud mayor

$$L[f_{\min} + \varepsilon \delta f] \geq L[f_{\min}]$$

Nótese que, si bien $L[f]$ es una funcional de la función f , una vez que elegimos las funciones f y δf la expresión $L[f_{\min} + \varepsilon \delta f]$ es una función de la variable ε en el sentido usual: devuelve un número para cada valor de ε . Esta observación nos permite usar lo que sabemos de los cursos de cálculo, expandiendo dicha función en potencias de ε a primer orden. Obtenemos

$$L[f_{\min}] + \varepsilon \delta L \geq L[f_{\min}]$$

donde hemos definido el coeficiente δL según

$$\delta L = \left. \frac{\partial}{\partial \varepsilon} L[f_{\min} + \varepsilon \delta f] \right|_{\varepsilon=0}$$

En la expansión de arriba, si el término lineal en ε no es nulo, entonces no estamos evaluando la funcional en su mínimo. En efecto, si una perturbación parametrizada por ε hace que el término lineal sea positivo, entonces la perturbación dada por $-\varepsilon$ lo hará negativo, resultando en una longitud menor para la curva perturbada. En conclusión, la única manera de que f_{\min} corresponda realmente a la longitud mínima es que el término lineal se anule. Para encontrarlo explícitamente escribimos

$$L[f_{\min} + \varepsilon \delta f] = \int_{x_0}^{x_1} \sqrt{1 + (f'_{\min}(x) + \varepsilon \delta f'(x))^2} dx$$

Expandiendo el integrando en potencias de ε y reordenando, obtenemos

$$L[f_{\min} + \varepsilon \delta f] = \int_{x_0}^{x_1} \sqrt{1 + f'^2_{\min}(x)} dx + \varepsilon \int_{x_0}^{x_1} \frac{f'_{\min}(x)}{\sqrt{1 + f'^2_{\min}(x)}} \delta f'(x) dx + \mathcal{O}(\varepsilon)^2$$

Esto nos permite tomar la derivada respecto de ε para identificar δL en la forma

$$\delta L = \int_{x_0}^{x_1} \frac{f'_{\min}(x)}{\sqrt{1 + f'^2_{\min}(x)}} \delta f'(x) dx$$

En esta expresión podemos integrar por partes para obtener

$$\delta L = \left(\frac{f'_{\min}(x)}{\sqrt{1 + f'^2_{\min}(x)}} \delta f(x) \right) \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \left(\frac{f'_{\min}(x)}{\sqrt{1 + f'^2_{\min}(x)}} \right)' \delta f(x) dx$$

Ahora bien, queremos comparar longitudes de diferentes curvas que unen los mismos puntos (x_0, y_0) y (x_1, y_1) . Para esto necesitamos que las curvas parametrizadas por $y = f_{\min}(x)$ y $y = f_{\min}(x) + \varepsilon \delta f(x)$ empiecen y terminen en (x_0, y_0) y (x_1, y_1) , lo que necesariamente implica la condición $\delta f(x_0) = \delta f(x_1) = 0$. Con esto, el primer término se anula y la expresión se reduce a

$$\delta L = - \int_{x_0}^{x_1} \left(\frac{f'_{\min}(x)}{\sqrt{1 + f'^2_{\min}(x)}} \right)' \delta f(x) dx$$

Es decir que hemos identificado el término lineal δL que escribimos más arriba. Como ya discutimos, si la curva $y = f_{\min}(x)$ es realmente la de menor longitud, tiene que

cumplirse que tal término lineal se anule. Dado que es una integral que tiene que anularse para cualquier función $\delta f(x)$, es necesario que el factor en el integrando se anule. En otras palabras

$$\frac{\delta L}{\delta f} = - \left(\frac{f'_{\min}}{\sqrt{1 + f'^2_{\min}}} \right)' = 0$$

Donde hemos definido la *derivada variacional* $\delta L/\delta f$ de la funcional $L[f]$, que es una función de la variable x . De aquí se puede deducir

$$\frac{f'_{\min}}{\sqrt{1 + f'^2_{\min}}} = \text{constante}$$

y por lo tanto $f'_{\min} = \text{constante}$. Esto implica necesariamente que la curva $y = f_{\min}(x)$ es una recta, con lo que hemos probado la afirmación de que *la distancia más corta entre dos puntos en el plano es la recta*.

Claro que en realidad con este cálculo sólo hemos demostrado que la longitud es *estacionaria*. Es decir, nuestros razonamientos también funcionan si se tratara de un máximo. Para determinar que lo que hemos encontrado es realmente de un mínimo, se deja como ejercicio demostrar que la contribución cuadrática en ε es siempre positiva.

En el ejemplo anterior fuimos capaces de encontrar, para la funcional longitud de arco, una derivada funcional que se comporta de un modo análogo a la derivada de una función, anulándose al evaluarla en el punto estacionario de la funcional.

Así como no todas las funciones son derivables, lo mismo sucede para las funcionales y no cualquier funcional posee una derivada funcional. Sin embargo, para la clase particular de funcionales que definimos más arriba, el problema tiene una solución sencilla que exploraremos a continuación.

Veamos como cambia una funcional de esa clase cuando nos movemos de una función f_i a una función ligeramente perturbada $f_i + \varepsilon \delta f_i$. Nótese que tanto la función inicial como la perturbada deben estar en el dominio de la funcional, esto significa que deben tomar los valores y_i^0 e y_i^1 en los puntos x_0 y x_1 respectivamente. Por lo tanto se debe cumplir la restricción para las variaciones $\delta f_i(x_0) = \delta f_i(x_1) = 0$. Reemplazando la función perturbada en la funcional, tenemos

$$F[f_i + \varepsilon \delta f_i] = \int_{x_0}^{x_1} g(f'_i(x) + \varepsilon \delta f'_i(x), f_i(x) + \varepsilon \delta f_i(x), x) dx$$

Dado que ε es por definición un parámetro pequeño y que las perturbaciones δf_i son acotadas y suaves, podemos expandir esta expresión a orden lineal en ε para obtener

$$\begin{aligned} F[f_i + \varepsilon \delta f_i] &= \int_{x_0}^{x_1} g(f'_i(x), f_i(x), x) dx + \\ &+ \varepsilon \int_{x_0}^{x_1} \left(\frac{\partial g}{\partial f'_j}(f'_i(x), f_i(x), x) \delta f'_j(x) + \frac{\partial g}{\partial f_j}(f'_i(x), f_i(x), x) \delta f_j(x) \right) dx \end{aligned}$$

En la segunda línea de esta expresión, podemos identificar el coeficiente del término lineal en ε como la integral

$$\delta F = \int_{x_0}^{x_1} \left(\frac{\partial g}{\partial f'_j}(f'_i(x), f_i(x), x) \delta f'_j(x) + \frac{\partial g}{\partial f_j}(f'_i(x), f_i(x), x) \delta f_j(x) \right) dx$$

Integrando por partes en el primer término para eliminar las derivadas de δf_j obtenemos

$$\begin{aligned} \delta F = \int_{x_0}^{x_1} \left(\frac{\partial g}{\partial f_j}(f'_i(x), f_i(x), x) - \frac{d}{dx} \left(\frac{\partial g}{\partial f'_j}(f'_i(x), f_i(x), x) \right) \right) \delta f_j(x) dx + \\ + \left(\frac{\partial g}{\partial f'_j}(f'_i(x), f_i(x), x) \delta f_j(x) \right) \Big|_{x_0}^{x_1} \end{aligned}$$

La segunda línea se anula dada la restricción que impusimos para que las funciones perturbadas se mantengan en el dominio de nuestra funcional $\delta f_i(x_1) = \delta f_i(x_2) = 0$. Es decir que la contribución lineal a la variación de nuestra funcional vendrá dada por la primera línea. Para que esta contribución se anule cualquiera sea la variación δf_i , es condición necesaria que el integrando se anule, es decir

$$\frac{\delta F}{\delta f_j} = \frac{\partial g}{\partial f_j}(f'_i, f_i, x) - \frac{d}{dx} \left(\frac{\partial g}{\partial f'_j}(f'_i, f_i, x) \right) = 0$$

Donde hemos identificado la derivada variacional de nuestra funcional. Nótese la similitud de esta expresión con las ecuaciones de Lagrange, algo que explotaremos un poco más adelante en esta clase.

Ejemplo: distancia más corta entre dos puntos en el espacio tridimensional

Probaremos ahora que la distancia más corta entre dos puntos en el espacio tridimensional también es una recta.

Dada una curva cualquiera que une los puntos $\vec{r}_0 = (x_0, y_0, z_0)$ y $\vec{r}_1 = (x_1, y_1, z_1)$ en el espacio, la podemos parametrizar como $\vec{r} = (x(\tau), y(\tau), z(\tau))$. Vamos a limitarnos a curvas suaves y que no se vayan a infinito en ningún punto, con lo que x, y, z son funciones continuas y finitas con derivada finita del parámetro τ . Además, deben cumplir que $x(\tau_0) = x_0, y(\tau_0) = y_0, z(\tau_0) = z_0$ y $x(\tau_1) = x_1, y(\tau_1) = y_1, z(\tau_1) = z_1$. Por otro lado, usando el teorema de Pitágoras en tres dimensiones, tenemos que

$$dl = \sqrt{x'^2 + y'^2 + z'^2} d\tau$$

Entonces podemos escribir la longitud de una curva cualquiera como

$$L[x, y, z] = \int_{\tau_0}^{\tau_1} \sqrt{x'^2(\tau) + y'^2(\tau) + z'^2(\tau)} d\tau$$

donde vemos que se obtiene una funcional del tipo especial para el cual pudimos definir una derivada variacional más arriba, y que estamos evaluándola en un conjunto de funciones con las limitaciones de nuestro caso. Aquí la variable x de nuestra fórmula general corresponde al parámetro τ , y las funciones f_i son tres y están dadas por x, y, z , mientras que la función auxiliar está dada por $g(x', y', z') = \sqrt{x'^2 + y'^2 + z'^2}$.

Con esto, tenemos que la contribución lineal a la variación de nuestra funcional se anulará cuando se anulen las derivadas variacionales

$$\frac{\delta L}{\delta x} = \frac{d}{d\tau} \left(\frac{\partial g}{\partial x'} \right) - \frac{\partial g}{\partial x} = \left(\frac{x'}{\sqrt{x'^2 + y'^2 + z'^2}} \right)' = 0$$

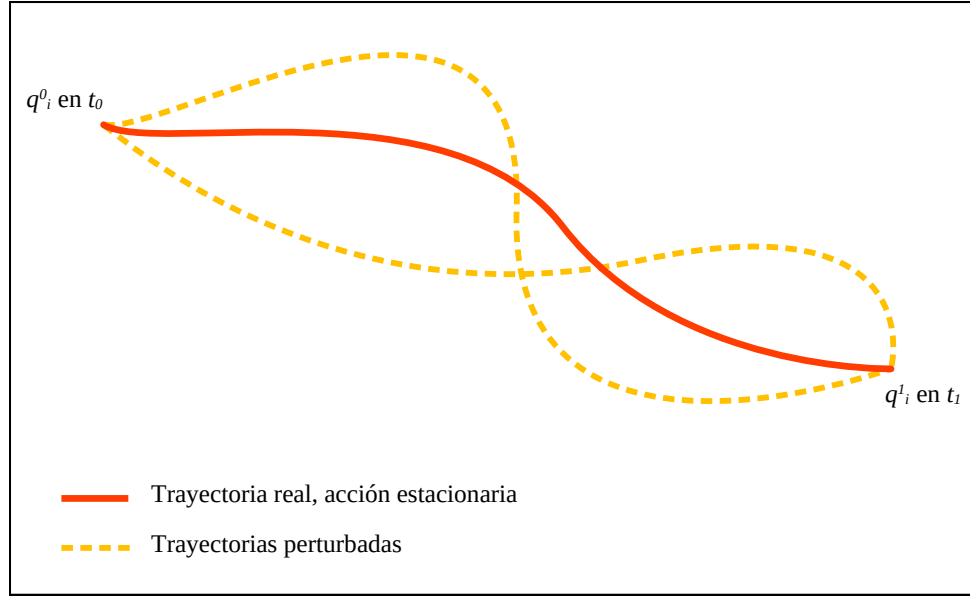


Figura 13.2.2: Principio de acción estacionaria. La trayectoria real del sistema entre un punto q_i^0 en el instante t_0 y un punto q_i^1 en el instante t_1 corresponde de entre todas las trayectorias posibles a aquella que hace que la acción sea estacionaria.

$$\frac{\delta L}{\delta y} = \frac{d}{d\tau} \left(\frac{\partial g}{\partial y'} \right) - \frac{\partial g}{\partial y} = \left(\frac{y'}{\sqrt{x'^2 + y'^2 + z'^2}} \right)' = 0$$

$$\frac{\delta L}{\delta z} = \frac{d}{d\tau} \left(\frac{\partial g}{\partial z'} \right) - \frac{\partial g}{\partial z} = \left(\frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} \right)' = 0$$

Esto implica que las cantidades entre paréntesis deben ser constantes, y por lo tanto las magnitudes x', y', z' también lo son. Esto necesariamente resulta en que la curva en cuestión es una recta que une los puntos inicial y final.

Nuevamente, esto solo prueba que para la recta la distancia es estacionaria, pero no demuestra que se trate realmente de un mínimo, podría ser un máximo o un punto silla. Para terminar la demostración se deja como ejercicio escribir la contribución cuadrática y probar que es positiva.

Ejercicio:

Encontrar la función θ que hace estacionaria la funcional que mide la longitud de un arco de una curva dibujada sobre una esfera, que se define como

$$L[\theta] = \int_{\phi_0}^{\phi_1} \sqrt{1 + \sin^2 \phi \theta'^2(\phi)} d\phi$$

con las condiciones $\theta(\phi_0) = \theta_0$ y $\theta(\phi_1) = \theta_1$. Identifique entre las soluciones cuales son los máximos y cuales los mínimos de la funcional ¿a qué curvas corresponden sobre la esfera?

13.3 Principio de Hamilton de acción estacionaria

Hasta ahora hemos escrito las ecuaciones de Lagrange a partir de un lagrangiano $L(\dot{q}_i, q_i, t)$ siguiendo la regla que obtuvimos en las primeras clases

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

Una pregunta natural que podríamos hacernos es ¿cuál es el origen de esta regla? ¿Existe acaso algún principio más fundamental que desemboque en esa combinación particular de derivadas del lagrangiano y no en otra?

Para tratar de responder esta cuestión, comenzamos definiendo la *acción* $S[q_i]$ de un sistema mecánico que se mueve entre la posición q_i^0 en el instante t_0 y la posición q_i^1 en el instante t_1 , según la integral

$$S[q_i] = \int_{t_0}^{t_1} L(\dot{q}_i(t), q_i(t), t) dt$$

Esta integral será un número real, que será diferente para diferentes curvas $q_i(t)$ que unan los mismos puntos q_i^0 y q_i^1 . Si identificamos el tiempo con nuestra variable x , las coordenadas generalizadas $q_i(t)$ con las funciones $f_i(x)$, y el lagrangiano $L(\dot{q}_i, q_i, t)$ con nuestra función auxiliar $g(f_i', f_i, x)$, la acción es una funcional de la forma general que estudiamos más arriba. Su derivada variacional toma entonces la forma

$$\frac{\delta S}{\delta q_i} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i}$$

De donde vemos que las ecuaciones de Lagrange no son sino la condición de que la derivada variacional se anule. En otras palabras, *los sistemas físicos siguen las trayectorias que hacen que su acción sea estacionaria*.

En una primera impresión, el principio de acción estacionaria puede parecer un poco anti-intuitivo. En efecto, para seleccionar de entre todas las trayectorias aquella que hace que la acción sea estacionaria, el sistema debería de algún modo tener acceso a la información sobre el valor de la acción en todas las trayectorias posibles, incluso aquellas que no recorrió.

Para explorar un poco más esta paradoja, analicemos otra instancia de la física en la cual aparece un principio similar.

Ejemplo: el principio de Fermat del tiempo mínimo

En óptica geométrica existe una ley que reza que *un rayo de luz selecciona el camino que minimiza el tiempo que le toma llegar desde su origen hasta su destino*.

La velocidad la luz en un medio material está dada por $v = c/n(\vec{r})$ donde $n(\vec{r})$ es el índice de refracción, que puede en principio variar de punto a punto. Tenemos entonces que el tiempo T que le toma a un rayo de luz llegar desde un punto \vec{r}_0 a un punto \vec{r}_1 estará dado por

$$T = \int_{\vec{r}_0}^{\vec{r}_1} \frac{d\ell}{v}$$

Si parametrizamos la trayectoria del rayo como $\vec{r}(s)$ en términos de un parámetro real s , entonces el tiempo pasa a ser una funcional de $\vec{r}(s)$ con la forma

$$T[\vec{r}] = \int_{s_0}^{s_1} \frac{1}{c} n(\vec{r}(s)) |\vec{r}'(s)| ds$$

Vemos entonces que, al pedir que el tiempo sea mínimo, las ecuaciones resultantes no serán otra cosa que la condición de que se anule la derivada funcional de $T[\vec{r}]$.

Ejemplo: ley de Snell

Una de las consecuencias del principio del tiempo mínimo, para la forma particular de la función $n(\vec{r})$ que describe una interface entre dos medios, es lo que en los cursos básicos se estudia como *ley de Snell*. Para demostrarla, escribamos

$$n(x, y, z) = \begin{cases} n(x, y, z) = n_0 & \text{para } z > 0 \\ n(x, y, z) = n_1 & \text{para } z < 0 \end{cases}$$

En este caso la funcional tiempo que definimos arriba se puede escribir

$$T[\vec{r}(s)] = \frac{n_0}{c} \int_{s_0}^{s_{\text{int}}} |\vec{r}'(s)| ds + \frac{n_1}{c} \int_{s_{\text{int}}}^{s_1} |\vec{r}'(s)| ds$$

Donde s_{int} es el valor del parámetro s para el cual el rayo cruza desde la región superior $z > 0$ a la inferior $z < 0$. Comparando con la fórmula del ejemplo anterior para la longitud de una curva en el espacio tridimensional, vemos que

$$T[\vec{r}(s)] = \frac{n_0}{c} L(s_0, s_{\text{int}}) + \frac{n_1}{c} L(s_{\text{int}}, s_1)$$

Siendo $L(s, \tilde{s})$ la longitud de la curva como función de los valores extremos s y \tilde{s} del parámetro. Ambos términos serán mínimos cuando la curva que va de $\vec{r}(s)$ a $\vec{r}(\tilde{s})$ sea una recta. Esto nos da

$$T(\vec{r}_{\text{int}}) = \frac{n_0}{c} |\vec{r}_0 - \vec{r}_{\text{int}}| + \frac{n_1}{c} |\vec{r}_1 - \vec{r}_{\text{int}}|$$

Donde al reemplazar los tramos $z > 0$ y $z < 0$ con rectas, hemos reducido la funcional tiempo a una función del punto intermedio \vec{r}_{int} . Eligiendo el sistema de coordenadas tal que el plano yz corresponde al que determinan los tres puntos \vec{r}_0, \vec{r}_1 y \vec{r}_{int} tenemos

$$T(y_{\text{int}}) = \frac{n_0}{c} \sqrt{(y_0 - y_{\text{int}})^2 + z_0^2} + \frac{n_1}{c} \sqrt{(y_1 - y_{\text{int}})^2 + z_1^2}$$

Esto será un mínimo cuando se anule su derivada

$$\frac{dT}{dy_{\text{int}}} = \frac{n_0(y_0 - y_{\text{int}})}{c \sqrt{(y_0 - y_{\text{int}})^2 + z_0^2}} + \frac{n_1(y_1 - y_{\text{int}})}{c \sqrt{(y_1 - y_{\text{int}})^2 + z_1^2}} = 0$$

Usando un poco de trigonometría esto se puede reescribir como

$$\frac{dT}{dy_{\text{int}}} = \frac{1}{c} (n_0 \sin \theta_0 - n_1 \sin \theta_1) = 0$$

de donde inmediatamente se deduce la ley de Snell.

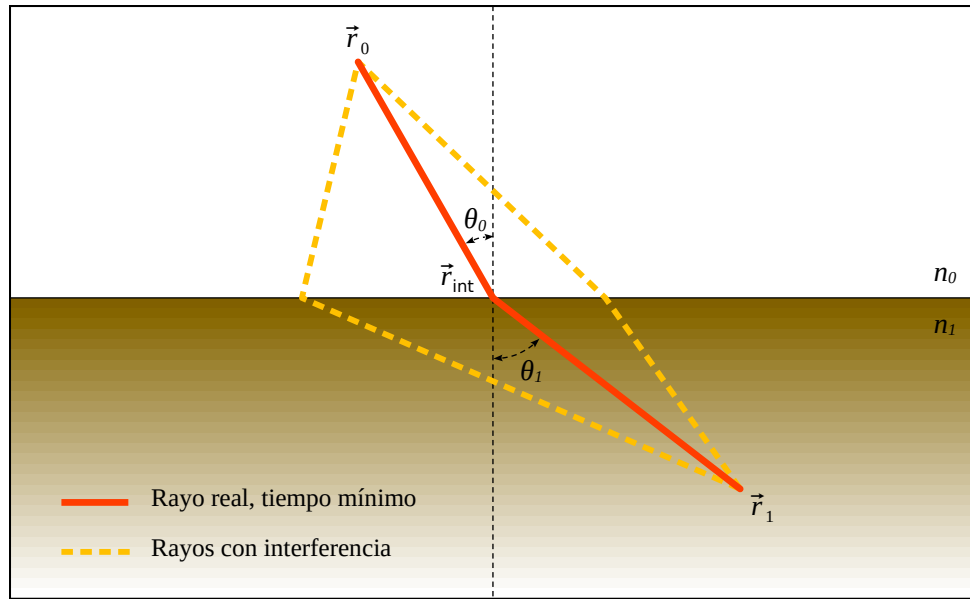


Figura 13.3.1: Principio de Fermat y ley de Snell. De entre todos los rayos de luz que pueden unir el punto \vec{r}_0 con el punto \vec{r}_1 , la mínima interferencia tiene lugar a lo largo del que cumple la ley de Snell, que es el que toma tiempo mínimo según el principio de Fermat.

Nuevamente tenemos la situación un tanto extraña de que un sistema físico (en este caso el rayo de luz) parece conocer el valor que tomaría el tiempo en trayectorias que no ha recorrido, de manera de seleccionar como su trayectoria real aquella donde su valor es mínimo.

Sin embargo, en el caso de la luz sabemos que la óptica geométrica, que nos permite hablar de *rayos de luz* que cumplen el principio de Fermat, es sólo una aproximación. En realidad, la luz es una onda que se propaga en el campo electromagnético que llena todos los puntos del espacio, siendo el rayo de luz una aproximación a la línea de mínima interferencia, que sólo es válida cuando la longitud de onda es muy pequeña comparada con los obstáculos del entorno.

Por lo anterior, podemos decir que la luz ha realmente explorado todas las trayectorias posibles para determinar cuál es la del tiempo mínimo: se mueve como una onda que llena el espacio y que interfiere destructivamente en todos lados menos a lo largo de dicha trayectoria. Para entender de qué modo se realiza dicha exploración, debemos realizar la *aproximación iconal*.

Ejemplo: aproximación iconal

Imaginemos que queremos resolver la ecuación de onda en una región del espacio donde la velocidad de la luz v es constante. Podemos escribir naturalmente una solución con forma de onda plana

$$\Phi(\vec{r}, t) = e^{-i(\omega t - I(\vec{r}))}$$

donde Φ es el potencial eléctrico, y $I(\vec{r})$ es una función conocida como *iconal*, que en el caso de la onda plana se escribe en términos del vector de onda como $I(\vec{r}) = \vec{k} \cdot \vec{r}$. Nótese que en este caso se cumple que $\vec{\nabla} I = \vec{k}$ y para la segunda derivada tenemos

$\nabla^2 I = 0$ porque \vec{k} es constante.

Ahora bien, si la velocidad de la luz no es constante porque el medio cambia suavemente sus propiedades de punto a punto, entonces la relación $v|k| = \omega$ implica que el vector de onda $\vec{k} = \vec{\nabla} I$ también cambiará suavemente $\vec{\nabla} \cdot \vec{k} \ll k^2$. En otras palabras, tendremos un iconal cuyo laplaciano no será nulo sino muy pequeño $\nabla^2 I \ll (\nabla I)^2$. Esto implica que podemos escribir

$$\vec{\nabla} \Phi = -i\Phi \vec{\nabla} I$$

$$\nabla^2 \Phi = \Phi (-i\nabla^2 I - (\nabla I)^2) \approx -(\nabla I)^2 \Phi$$

Reemplazando estas derivadas en la ecuación de onda, obtenemos

$$\partial_t^2 \Phi - v^2 \nabla^2 \Phi = (-\omega^2 + v^2 (\nabla I)^2) \Phi = 0$$

Una solución de lo cual es

$$I(\vec{r}) = \omega \int_{\vec{r}} \frac{dl}{v} = \omega T$$

donde la integral se calcula a lo largo de una curva cualquiera que une el punto inicial y el punto final \vec{r} . La solución general de la ecuación incluye una superposición sobre todas las soluciones que corresponden a las diferentes curvas. Sin embargo, todas estas curvas interferirán destructivamente entre sí, con más interferencia cuanto mayor sea su diferencia de camino óptico, que en nuestra solución está dada por la diferencia entre sus correspondientes iconales. Para ver esto, escribamos

$$\Phi = e^{-i(\omega t - I_1(r))} + e^{-i(\omega t - I_2(r))} + \dots$$

Donde cada término está calculado según la fórmula de arriba usando curvas distintas, y los puntos suspensivos recorren todas las posibles curvas. Es fácil ver que para cualquier par de términos tenemos $I_2 - I_1 = \Delta I$ siendo ΔI la diferencia de camino óptico entre las curvas. Entonces podemos escribir

$$\Phi = e^{-i(\omega t - I_{\min}(r))} (1 + e^{i\Delta I} + \dots)$$

Donde I_{\min} es la solución para la cual el iconal es mínimo, y la suma recorre los diferentes ΔI entre cualquier otra curva y la que corresponde a I_{\min} . Todos estos términos tienen una fase no nula y por lo tanto contribuyen a la suma con un número de módulo menor que uno y que cambia de signo a medida que recorremos diferentes curvas. Esto resulta en una cancelación que solo deja el prefactor que contiene I_{\min} .

Es decir que la solución que tendrá menos interferencia será aquella cuyo iconal sea más pequeño, de donde inmediatamente se deduce el principio de Fermat del tiempo mínimo.

Volviendo a la mecánica, la consecuencia de la discusión es inmediata: el principio de acción estacionaria nos sugiere que tiene que existir algún tipo de onda asociada a cualquier sistema mecánico. Esta onda se propaga en un medio que llena todo el espacio, de modo tal que su interferencia es mínima a lo largo de las trayectorias que minimizan la acción. Es fácil imaginar que se trata de la función de onda de Schrödinger.

Es decir que, con conocimientos que ya estaban disponibles durante el siglo XIX, era posible inferir la existencia de una onda asociada a las partículas materiales. La Mecánica Clásica nos advierte que debe existir una Mecánica Cuántica.

Nota

La película *Arrival* (2016) está basada en el maravilloso cuento del autor estadounidense Ted Chiang titulado *La historia de tu vida*. En él, una raza extraterrestre llega a la Tierra, y una lingüista y un físico intentan entablar comunicación con los visitantes.

Los primeros intentos resultan infructuosos, en parte debido al curioso sistema de escritura de los extraterrestres, donde la frase parece tomar sentido solo después de terminada, y de la cual nada dice una lectura parcial. Más aún, al intentar intercambiar conceptos básicos de física, no parece haber entendimiento alguno. La situación parece estancada, hasta que el físico menciona el principio de Fermat, el cual despierta inmediatamente el interés de los visitantes.

Poco a poco, los investigadores se dan cuenta de que los extraterrestres no perciben el tiempo como un devenir, sino como una imagen completa y simultánea de toda la historia de vida de cada individuo.

Al aprender el lenguaje de los extraterrestres, la protagonista comienza a pensar como ellos, en una realización de la tesis lingüística conocida como *hipótesis de Whorf-Sapir*. Según esta idea, el aprender un idioma nos nutre de una visión del mundo que estaba imbuída en el mismo. Así, la lingüista del cuento experimenta en tiempo presente la interacción con una hija que aún no ha nacido.

De ese relato proviene esta hermosa cita, que resume en pocas líneas el sentimiento de impotencia que genera un hijo que aprende a caminar:

«Cuando aprendas a caminar tendré una demostración cotidiana de la asimetría de nuestra relación. Correrás incesantemente de un lado para otro, y cada vez que choques contra el marco de una puerta o te hagas un arañazo en la rodilla, sentiré el dolor como si fuera mío. Será como si me creciera un miembro errante, una extensión de mí misma cuyos nervios sensores transmiten el dolor perfectamente, pero cuyos nervios motores no obedecen en absoluto mis órdenes. Es tan injusto: voy a dar a luz un muñeco vudú de mí misma que está dotado de vida. No vi esto en el contrato cuando me apunté. ¿Era esto parte del trato?»

Nota:

Entre sus invaluables contribuciones a la Mecánica, William Rowan Hamilton se ocupó de la descripción de las rotaciones en términos de *cuaterniones*, que hoy es de gran utilidad en la manipulación de gráficos tridimensionales en los juegos de computadora.

Cuenta la historia que la solución al problema de multiplicación de cuaterniones se le ocurrió a Hamilton mientras paseaba con su esposa por la campiña irlandesa y, para no olvidarla, la talló con su navaja en el puente de Brougham.

Una primera consecuencia del principio de mínima acción es que las ecuaciones de movimiento no cambian si sumamos al lagrangiano la derivada total respecto del tiempo de una función arbitraria de las coordenadas y el tiempo. Es decir si escribimos

$$\tilde{L}(\dot{q}_i, q_i, t) = L(\dot{q}_i, q_i, t) + \frac{d}{dt}F(q_i, t)$$

con $F(q_i, t)$ una función cualquiera. Se puede probar directamente que las ecuaciones de Lagrange resultantes para \tilde{L} son iguales a las que teníamos para L . Sin embargo, es mucho más sencillo hacerlo utilizando la acción

$$\tilde{S}[q_i] = \int_{t_0}^{t_1} \tilde{L}(\dot{q}_i, q_i, t) dt = \int_{t_0}^{t_1} \left(L(\dot{q}_i, q_i, t) + \frac{d}{dt}F(q_i, t) \right) dt = S[q_i(t)] + F(q_i, t) \Big|_{t_0}^{t_1}$$

Donde vemos que al evaluar la acción en $q_i + \varepsilon \delta q_i$ tendremos que

$$\tilde{S}[q_i + \varepsilon \delta q_i] = S[q_i + \varepsilon \delta q_i] + F(q_i, t) \Big|_{t_0}^{t_1}$$

ya que los términos con F están evaluados en los extremos de la trayectoria donde $\delta q_i = 0$. Esto implica que las derivadas funcionales, que se obtienen del término lineal en ε , serán iguales. Luego las ecuaciones de Lagrange que se derivan de $\tilde{S}[q_i]$ coincidirán con las que se obtienen de $S[q_i]$.

13.3.1 Acción para las ecuaciones de Hamilton

Como vimos en las clases precedentes, el movimiento de un sistema mecánico puede también ser descrito en términos de ecuaciones de Hamilton, según

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

Donde $H(p_i, q_i, t)$ es el hamiltoniano del sistema, a partir del cual podemos obtener un lagrangiano como la transformada de Legendre del respecto de los impulsos. Para esto, invertimos la primera ecuación para obtener $p_i = w_i(\dot{q}_j, q_j, t)$ y escribimos

$$L(\dot{q}_j, q_j, t) = \dot{q}_i w_i(\dot{q}_j, q_j, t) - H(w_i(\dot{q}_j, q_j, t), q_i, t)$$

Por supuesto, al usar este lagrangiano en el principio de acción estacionaria obtendremos las ecuaciones de Lagrange correspondientes. Sin embargo, es natural preguntarnos si podríamos obtener directamente las ecuaciones de Hamilton a partir de una condición de acción estacionaria

para algún lagrangiano adecuado. Para probar esto, modificamos la fórmula anterior definiendo un nuevo *lagrangiano de primer orden*, según

$$L_{1er}(p_j, \dot{q}_j, q_j, t) = \dot{q}_i p_i - H(p_i, q_i, t)$$

En esta nueva expresión, los impulsos no son funciones de las coordenadas y velocidades sino que son magnitudes independientes. Podemos usar este lagrangiano en una acción, según

$$S_{1er}[p_i, q_i] = \int_{t_0}^{t_1} dt L_{1er}(p_j, \dot{q}_j, q_j, t)$$

Para obtener las derivadas funcionales de esta acción, tenemos que perturbar no solo las coordenadas $q_i + \varepsilon \delta q_i$ sino también los impulsos $p_i + \varepsilon \delta p_i$, de modo de poder escribir

$$S_{1er}[p_i + \varepsilon \delta p_i, q_i + \varepsilon \delta q_i] = S_{1er}[p_i, q_i] + \varepsilon \int_{t_0}^{t_1} dt \left(\left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \delta q_i \right) + \delta q_i \Big|_{t_0}^{t_1}$$

Si asumimos que $\delta q_i(t_0) = \delta q_i(t_1) = 0$, obtenemos la condición de acción estacionaria en la forma de las ecuaciones de Hamilton.

13.3.2 Generalización de los sistemas mecánicos

Hasta este punto, hemos construido sistemas mecánicos a partir de un conjunto de partículas que interactúan mediante fuerzas bastante generales, y eventualmente imponiendo vínculos sobre el Lagrangiano resultante. Sin embargo, es lícito hacerse la pregunta de si esos son todos los sistemas mecánicos que pueden existir.

En esta clase hemos aprendido que las ecuaciones de Lagrange que obedecen la totalidad de los sistemas que conocemos, pueden obtenerse a partir de un principio de acción estacionaria. En estos sistemas, la acción está dada por una funcional de las trayectorias de un tipo bastante especial.

Llegamos entonces a la siguiente cuestión ¿qué pasa si definimos la acción en términos de funcionales más generales? ¿obtendremos tal vez sistemas mecánicos que no correspondan a sistemas newtonianos de partículas componentes?

En los ejemplos que siguen veremos algunas posibilidades, junto con sus dificultades y virtudes.

Ejemplo: sistemas en derivadas mayores.

El principio de acción estacionaria permite definir sistemas mecánicos cuyas ecuaciones de movimiento contienen derivadas temporales de orden mayor al segundo. Por ejemplo, podemos estudiar un sistema cuya acción se lee

$$S[q_i] = \int_{t_0}^{t_1} dt L(\ddot{q}_i, \dot{q}_i, q_i, t)$$

Si tuviéramos que escribir las ecuaciones de movimiento para un lagrangiano dependiente de \ddot{q}_i no sabríamos cómo hacerlo. Sin embargo, con el principio de acción

estacionaria el problema se vuelve sencillo: hay que obtener el cambio de la acción a primer orden bajo la variación $q_i(t) + \varepsilon \delta q_i(t)$ y pedir que sea nulo. Para esto hacemos

$$S[q_i + \varepsilon \delta q_i] = S[q_i] + \varepsilon \int_{t_0}^{t_1} dt \left(\frac{\partial L}{\partial \ddot{q}_i} \delta \ddot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i + \frac{\partial L}{\partial q_i} \delta q_i \right)$$

De aquí podemos identificar el coeficiente δS del término lineal en ε , el cual después de algunas integraciones por partes queda escrito en la forma

$$\delta S = \int_{t_0}^{t_1} dt \left(\frac{d^2}{dt^2} \left(\frac{\partial L}{\partial \ddot{q}_i} \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial q_i} \right) \delta q_i$$

Donde hemos usado que $\delta q_i(t_1) = \delta q_i(t_2) = 0$ y hemos impuesto la condición adicional $\delta \dot{q}_i(t_1) = \delta \dot{q}_i(t_2) = 0$. Las ecuaciones de movimiento resultantes son entonces

$$\frac{d^2}{dt^2} \left(\frac{\partial L}{\partial \ddot{q}_i} \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial q_i} = 0$$

Esto se puede generalizar sin dificultad para sistemas de orden arbitrario.

Sin embargo, un punto interesante es que todos los sistemas de orden superior tienen una inestabilidad conocida como *fantasma de Ostrogradsky* que los vuelve poco útiles para representar situaciones físicas realistas. No vamos a exponer aquí una demostración general, sino solo un ejemplo: sea el sistema

$$S = \int_{t_0}^{t_1} dt \left(\frac{1}{2} \ddot{q}^2 - \frac{1}{2} \alpha^4 q^2 \right)$$

Su ecuación de movimiento será, de acuerdo a lo que vimos más arriba

$$\frac{d^4 q}{dt^4} - \alpha^4 q = 0$$

Complexificando esta ecuación y proponiendo la solución de prueba $q(t) = e^{i\omega t}$, obtenemos $\omega^4 = \alpha^4$ o en otras palabras $\omega = \pm \alpha$ y $\omega = \pm i\alpha$. Es decir que, independientemente de que α sea real o imaginario, tendremos soluciones exponenciales crecientes, que corresponden a una inestabilidad. Esto sucede incluso si cambiamos el signo delante de α^4 en el lagrangiano.

Se puede probar, realizando una expansión del lagrangiano en torno a una solución cualquiera en el estilo de las que investigamos en el capítulo 12, que esta inestabilidad aparece de modo completamente general de este tipo de sistemas.

Ejemplo: sistemas no locales en el tiempo.

Otro tipo de sistemas cuyas ecuaciones de movimiento podemos obtener a partir del principio de mínima acción, son sistemas no locales en el tiempo. Supongamos una acción del tipo

$$S[q_i] = \iint dt dt' L(\dot{q}_i(t), \dot{q}_i(t'), q_i(t), q_i(t'), t, t')$$

Aquí el lagrangiano depende del estado del sistema en dos tiempos diferentes. De nuevo, no podríamos escribir sus ecuaciones de Lagrange sin contrar con algún

principio adicional que nos ayude a determinar su forma. Con el principio de acción estacionaria en cambio, esto es relativamente sencillo, resultando en las ecuaciones de movimiento

$$\int_{t_0}^{t_1} dt' \left(\frac{\partial \tilde{L}}{\partial q_i(t)} - \frac{d}{dt} \left(\frac{\partial \tilde{L}}{\partial \dot{q}_i(t)} \right) + \frac{\partial L}{\partial q_i(t)} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i(t)} \right) \right) = 0$$

Donde en $\tilde{L} = L(\dot{q}_i(t'), \dot{q}_i(t), q_i(t'), q_i(t), t', t)$ se han intercambiado t y t' .

Sin embargo, estos sistemas tienen problemas *de causalidad*. Esto se puede ver con un ejemplo sencillo definido en términos de una función $K(t, t')$ según

$$S = \iint dt dt' \left(\frac{1}{2} \dot{q}^2 - \frac{1}{2} K(t, t') q(t) q(t') \right)$$

La ecuación de movimiento que se obtiene de aquí es, usando la fórmula de más arriba

$$\ddot{q}(t) = \frac{1}{2} \int dt' (K(t, t') q(t') + K(t', t) q(t))$$

Aquí vemos que la aceleración de q en el tiempo t depende no sólo de la historia pasada del sistema, es decir de $t' < t$, sino también de su historia futura $t' > t$. Esto no se puede evitar eligiendo de ningún modo la función $K(t, t')$. Este es un problema que aparece de manera general en estos sistemas.

Vemos que el principio de acción estacionaria nos permite definir sistemas mecánicos mucho más generales que los que veníamos discutiendo, pero que esta libertad trae en general aparejados diversos problemas que pueden ser difíciles de controlar.

Vamos a ver a continuación varios ejemplos en los que la generalización sí funciona, y nos abre la puerta a un universo de teorías físicas que se pueden expresar como sistemas mecánicos: las teorías clásicas de campos.

Ejemplo: ecuación de onda

Supongamos que queremos analizar un sistema que obedece a la ecuación de ondas en una dimensión

$$c^2 u'' - \ddot{u} = 0$$

donde c es la velocidad de la onda. Podría tratarse por ejemplo de las ondas en una cuerda estirada entre dos soportes, o de las ondas de sonido en un medio material. Con esta ecuación podemos conocer el estado del sistema $u(x, t)$ a partir de su estado inicial $u(x, 0)$. Podríamos entonces imaginarnos que se trata de un sistema mecánico con coordenadas

$$u(x, t) = q_x(t)$$

Es decir que el índice de coordenada es ahora un índice continuo. Esta interpretación hace natural que nos preguntemos ¿podremos escribir una acción para la ecuación de

ondas? Deberíamos tener algo con la forma

$$S[q_x] = S[u] = \int_{t_0}^{t_1} dt L$$

Dado que la acción no puede tener dependencia en x , debería haber una integral en esta variable, algo así como

$$S[u] = \int_{t_0}^{t_1} dt \underbrace{\int dx \mathcal{L}}_L$$

Donde hemos escrito al lagrangiano en forma de la integral espacial de una *densidad lagrangiana* \mathcal{L} . Para intentar darle una forma concreta a esta magnitud, hagamos el siguiente análisis:

- Las ecuaciones se obtienen tomando derivadas del lagrangiano respecto de las coordenadas y velocidades generalizadas. Es decir que si la ecuación es lineal, la acción no puede tener potencias superiores a la cuadrática.
- Todos los términos de la acción deben tener las mismas unidades, es decir que si incluimos un término con \dot{u}^2 debe estar combiando con $c^2 u'^2$.
- La ecuación no tiene un términos sin derivadas, por lo que no debemos incluir en la acción términos que dependan sólo de u .
- Si pusiéramos términos cruzados $u'u$ o $u\dot{u}$ o $u'\dot{u}$ aparecerían terminos de primer orden en derivadas o con derivadas segundas.

Con todo esto, podemos proponer la siguiente forma para la acción

$$S[u] = \frac{1}{2} \int_{t_1}^{t_2} dt dx \underbrace{(\dot{u}^2 - c^2 u'^2)}_{\mathcal{L}}$$

Para comprobar que de su variación se obtiene la ecuación de onda, la evaluamos en $u + \varepsilon \delta u$ y desarrollamos en potencias de ε , obteniendo

$$S[u + \varepsilon \delta u] = S[u] + \varepsilon \int_{t_0}^{t_1} dt dx (\dot{u} \delta \dot{u} - c^2 u' \delta u')$$

con lo que podemos aislar el término lineal en la forma

$$\delta S = \int_{t_0}^{t_1} dt dx (\dot{u} \delta \dot{u} - c^2 u' \delta u')$$

Integrando por partes en ambos términos, obtenemos

$$\delta S = \int_{t_0}^{t_1} dt dx (c^2 u'' - \ddot{u}) \delta u + \int dx (\dot{u} \delta u)|_{t_0}^{t_1} + c^2 \int dt (u' \delta u)|_{x_0}^{x_1}$$

y usando el hecho de que las variaciones deben anularse en el instante inicial y final, es decir que $\delta u(x, t_0) = \delta u(x, t_1) = 0$ para cualquier valor de x , tenemos

$$\delta S = \int_{t_0}^{t_1} dt dx (c^2 u'' - \ddot{u}) \delta u + c^2 \int dt (u' \delta u)|_{x_0}^{x_1} = 0$$

Pero dado que δu es una función arbitraria, esto implica que para que δS se anule deben anularse los factores dentro de cada integrando. En otras palabras

$$c^2 u'' - \ddot{u} = 0$$

$$(u' \delta u)|_{x_0}^{x_1} = 0$$

Con lo que hemos tenido éxito en obtener la ecuación de onda a partir de nuestra acción, pero además obtuvimos una condición *de contorno* en los extremos de la cuerda en la cual se propaga la onda. Esta condición se puede resolver pidiendo que se cumpla

$$u'(x_0) = u'(x_1) = 0$$

lo que corresponde a imponer condiciones de contorno de Neumann. Otra opción es pedir que las variaciones δu cumplan la restricción

$$\delta u(x_0) = \delta u(x_1) = 0$$

lo que se resuelve en términos de dos constantes u_0 y u_1 en la forma

$$u(x_0) = u_0 \quad u(x_1) = u_1$$

con lo que vemos que no es otra cosa que imponer condiciones de contorno de Dirichlet. En otras palabras, la condición de acción estacionaria para la acción que hemos propuesto más arriba no solo resulta en la ecuación de onda, sino también sus posibles condiciones de contorno. Nótese que hubiéramos podido incluir términos de borde en nuestra acción original, que solo dependieran de los valores de u en x_0 y x_1 , y en tal caso hubiéramos obtenido condiciones de contorno diferentes.

Ejercicio: ecuación de onda en tres dimensiones

Pruebe que la ecuación de onda en tres dimensiones se puede obtener de la acción

$$S[u] = \frac{1}{2} \int_{\Omega} d^3x dt \left(\dot{u}^2 - c^2 (\vec{\nabla} u)^2 \right)$$

La integral espacial está calculada dentro de una región del espacio Ω ¿cuáles serán las condiciones de contorno resultantes de esta acción en el borde $\partial\Omega$ de dicha región espacial?

Ejemplo: acción para las ecuaciones de Maxwell

El tipo de acción que escribimos para la ecuación de onda se puede formular también para sistemas continuos más generales. Por ejemplo, para el electromagnetismo podemos proponer una acción de la forma

$$S[\Phi, \vec{A}] = \frac{1}{2} \int_{\Omega} d^3x dt \left(\vec{E}^2 - c \vec{B}^2 \right) = \frac{1}{2} \int d^3x dt \left(\left(-\vec{\nabla} \Phi + \frac{1}{c} \partial_t \vec{A} \right)^2 - \frac{1}{c} (\vec{\nabla} \times \vec{A})^2 \right)$$

Donde estamos integrando en una región tridimensional del espacio que hemos llamado Ω . Para encontrar el punto estacionario de esta funcional escribimos $\vec{A} + \varepsilon \delta \vec{A}$ y $\Phi + \varepsilon \delta \Phi$, con lo cual

$$\begin{aligned}\delta S &= \int d^3x dt \left(\vec{E} \cdot \delta \vec{E} - c \vec{B} \cdot \delta \vec{B} \right) \\ &= \int d^3x dt \left(\vec{E} \cdot \left(-\vec{\nabla} \delta \Phi + \frac{1}{c} \partial_t \delta \vec{A} \right) - \vec{B} \cdot (\vec{\nabla} \times \delta \vec{A}) \right)\end{aligned}$$

Integrando por partes todas las derivadas, y usando el hecho de que $\delta \vec{A}$ se anula en los instantes inicial y final, tenemos

$$\delta S = \int d^3x dt \left(\vec{\nabla} \cdot \vec{E} \delta \Phi + \delta \vec{A} \cdot \left(\frac{1}{c} \partial_t \vec{E} - \vec{\nabla} \times \vec{B} \right) \right) + \int_{\partial \Omega} dt d\vec{S} \cdot \left(\vec{E} \delta \Phi + \delta \vec{A} \times \vec{B} \right)$$

donde la notación $\partial \Omega$ significa el borde de la región Ω . De aquí se deducen las ecuaciones de Maxwell

$$\vec{\nabla} \cdot \vec{E} = 0 \quad \vec{\nabla} \times \vec{B} - \frac{1}{c} \partial_t \vec{E} = 0$$

junto con las condiciones de contorno

$$\vec{E} \cdot \vec{n} \delta \Phi|_{\partial \Omega} = 0 \quad (\vec{B} \times \vec{n}) \cdot \delta \vec{A}|_{\partial \Omega} = 0$$

donde \vec{n} es un vector unitario perpendicular al borde $\partial \Omega$. Éstas se pueden traducir en condiciones de conductor perfecto

$$\Phi|_{\partial \Omega} = \Phi_0 \quad \vec{A}|_{\partial \Omega} = \vec{A}_0$$

o de dieléctrico perfecto

$$\vec{E} \cdot \vec{n}|_{\partial \Omega} = 0 \quad (\vec{B} \times \vec{n})|_{\partial \Omega} = 0$$

En otras palabras, el principio de acción estacionaria nos permite considerar al electromagnetismo como un sistema mecánico con coordenadas $\{q_x\} = \{\Phi(x), A_a(x)\}$.

Ejercicio: densidades de carga y corriente

Pruebe que si quisiéramos incluir una densidad de carga ρ y una densidad de corriente \vec{j} en las ecuaciones de Maxwell, la acción correcta vendría dada por

$$S = \frac{1}{2} \int d^3x dt \left(\vec{E}^2 - c \vec{B}^2 + \Phi \rho - \frac{1}{c} \vec{A} \cdot \vec{j} \right)$$

Ejemplo: ecuación de Schrödinger libre

Otro ejemplo de un sistema continuo es el aquél cuyas ecuaciones de movimiento se obtienen a partir de la condición de punto estacionario de la acción

$$S = \int d^3x dt \left(\frac{\hbar^2}{2m} \vec{\nabla} \Psi \cdot \vec{\nabla} \Psi^* - \hbar i \partial_t \Psi \Psi^* \right)$$

donde Ψ es una coordenada generalizada compleja. La acción debe derivarse respecto de la parte real y la parte imaginaria de Ψ , o equivalentemente respecto de Ψ y de Ψ^* . Esto resulta en

$$\delta S = \int d^3x dt \left(\frac{\hbar^2}{2m} \left(\vec{\nabla} \Psi \cdot \vec{\nabla} \delta \Psi^* + \vec{\nabla} \delta \Psi \cdot \vec{\nabla} \Psi^* \right) - \hbar i (\delta \partial_t \Psi \Psi^* + \partial_t \Psi \delta \Psi^*) \right)$$

Integrando por partes y reordenando

$$\delta S = \int d^3x dt \left(\left(-\frac{\hbar^2}{2m} \nabla^2 \Psi - \hbar i \partial_t \Psi \right) \delta \Psi^* + \left(-\frac{\hbar^2}{2m} \nabla^2 \Psi^* + \hbar i \partial_t \Psi^* \right) \delta \Psi \right) + \int_{\partial\Omega} dt d\vec{S} \cdot \left(\delta \Psi \vec{\nabla} \Psi^* + \delta \Psi^* \vec{\nabla} \Psi \right)$$

De donde sale la ecuación de Schrödinger para una partícula libre

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi = \hbar i \partial_t \Psi$$

(y su conjugada compleja) junto con las condiciones de contorno

$$\delta \Psi \vec{\nabla} \Psi^* \cdot \vec{n}|_{\partial\Omega} = 0$$

(y su conjugada compleja). Esta condición se puede escribir como la que corresponde a una pared infinita de potencial

$$\Psi|_{\partial\Omega} = 0$$

En otras palabras la ecuación de Schrödinger, que describe la función de onda cuántica de una partícula, es en sí misma un sistema mecánico clásico. Esta observación, que puede en principio parecer un poco paradójica, da origen a la Teoría Cuántica de Campos, que se aplica extensivamente en física de partículas y en el área de materia condensada, para describir sistemas cuánticos locales.

Ejercicio: ecuación de Schrödinger con un potencial

Pruebe que la ecuación de Schrödinger en presencia de un potencial $V(\vec{r})$ se obtiene a partir de la acción

$$S = \int d^3x dt \left(\frac{\hbar^2}{2m} \vec{\nabla} \Psi \cdot \vec{\nabla} \Psi^* - \hbar i \partial_t \Psi \Psi^* + V(\vec{r}) \Psi \Psi^* \right)$$

Si estamos describiendo una partícula cargada tenemos que $V(\vec{r}) = e\Phi(\vec{r})$ siendo e la carga de la partícula y $\Phi(\vec{r})$ el potencial eléctrico ¿qué sucede si sumamos esta acción a la que escribimos más arriba para las ecuaciones de Maxwell?

Todos los ejemplos anteriores corresponden a *sistemas continuos*, es decir aquéllos donde las coordenadas son *campos* que toman un valor en cada punto del espacio. En el caso general, si tenemos un conjunto arbitrario de campos $q_n(\vec{r}, t)$, la acción se escribirá

$$S[q_n] = \int_{\Omega} d^3x dt \mathcal{L}(q_n(\vec{r}, t), \partial_t q_n(\vec{r}, t), \vec{\nabla} q_n(\vec{r}, t), \vec{r}, t)$$

en términos de una densidad lagrangiana $\mathcal{L}(q_n, \partial_t q_n, \vec{\nabla} q_n, \vec{r}, t)$, integrada en la región espacial Ω en la que está definido el sistema. Las ecuaciones de Lagrange resultantes del principio de acción estacionaria para este sistema toman la forma

$$\frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial_t q_n)} \right) + \frac{\partial}{\partial x_a} \left(\frac{\partial \mathcal{L}}{\partial (\partial_a q_n)} \right) - \frac{\partial \mathcal{L}}{\partial q_n} = 0$$

y deben ser suplementadas con las condiciones de contorno

$$\vec{n}_a \frac{\partial \mathcal{L}}{\partial (\partial_a q_n)} \delta q_n \Big|_{\partial \Omega} = 0$$

donde \vec{n} es un vector normal al borde de la región Ω .

Ejercicio: condiciones de borde arbitrarias

Pruebe que sumando a la acción un término evaluado en el borde

$$\Delta S = \int_{\partial \Omega} dt dS_a (\mathcal{K}_{ab}(q_n(\vec{r}, t), \vec{r}, t) \partial_b q_n(\vec{r}, t) + \mathcal{S}_a(q_n(\vec{r}, t), \vec{r}, t))$$

con $\mathcal{K}_{ab}(q_n, \vec{r}, t)$ y $\mathcal{S}_a(q_n, \vec{r}, t)$ funciones de los campos, las coordenadas y el tiempo, las condiciones de contorno se modifican. Escriba la forma resultante de las mismas.

13.4 Resumen

Hemos investigado el principio de acción estacionaria, que permite encontrar las ecuaciones de Lagrange como las que corresponden al punto estacionario de una funcional acción, que está definida como la integral temporal del lagrangiano.

Vimos que este principio hace evidente que la mecánica de Newton debe necesariamente estar contenida dentro de una teoría de ondas, como el *límite de óptica geométrica* en el que las trayectorias de las partículas corresponden a los rayos.

Usamos el principio de acción estacionaria para proponer generalizaciones de la mecánica más allá de la mecánica de partículas. Vimos que varias de esas generalizaciones conducen a situaciones físicamente irreales. Sin embargo, vimos encontramos también una de ellas, la generalización a sistemas continuos, que nos permite englobar dentro de los sistemas mecánicos una enorme variedad de teorías físicas.

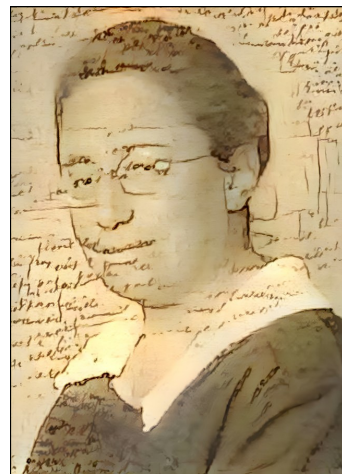
14. Simetrías

14.1 Objetivos

En esta clase definiremos la idea de *simetría* de un sistema mecánico y estudiaremos sus consecuencias.

En particular, estaremos interesados en formalizar la relación entre simetrías y cantidades conservadas que mencionamos en varias ocasiones en las clases previas. Para esto, formularemos y aplicaremos el *teorema de Noether*, que nos permitirá comprender las leyes de conservación de la naturaleza desde una óptica mucho más general.

Los resultados que presentaremos aquí se encuentran entre los más importantes de la física teórica, trascendiendo ampliamente el área de la Mecánica y aplicándose en contextos mucho más generales.



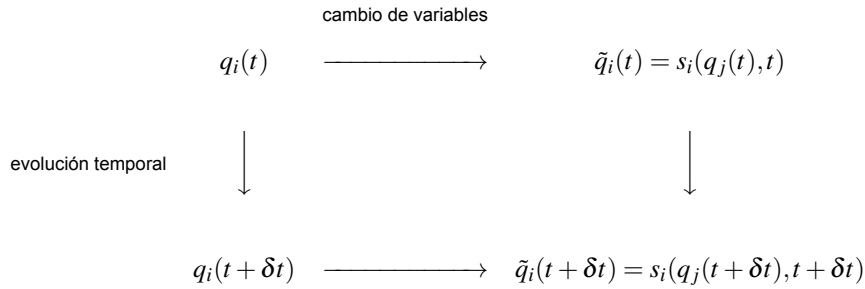
Emmy Noether

14.2 Simetrías

Motivemos la idea intuitiva de simetría con una situación hipotética sencilla. Un jugador de fútbol patea una pelota en dirección al arco. Si la pelota es esférica, esperamos que el resultado sea el mismo independientemente de si la hacemos girar o no unos pocos grados antes del puntapié. Si el resultado fuera diferente luego de tal rotación, diríamos que la pelota está ovalada, es decir que ha perdido su carácter simétrico.

El ejemplo pone de manifiesto que la idea de simetría está asociada a la insensibilidad de un sistema físico frente a alguna transformación: el sistema evoluciona exactamente del mismo modo, independientemente de si aplicamos o no la transformación antes del inicio del movimiento.

Esto se puede formalizar como sigue: una *simetría* de un sistema mecánico es un cambio de coordenadas $q_i(t) \rightarrow \tilde{q}_i(t) = s_i(q_j(t), t)$ que *conmuta* con la evolución temporal. Esta definición puede entenderse en el siguiente esquema



O sea que si hacemos el cambio de variables en el instante t , y luego dejamos evolucionar el sistema, obtenemos después de un cierto tiempo δt el mismo valor para las nuevas variables que hubiéramos obtenido de transformar las variables originales directamente en ese instante. Podríamos decir que el sistema no reacciona al cambio de variables, resultando en la misma configuración tanto si lo hemos realizado antes de la evolución, como si lo hacemos después.

En términos de las ecuaciones de movimiento, la definición implica que las coordenadas transformadas \tilde{q}_i deben evolucionar con las mismas ecuaciones que las coordenadas originales q_i . Sabemos de la clase previa que esto sucede cuando el lagrangiano para las \tilde{q}_i difiere del lagrangiano para las q_i en una derivada total

$$L(\dot{\tilde{q}}_i, \tilde{q}_i, t) = L(\dot{q}_i, q_i, t) + \frac{d}{dt}F(q_i, t)$$

Siendo $F(q_i, t)$ alguna función que dependerá del cuál sea el cambio de variables en cuestión. En términos de la acción, se debe cumplir que

$$S[\tilde{q}_i] = S[q_i] + F(q_i(t), t)|_{t_1}^{t_2}$$

Las simetrías pueden ser *discretas*, es decir definidas a través de un cambio de variables del tipo

$$q_i \rightarrow \tilde{q}_i = s_i(q_i, t)$$

donde $s_i(q_i, t)$ es una función determinada. Las simetrías también pueden ser *continuas*, si el cambio de variables tiene la forma

$$q_i \rightarrow \tilde{q}_i = s_i(q_i, t; \varepsilon)$$

donde ahora la función $s_i(q_i, t; \varepsilon)$ depende de un parámetro real ε . Si para un valor dado del parámetro, que sin pérdida de generalidad podemos elegir como $\varepsilon = 0$, la transformación es la

identidad $s_i(q_i, t; 0) = q_i$ entonces decimos que la transformación dada por $s_i(q_i, t; \varepsilon)$ es *conexa con la identidad*. Esto nos permite desarrollar en potencias del parámetro ε

$$q_i \rightarrow \tilde{q}_i = q_i + \varepsilon g_i(q_i, t) + \mathcal{O}(\varepsilon)^2$$

En esta expresión, las funciones $g_i(q_i, t)$ que determinan el primer orden del desarrollo se llaman los *generadores* de la simetría.

14.2.1 Teorema de Noether

En el caso de las simetrías continuas, el hecho de que la acción es invariante a menos de una derivada total tiene consecuencias muy importantes y útiles. En efecto, si escribimos

$$S[\tilde{q}_i] = S[q_i + \varepsilon g_i] = \int_{t_0}^{t_1} L(\dot{q}_i + \varepsilon \dot{g}_i(q_i, t), q_i + \varepsilon g_i(q_i, t), t)$$

En esta expresión podemos expandir el lagrangiano a primer orden en el parámetro infinitesimal ε , para obtener

$$S[\tilde{q}_i] = \int_{t_0}^{t_1} L(\dot{q}_i, q_i, t) + \varepsilon \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial \dot{q}_i} \dot{g}_i(q_i, t) + \frac{\partial L}{\partial q_i} g_i(q_i, t) \right)$$

El primer término no es otra cosa que $S[q_i]$. En cuanto al segundo, podemos integrarlo por partes para eliminar la derivada temporal que actúa sobre los generadores. Esto resulta en la expresión

$$S[\tilde{q}_i] = S[q_i] + \varepsilon \int_{t_1}^{t_2} \left(-\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial q_i} \right) g_i(q_i, t) + \varepsilon \left(\frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) \right) \Big|_{t_1}^{t_2}$$

Observemos aquí que la expresión en paréntesis bajo el signo integral se anula cuando se satisfacen las ecuaciones de movimiento. Por lo tanto, evaluada en las soluciones la expresión toma la forma

$$S[\tilde{q}_i] = S[q_i] + \varepsilon \left(\frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) \right) \Big|_{t_1}^{t_2}$$

Comparando con la expresión que habíamos escrito más arriba para la acción transformada, podemos identificar los términos de borde que antes habíamos expresado en términos de nuestra función $F(q_i, t)$. Tenemos que

$$\varepsilon \left(\frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) \right) \Big|_{t_1}^{t_2} = F(q_i, t) \Big|_{t_1}^{t_2}$$

lo que se puede reordenar como

$$\left(\varepsilon \frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) - F(q_i, t) \right) \Big|_{t_1} = \left(\varepsilon \frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) - F(q_i, t) \right) \Big|_{t_2}$$

Es decir que la magnitud entre paréntesis se conserva durante la evolución temporal. Esta magnitud se llama *carga conservada asociada a la simetría* $g_i(q_i, t)$, y se denota como

$$C = \varepsilon \frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) - F(q_i, t)$$

La construcción que hemos realizado se denomina *teorema de Noether* y está entre los resultados más trascendentes del curso de Mecánica Analítica. Lo que dice este teorema es que cada vez que haya una simetría continua en un sistema físico, existirá una magnitud que se conserva a lo largo de toda su evolución temporal.

Vamos ahora a enumerar algunos ejemplos de aplicación del teorema de Noether, que nos llevarán a revisar resultados de los cursos básicos de física bajo esta nueva óptica.

Dado un sistema físico definido a partir de su lagrangiano, la receta general consiste en identificar sus simetrías continuas y escribir su forma infinitesimal. De este modo podemos aislar sus generadores $g_i(q_i, t)$ y observar qué forma toma el término de borde $F(q_i, t)$. Con estos dos elementos, estamos en condiciones de escribir la cantidad conservada C usando la fórmula de más arriba.

Ejemplo: invarianza traslacional

Supongamos una partícula no relativista descrita por las coordenadas cartesianas x_a que se mueve de acuerdo a un lagrangiano de la forma

$$L = \frac{1}{2} m \dot{x}_a^2 - V(x_a)$$

Vamos a suponer que el lagrangiano es invariante bajo una transformación de traslación

$$x_a \rightarrow \tilde{x}_a = x_a + \varepsilon \tilde{n}_a$$

donde \tilde{n} es un vector unitario que nos indica la dirección de la traslación, y ε es un parámetro que toma valores reales arbitrarios.

Esta invarianza es inmediatamente cierta para el término cinético, ya que la velocidad no cambia al sumar a x_a una constante $\dot{\tilde{x}}_a = \dot{x}_a$, y para el potencial implica que $V(x_a + \varepsilon \tilde{n}_a) = V(x_a)$. Para entender el significado de esta condición, escribamos

$$V(x_a + \varepsilon \tilde{n}_a) - V(x_a) = \varepsilon \tilde{n}_a \frac{\partial V}{\partial x_a} = -\varepsilon \tilde{n} \cdot \vec{F} = 0$$

En otras palabras, que el potencial sea invariante frente a traslaciones en la dirección \tilde{n} implica que la fuerza en esa dirección se anula.

Si ahora reemplazamos el cambio de coordenadas directamente en la acción, vemos que se cumple

$$S[x_a] = S[\tilde{x}_a]$$

Por lo tanto tenemos el caso particular de una simetría para la cual el término de borde se anula $F(x_a, t) = 0$.

Para obtener explícitamente el generador, escribimos la forma infinitesimal de la transformación según

$$\tilde{x}_a = x_a + \varepsilon g_a(x_b, t) \quad \text{con} \quad g_a(x_b, t) = \tilde{n}_a$$

Con esto se puede escribir inmediatamente la carga conservada de acuerdo a la fórmula que obtuvimos más arriba, en la forma

$$C = \varepsilon \frac{\partial L}{\partial \dot{x}_a} g_a(x_b, t) - F(x_b, t) = \varepsilon \check{n}_a m \dot{x}_a = \varepsilon \check{n} \cdot \vec{p}$$

de donde vemos que hemos recuperado la ley de conservación del momento lineal: cuando la fuerza en una dirección cartesiana se anula, el momento lineal en esa dirección se conserva.

Nótese que la reformulación que hemos obtenido tiene un significado mucho más abarcativo: relaciona la conservación del impulso lineal en una dada dirección del espacio con la invarianza de traslaciones en esa misma dirección. Se deduce inmediatamente que cada vez que exista invarianza de traslaciones, independientemente del sistema físico particular que estamos analizando, existirá un análogo al impulso lineal para ese sistema, que se conservará a medida que transcurre el movimiento.

Ejemplo: coordenadas cíclicas

El resultado anterior se puede generalizar de un modo muy sencillo para dar cuenta de la conservación del momento generalizado asociado a una coordenada generalizada que es cíclica. En efecto, una traslación en una coordenada generalizada q_i se puede escribir como

$$q_i \rightarrow \tilde{q}_i = q_i + \varepsilon$$

Considerar ε como una variable infinitesimal nos permite identificar el generador, de acuerdo a

$$\tilde{q}_i = q_i + \varepsilon g_i(q_i, t) \quad \text{con} \quad g_i(q_i, t) = 1$$

Ahora bien, dado que la coordenada es cíclica, entonces el lagrangiano es invariante bajo esta transformación. Esto resulta en que la acción también es invariante, por lo que la función $F(q_i, t)$ será nula. Con esto, la carga conservada asociada a esta invarianza será

$$C = \varepsilon \frac{\partial L}{\partial \dot{q}_i} g_i(q_i, t) = \varepsilon p_i$$

En otras palabras: cuando una coordenada generalizada es cíclica, el momento generalizado correspondiente se conserva.

Conocíamos este resultado de las clases anteriores. El punto importante es que aquí lo hemos reobtenido no mediante examen de las ecuaciones de Lagrange, sino a partir del teorema de Noether. Es decir, ahora sabemos que la presencia de una coordenada cíclica implica una simetría frente a traslaciones en esa coordenada, lo que resulta en una carga conservada.

Ejemplo: invarianza rotacional

Según vimos oportunamente, una rotación en ángulo θ alrededor del eje determinado por un versor \hat{n} se escribe usando la fórmula

$$\vec{r} \rightarrow \vec{r}' = (\vec{r} \cdot \hat{n}) \hat{n} + \sin \theta \hat{n} \times \vec{r} + \cos \theta (\vec{r} - (\vec{r} \cdot \hat{n}) \hat{n})$$

Lo cual, aplicado a una partícula descrita en coordenadas cartesianas por el vector \vec{r} , implica para su velocidad

$$\dot{\vec{r}}' = (\dot{\vec{r}} \cdot \hat{n}) \hat{n} + \sin \theta \hat{n} \times \dot{\vec{r}} + \cos \theta (\dot{\vec{r}} - (\dot{\vec{r}} \cdot \hat{n}) \hat{n})$$

Escribiendo el lagrangiano de dicha partícula en la forma

$$L = \frac{1}{2} m \dot{\vec{r}}^2 - V(\vec{r})$$

el término de energía cinética queda inmediatamente invariante $m \dot{\vec{r}}^2 / 2 = m \dot{\vec{r}}'^2 / 2$. Por otro lado, si se cumple que el potencial también queda invariante $V(\vec{r}') = V(\vec{r})$ podemos decir que las rotaciones en torno al eje \hat{n} son una simetría del sistema. En particular esto resulta en que tenemos

$$S[\vec{r}'] = S[\vec{r}]$$

por lo que nuevamente se trata de una simetría con un término de borde nulo $F(\vec{x}, t) = 0$.

Para encontrar el generador, desarrollamos la fórmula que escribimos más arriba en potencias del ángulo $\theta = 0 + \varepsilon$, obteniendo a primer orden

$$\tilde{x}_a = x_a + \varepsilon \varepsilon_{abc} n_b x_c = x_a + \varepsilon g_a(x)$$

En términos del generador $g_a(x) = \varepsilon_{abc} n_b x_c$. Con esto, podemos escribir la cantidad conservada, dada por

$$C_{\hat{n}} = \varepsilon \frac{\partial L}{\partial \dot{x}_a} g_a(x) = \varepsilon m \dot{x}_a \varepsilon_{abc} n_b x_c = \varepsilon n_b \underbrace{\varepsilon_{abc} x_c p_a}_{(\vec{x} \times \vec{p})_b = L_b} = \varepsilon \hat{n} \cdot \vec{\ell}$$

Vemos que corresponde a la componente del momento angular a lo largo del eje de rotación.

Para entender lo que esto significa, reescribamos la condición de invarianza del potencial para una rotación infinitesimal

$$V(\vec{r}') - V(\vec{r}) = \varepsilon n_b \underbrace{\varepsilon_{abc} x_c \frac{\partial V}{\partial x_a}}_{\tau_b} = \varepsilon \hat{n} \cdot \vec{\tau} = 0$$

donde $\vec{\tau}$ es el torque respecto del origen. Es decir que un potencial invariante frente a rotaciones alrededor de \hat{n} implica que el torque en esa dirección se anula.

Esto nos dice que hemos reobtenido la ley de conservación del momento angular. De nuevo, el significado es ahora mucho más profundo: la conservación del momento angular se relaciona con la invarianza de rotaciones, y habrá un momento angular conservado en cualquier sistema que tenga tal invarianza, independientemente de la forma particular de su lagrangiano.

Ejemplo: invarianza frente a traslaciones temporales

La transformación de traslación temporal se define como

$$t \rightarrow \tilde{t} = t + \varepsilon$$

donde ε puede tomar cualquier valor real. Nótese que no se trata de una transformación de las coordenadas generalizadas del sistema, sino de la variable independiente t , que en la acción cumple el rol de la variable de integración.

Por lo tanto, no estamos en las condiciones supuestas más arriba para la aplicación del teorema de Noether. En efecto, habíamos definido una simetría como un cambio en las coordenadas generalizadas que puede eventualmente depender del tiempo, pero no habíamos incluido la posibilidad de un cambio en el tiempo.

Para resolver esto, podemos promover t a una coordenada generalizada si hacemos un cambio de variables $t(\tau)$ en términos de un nuevo parámetro τ que caracteriza la evolución temporal y que cumplirá el rol de variable de integración. La acción se reescribe como

$$S[q_i, t] = \int_{\tau_1}^{\tau_2} d\tau t'(\tau) L \left(\underbrace{\frac{1}{t'} q'_i(\tau), q_i(\tau), t(\tau)}_{\hat{L}(q'_i(\tau), q_i(\tau), t'(\tau), t(\tau))} \right) = \int_{\tau_1}^{\tau_2} d\tau \hat{L}(q'_i(\tau), q_i(\tau), t'(\tau), t(\tau))$$

Con lo que obtenemos un nuevo lagrangiano \hat{L} , donde ahora t entra en pie de igualdad con las coordenadas generalizadas q_i . Es fácil demostrar que a partir de las ecuaciones de Lagrange de este lagrangiano

$$\frac{d}{d\tau} \frac{\partial \hat{L}}{\partial q'_i} - \frac{\partial \hat{L}}{\partial q_i} = 0 \qquad \frac{d}{d\tau} \frac{\partial \hat{L}}{\partial t'} - \frac{\partial \hat{L}}{\partial t} = 0$$

se recuperan las ecuaciones originales en términos de t .

Con \hat{L} podemos construir la carga conservada asociada a la traslación temporal aplicando el procedimiento de Noether. Primero identificamos el generador dada la regla de transformación

$$\tilde{t} = t + \varepsilon$$

de donde se ve que $g(q_i, t) = 1$. Por otro lado el lagrangiano \hat{L} será invariante siempre que el lagrangiano original L sea independiente del tiempo, con lo que obtenemos una vez más $F(q_i, t) = 0$. Con esto tenemos para la carga conservada la expresión

$$C = \varepsilon \frac{\partial \hat{L}}{\partial t'} = \varepsilon \left(L + t' \frac{\partial L}{\partial t'} \right) = \varepsilon \left(L - \frac{1}{t'} \frac{\partial L}{\partial \dot{q}_i} q'_i \right)$$

La última expresión se puede reescribir como

$$C = \varepsilon \left(L - \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) = \varepsilon (L - p_i \dot{q}_i) = -\varepsilon E$$

Con lo que la carga conservada es la energía mecánica del sistema, según la definición general que habíamos dado en la sección 2.2.4

De nuevo, un teorema de conservación que habíamos encontrado simplemente explorando las ecuaciones de movimiento, ahora aparece claramente ligado a una simetría del sistema mecánico.

Nota:

Según la visión moderna de la naturaleza, si conocemos las coordenadas generalizadas de un sistema físico y las simetrías que las transforman, estaremos en condiciones de construir un lagrangiano que describa completamente la dinámica, bajo el único requerimiento de que sea invariante frente a las simetrías.

En este sentido, la descripción más precisa que tenemos hoy en día de la estructura del universo dice que debe ser invariante localmente frente a rotaciones, traslaciones y transformaciones de Lorentz. Estas simetrías constituyen el grupo conocido como $SO(3,1)$ o *grupo de Poincaré*.

Por otro lado, los experimentos con aceleradores nos permiten elucidar que la física a escala nuclear es invariante frente al grupo $SU(3) \times SU(2) \times U(1)$.

Combinando esa información se construye el lagrangiano para el *Modelo Standard de las Interacciones Fundamentales* que permite hacer predicciones muy precisas sobre la física fundamental.

14.3 Resumen

En esta clase formulamos la idea de simetría de un sistema mecánico y demostramos el teorema de Noether, que dice que cada vez que existe una simetría continua hay una magnitud o *carga* que se conserva a lo largo de todo el movimiento.

Este teorema es probablemente el resultado más trascendente que se aprende durante la Licenciatura. Se generaliza a sistemas continuos dando lugar a la idea de *corrientes conservadas*. Es un teorema que sigue valiendo a nivel cuántico, y que yace en la base de la formulación de todas las teorías físicas modernas. Se debe a la matemática alemana Emmy Noether, de quien Albert Einstein dijo

«La señorita Noether fue el genio matemático creativo más importante que haya existido desde que comenzó la educación superior para las mujeres.»

A pesar de esto, por alguna razón Emmy Noether es injustamente olvidada con demasiada frecuencia en los recuentos del aporte de las mujeres al conocimiento científico.

IV Formulación hamiltoniana

Clase 15 Reversibilidad 215

- 15.1 Objetivos
- 15.2 Sistemas con configuraciones discretas
- 15.3 Sistemas con configuraciones continuas
- 15.4 Resumen

Clase 16 Paréntesis de Poisson 235

- 16.1 Objetivos
- 16.2 Paréntesis de Poisson
- 16.3 Resumen

Clase 17 Transformaciones canónicas 245

- 17.1 Objetivos
- 17.2 Transformaciones canónicas
- 17.3 Resumen

Clase 18 Ecuación de Hamilton-Jacobi 257

- 18.1 Objetivos
- 18.2 La ecuación de Hamilton Jacobi
- 18.3 Separación de variables
- 18.4 Resumen

Clase 19 Vínculos hamiltonianos 265

- 19.1 Objetivos
- 19.2 Vínculos hamiltonianos
- 19.3 Resumen

15. Reversibilidad

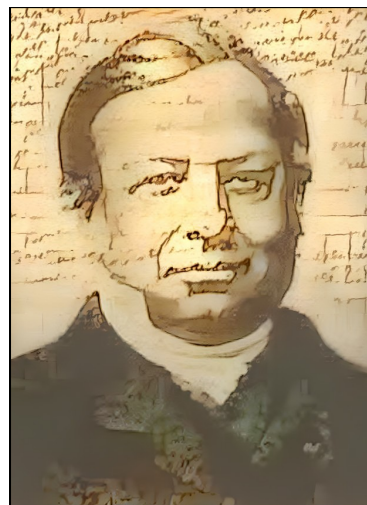
15.1 Objetivos

Comenzamos este curso analizando un sistema de N partículas en coordenadas generalizadas. Vimos que cuando las interacciones se restringen a fuerzas que se pueden derivar de un potencial, eventualmente dependiente del tiempo y de las velocidades, el movimiento del sistema se describe con las ecuaciones de Lagrange.

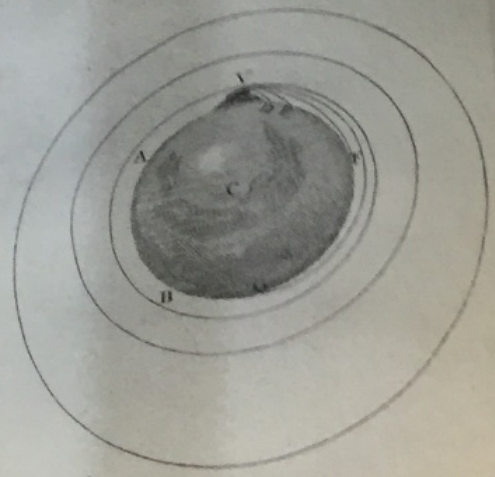
Vimos más adelante que las coordenadas generalizadas nos dejan introducir fácilmente vínculos entre las partículas, permitiéndonos entonces describir sistemas más generales, tales como los cuerpos rígidos.

Al estudiar las ecuaciones de Lagrange, encontramos que se pueden obtener a partir del principio de acción estacionaria. Pudimos usar este principio para generalizar los sistemas a los que podemos aplicar las leyes de la mecánica. Sin embargo, algunas de estas generalizaciones dieron lugar a sistemas físicamente poco realistas (sistemas en derivadas mayores, sistemas no locales), mientras que otras funcionaron correctamente para describir sistemas reales (sistemas continuos).

Estamos entonces en el punto en el que resulta natural preguntarnos ¿cuál es la característica definitoria de un sistema mecánico, que nos permitirá formular modelos completamente generales capaces de describir cualquier sistema físico? En esta clase nos concentraremos en responderla, analizando primero los sistemas más sencillos posibles, que son aquéllos cuya evolución es discreta, y luego avanzando hacia sistemas más realistas con evolución continua.



Joseph Liouville



15.2 Sistemas con configuraciones discretas

Para preguntarnos acerca de las características generales de las leyes de evolución, comenzaremos en esta sección analizando los sistemas más simples posibles, es decir aquéllos que tienen un conjunto contable de configuraciones discretas.

15.2.1 Leyes de evolución

Un sistema físico cualquiera tiene un conjunto de configuraciones posibles que podemos observar. Supongamos que ponemos una etiqueta $\{a, b, c, \dots\}$ a cada una de ellas, podemos definir su *espacio de configuración* \mathcal{C} como el conjunto de todas las etiquetas posibles

$a \ b \ c \ \dots$

Queremos hablar de la evolución temporal, porque después de todo de eso se trata la Mecánica. Es decir, queremos ver cómo cambia el sistema de una configuración a otra a medida que transcurre el *tiempo*.

Sin pretender plantearnos una pregunta filosóficamente tan complicada como *¿qué es el tiempo?*, lo mínimo que necesitamos es saber cómo medirlo. Para esto requerimos un *reloj*, que es un sistema físico cuyas configuraciones están ordenadas de alguna manera. Por ejemplo, en un reloj digital las diferentes configuraciones son los valores numéricos que se ven en la pantalla, y su orden está dado por el orden natural de esos números.

$15:01 \ 15:02 \ 15:03 \ \dots$

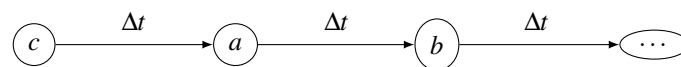
En un reloj analógico, el orden estará dado por las sucesivas posiciones de las agujas. Lo importante para esta discusión es que cualquier sistema cuyas configuraciones estén ordenadas puede cumplir el rol de un reloj.

Estudiar entonces la evolución temporal de nuestro sistema, es establecer la correlación entre las configuraciones del sistema y las configuraciones del reloj. Podemos resumirla en un cuadro

Configuración del reloj	Configuración del sistema
15:01	c
15:02	a
\vdots	\vdots

Este cuadro puede leerse como “cuando el reloj marca 15:01 el sistema está en la configuración c , cuando el reloj marca 15:02 el sistema está en la configuración a , ...”.

Sin embargo, es fácil notar que este cuadro contiene información redundante: la columna de la izquierda contiene las configuraciones del reloj ordenadas de acuerdo a su orden natural. Es decir que no necesitamos escribirla, basta con saber cuál es el orden de las configuraciones del sistema para poder recuperar la correlación. Es decir que podemos representar la evolución temporal de la siguiente manera



Donde cada globo representa la configuración del sistema cuando el reloj está en cada una de sus configuraciones ordenadas, y la flecha resume la evolución de una configuración a la siguiente. El símbolo Δt nos dice que durante esa evolución el reloj “hizo un tic” avanzando a su configuración vecina.

Esta evolución temporal, que para cada configuración del reloj nos da la configuración del sistema, se denomina *trayectoria* y se puede expresar formalmente como una función

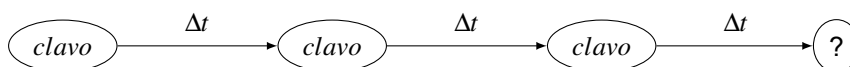
$$\text{Configuración del sistema} = f(\text{Configuración del reloj})$$

Con estas definiciones, estamos en condiciones de ponernos a discutir los elementos básicos de un sistema mecánico.

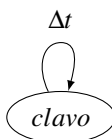
Podemos comenzar preguntándonos ¿cuál es el sistema mecánico más sencillo que podemos imaginar? La respuesta evidente es un sistema que sólo puede estar en una configuración. Un buen ejemplo de eso es un clavo en una pared. Si etiquetamos las configuraciones posibles del clavo con un nombre para cada una, tendremos solamente la configuración “clavo” definida como:

- *clavo* = El clavo está clavado en la pared

Supongamos que queremos estudiar la correlación entre las configuraciones de este sistema y las del reloj. ¿Cómo será la evolución resultante? Es decir ¿qué trayectorias se pueden presentar? bueno, eso tiene una respuesta muy sencilla:



Es fácil determinar qué encontraremos en el lugar de “?” en la siguiente observación, nada menos que la configuración *clavo*. Por lo tanto, en este sistema sólo hay una *regla de evolución* posible, que podemos resumir con el gráfico

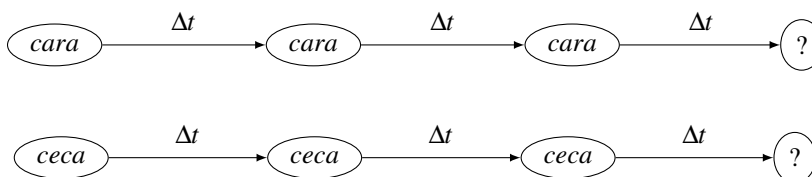


Cada vez que el reloj hace un tic, el sistema pasa de la configuración *clavo* a la configuración *clavo*. No parece ser un sistema muy divertido, si bien podemos identificar en él los elementos fundamentales para generalizar la discusión.

Intentemos definir ahora un sistema un poco más variado, que tenga al menos dos configuraciones posibles. Por ejemplo una moneda sobre una mesa, para la cual podemos definir las configuraciones

1. *cara* = La moneda tiene la cara hacia arriba
2. *ceca* = La moneda tiene la cara hacia abajo

Una regla evolución posible para este sistema es la más aburrida: que a medida que recorremos las configuraciones del reloj, la moneda se quede como está. En ese caso, las trayectorias serán



Entonces la regla de evolución será: si la moneda muestra la cara, en el siguiente instante mostrará la cara, si no la muestra, en el siguiente instante tampoco lo hará. Esto se puede simbolizar según

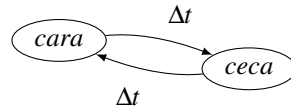
Regla 1:



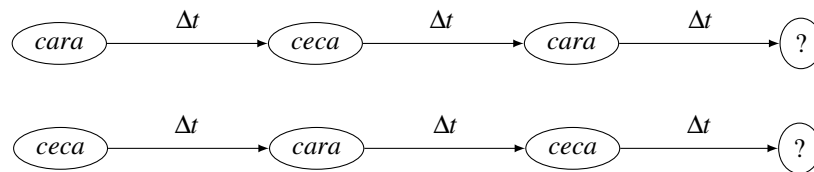
Lo que de hecho sigue siendo una regla de evolución extremadamente sencilla.

Sin embargo, con este sistema que tiene dos configuraciones, podemos imaginar reglas de evolución más complicadas. Por ejemplo, la que viene dada por el siguiente gráfico

Regla 2:



la cual implicaría trayectorias de este tipo

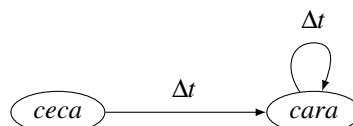


En otras palabras, la regla de evolución dice que: si la moneda muestra cara, pasa a ocultarla, y si la oculta, pasa a mostrarla. Sigue siendo una regla simple, pero ya empezamos a ver un poco más de estructura.

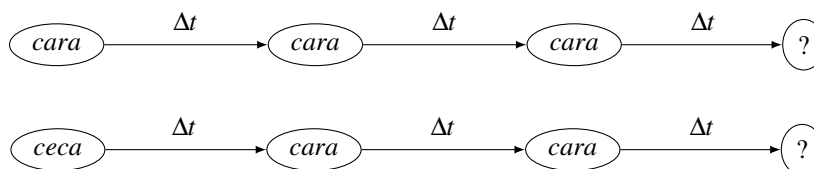
15.2.2 Reversibilidad y disipación

Hay otra regla de evolución posible para el sistema de la moneda sobre la mesa, a saber

Regla 3:

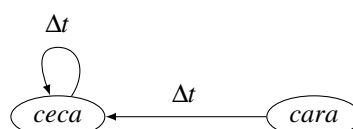


O sea, si la moneda muestra la cara sigue mostrando la cara, si no la muestra pasa a mostrarla. Esto resulta en las siguientes trayectorias posibles

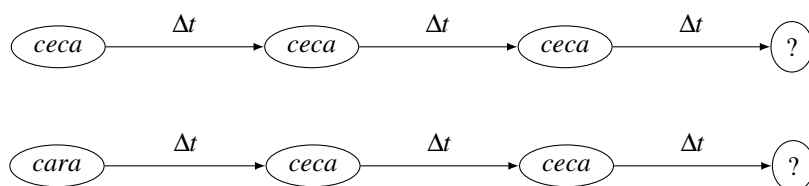


Y hay por supuesto una posibilidad complementaria, a saber

Regla 4:



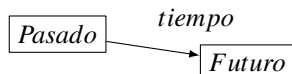
que resulta en las trayectorias



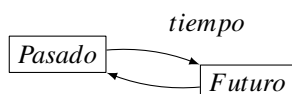
Sin embargo, estas dos últimas reglas de evolución tienen una característica fundamental que las diferencia de las dos anteriores: si miramos la configuración de la moneda en un dado instante, no podemos decir en qué configuración estaba en el instante anterior.

Por ejemplo, supongamos que nos concentramos en el último caso, la regla 4 dice que si la moneda oculta la cara, queda como está, pero si la muestra, entonces en el siguiente instante la ocultará. Pero entonces si observamos la moneda en la configuración *ceca*, no podemos afirmar si en el instante anterior estaba en la configuración *ceca* o en la configuración *cara*.

En otras palabras, usando estas dos últimas reglas de evolución podemos predecir, pero no podemos *retrodecir*, no es posible conocer la configuración anterior del sistema a partir del estado presente. Es decir que *se pierde información*, aparece una *flecha del tiempo* que nos dice hacia donde está el futuro y hacia dónde está el pasado.



Este no es el tipo de evolución temporal que presenta la naturaleza a nivel microscópico. En todos los sistemas que hemos estudiado la información se conserva, no hay una flecha del tiempo, y es posible tanto predecir el estado futuro a partir del presente, como retrodecir la configuración pasada que trajo al sistema hasta el presente.



En otras palabras, de las posibles reglas de evolución que hemos propuesto para la moneda, sólo las reglas 1 y 2 podrían corresponder a un sistema mecánico microscópico. Las reglas 3 y 4 son los ejemplos más sencillos del fenómeno de *disipación*, en donde la información se pierde en la evolución temporal. Este fenómeno sólo se observa en la naturaleza a nivel macroscópico.

Una cosa interesante para notar de las cuatro reglas arriba estudiadas, es que podrían formularse en términos matemáticos como sigue. Definamos la variable $q = \pm 1$ con $q = 1$ cuando la moneda está en la configuración *cara* y $q = -1$ cuando la moneda está en la configuración *ceca*. En esta notación t denota el estado del reloj, y denotaremos el estado siguiente como $t + \Delta t$. En ese caso

las reglas de evolución toman la forma:

Regla 1: $q(t+\Delta t) = q(t)$

Regla 2: $q(t+\Delta t) = -q(t)$

Regla 3: $q(t+\Delta t) = q(t)^2$

Regla 4: $q(t+\Delta t) = -q(t)^2$

En cada caso, el valor de la variable en $t + \Delta t$ es una función del valor de la variable en t .

$$q(t+\Delta t) = f(q(t))$$

En los casos 3 y 4 donde la evolución era disipativa, la función no es invertible. Estas últimas observaciones implican que, si bien no es necesario, lo que hemos discutido en esta sección y lo que veremos en las que siguen puede codificarse en fórmulas además de en diagramas.

Ejercicio:

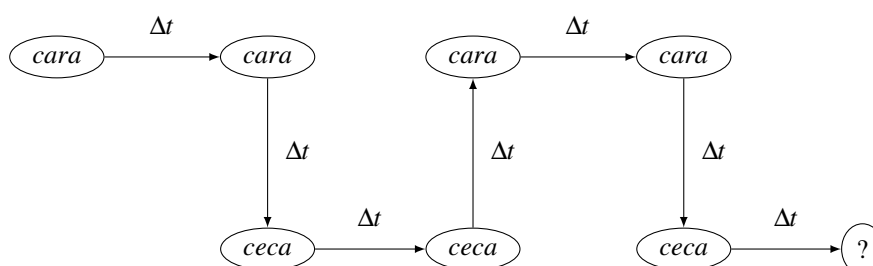
Describa el sistema físico consistente en un dado cúbico. Identifique las configuraciones posibles y enumere las posibles reglas de evolución ¿puede contar cuántas reglas diferentes hay?

De entre todas las posibles reglas de evolución, imagine dos que sean retrodecibles, es decir que permitan recorrer hacia atrás la evolución del dado para determinar el estado inicial. Dibuje los correspondientes diagramas para las trayectorias posibles y para la regla de evolución. Encuentre también dos que sean disipativos, o sea que no permitan recuperar el estado inicial, y dibuje también los mencionados diagramas.

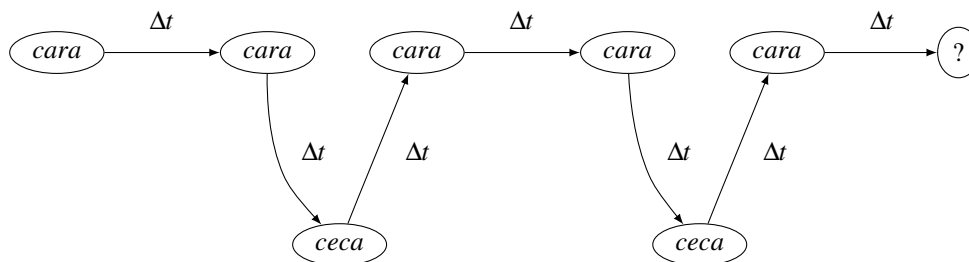
15.2.3 Configuraciones y estados

En este punto, uno podría preguntarse ¿no hay acaso mayor riqueza en el sistema que estamos describiendo? No parece evidente que todas las reglas de evolución posibles se agoten en las cuatro que acabamos de enumerar, cuando con casi ningún esfuerzo podemos imaginar trayectorias más elaboradas.

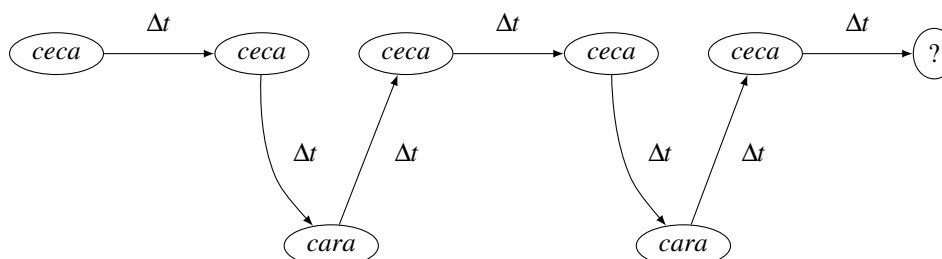
Por ejemplo, uno podría imaginar la siguiente trayectoria



donde la regla sería: dos caras y luego dos cecas. También podríamos tener una trayectoria como la que sigue



donde la regla es: dos caras y luego una ceca. O bien la que se obtiene a partir de la anterior intercambiando las configuraciones



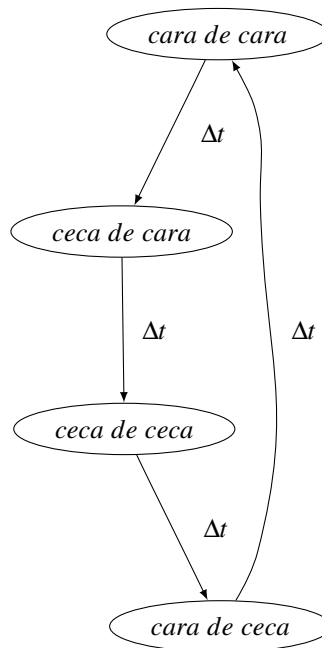
Es decir: una cara y luego dos cecas. Hay varias otras reglas similares que podemos inventar. ¿Qué pasa con estas trayectorias? ¿Existen acaso reglas de evolución capaces de originarlas?

La cosa se aclara cuando intentamos escribir la regla de evolución más claramente. Tomemos por ejemplo la primera, la regla de evolución “dos caras y luego dos cecas” puede escribirse como: si la moneda muestra la cara, y en el instante anterior no lo hacía, entonces en el instante siguiente seguirá mostrando la cara; en cambio si la moneda muestra la cara y en el instante anterior también lo hacía, entonces en el instante siguiente no lo hará; y lo mismo pasa con la ceca. Es decir que lo que va a pasar en el siguiente instante depende de lo que pasó en el instante previo y en el anterior.

Esta observación nos lleva a concluir que para poder escribir un gráfico que describa esta regla de evolución, tenemos que definir los *estados* del sistema de manera de contener toda la información relevante

1. *cara de cara* = la moneda muestra cara y en el instante anterior también lo hacía
2. *cara de ceca* = la moneda muestra cara y en el instante anterior no lo hacía
3. *ceca de cara* = la moneda no muestra cara y en el instante anterior sí lo hacía
4. *ceca de ceca* = la no moneda muestra cara y en el instante anterior tampoco lo hacía

En ese caso la regla de evolución se puede representar en el siguiente gráfico

Regla 5:

Con esta regla de evolución generamos la trayectoria de dos en dos antes mencionada.

Si quisieramos escribirla en términos de una ecuación, tendríamos que definir una variable para el estado, que contenga la configuración presente y la anterior $\eta(t) = (q(t - \Delta t), q(t))$, donde $t - \Delta t$ representa la configuración del reloj que precede a la denotada por t . Con lo que la ecuación quedaría

$$\text{Regla 5: } \eta(t + \Delta t) = f(\eta(t)) \quad \text{siendo } f(q_1, q_2) = (q_2, -q_1)$$

Esta función es invertible, lo que coincide con la propiedad evidente de la trayectoria que escribimos más arriba de poder recuperar cualquier estado pasado a partir del estado presente. Es decir que esta regla no presenta disipación.

Ejercicio:

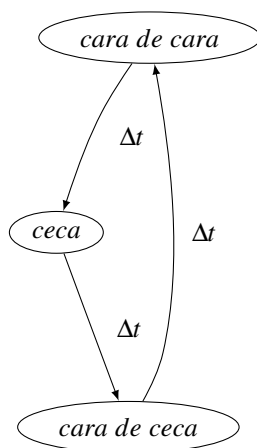
Imagine reglas no disipativas a partir de los estados que hemos definido para escribir la Regla 5 ¿cuántas reglas no disipativas diferentes se pueden escribir? En otras palabras ¿cuántas funciones invertibles diferentes se pueden escribir de un conjunto de cuatro elementos en sí mismo?

Para lograr una descripción similar de la trayectoria “dos caras y luego una ceca” se requiere en cambio definir los estados de la siguiente manera

1. *cara de cara* = la moneda muestra cara y en el instante anterior también lo hacía
2. *cara de ceca* = la moneda muestra cara y en el instante anterior no lo hacía
3. *ceca* = la moneda no muestra cara

Y entonces la regla de evolución se puede representar en el siguiente gráfico

Regla 6:



con la cual generamos la correspondiente trayectoria muy fácilmente. La ecuación correspondiente a esta regla se puede escribir en la forma

$$\text{Regla 6: } \eta(t+\Delta t) = f(\eta(t)) \quad \text{siendo } f(q_1, q_2) = (q_2, -q_1 q_2)$$

Nuevamente se trata de una función invertible que permite volver hacia el pasado para recuperar el estado de partida.

Ejercicio:

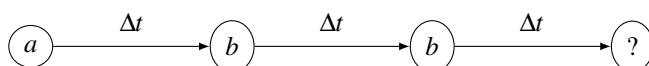
Encuentre los estados necesarios para definir la regla de evolución para la trayectoria “una cara y luego dos cecas” en forma de un diagrama como el anterior.

Escriba también una función que permita codificar la mencionada regla correctamente en una fórmula ¿Es una función invertible?

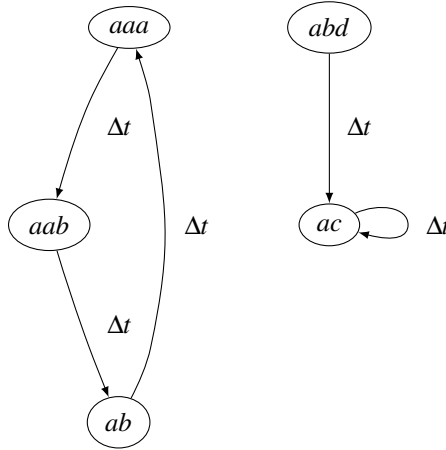
Lo que los dos ejemplos anteriores muestran es que hay una diferencia entre las *configuraciones* posibles del sistema, que en el caso de la moneda serían solamente *cara* y *ceca*, y sus *estados* posibles, que en el último caso de la regla 6 vendrían dados por *cara de ceca*, *cara de cara* y *ceca*, y en el caso anterior de la regla 5 por *cara de cara*, *cara de ceca*, *ceca de cara* y *ceca de ceca*.

El *estado* del sistema se define como la cantidad mínima de información que necesitamos para predecir la evolución temporal. Hicimos esto evidente en las ecuaciones de evolución, en las cuales para predecir el siguiente instante fue necesario definir una variable η que contuviera $q(t)$ y $q(t-\Delta t)$.

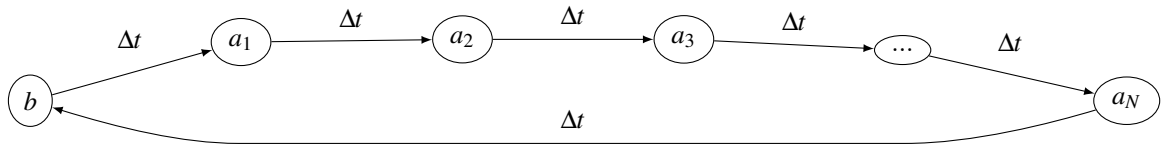
En general, este tipo de sistemas físicos *discretos* tienen un conjunto finito de configuraciones distinguibles que definen su *espacio de configuración* $\mathcal{C} = \{a, b, c, d \dots\}$. Las *trayectorias* son una sucesión de configuraciones del tipo



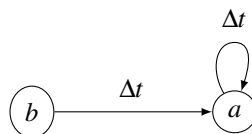
Las *reglas de evolución* se formulan haciendo referencia a sucesiones finitas de configuraciones que forman lo que se denomina el *espacio de estados* del sistema $\mathcal{E} = \{aaa, aab, ab, ac, abc, \dots\}$. Tales reglas se pueden representar según



Estas reglas pueden ser *disipativas* si se pierde información a lo largo de la evolución, y en ese caso no representan la evolución real de ningún sistema físico microscópico. Sin embargo, sabemos que las reglas disipativas sí aparecen a nivel macroscópico, lo que se puede entender con el siguiente ejemplo. Supongamos un espacio de configuración $\mathcal{C} = \{a_1, a_2, \dots, a_N, b\}$ y supongamos la regla de evolución



Esta regla de evolución no presenta flecha del tiempo y es perfectamente aceptable a nivel microscópico. Supongamos que las configuraciones a_1, a_2, \dots, a_N son indistinguibles a nivel macroscópico, llamémoslas a todas a . Entonces veríamos para tiempos cortos, la siguiente regla *efectiva* de evolución



Por supuesto, si esperamos un tiempo lo bastante largo, conocido como *tiempo de recurrencia de Poincaré*, el sistema finalmente llegará a la flecha larga del diagrama anterior, y volverá a su estado inicial b , con lo que veríamos que la dinámica no es realmente disipativa.

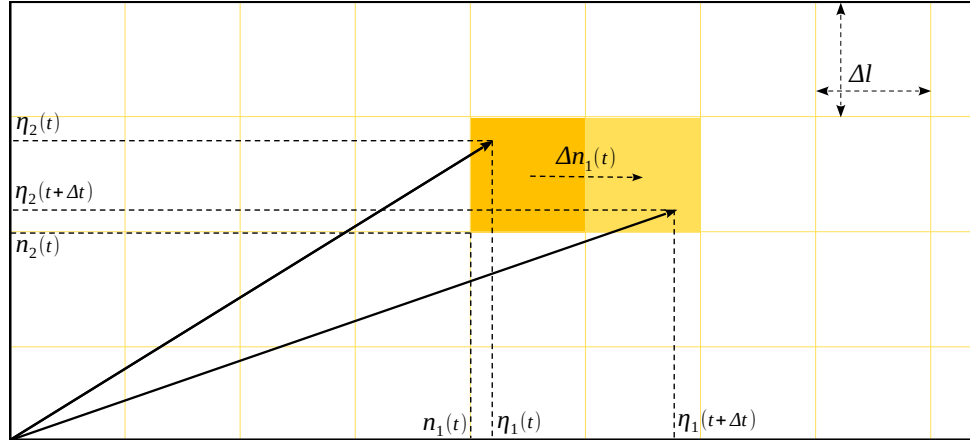


Figura 15.2.1: Discretización de un sistema con configuraciones continuas. El espacio se cuadrícula en hipercubos de lado Δl y la evolución temporal se caracteriza en términos de un $\Delta n(t)$ que parametriza el salto a un hipercubo vecino.

15.3 Sistemas con configuraciones continuas

Ahora imaginemos que queremos aplicar lo que aprendimos en las secciones anteriores a un sistema cuyas configuraciones son continuas, es decir que son puntos en un espacio de estados \mathcal{E} descrito por d variables reales $\{\eta_\alpha\}$.

Podemos primero transformarlo en un sistema discreto, cuadrículando el espacio \mathbb{R}^d en pequeños hipercubos de lado Δl situados en $\eta_\alpha = n_\alpha \Delta l$, donde n_α con $\alpha \in \{1, \dots, d\}$ son números enteros. Con esto, podemos caracterizar la configuración del sistema diciendo en cuál cubo está en cada instante, es decir reemplazando η_α por n_α . Así, hemos construido un sistema discreto del tipo estudiado anteriormente, y la regla de evolución puede entonces escribirse según

$$n_\alpha(t + \Delta t) = f_\alpha(n_\beta(t))$$

Si hacemos que Δt sea lo suficientemente pequeño, es natural imaginar que el sistema solo pudo moverse a cubos cercanos. En otras palabras, podemos escribir

$$n_\alpha(t + \Delta t) = n_\alpha(t) + \Delta n_\alpha(n_\beta(t))$$

Donde la variable $\Delta n_\alpha(n_\beta)$ en cada sitio vale cero para todas las direcciones α excepto una, para la cual vale uno, indicando hacia dónde saltó el sistema. Esto se puede reescribir como

$$\frac{\Delta l n_\alpha(t + \Delta t) - \Delta l n_\alpha(t)}{\Delta t} = \Delta n_\alpha(n_\beta(t)) \frac{\Delta l}{\Delta t}$$

Ahora si hacemos a la vez que Δt sea muy pequeño y refinamos nuestra rejilla Δl , tenemos que $\eta_\alpha \approx n_\alpha \Delta l$ y $u_\alpha \approx \Delta n_\alpha \Delta l / \Delta t$ y obtenemos

$$\dot{\eta}_\alpha(t) = u_\alpha(\eta_\beta(t))$$

donde u_α es algún vector en \mathbb{R}^d que es función de η_α . Es decir que una descripción de la regla de evolución de nuestro sistema viene dada por un sistema autónomo de ecuaciones diferenciales ordinarias de primer orden en el tiempo.

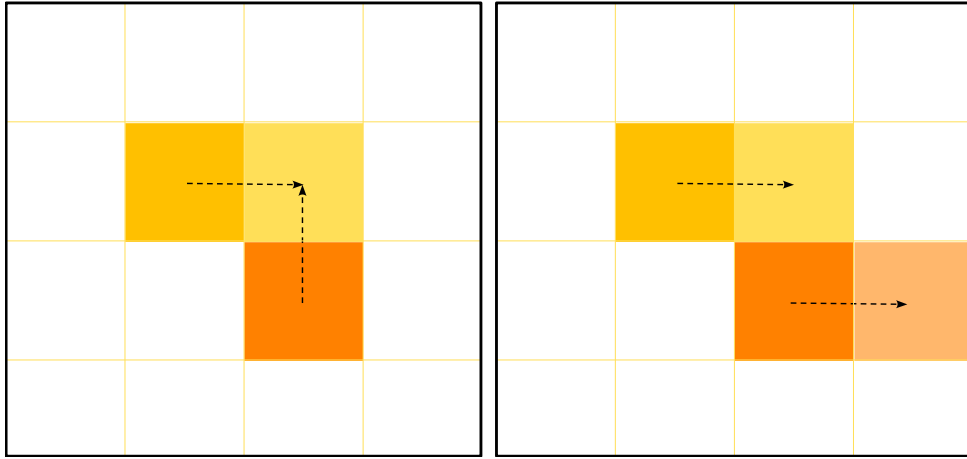


Figura 15.3.1: Evolución en el espacio discretizado. Si la evolución es disipativa (izquierda) a un dado hipercubo del espacio de estados se puede llegar desde dos hipercubos diferentes. Caso contrario la evolución no es disipativa (derecha)

Un punto importante a resaltar es que supusimos que la configuración η_α determina el estado del sistema. Si el estado y la configuración fueran diferentes, necesitaríamos escribir sistemas ecuaciones de orden mayor.

Sin embargo, si queremos describir sistemas realistas a nivel microscópico, necesitamos que las ecuaciones resultantes no determinen una flecha del tiempo, es decir que el sistema no sea disipativo. Para asegurarnos que se cumpla esto, volvamos por un segundo a nuestra rejilla: si el sistema está en el cubo caracterizado por los índices n_α , es necesario estar seguros de que no pudo haber llegado allí desde dos cubos diferentes n_α^1 y n_α^2 . En otras palabras, el contenido de los cubos n_α^1 y n_α^2 no puede haber fluido con la evolución temporal dentro del mismo cubo n_α . El sistema no puede comprimirse, debe *preservar el volumen del espacio de estados*. ¿Puede una ecuación diferencial hacer esto? Veamos eso en algunos ejemplos

Ejemplo: caso unidimensional

Supongamos que η_α tiene una sola componente η , es decir que el espacio de estados es unidimensional. En ese caso, la evolución temporal vendrá dada por la ecuación

$$\dot{\eta} = u(\eta)$$

Definamos ahora una pequeña región del espacio de estados, limitada por η y $\eta + \delta\eta$ ¿adonde se mapeará esa región después de un tiempo infinitesimal dt ? Tenemos que

$$\eta(t+dt) = \eta(t) + u(\eta(t))dt$$

$$(\eta + \delta\eta)(t+dt) = (\eta + \delta\eta)(t) + u(\eta + \delta\eta)dt$$

restando tenemos

$$\delta\eta(t+dt) = \delta\eta(t) + (u(\eta + \delta\eta) - u(\eta))dt$$

Por lo tanto la longitud $d\eta$ de un pequeño intervalo en el espacio de estados necesariamente cambia salvo que $u(\eta)$ sea constante. Es decir que, excepto el sistema trivial dado por u constante, cuya trayectoria es $\eta = u(t-t_0) + \eta_0$, no existen sistemas mecánicos no disipativos cuyo espacio de estados tenga una sola dimensión.

Este resultado podría inducir a confusión ¿qué sucede por ejemplo con una partícula que se mueve en una dimensión? ¿No es acaso un sistema no disipativo bien definido? Por supuesto que sí, pero el punto crucial es que si bien este sistema tiene un espacio de configuración \mathcal{C} que es unidimensional, su espacio de estados \mathcal{E} es bidimensional.

Ejemplo: caso bidimensional

Vamos ahora al caso de un espacio de estados bidimensional. Tendremos que $\eta_\alpha = (\eta_1, \eta_2) \in \mathbb{R}^2$ y entonces la evolución temporal estará dada por

$$\dot{\eta}_\alpha = u_\alpha(\eta_\beta)$$

para un cambio temporal dt tendremos que

$$\eta_\alpha(t+dt) = \eta_\alpha(t) + u_\alpha(\eta_\beta(t))dt$$

o en otras palabras

$$\eta_1(t+dt) = \eta_1(t) + u_1(\eta_\beta(t))dt \quad \eta_2(t+dt) = \eta_2(t) + u_2(\eta_\beta(t))dt$$

Esto puede considerarse como un cambio de coordenadas en el espacio de configuración, que antes describíamos con coordenadas $\eta_\alpha(t)$ y ahora con coordenadas $\eta_\alpha(t+dt)$. Entonces el área cambiará con el jacobiano, según

$$\delta A(t+dt) = \text{Det} \left(\frac{\partial \eta_\alpha(t+dt)}{\partial \eta_\beta(t)} \right) \delta A(t)$$

Esto que se puede escribir como

$$\delta A(t+dt) = \text{Det} \left(\delta_{\alpha\beta} + \frac{\partial u_\alpha}{\partial \eta_\beta} \Big|_t dt \right) \delta A(t) \approx \left(1 + \frac{\partial u_\alpha}{\partial \eta_\alpha} \Big|_t dt \right) \delta A(t)$$

Donde en la segunda igualdad usamos la relación $\text{Det}(I+M) = 1 + \text{Tr}M + \mathcal{O}(M)^2$ para una matriz infinitesimal M . Es decir que para que el área no cambie, necesitamos que se anule la divergencia del vector u_α

$$\frac{\partial u_\alpha}{\partial \eta_\alpha} \Big|_t = 0$$

Sabemos de los cursos de Análisis que esto pasa cuando existe alguna función $H(\eta_\alpha)$ tal que el vector u_α se escribe como

$$u_\alpha = \epsilon_{\alpha\beta} \frac{\partial H}{\partial \eta_\beta}$$

o en componentes

$$u_1 = \frac{\partial H}{\partial \eta_2} \quad u_2 = -\frac{\partial H}{\partial \eta_1}$$

Si tomamos las derivadas de esta expresión, tenemos que

$$\frac{\partial u_1}{\partial \eta_1} = \frac{\partial^2 H}{\partial \eta_1 \partial \eta_2} \quad \frac{\partial u_2}{\partial \eta_2} = -\frac{\partial^2 H}{\partial \eta_2 \partial \eta_1}$$

de donde inmediatamente vemos que, en efecto, la divergencia se anula

$$\frac{\partial u_1}{\partial \eta_1} + \frac{\partial u_2}{\partial \eta_2} = 0$$

Esto conduce a las ecuaciones de movimiento

$$\dot{\eta}_1 = \frac{\partial H}{\partial \eta_2} \quad \dot{\eta}_2 = -\frac{\partial H}{\partial \eta_1}$$

que no son sino ecuaciones de Hamilton del tipo que estudiamos en la clase 10, si hacemos las identificaciones $\eta_1 = q$ y $\eta_2 = p$.

En conclusión, un sistema mecánico bidimensional con reglas de evolución no disipativas, debe necesariamente cumplir las ecuaciones de Hamilton con algún hamiltoniano H . Por supuesto, podríamos ahora hacer una transformación de Legendre para eliminar los impulsos canónicos en favor de las velocidades generalizadas, y obtendríamos una descripción lagrangiana del sistema en términos de ecuaciones de Lagrange.

Nótese que se trata de un resultado de enorme generalidad. Hemos arribado a él sin hacer ningún tipo de suposición previa sobre la dinámica. Es decir, no asumimos que se trata de un sistema formado por partículas, ni que satisface vínculos de un tipo u otro, etc. Sólo impusimos la condición de que el sistema evolucione en el tiempo de manera no disipativa, es decir sin perder información.

Ejemplo: flujo hamiltoniano bidimensional

En el caso bidimensional, los sistemas no disipativos tienen una ventaja adicional que permite describir su movimiento con facilidad. Para identificarla, calculamos el producto escalar del vector de flujo $(\dot{\eta}_1, \dot{\eta}_2)$, que nos dice hacia donde se mueve el sistema en cada punto del espacio de estados, con el gradiente del hamiltoniano

$$\dot{\eta}_1 \frac{\partial H}{\partial \eta_1} + \dot{\eta}_2 \frac{\partial H}{\partial \eta_2} = \frac{\partial H}{\partial \eta_2} \frac{\partial H}{\partial \eta_1} - \frac{\partial H}{\partial \eta_1} \frac{\partial H}{\partial \eta_2} = 0$$

donde en la primera igualdad utilizamos las ecuaciones de Hamilton. Este resultado implica que el flujo es perpendicular al gradiente del Hamiltoniano, es decir que apunta a lo largo de sus curvas de nivel.

En otras palabras, la evolución temporal se produce siguiendo las curvas de nivel del hamiltoniano. Es decir que para tener una descripción de sus trayectorias basta con dibujar dichas curvas de nivel.

Ejercicio:

Obtenga el hamiltoniano de un péndulo a partir de la transformación de Legendre de su lagrangiano, y dibuje sus curvas de nivel con algún software de ploteo como Mathematica, Maple o Geogebra.

Analice cada una de estas curvas como una posible evolución temporal del sistema ¿qué significan las curvas cerradas? ¿qué significan las curvas abiertas?

Ejercicio:

A partir del sistema unidimensional equivalente que encontramos para describir la nutación de un trompo en la sección 9.2.3, encuentre el hamiltoniano correspondiente a través de la transformación de Legendre.

Dibuje sus curvas de nivel y analice el movimiento resultante ¿puede identificar qué curvas corresponden a una precesión cuasiperiódica y cuáles a una precesión alternada?

Podemos probar que el resultado que acabamos de obtener para el caso bidimensional se replica de manera similar para todos los sistemas mecánicos no disipativos con un espacio de estados de dimensión par. En efecto, si tenemos $2D$ variables η_α con ecuación de movimiento

$$\dot{\eta}_\alpha = u_\alpha(\eta_\beta)$$

podemos escribir la evolución para un tiempo infinitesimal en la forma

$$\eta_\alpha(t+dt) = \eta_\alpha(t) + u_\alpha(\eta_\beta(t))dt$$

Si tomamos esta expresión como un cambio de variables en el espacio de estados, el cambio en el volumen se podrá escribir como

$$dV(t+dt) = \text{Det} \left(\frac{\partial \eta_\alpha(t+dt)}{\partial \eta_\beta(t)} \right) dV(t) = \text{Det} \left(\delta_{\alpha\beta} + \frac{\partial u_\alpha}{\partial \eta_\beta} \Big|_t \Delta t \right) dV(t) \approx \left(1 + \frac{\partial u_\alpha}{\partial \eta_\alpha} \Big|_t \Delta t \right) dV(t)$$

donde hemos repetido los pasos del caso bidimensional. Por lo tanto, que se preserve el volumen en el espacio de estados implica que debe anularse la divergencia

$$\frac{\partial u_\alpha}{\partial \eta_\alpha} \Big|_t dt = 0$$

Es fácil ver que esto siempre se puede satisfacer eligiendo una función H tal que

$$u_\alpha = J_{\alpha\beta} \frac{\partial H}{\partial \eta_\beta}$$

donde $J_{\alpha\beta}$ es una matriz o *métrica simpléctica* que cumple $J_{2i-1,2i} = -J_{2i,2i-1} = 1$ y $J_{\alpha\beta} = 0$ en cualquier otro caso. En otras palabras

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ -1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

O sea que la forma de más arriba para u_α implica

$$u_{2i-1} = \frac{\partial H}{\partial \eta_{2i}} \quad u_{2i} = -\frac{\partial H}{\partial \eta_{2i-1}}$$

Para $i \in \{1, \dots, D\}$. Por lo tanto las ecuaciones de movimiento se separan en pares según

$$\dot{\eta}_{2i-1} = \frac{\partial H}{\partial \eta_{2i}} \quad \dot{\eta}_{2i} = -\frac{\partial H}{\partial \eta_{2i-1}}$$

Estas también son ecuaciones de Hamilton, ahora para un sistema en dimensión par arbitraria, si hacemos las identificaciones $q_i = \eta_{2i-1}$ y $p_i = \eta_{2i}$.

Una vez más, hemos llegado a un sistema descrito en términos de ecuaciones de Hamilton a partir del sólo requerimiento de que su evolución sea no disipativa en el sentido de que no se pierda información. De nuevo, podríamos a continuación practicar una transformación de Legendre para describir el sistema en términos de ecuaciones de Lagrange.

Un punto importante a atender es que en estos ejemplos de dimensión par el lagrangiano resultante es independiente del tiempo. Sin embargo, hemos usado las ecuaciones de Lagrange para sistemas con fuerzas monogénicas, cuyo lagrangiano depende del tiempo. En esta instancia tendríamos que preguntarnos ¿serán esos sistemas no disipativos? Veremos la respuesta en el ejemplo que sigue.

Ejemplo: caso tridimensional

En el caso de que el espacio de estados es tridimensional, escribamos

$$\dot{\eta}_\alpha = u_\alpha(\eta_\beta)$$

o, en coordenadas

$$\dot{\eta}_1 = u_1(\eta_\alpha) \quad \dot{\eta}_2 = u_2(\eta_\alpha) \quad \dot{\eta}_3 = u_3(\eta_\alpha)$$

Primero hacemos una *transformación hodográfica* para eliminar el tiempo en favor de la coordenada η_3 . En otras palabras, estamos cambiando la variable independiente para escribir las ecuaciones en términos de un nuevo tiempo η_3 . Obtenemos

$$\eta_1' = \frac{u_1(\eta_\alpha)}{u_3(\eta_\alpha)} \equiv v_1 \quad \eta_2' = \frac{u_2(\eta_\alpha)}{u_3(\eta_\alpha)} \equiv v_2 \quad \eta_3' = 1 \equiv v_3$$

Donde ' significa derivada con respecto a η_3 . Este sistema es completamente equivalente al anterior, sólo que hemos cambiado de reloj. Podemos entonces escribir

$$\eta'_\alpha = v_\alpha(\eta_\beta)$$

Con esto, los cálculos son iguales a los casos de las secciones anteriores, hasta el punto en el que escribimos la divergencia del vector, que en este caso será v_α

$$\left. \frac{\partial v_\alpha}{\partial \eta_\alpha} \right|_t = 0$$

En tres dimensiones, que la divergencia de un vector se anule implica que el vector debe ser el rotor de otro vector h_α , como sigue

$$v_\alpha = \varepsilon_{\alpha\beta\gamma} \frac{\partial h_\gamma}{\partial \eta_\beta}$$

o en componentes

$$v_1 = \frac{\partial h_3}{\partial \eta_2} - \frac{\partial h_2}{\partial \eta_3} \quad v_2 = \frac{\partial h_1}{\partial \eta_3} - \frac{\partial h_3}{\partial \eta_1} \quad v_3 = \frac{\partial h_2}{\partial \eta_1} - \frac{\partial h_1}{\partial \eta_2} = 1$$

Las ecuaciones de movimiento entonces quedan

$$\eta'_1 = \frac{\partial h_3}{\partial \eta_2} - \frac{\partial h_2}{\partial \eta_3} \quad \eta'_2 = -\frac{\partial h_1}{\partial \eta_3} + \frac{\partial h_3}{\partial \eta_1} \quad \eta'_3 = 1$$

Nótese que el vector h_α está definido a menos de un gradiente, por lo que siempre podemos hacer $h_\alpha = h'_\alpha + \partial_\alpha g$ con una función $g(\eta_\alpha)$ arbitraria, y las ecuaciones tomarán la misma forma. Esto se puede usar para obtener un h'_α que cumpla

$$\frac{\partial h'_2}{\partial \eta_1} = 0 \quad \frac{\partial h'_3}{\partial \eta_1} = 0$$

Lo que al reemplazar en las ecuaciones implica

$$\eta'_1 = \frac{\partial h'_3}{\partial \eta_2} \quad \eta'_2 = -\frac{\partial h'_3}{\partial \eta_1} \quad \eta'_3 = 1$$

Es decir que, si identificamos h'_3 con el hamiltoniano H , las variables $q = \eta_1$ y $p = \eta_2$ y el tiempo $t = \eta_3$, hemos llegado a un sistema de ecuaciones de Hamilton en términos de un hamiltoniano que depende del tiempo.

En conclusión, un sistema tridimensional no disipativo resulta en un sistema mecánico bidimensional cuyo hamiltoniano depende del tiempo. Ahora podríamos transformar Legendre para obtener un lagrangiano, que será del tipo de los que obtuvimos para las fuerzas monogénicas.

Ejercicio: sistemas generales en dimensión impar

Es fácil generalizar lo anterior a sistemas cuyo espacio de estados tiene dimensión impar arbitraria, seleccionando una de las coordenadas como el nuevo tiempo, y obteniendo un sistema mecánico no autónomo con un espacio de configuración de dimensión mitad. Las ecuaciones de Hamilton serán

$$\dot{\eta}_{2i-1} = \frac{\partial H}{\partial \eta_{2i}} \quad \dot{\eta}_{2i} = -\frac{\partial H}{\partial \eta_{2i-1}}$$

Con $i \in \{1, \dots, D\}$. Dado que la dimensión es ahora impar, en estas ecuaciones queda excluida la última variable de la lista η_d , que cumple el rol del tiempo. Complete el cálculo correspondiente.

En conclusión, hemos probado que los sistemas no disipativos están descritos por ecuaciones de primer orden que toman la forma de ecuaciones de Hamilton. Las variables que describen la configuración del sistema corresponden a las variables canónicas del sistema hamiltoniano $\{\eta_\alpha\} = \{(p_i, q_i)\}$. Si la dimensión del espacio de estados es impar, una de las variables queda fuera de redefinición anterior y cumple el rol del tiempo. El sistema se puede transformar en un sistema de ecuaciones de Lagrange a través de una transformación de Legendre.

Este resultado, que históricamente fuera probado en la dirección inversa, se conoce como *teorema de Liouville*.

15.4 Resumen

Analizando sistemas discretos, hemos definido las nociones de reloj, espacio de configuraciones, trayectoria, leyes de evolución, y espacio de estados. Hemos concluido que la característica definitoria de un sistema mecánico general es la de ser no disipativo, es decir la de no perder información en la evolución temporal. Esto permite no solo la predicción, sino también la retrodicción.

Aplicando esta idea a los sistemas cuyo espacio de configuración es continuo, hemos descubierto que la evolución temporal está descrita por ecuaciones de primer orden con la forma de Hamilton. Hemos también demostrado que estas ecuaciones se pueden transformar en ecuaciones de Lagrange donde el lagrangiano depende de las coordenadas, las velocidades y el tiempo.

Esto revierte de una enorme generalidad a las leyes de la Mecánica que venimos aprendiendo a lo largo del presente curso, extendiendo su rango de validez más allá de los sistemas de partículas con vínculos con los que empezamos, para hacerlas aplicables a cualquier sistema no disipativo.

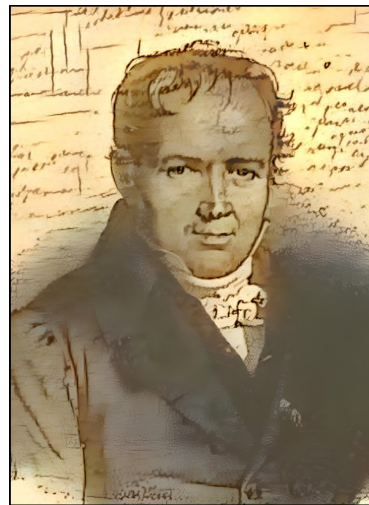
16. Paréntesis de Poisson

16.1 Objetivos

Aprendimos en la clase anterior que cualquier sistema no disipativo, es decir que evoluciona en el tiempo sin perder información, tiene una descripción en términos de ecuaciones de Hamilton.

En un punto crucial en la demostración de este resultado, las variables que parametrizan el espacio de estados se identifican como variables canónicas, separándose en dos subconjuntos diferentes que representan respectivamente las coordenadas y los impulsos.

En esta clase, vamos reparametrizar el espacio de estados utilizando variables arbitrarias no necesariamente canónicas, es decir que no distinguen entre coordenadas e impulsos. En el camino, definiremos un concepto extremadamente útil: el de los paréntesis de Poisson.



Siméon Denis Poisson

16.2 Paréntesis de Poisson

El espacio de estados \mathcal{E} de sistema mecánico puede describirse en términos de un conjunto de $2D$ variables canónicas $\{(p_i, q_i)\}$. Como vimos en las clases previas, sus ecuaciones de movimiento pueden obtenerse a partir de un hamiltoniano $H(p_i, q_i, t)$

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Estas ecuaciones permiten calcular la derivada temporal de una función arbitraria F definida sobre el espacio de estados. Para hacer esto la escribimos como función de las variables canónicas en la forma $F(q_i, p_i, t)$, y luego usamos la regla de la cadena

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial p_i} \dot{p}_i$$

En esta expresión podemos reemplazar las derivadas temporales de las variables canónicas $\{(\dot{q}_i, \dot{p}_i)\}$ haciendo uso de las ecuaciones de Hamilton que escribimos más arriba. Esto da como resultado

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q_i} = \frac{\partial F}{\partial t} + \{F, H\}$$

En la segunda igualdad hemos definido la operación $\{F, H\}$ entre nuestra función F sobre el espacio de estados y el hamiltoniano H . Ésta se denomina *paréntesis de Poisson* y su definición se puede extender para un par de funciones cualesquiera sobre el espacio de estados F y G según

$$\{F, G\} = \frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i}$$

Como veremos en lo que sigue, la distinción de las variables canónicas entre coordenadas y momentos está codificada en los paréntesis de Poisson. Una descripción del movimiento en términos de paréntesis de Poisson mantiene tal información, mientras a la vez permite parametrizar el espacio de fases con variables arbitrarias no necesariamente canónicas.

Ejemplo: ecuaciones de Hamilton

Por consistencia, tendría que ser posible reobtener las ecuaciones de Hamilton usando paréntesis de Poisson para calcular las derivadas temporales de las variables canónicas.

Para ver esto, escribimos

$$\dot{q}_j = \{q_j, H\} \quad \dot{p}_j = \{p_j, H\}$$

y usamos la forma explícita de los paréntesis de Poisson, para deducir

$$\dot{q}_j = \frac{\partial H}{\partial p_i} \overbrace{\frac{\partial q_j}{\partial q_i}}^{\delta_{ij}} - \overbrace{\frac{\partial q_j}{\partial p_i}}^0 \frac{\partial H}{\partial q_i} = \frac{\partial H}{\partial p_j} \quad \dot{p}_j = \frac{\partial H}{\partial p_i} \underbrace{\frac{\partial p_j}{\partial q_i}}_0 - \underbrace{\frac{\partial p_j}{\partial p_i}}_{\delta_{ij}} \frac{\partial H}{\partial q_i} = -\frac{\partial H}{\partial q_j}$$

con lo que hemos recuperado las ecuaciones de Hamilton, como esperábamos.

Los paréntesis de Poisson entre las variables canónicas $\{(p_i, q_i)\}$ resultan de especial interés para la descripción de la dinámica. Los podemos calcular de manera muy simple usando la definición de más arriba, según

$$\begin{aligned} \{q_i, q_j\} &= \frac{\partial q_i}{\partial q_k} \frac{\partial q_j}{\partial p_k} - \frac{\partial q_j}{\partial q_k} \frac{\partial q_i}{\partial p_k} = 0 & \{p_i, p_j\} &= \frac{\partial p_i}{\partial q_k} \frac{\partial p_j}{\partial p_k} - \frac{\partial p_j}{\partial q_k} \frac{\partial p_i}{\partial p_k} = 0 \\ \{q_i, p_j\} &= \frac{\partial q_i}{\partial q_k} \frac{\partial p_j}{\partial p_k} - \frac{\partial p_j}{\partial q_k} \frac{\partial q_i}{\partial p_k} = \delta_{ik} \delta_{jk} = \delta_{ij} \end{aligned}$$

Estos resultados se pueden escribir de forma resumida si utilizamos la notación que introdujimos en la clase anterior para las variables canónicas $\{\eta_\alpha\} = \{(q_i, p_i)\}$ con $\alpha \in \{1, \dots, 2D\}$. En ese caso, las tres fórmulas previas se resumen en una sola línea, según

$$\{\eta_\alpha, \eta_\beta\} = J_{\alpha\beta}$$

siendo J la métrica simpléctica ya definida. Como vemos, estos *paréntesis de Poisson fundamentales* contienen la información sobre la manera en que las variables canónicas se separan en coordenadas e impulsos.

Ejemplo: paréntesis de Poisson en términos de la métrica simpléctica

La matriz J también se puede usar para reescribir la definición del paréntesis de Poisson entre dos funciones cualesquiera sobre el espacio de estados F y G . En efecto, es fácil ver que de la combinación particular de derivadas corresponde a la expresión

$$\{F, G\} = \frac{\partial F}{\partial \eta_\alpha} J_{\alpha\beta} \frac{\partial G}{\partial \eta_\beta} = \frac{\partial F}{\partial \eta_\alpha} \{\eta_\alpha, \eta_\beta\} \frac{\partial G}{\partial \eta_\beta}$$

Aquí en la segunda igualdad hemos reemplazado $J_{\alpha\beta}$ por los paréntesis de Poisson fundamentales $\{\eta_\alpha, \eta_\beta\}$ entre las variables canónicas $\{\eta_\alpha\}$.

Ejemplo: ecuaciones de Hamilton

En términos de los paréntesis de Poisson fundamentales $\{\eta_\alpha, \eta_\beta\} = J_{\alpha\beta}$, las ecuaciones de Hamilton pueden escribirse como

$$\dot{\eta}_\alpha = J_{\alpha\beta} \frac{\partial H}{\partial \eta_\beta} = \{\eta_\alpha, \eta_\beta\} \frac{\partial H}{\partial \eta_\beta}$$

donde la primera igualdad contiene la expresión que vimos en la clase previa, y en la segunda reemplazamos $J_{\alpha\beta}$ por $\{\eta_\alpha, \eta_\beta\}$.

Lo que los ejemplos anteriores evidencian es que, dado un conjunto de variables canónicas $\{\eta_\alpha\}$, podemos reescribir las ecuaciones de Hamilton y el paréntesis de Poisson de dos funciones cualesquiera sin necesidad de identificar cuáles de las variables son coordenadas y cuáles impulsos. Lo único que se requiere son paréntesis de Poisson fundamentales $\{\eta_\alpha, \eta_\beta\}$. Como veremos, esta propiedad se extiende a un conjunto arbitrario de variables.

Nos proponemos describir nuestro espacio de estados \mathcal{E} en términos de un nuevo conjunto de variables $\{\zeta_\alpha\}$, no necesariamente canónicas, definidas por medio del cambio de variables $\zeta_\alpha = \zeta_\alpha(\eta_\beta, t)$. Dos funciones cualesquiera del espacio de estados F y G se pueden escribir entonces como $F(\zeta_\alpha, t)$ y $G(\zeta_\beta, t)$. Si calculamos su paréntesis de Poisson

$$\{F, G\} = \frac{\partial G}{\partial p_i} \frac{\partial F}{\partial q_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i}$$

podemos usar la regla de la cadena para escribirlo en la forma

$$\{F, G\} = \frac{\partial G}{\partial \zeta_\alpha} \frac{\partial \zeta_\alpha}{\partial p_i} \frac{\partial F}{\partial \zeta_\beta} \frac{\partial \zeta_\beta}{\partial q_i} - \frac{\partial F}{\partial \zeta_\alpha} \frac{\partial \zeta_\alpha}{\partial p_i} \frac{\partial G}{\partial \zeta_\beta} \frac{\partial \zeta_\beta}{\partial q_i} = \left(\frac{\partial \zeta_\alpha}{\partial p_i} \frac{\partial \zeta_\beta}{\partial q_i} - \frac{\partial \zeta_\alpha}{\partial p_i} \frac{\partial \zeta_\alpha}{\partial q_i} \right) \frac{\partial G}{\partial \zeta_\beta} \frac{\partial F}{\partial \zeta_\beta}$$

o en otras palabras

$$\{F, G\} = \frac{\partial F}{\partial \zeta_\alpha} \{\zeta_\alpha, \zeta_\beta\} \frac{\partial G}{\partial \zeta_\beta}$$

Es decir que el paréntesis de Poisson de F y G en las nuevas variables $\{\zeta_\alpha\}$ queda escrito en términos de los paréntesis de Poisson fundamentales $\{\zeta_\alpha, \zeta_\beta\}$ de las mismas. En el caso en que las coordenadas son canónicas, tenemos que $\{\eta_\alpha, \eta_\beta\} = J_{\alpha\beta}$ y recuperamos la expresión del ejemplo de más arriba.

La regla que hemos encontrado se puede aplicar para obtener las ecuaciones de movimiento en términos de las nuevas variables $\{\zeta_\alpha\}$. En efecto, sabemos que la derivada temporal de cualquier función del espacio de estados viene dada por su paréntesis de Poisson con el hamiltoniano, según

$$\dot{\zeta}_\alpha = \{\zeta_\alpha, H\} + \frac{\partial \zeta_\alpha}{\partial t} = \frac{\partial \zeta_\alpha}{\partial \zeta_\beta} \{\zeta_\beta, \zeta_\gamma\} \frac{\partial H}{\partial \zeta_\gamma} + \frac{\partial \zeta_\alpha}{\partial t} = \delta_{\alpha\beta} \{\zeta_\beta, \zeta_\gamma\} \frac{\partial H}{\partial \zeta_\gamma} + \frac{\partial \zeta_\alpha}{\partial t}$$

o en otras palabras

$$\dot{\zeta}_\alpha = \{\zeta_\alpha, \zeta_\beta\} \frac{\partial H}{\partial \zeta_\beta} + \frac{\partial \zeta_\alpha}{\partial t}$$

Con lo que hemos logrado escribir las ecuaciones de Hamilton para un conjunto de variables no necesariamente canónicas que describen el espacio de estados.

Lo que los cálculos que hemos desarrollado demuestran, es que podemos describir el espacio de estados en términos de un conjunto de variables arbitrario, siempre que tengamos la información de la derivada temporal parcial de cada variable $\partial \zeta_\alpha / \partial t$ y de los paréntesis de Poisson fundamentales $\{\zeta_\alpha, \zeta_\beta\}$. Con estos elementos, podemos calcular el paréntesis de Poisson entre dos funciones arbitrarias sobre el espacio de estados, y podemos describir el movimiento del sistema.

Ejemplo: oscilador armónico

Sea el hamiltoniano de un oscilador armónico unidimensional

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

escrito en términos del par de variables canónicas (p, q) . Definamos un nuevo par de variables (ρ, α) en el espacio de estados, que no necesariamente es canónico, y que está dado por

$$\rho = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

$$\alpha = \arctan \left(\sqrt{\frac{1}{km}} \frac{p}{q} \right)$$

El paréntesis de Poisson fundamental entre estas nuevas variables se puede escribir como

$$\{\rho, \alpha\} = \frac{\partial \rho}{\partial q} \frac{\partial \alpha}{\partial p} - \frac{\partial \rho}{\partial p} \frac{\partial \alpha}{\partial q} = \sqrt{\frac{k}{m}}$$

Es obvio entonces que no se trata de variables canónicas, ya que si lo fueran deberíamos haber obtenido ± 1 . Es evidente también que $\{\rho, \rho\} = \{\alpha, \alpha\} = 0$ y $\{\alpha, \rho\} = -\sqrt{k/m}$. En términos de las nuevas coordenadas el hamiltoniano se escribe

$$H = \rho$$

Con esto, estamos en condiciones de escribir las ecuaciones de Hamilton para las nuevas variables, en la forma

$$\dot{\rho} = \{\rho, \alpha\} \frac{\partial H}{\partial \alpha} = 0 \quad \dot{\alpha} = \{\alpha, \rho\} \frac{\partial H}{\partial \rho} = -\sqrt{\frac{k}{m}}$$

Estas ecuaciones se pueden resolver muy fácilmente, según

$$\rho = \rho_0$$

$$\alpha = \alpha_0 - \sqrt{\frac{k}{m}}(t - t_0)$$

Con lo que hemos resuelto el problema.

Nótese que el cálculo de las ecuaciones de movimiento a partir del hamiltoniano no requirió del conocimiento del cambio de variables, si asumimos conocido el paréntesis de Poisson $\{\rho, \alpha\}$. En otras palabras, dado el paréntesis de Poisson fundamental, el hamiltoniano contiene la información dinámica del sistema sin necesidad de recurrir a coordenadas canónicas.

El resultado de arriba se puede invertir para reobtener la solución en términos de las coordenadas originales usando la transformación inversa

$$p = \sqrt{2m\rho} \cos \alpha$$

$$q = \sqrt{\frac{2\rho}{k}} \sin \alpha$$

con lo que recuperamos la solución para el oscilador armónico

$$p = \sqrt{2m\rho_0} \cos \left(\sqrt{\frac{k}{m}}(t - t_0) - \alpha_0 \right)$$

$$q = \sqrt{\frac{2\rho_0}{k}} \sin \left(\sqrt{\frac{k}{m}}(t - t_0) - \alpha_0 \right)$$

16.2.1 Álgebra de observables

El estado de un sistema mecánico está completamente determinado por el valor de sus variables canónicas. Esto significa que cualquier magnitud que se pueda medir sobre el sistema debe ser una función de tales variables. En otras palabras, las diferentes funciones sobre el espacio de estados representan a los distintos *observables* físicos del sistema.

La suma de dos funciones sobre el espacio de estados es una nueva función, y lo mismo sucede con el producto de una función por un número real. Esto significa que el conjunto de observables

tiene la estructura de un *espacio vectorial*. Más aún, el paréntesis de Poisson toma dos funciones sobre el espacio de estados, y nos devuelve como resultado una tercera. Esto dota al conjunto de observables con una estructura de *álgebra*.

La operación definida por el paréntesis de Poisson cumple con una serie de propiedades, entre las cuales resultan especialmente importantes las que se enumeran a continuación:

- **Antisimetría:** es evidente a partir de su definición que el paréntesis de Poisson no es una operación conmutativa, sino que cumple la propiedad

$$\{F, G\} = -\{G, F\}$$

La prueba es trivial, a saber

$$\{F, G\} = \frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} = - \left(\frac{\partial G}{\partial q_i} \frac{\partial F}{\partial p_i} - \frac{\partial G}{\partial p_i} \frac{\partial F}{\partial q_i} \right) = -\{G, F\}$$

- **Bilinealidad:** dadas dos funciones del tiempo α, β , se puede probar muy fácilmente que para cualquier conjunto de tres funciones F, G y H en el espacio de estados, se cumple la propiedad

$$\{F, \alpha G + \beta H\} = \alpha \{F, G\} + \beta \{F, H\}$$

La prueba es nuevamente muy simple, escribiendo

$$\{F, \alpha G + \beta H\} = \frac{\partial F}{\partial q_i} \frac{\partial (\alpha G + \beta H)}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial (\alpha G + \beta H)}{\partial q_i}$$

y usando el hecho de que ni α ni β dependen de las coordenadas $\{q_i\}$ y momentos $\{p_i\}$ podemos reacomodar las derivadas en la forma

$$\{F, \alpha G + \beta H\} = \frac{\partial F}{\partial q_i} \left(\alpha \frac{\partial G}{\partial p_i} + \beta \frac{\partial H}{\partial p_i} \right) - \frac{\partial F}{\partial p_i} \left(\alpha \frac{\partial G}{\partial q_i} + \beta \frac{\partial H}{\partial q_i} \right)$$

Lo que nos permite entonces obtener

$$\begin{aligned} \{F, \alpha G + \beta H\} &= \alpha \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} \right) + \beta \left(\frac{\partial F}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = \\ &= \alpha \{F, G\} + \beta \{F, H\} \end{aligned}$$

La misma propiedad se puede demostrar en la primera componente, a saber

$$\{\alpha F + \beta G, H\} = \alpha \{F, H\} + \beta \{G, H\}$$

- **Regla de Leibniz:** Si tenemos tres funciones F, G y H en el espacio de fases, se cumple que

$$\{F, GH\} = \{F, G\}H + G\{F, H\}$$

Lo que implica que la operación $\{F, \cdot\}$ actúa como una derivada en las funciones del espacio de estados. Para probar esta regla, escribimos la forma explícita de la expresión de arriba

$$\{F, GH\} = \frac{\partial F}{\partial q_i} \frac{\partial (GH)}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial (GH)}{\partial q_i}$$

Expandiendo el producto de derivadas, reescribimos

$$\{F, GH\} = \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} \right) H + G \left(\frac{\partial F}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = \{F, G\}H + G\{F, H\}$$

con lo que hemos demostrado la propiedad.

- Identidad de Jacobi: dadas tres funciones cualesquiera en el espacio de estados F, G y H , el paréntesis de Poisson verifica

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$$

Esto significa que no se trata de una operación asociativa.

Ejercicio:

Pruebe la identidad de Jacobi. Se trata de una demostración un poco más larga que las anteriores, ya que involucra derivadas segundas. Escribiremos aquí los primeros pasos, comenzando con

$$\{F, \{G, H\}\} = \frac{\partial F}{\partial q_i} \frac{\partial \{G, H\}}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial \{G, H\}}{\partial q_i}$$

o más explícitamente

$$\{F, \{G, H\}\} = \frac{\partial F}{\partial q_i} \frac{\partial}{\partial p_i} \left(\frac{\partial G}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial G}{\partial p_j} \frac{\partial H}{\partial q_j} \right) - \frac{\partial F}{\partial p_i} \frac{\partial}{\partial q_i} \left(\frac{\partial G}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial G}{\partial p_j} \frac{\partial H}{\partial q_j} \right)$$

y actuando con las derivadas en las expresiones entre paréntesis, obtenemos ocho términos con derivadas primeras y segundas. En estos términos hay que permutar F, G y H para obtener los otros dos ordenamientos de la identidad de Jacobi. En la suma resultante, cada expresión aparece dos veces con diferente signo y por lo tanto se cancela.

Estas propiedades nos permiten abstraer la estructura del *álgebra de observables* de un sistema mecánico. En efecto, podemos afirmar que dos observables cualesquiera se componen para dar un tercer observable, cumpliendo con una ley que es antisimétrica, bilineal, y satisface la identidad de Jacobi y la regla de Leibniz.

Esta caracterización de las propiedades de los observables nos permite generalizar las leyes de la mecánica, para extenderlas incluso más allá de los sistemas no disipativos con espacio de configuración continuo que vimos en el capítulo anterior. Podemos considerar otros sistemas cuyos observables no sean funciones sobre un espacio de estados, sino algún otro tipo de objeto que cumpla con un álgebra similar.

Ejemplo: mecánica de matrices

El conmutador de matrices, que toma dos matrices y nos devuelve una tercera de acuerdo a la regla

$$[A, B] = A \cdot B - B \cdot A$$

Es una aplicación del espacio de matrices en sí mismo que cumple propiedades análogas a las del paréntesis de Poisson, a saber

- Antisimetría: $[A, B] = -[B, A]$
- Bilinealidad: $[A, \alpha B + \beta C] = \alpha [A, B] + \beta [A, C]$, donde α y β ahora son números.
- Identidad de Jacobi: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$
- Regla de Leibniz: $[A, B \cdot C] = [A, B] \cdot C + B \cdot [A, C]$

Esto nos permite definir una *mecánica de matrices*, en la que los observables son matrices y existe una matriz hamiltoniana que describe la evolución temporal.

Ejemplo: mecánica de operadores

Un operador diferencial lineal \hat{L} que actúa sobre las funciones del espacio de configuración se define según

$$\hat{L}\Psi(q_i) = \sum_o c_{i_1 \dots i_o} \frac{\partial^o \Psi}{\partial q_{i_1} \dots \partial q_{i_o}}$$

donde la suma corre sobre los diferentes órdenes o de derivación, y los coeficientes $c_{i_1 \dots i_o}$ son funciones de q_i

El producto $\hat{M}\hat{L}$ de operadores diferenciales lineales \hat{M} y \hat{L} es otro operador diferencial lineal, cuya acción sobre una función del espacio de configuración $\Psi(q_i)$ se define como la aplicación sucesiva $(\hat{M}\hat{L})\Psi = \hat{M}(\hat{L}\Psi)$. Con esto definimos

$$[\hat{L}, \hat{M}] = \hat{L}\hat{M} - \hat{M}\hat{L}$$

Este conmutador cumple propiedades análogas a las enumeradas más arriba para el paréntesis de Poisson. Por lo tanto, podemos definir una *mecánica de operadores* en la que los observables sean operadores lineales sobre un espacio de configuración, uno de los cuales es un operador hamiltoniano que determina la evolución temporal.

Este ejemplo resulta particularmente interesante, ya que sabemos que en la Mecánica Cuántica los observables se representan como operadores diferenciales lineales sobre la función de onda del sistema, que es una función sobre el espacio de configuración. En este sentido, la Mecánica Cuántica es una generalización de la Mecánica.

Ejemplo: cuantización canónica

El procedimiento de *cuantización* consiste en encontrar un sistema cuántico a partir de su límite clásico. Por supuesto este proceso no es unívoco, ya que puede existir más de un sistema cuántico con el mismo límite clásico.

Se trata de comenzar con un sistema clásico donde los observables son funciones sobre el espacio de estados F , y transformarlo en un sistema cuántico donde son

operadores diferenciales lineales \hat{F} que actúan sobre las funciones del espacio de configuración.

La similitud de los paréntesis de Poisson con los conmutadores de operadores diferenciales en el espacio de configuración nos permite definir la *regla de cuantización canónica*. Se trata de una aplicación de cuantización $\hat{}$ que transforma una función del espacio de estados en un operador sobre funciones del mismo, y cumple que

$$[\hat{F}, \hat{G}] = i\hbar \widehat{\{F, G\}}$$

Es decir, el operador que corresponde al paréntesis de Poisson de dos observables, está dado por el conmutador de los operadores que corresponden a cada uno de los observables.

La regla propuesta se verifica en el caso particular de las coordenadas y los impulsos, si definimos sus operadores asociados según

$$\hat{p}_i = \frac{\hbar}{i} \frac{\partial}{\partial q_i} \quad \hat{q}_i = q_i$$

Para ver esto, hacemos actuar al conmutador de operadores sobre una función cualquiera $\Psi(q_i)$ del espacio de configuración

$$[\hat{q}_j, \hat{p}_i]\Psi = \hat{q}_j(\hat{p}_i\Psi) - \hat{p}_i(\hat{q}_j\Psi) = \frac{\hbar}{i} q_j \frac{\partial}{\partial q_i} \Psi - \frac{\hbar}{i} \frac{\partial}{\partial q_i} (q_j \Psi) = i\hbar \delta_{ij} \Psi$$

con lo que vemos que se cumple el requerimiento $[\hat{q}_j, \hat{p}_i] = i\hbar \widehat{\{q_j, p_i\}}$.

16.3 Resumen

En esta clase, definimos la operación matemática conocida como parentesis de Poisson y estudiamos sus propiedades y aplicaciones.

Aplicamos el paréntesis de Poisson para la descripción del movimiento en términos de un conjunto de variables cualesquiera no necesariamente canónicas que parametrizan el espacio de estados. Vimos que la información sobre la distinción entre coordenadas e impulsos se almacena en los paréntesis de Poisson fundamentales los cuales, junto con la derivada temporal parcial de cada coordenada, permiten escribir las ecuaciones de movimiento del sistema.

Aprendimos que el paréntesis de Poisson es una operación binaria, que dota al conjunto de los observables del sistema de una estructura de álgebra. Esta idea nos permitió abstraer la noción de sistema mecánico para poder generalizarla, incluyendo a los sistemas cuánticos. Vimos que el paréntesis de Poisson resulta útil para encontrar la versión cuántica de un dado sistema clásico a través de la regla de cuantización canónica.

17. Transformaciones canónicas

17.1 Objetivos

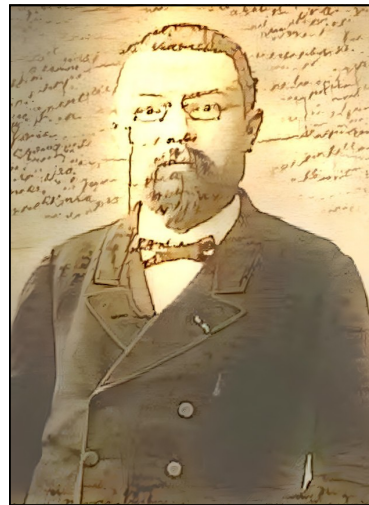
En la clase previa estudiamos la posibilidad de reparametrizar el espacio de estados mediante un cambio de variables completamente general. Encontramos la forma de escribir las ecuaciones de movimiento de las nuevas variables, que no necesariamente eran canónicas. Esto requería el conocimiento de los paréntesis de Poisson fundamentales entre las variables, y de sus derivadas temporales parciales.

Sin embargo, entre todos los cambios de variables posibles existe un subconjunto particular compuesto por aquéllos que transforman variables canónicas en variables canónicas. Se trata de transformaciones que preservan los paréntesis de Poisson, y que resultan particularmente útiles para estudiar la evolución temporal y las simetrías en el contexto de la formulación hamiltoniana.

En esta clase nos concentraremos en la obtención, las propiedades, y la aplicación de tales cambios de variables.

17.2 Transformaciones canónicas

Un cambio de variables en el espacio de estados se denomina una *transformación canónica*, cuando cumple que a partir de un conjunto de variables canónicas, es decir uno en el que se pueden distinguir coordenadas de impulsos $\{(p_i, q_i)\}$, devuelve otro conjunto de variables que



Henri Poincaré

también son canónicas $\{(P_i, Q_i)\}$.

Aprendimos antes que las variables canónicas tienen paréntesis de Poisson fundamentales dados por las expresiones

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

Si las nuevas coordenadas también son canónicas, deben cumplir fórmulas similares

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

Dado el cambio de variables en la forma $P_i = P_i(p_j, q_j, t)$ y $Q_i = Q_i(p_j, q_j, t)$, la primera condición puede escribirse explícitamente como

$$\frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial P_j}{\partial q_k} = \delta_{ij}$$

Este es un conjunto de ecuaciones diferenciales para las funciones $P_i(p_j, q_j, t)$ y $Q_i(p_j, q_j, t)$, que podríamos intentar resolver para obtener la transformación canónica más general. Una manera de hacerlo es invertir la relación $Q_i = Q_i(p_j, q_j, t)$ para obtener $p_i = \chi_i(Q_j, q_j, t)$, con lo que también podemos escribir $P_i = P_i(\chi_j(Q_k, q_k), q_j, t) = \xi_i(Q_j, q_j, t)$. Es evidente que, dada la definición de la función χ_i , se cumple la identidad

$$Q_i(\chi_j(Q_k, q_k), q_j, t) = Q_i$$

que se puede derivar respecto de Q_j a ambos lados para obtener

$$\frac{\partial Q_i}{\partial p_k} \frac{\partial \chi_k}{\partial Q_j} = \delta_{ij}$$

Por otro lado, reemplazando la función $\xi_i(Q_j, q_j)$ en el paréntesis de Poisson

$$\frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial P_j}{\partial q_k} = \frac{\partial Q_i}{\partial q_k} \frac{\partial \xi_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial \xi_j}{\partial q_k} = \frac{\partial Q_i}{\partial q_k} \frac{\partial \xi_j}{\partial Q_l} \frac{\partial Q_l}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \left(\frac{\partial \xi_j}{\partial q_k} + \frac{\partial \xi_j}{\partial Q_l} \frac{\partial Q_l}{\partial q_k} \right)$$

lo que se puede reordenar

$$\frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial P_j}{\partial q_k} = \frac{\partial \xi_j}{\partial Q_l} \left(\frac{\partial Q_i}{\partial q_k} \frac{\partial Q_l}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial Q_l}{\partial q_k} \right) - \frac{\partial Q_i}{\partial p_k} \frac{\partial \xi_j}{\partial q_k} = \frac{\partial \xi_j}{\partial Q_l} \{Q_i, Q_l\} - \frac{\partial Q_i}{\partial p_k} \frac{\partial \xi_j}{\partial q_k}$$

Dado que queremos que las nuevas variables sean canónicas, se debe cumplir también que $\{Q_i, Q_l\} = 0$. Imponiendo esa condición en la última ecuación, obtenemos

$$-\frac{\partial Q_i}{\partial p_k} \frac{\partial \xi_j}{\partial q_k} = \delta_{ij}$$

Pero más arriba habíamos expresado la delta de Kronecker δ_{ij} en términos de derivadas de $\chi_i(Q_j, q_j)$, con lo que podemos establecer la igualdad

$$-\frac{\partial Q_i}{\partial p_k} \frac{\partial \xi_j}{\partial q_k} = \frac{\partial Q_i}{\partial p_k} \frac{\partial \chi_k}{\partial Q_j}$$

o en otras palabras

$$-\frac{\partial \xi_j}{\partial q_k} = \frac{\partial \chi_k}{\partial Q_j}$$

Esta ecuación se resuelve fácilmente en términos de una función $F_1(q_i, Q_i, t)$ en la forma

$$\chi_k = \frac{\partial F_1}{\partial q_k} \quad \xi_j = -\frac{\partial F_1}{\partial Q_j}$$

donde hemos asumido las condiciones de regularidad necesarias para que las derivadas conmuten, de manera de resolver la ecuación. Con esto, para cada función $F_1(q_i, Q_i, t)$ podremos definir una transformación canónica, según las ecuaciones

$$p_i(q_j, Q_j, t) = \frac{\partial F_1}{\partial q_i} \quad P_i(q_j, Q_j, t) = -\frac{\partial F_1}{\partial Q_i}$$

La forma explícita de la transformación $P_i = P_i(p_j, q_j, t)$, $Q_i = Q_i(p_j, q_j, t)$ se obtiene invirtiendo estas ecuaciones. La función $F_1(q_i, Q_i, t)$ se conoce como *función generatriz* de la transformación canónica.

Ejemplo: dualidad de Poincaré

Supongamos que tenemos la función generatriz más simple posible que depende de las variables q_i y Q_i , dada por

$$F_1 = q_i Q_i$$

Si la usamos para escribir una transformación canónica según la fórmula de más arriba, obtenemos

$$p_i(q_j, Q_j, t) = \frac{\partial F_1}{\partial q_i} = Q_i \quad P_i(q_j, Q_j, t) = -\frac{\partial F_1}{\partial Q_i} = -q_i$$

Es decir, una transformación canónica que intercambia coordenadas con momentos. Esto se conoce como *dualidad de Poincaré*.

El que hemos encontrado más arriba no es el único tipo de función generatriz que podemos definir. En efecto, podemos hacer transformaciones de Legendre de F_1 para obtener otros tipos de funciones generatrices. Por ejemplo, dado

$$P_i(q_j, Q_j, t) = -\frac{\partial F_1}{\partial Q_i}$$

podemos invertir la ecuación para tener

$$Q_i = \gamma_i(P_j, q_j, t)$$

Y con esto hacer una transformación de Legendre según

$$F_2(P_i, q_i, t) = P_i \gamma_i(P_j, q_j, t) + F_1(\gamma_k(P_j, q_j, t), q_k, t)$$

Como aprendimos cuando estudiamos transformaciones de Legendre, se cumple que

$$\frac{\partial F_2}{\partial P_i} = Q_i(P_i, q_i, t) \quad \frac{\partial F_2}{\partial q_i} = p_i(P_i, q_i, t)$$

Esto quiere decir que para cualquier función $F_2(P_i, q_i, t)$ podemos definir una transformación canónica invirtiendo las fórmulas de más arriba para obtener $P_i = P_i(p_j, q_j, t)$, $Q_i = Q_i(p_j, q_j, t)$.

Ejemplo: transformación identidad

El caso más simple para las funciones generatrices de tipo $F_2(P_i, q_i, t)$ está dado por

$$F_2(P_i, q_i, t) = q_i P_i$$

de la cual se deduce la ley de transformación

$$Q_i(P_j, q_j, t) = \frac{\partial F_2}{\partial P_i} = q_i \quad p_i(P_j, q_j, t) = \frac{\partial F_2}{\partial q_i} = P_i$$

Es decir que se trata de la transformación identidad.

Ejemplo: cambios de coordenadas

Sabemos que la descripción lagrangiana nos permite usar coordenadas generalizadas arbitrarias. Es natural entonces preguntarse si un cambio arbitrario de las coordenadas es una transformación canónica. Para responder esto, escribimos la función generatriz

$$F_2(P_i, q_i, t) = s_i(q_j, t) P_i$$

siendo $h_i(q_j, t)$ un conjunto arbitrario de tantas funciones como coordenadas. Usando las derivadas como se explicó más arriba, tenemos

$$Q_i(P_j, q_j, t) = \frac{\partial F_2}{\partial P_i} = s_i(q_j, t) \quad p_i(P_j, q_j, t) = \frac{\partial F_2}{\partial q_i} = \frac{\partial s_j}{\partial q_i} P_j$$

Con lo que hemos realizado un cambio arbitrario de coordenadas con una transformación canónica. Puede probarse también que el cambio en los impulsos es precisamente el que corresponde en el formalismo lagrangiano, si recordamos la identidad que probamos en los primeros capítulos $\partial s_j / \partial q_i = \partial \dot{s}_j / \partial \dot{q}_i$.

Por supuesto, tenemos la opción alternativa de hacer la transformación de Legendre respecto de la otra variable. Es decir, usando que

$$p_i(q_j, Q_j, t) = \frac{\partial F_1}{\partial q_i}$$

e invirtiendo para obtener

$$q_i = \mu_i(Q_j, p_j, t)$$

Con esa función escribimos la transformación de Legendre según

$$F_3(p_i, Q_i, t) = p_i \mu_i(Q_j, p_j, t) - F_1(\mu_k(Q_j, p_j, t), Q_k, t)$$

Es evidente entonces a partir de las propiedades de las transformadas de Legendre que se cumple que

$$\frac{\partial F_3}{\partial p_i} = q_i(p_i, Q_i, t) \quad \frac{\partial F_3}{\partial Q_i} = P_i(p_i, Q_i, t)$$

En otras palabras, para cualquier función $F_3(P_i, q_i, t)$ podemos obtener una transformación canónica invirtiendo las fórmulas de arriba.

Ejercicio:

Dada la forma más simple de la función $F_3(P_i, q_i, t)$

$$F_3(p_i, Q_i, t) = p_i Q_i$$

obtener la forma de la transformación canónica resultante, demostrando que se trata nuevamente de la transformación identidad.

También podemos hacer la transformación de Legendre de $F_1(q_i, Q_i, t)$ respecto de todas sus variables, en la forma

$$F_4(p_i, P_i, t) = p_i \tilde{\mu}_i(P_j, p_j, t) - P_i \tilde{\gamma}_i(P_j, p_j, t) - F_1(\tilde{\mu}_i(P_j, p_j, t), \tilde{\gamma}_i(P_j, p_j, t), t)$$

donde las funciones $q_i = \tilde{\mu}_i(P_j, p_j, t)$ y $Q_i = \tilde{\gamma}_i(P_j, p_j, t)$ se obtienen invirtiendo las derivadas de $F_1(q_i, Q_i, t)$. Esto resulta en las siguientes relaciones

$$\frac{\partial F_4}{\partial p_i} = q_i(p_i, Q_i, t) \quad \frac{\partial F_4}{\partial P_i} = -Q_i(p_i, Q_i, t)$$

cuyas funciones inversas proveen la transformación canónica en su forma explícita.

Ejercicio:

Dado el ejemplo trivial para una función del tipo anterior

$$F_4(p_i, P_i, t) = p_i P_i$$

obtener la correspondiente transformación canónica, demostrando que se reobtiene la dualidad de Poincaré.

Con esto, estamos en condiciones de establecer las ecuaciones de movimiento para las variables canónicas transformadas. Éstas se obtienen utilizando los paréntesis de Poisson, en la forma

$$\dot{Q}_i = \{Q_i, H\} + \frac{\partial Q_i}{\partial t} = \frac{\partial H}{\partial P_i} + \frac{\partial Q_i}{\partial t} \quad \dot{P}_i = \{P_i, H\} + \frac{\partial P_i}{\partial t} = -\frac{\partial H}{\partial Q_i} + \frac{\partial P_i}{\partial t}$$

Extrañamente, estas expresiones no parecen tener la forma de ecuaciones de Hamilton, que es lo que esperaríamos si las variables $\{(P_i, Q_i)\}$ fueran realmente canónicas. Para resolver esto, comencemos suponiendo que la transformación canónica está generada por una función de tipo F_1 , lo que implica

$$p_i = \frac{\partial F_1}{\partial q_i} \quad P_i = -\frac{\partial F_1}{\partial Q_i}$$

Por otro lado, también podríamos describirla en términos de la transformada de Legendre F_2 , lo que nos permite escribir

$$p_i = \frac{\partial F_2}{\partial q_i} \quad Q_i = \frac{\partial F_2}{\partial P_i}$$

De este par de ecuaciones y del anterior, extraemos la segunda ecuación para reescribir las derivadas temporales en la forma

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} + \frac{\partial^2 F_2}{\partial t \partial P_i} = \frac{\partial}{\partial P_i} \left(H + \frac{\partial F_2}{\partial t} \right) \quad \dot{P}_i = -\frac{\partial H}{\partial Q_i} - \frac{\partial^2 F_1}{\partial t \partial Q_i} = -\frac{\partial}{\partial Q_i} \left(H + \frac{\partial F_1}{\partial t} \right)$$

Ahora bien, si recordamos que $F_1(Q_i, q_i, t)$ y $F_2(P_i, q_i, t)$ están relacionadas por una transformación de Legendre, sabemos que debe cumplirse la identidad $\partial F_2 / \partial t = \partial F_1 / \partial t$, con lo que las derivadas temporales de las nuevas variables canónicas pueden reescribirse como ecuaciones de Hamilton

$$\dot{Q}_i = \frac{\partial}{\partial P_i} \left(H + \frac{\partial F_1}{\partial t} \right) = \frac{\partial K}{\partial P_i} \quad \dot{P}_i = -\frac{\partial}{\partial Q_i} \left(H + \frac{\partial F_1}{\partial t} \right) = -\frac{\partial K}{\partial Q_i}$$

En términos del nuevo hamiltoniano

$$K(P_i, Q_i, t) = H(p_i(P_i, Q_i, t), q_i(P_i, Q_i, t), t) + \frac{\partial F_1}{\partial t}$$

En otras palabras, la dependencia temporal de una transformación canónica se refleja en un cambio en la forma del Hamiltoniano.

Con esto, hemos tenido éxito en construir cambios de variables que preservan el carácter canónico de las mismas. Estas transformaciones mantienen la forma de los paréntesis de Poisson fundamentales, y están definidas en términos de funciones generatrices de cuatro tipos posibles. La evolución de las nuevas variables se describe en términos de ecuaciones de Hamilton.

17.2.1 Transformaciones canónicas infinitesimales

Supongamos ahora que tenemos una transformación canónica que es cercana a la identidad, es decir que la diferencia entre las variables transformadas $\{(P_i, Q_i)\}$ y las originales $\{(p_i, q_i)\}$ es una magnitud muy pequeña. Es natural suponer que su función generatriz estará dada por una pequeña deformación de aquella que genera la transformación identidad. En otras palabras, podemos escribir

$$F_2(P_i, q_i, t) = q_i P_i + \varepsilon g(P_i, q_i, t)$$

donde ε es un parámetro infinitesimal, y la nueva función $g(P_i, q_i, t)$ se conoce como el *generador* de la transformación. Bajo la acción de esta transformación, las variables canónicas se verán modificadas en la forma

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i + \varepsilon \frac{\partial g}{\partial q_i} \quad Q_i = \frac{\partial F_2}{\partial P_i} = q_i + \varepsilon \frac{\partial g}{\partial P_i}$$

La segunda ecuación se puede reescribir usando la regla de la cadena según

$$Q_i = \frac{\partial F_2}{\partial P_i} = q_i + \varepsilon \frac{\partial g}{\partial p_j} \frac{\partial p_j}{\partial P_i}$$

reemplazando ahora la forma explícita de p_j en esta expresión, tenemos

$$Q_i = q_i + \varepsilon \frac{\partial g}{\partial p_j} \left(\delta_{ij} + \varepsilon \frac{\partial^2 g}{\partial q_j \partial P_i} \right) = q_i + \varepsilon \frac{\partial g}{\partial p_j} \delta_{ij} + \mathcal{O}(\varepsilon)^2$$

Volviendo a la regla de transformación de más arriba, vemos que al orden lineal en ε se puede reemplazar por

$$P_i = \frac{\partial F_2}{\partial q_i} = p_i - \varepsilon \frac{\partial g}{\partial q_i} \quad Q_i = \frac{\partial F_2}{\partial P_i} = q_i + \varepsilon \frac{\partial g}{\partial p_i}$$

Con lo que hemos escrito la regla de transformación de las variables canónicas bajo una transformación canónica infinitesimal en términos de derivadas del generador $g(P_i, q_i, t)$.

Nos preguntamos ahora cómo cambia una función cualquiera sobre el espacio de estados F bajo una transformación canónica. Para ver esto, escribimos

$$F(q_i, p_i, t) = F\left(Q_i - \varepsilon \frac{\partial g}{\partial p_i}, P_i + \varepsilon \frac{\partial g}{\partial q_i}, t\right)$$

lo que se puede expandir al orden lineal en el parámetro infinitesimal ε en la forma

$$F(q_i, p_i, t) = F(Q_i, P_i, t) - \varepsilon \frac{\partial F}{\partial q_j} \frac{\partial g}{\partial p_i} + \varepsilon \frac{\partial F}{\partial p_i} \frac{\partial g}{\partial q_i} = F(Q_i, P_i, t) + \varepsilon \{g, F\}$$

o en otras palabras

$$F(Q_i, P_i, t) = F(q_i, p_i, t) + \varepsilon \{F, g\}$$

Esto implica que el cambio en cualquier función bajo una transformación canónica infinitesimal estará dado por su paréntesis de Poisson con el correspondiente generador.

Generadores:

En el capítulo 14, escribimos un cambio de coordenadas generalizadas infinitesimal en la forma

$$Q_i = s_i(q_j, t) = q_i + \varepsilon g_i(q_j, t)$$

donde llamamos a las funciones $g_i(q_j, t)$ los generadores de la transformación. En este punto, quisiéramos saber qué relación tienen con el generador de la correspondiente transformación canónica infinitesimal. De acuerdo a lo que discutimos más arriba para escribir un cambio de coordenadas generalizadas en la forma de una transformación canónica, tenemos que la función generatriz será

$$F_2(P_i, q_i, t) = s_i(q_j, t)P_i = q_i P_i + \varepsilon p_i g_i(q_j, t)$$

donde en la segunda igualdad reemplazamos $s_i(q_i, t)$ por su forma infinitesimal, y expandimos a orden lineal en ε . Esto nos permite concluir que $g(p_i, q_i, t) = p_i g_i(q_i, t)$.

La forma de la transformación infinitesimal de las variables canónicas en términos de derivadas del correspondiente generador es reminiscente de las ecuaciones de Hamilton. Esto nos invita a pensar si la evolución temporal será también una transformación canónica. Si escribimos el cambio en las variables canónicas cuando transcurre un tiempo ε a primer orden, tenemos

$$Q_i(t) = q_i(t + \varepsilon) = q_i(t) + \varepsilon \frac{\partial H}{\partial p_i} \quad P_i(t) = p_i(t + \varepsilon) = p_i(t) - \varepsilon \frac{\partial H}{\partial q_i}$$

Con lo que verificamos que la evolución temporal tiene la forma de una transformación canónica en la cual el rol del generador $g(p_i, q_i, t)$ lo cumple el hamiltoniano $H(p_i, q_i, t)$.

Lo anterior se puede interpretar también en el sentido opuesto: cualquier transformación canónica que tenga una forma infinitesimal se puede pensar como la evolución temporal con un hamiltoniano dado por el correspondiente generador g . Una implicación inmediata es que, dado que la evolución temporal preserva el volumen en el espacio de estados, lo mismo hace una transformación canónica cualquiera. Para probar esto de manera directa, escribimos el diferencial de volumen en el espacio de estados en la forma

$$\prod_i dQ_i dP_i = \prod_i \left| \frac{\partial(Q_j, P_j)}{\partial(q_k, p_k)} \right| dq_i dp_i = \prod_i \begin{vmatrix} \delta_{jk} + \varepsilon \frac{\partial^2 g}{\partial q_i \partial p_i} & -\varepsilon \frac{\partial^2 g}{\partial q_i^2} \\ \varepsilon \frac{\partial^2 g}{\partial p_i^2} & \delta_{jk} - \varepsilon \frac{\partial^2 g}{\partial q_i \partial p_i} \end{vmatrix} dq_i dp_i$$

Es fácil ver que a orden lineal en ε el determinante en la última expresión de la derecha es la unidad, con lo que el volumen en el espacio de estados se conserva.

17.2.2 Versión hamiltoniana del teorema de Noether

Las transformaciones canónicas permiten rederivar el teorema de Noether de una manera extremadamente simple e iluminadora, y generalizarlo al caso de simetrías que mezclan coordenadas e impulsos.

Supongamos que tenemos una transformación canónica cuyo generador C es una magnitud conservada, es decir que cumple $\dot{C} = 0$. Podemos utilizar la regla que obtuvimos antes para describir cómo cambiará el hamiltoniano del sistema bajo esta transformación

$$H(P_i, Q_i, t) = H(p_i, q_i, t) + \varepsilon \{H, C\} = H(p_i, q_i, t) - \varepsilon \left(\dot{C} - \frac{\partial C}{\partial t} \right) = K(P_i, Q_i, t)$$

En la última igualdad reemplazamos $\dot{C} = 0$ y utilizamos la definición que dimos más arriba para el hamiltoniano transformado K . Expandiendo a orden lineal en ε a ambos lados de esta igualdad, obtenemos

$$K(p_i, q_i, t) = H(p_i, q_i, t) + \mathcal{O}(\varepsilon)^2$$

Concluimos que el hamiltoniano es invariante a orden ε frente a la transformación. Esto a su vez resulta en las que las ecuaciones de Hamilton no cambian, lo que significa que la transformación es una simetría. Con esto hemos probado que cada vez que haya una magnitud conservada, esta actúa como el generador de una transformación canónica que es una simetría del sistema.

El resultado anterior también se puede leer a la inversa: dada una transformación canónica con generador g , debe cumplirse la identidad

$$K(P_i, Q_i, t) = H(p_i, q_i, t) + \varepsilon \frac{\partial g}{\partial t}$$

Aplicando al hamiltoniano K la regla de transformación que derivamos más arriba para cualquier función sobre el espacio de estados, esta fórmula puede reescribirse como

$$K(p_i, q_i, t) + \varepsilon \{K, g\} = H(p_i, q_i, t) + \varepsilon \frac{\partial g}{\partial t}$$

Lo que puede reordenarse en la forma

$$\begin{aligned} K(p_i, q_i, t) &= H(p_i, q_i, t) + \varepsilon \left(\frac{\partial g}{\partial t} + \{f, H\} \right) + \mathcal{O}(\varepsilon)^2 \\ &= H(p_i, q_i, t) + \varepsilon \dot{g} + \mathcal{O}(\varepsilon)^2 \end{aligned}$$

Pero si el hamiltoniano es invariante $K(p_i, q_i, t) = H(p_i, q_i, t)$, entonces la segunda línea implica inmediatamente que $\dot{g} = 0$, es decir que g es una magnitud conservada. Por lo tanto cada vez que haya una simetría infinitesimal, habrá una magnitud conservada que es el generador de dicha simetría.

Con lo de arriba, hemos entonces recuperado el teorema de Nonther de una manera bastante sencilla en el formalismo hamiltoniano, usando paréntesis de Poisson.

Es importante señalar que las transformaciones que hemos incluido en la deducción pueden mezclar arbitrariamente impulsos y coordenadas. Esto implica que lo que demostramos es en realidad una generalización de la versión lagrangiana del teorema de Noether.

Ejemplo: rotaciones

Supongamos un sistema invariante rotacional, con hamiltoniano

$$H = \frac{p_a^2}{2m} + V(r)$$

Este sistema conserva el momento angular

$$\vec{\ell} = \vec{r} \times \vec{p}$$

cuyas componentes se escriben $\ell_a = \varepsilon_{abc} x_b p_c$. Esto implica que se cumple la regla para los paréntesis de Poisson

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

Por otro lado, también deducimos de aquí que los ℓ_a generarán simetrías del sistema.

Para ver cuáles son estas simetrías, hacemos

$$\begin{aligned} x'_a &= x_a + \varepsilon \frac{\partial \ell_d}{\partial p_a} = x_a + \varepsilon \varepsilon_{dbc} x_b \delta_{ac} = x_a + \varepsilon \varepsilon_{dba} x_b \\ p'_a &= p_a - \varepsilon \frac{\partial \ell_d}{\partial x_a} = p_a - \varepsilon \varepsilon_{dbc} p_c \delta_{ab} = p_a + \varepsilon \varepsilon_{dca} p_c \end{aligned}$$

lo que en términos vectoriales se escribe como

$$\vec{r}' = \vec{r} + \varepsilon \vec{r} \times \vec{n}_d \quad \vec{p}' = \vec{p} + \varepsilon \vec{p} \times \vec{n}_d$$

donde \vec{n}_d es un vector unitario a lo largo del eje x_d . En otras palabras, acabamos de demostrar que las componentes del momento angular generan las rotaciones a lo largo de cada uno de los ejes.

Ejemplo: vector de Laplace-Runge-Lenz

En el problema de Kepler, existe un vector adicional que se conserva además del momento angular. Este es el vector de Laplace-Runge-Lenz, definido como

$$\vec{V} = \vec{p} \times \vec{\ell} - k\mu \vec{r} = \vec{p} \times (\vec{r} \times \vec{p}) - k\mu \vec{r} = p^2 \vec{r} - (\vec{p} \cdot \vec{r}) \vec{p} - k\mu \vec{r}$$

Donde μ es la masa reducida y k es la constante de acoplamiento del potencial de Kepler. En componentes esto se lee

$$V_a = p_b(p_b x_a - p_a x_b) - k\mu \frac{x_a}{r}$$

Dado que se conserva, debe ser el generador de alguna simetría. Para ver cómo actúa esta simetría sobre las variables canónicas, escribimos

$$\begin{aligned} x'_a &= x_a + \varepsilon \frac{\partial V_d}{\partial p_a} = x_a + \varepsilon (2p_a x_d - p_d x_a - p_b x_b \delta_{ad}) \\ p'_a &= p_a - \varepsilon \frac{\partial V_d}{\partial x_a} = p_a - \varepsilon \left(p_b p_b \delta_{ad} - p_a p_d - \frac{k\mu}{r} \left(\delta_{ad} - \frac{x_d x_a}{r^2} \right) \right) \end{aligned}$$

Aquí se ve claramente que es una simetría que mezcla coordenadas e impulsos, algo que nunca habríamos podido escribir en el formalismo lagrangiano.

Los paréntesis de Poisson proporcionan una ventaja adicional. Supongamos que conocemos dos magnitudes conservadas C_1 y C_2 . Estas magnitudes cumplen que

$$\{C_1, H\} + \frac{\partial C_1}{\partial t} = 0 \quad \{C_2, H\} + \frac{\partial C_2}{\partial t} = 0$$

Entonces usando la identidad de Jacobi, tenemos que

$$\{\{C_2, C_1\}, H\} + \{\{C_1, H\}, C_2\} + \{\{H, C_2\}, C_1\} = 0$$

donde podemos reemplazar las expresiones de arriba para deducir

$$\{\{C_2, C_1\}, H\} + \left\{ \frac{\partial C_1}{\partial t}, C_2 \right\} - \left\{ \frac{\partial C_2}{\partial t}, C_1 \right\} = \{\{C_2, C_1\}, H\} + \frac{\partial}{\partial t} \{C_2, C_1\} = \frac{d}{dt} \{C_2, C_1\} = 0$$

Esto implica que la nueva magnitud $C_3 = \{C_2, C_1\}$ también será conservada $\dot{C}_3 = 0$. Esta operación puede resultar en magnitudes conservadas nuevas, o a veces simplemente nos devuelve funciones de las magnitudes que ya teníamos $C_3 = f(C_1, C_2)$.

El conjunto de todas las magnitudes conservadas $\{C_n\}$ de un sistema físico es un subconjunto de su álgebra de observables que estudiamos en la clase previa. La propiedad que acabamos de probar implica que todas estas cantidades satisfacen la siguiente regla

$$\{C_n, C_m\} = f_{nm}(C_k)$$

En otras palabras, el subconjunto de magnitudes conservadas forma en sí mismo un álgebra, que es una subálgebra del álgebra de observables. Se denomina *álgebra de cargas* o *álgebra de simetrías*, ya que con las $\{C_n\}$ podremos generar todas las simetrías del sistema. Las funciones $f_{nm}(C_k)$ se denominan las *funciones de estructura* del álgebra de simetrías. En el caso bastante frecuente en el que tales funciones son lineales $\{C_n, C_m\} = f_{mnk} C_k$, las constantes f_{mnk} se llaman *constantes de estructura* del sistema.

Ejercicio:

Volviendo a las componentes del momento angular, de las que aprendimos que son los generadores de las rotaciones alrededor de cada eje, demuestre que se cumple que

$$\{\ell_a, \ell_b\} = \varepsilon_{abc} \ell_c$$

Con lo que las constantes de estructura están dadas por el símbolo de Levi-Civita ε_{abc} .

La regla para el paréntesis de Poisson de ℓ_a es reminiscente de los generadores matriciales t_i que mencionamos en el capítulo sobre rotaciones. Estos satisfacen el álgebra de conmutadores

$$[t_a, t_b] = \varepsilon_{abc} t_c$$

La cuantización canónica debe respetar la regla escrita más arriba, por lo que los operadores de momento angular $\hat{\ell}$ deben cumplir

$$[\hat{\ell}_a, \hat{\ell}_b] = \hbar i \widehat{\{\ell_a, \ell_b\}} = \hbar i \varepsilon_{abc} \hat{\ell}_c$$

Es por esta razón que el spin de un electrón se representa con matrices de Pauli, de modo de satisfacer esta regla de conmutación.

Ejercicio:

Usando las propiedades de los paréntesis de Poisson, pruebe que el vector de Laplace-Runge-Lenz cumple con la siguiente álgebra

$$\{V_a, V_b\} = -2\mu H \varepsilon_{abc} \ell_c$$

donde H es el hamiltoniano del problema de Kepler. Pruebe que se cumple además que

$$\{\ell_a, V_b\} = \varepsilon_{abc} V_c$$

Lo que implica que \vec{V} transforma como un vector frente a rotaciones.

17.3 Resumen

Estudiamos en esta clase las transformaciones canónicas, que están dadas por aquéllos cambios de variables en el espacio de estados que preservan los paréntesis de Poisson fundamentales. Aprendimos que hay cuatro tipos de funciones que definen transformaciones canónicas, que se relacionan entre sí mediante transformaciones de Legendre.

Estudiamos también las transformaciones canónicas infinitesimales, y las usamos para demostrar que la evolución temporal es de hecho una transformación canónica. Más aún, entendimos que cualquier transformación canónica puede interpretarse como una evolución temporal con

un hamiltoniano dado por el generador de la transformación, y dedujimos que debe entonces preservar el volumen del espacio de estados.

Finalmente, generalizamos el teorema de Noether al caso hamiltoniano, incluyendo simetrías que mezclan coordenadas e impulsos. Vimos que cada vez que una transformación canónica es una simetría, su generador se conserva y recíprocamente, cada vez que hay una cantidad conservada, esta genera una simetría.

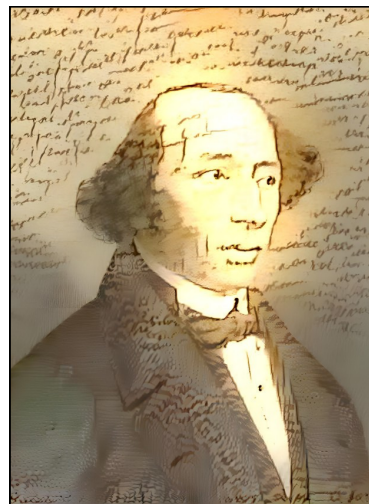
18. Ecuación de Hamilton-Jacobi

18.1 Objetivos

En esta clase, vamos a usar las transformaciones canónicas que estudiamos la clase pasada para reformular completamente la descripción del movimiento de un sistema mecánico en el espacio de estados.

Hasta ahora hemos usado ecuaciones diferenciales ordinarias. En el formalismo lagrangiano, usábamos ecuaciones diferenciales de segundo orden, que nos decían como evolucionaba el sistema en el espacio de configuración. En el formalismo hamiltoniano en cambio, utilizamos ecuaciones de primer orden para estudiar el movimiento en el espacio de estados. En ambos casos, el número de ecuaciones era igual al número de variables.

En esta clase, vamos a construir una ecuación diferencial en derivadas parciales para una sola magnitud, que contiene toda la información sobre la evolución temporal del sistema, y estudiaremos sus propiedades y algunas técnicas para su resolución.



Carl Gustav Jacob Jacobi

18.2 La ecuación de Hamilton Jacobi

Dado un sistema mecánico descrito por variables canónicas $\{(p_i, q_i)\}$, aprendimos en la clase anterior que podemos hacer transformaciones canónicas para describirlo con nuevas variables $\{(P_i, Q_i)\}$. Concentrémonos en particular en aquellas cuya función generatriz es del tipo $F_1(Q_i, q_i, t)$.

Para tales transformaciones, sabemos que las reglas de transformación se escriben

$$K(P_i, Q_i, t) = H(p_i, q_i, t) + \frac{\partial F_1}{\partial t}$$

$$p_i = \frac{\partial F_1}{\partial q_i} \quad P_i = \frac{\partial F_1}{\partial Q_i}$$

Una pregunta posible es si podemos encontrar una transformación canónica que simplifique la solución de las ecuaciones de Hamilton resultantes. Por ejemplo, si fuéramos capaces de construir una función F_1 particular tal que el nuevo hamiltoniano se anule

$$K(P_i, Q_i, t) = 0$$

entonces las nuevas ecuaciones de Hamilton resultarían triviales

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0 \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0$$

y se podrían resolver muy simplemente como $Q_i = \alpha_i$ y $P_i = \beta_i$, en términos de constantes de integración α_i, β_i .

Supongamos que la función generatriz que cumple esta propiedad es $F_1 = S(q_i, Q_i, t)$. La condición de que se anule el nuevo hamiltoniano puede entonces escribirse como

$$H\left(\frac{\partial S}{\partial q_i}, q_i, t\right) + \frac{\partial S}{\partial t} = 0$$

donde reemplazamos los impulsos utilizando la regla de transformación que escribimos más arriba. Hemos obtenido una ecuación diferencial en derivadas parciales de primer orden que se conoce como *ecuación de Hamilton-Jacobi*. Esta ecuación determina la *primera función principal de Hamilton* $S(q_i, Q_i, t)$ la cual, como veremos, contiene la solución completa del movimiento del sistema.

Ejemplo: partícula no-relativista en tres dimensiones

Dada una partícula no relativista que se mueve en el espacio tridimensional con coordenadas x_a , tenemos que su hamiltoniano se escribe como

$$H(x_a, p_a) = \frac{p_a^2}{2m} + V(x_a)$$

con lo que la ecuación de Hamilton-Jacobi se puede escribir inmediatamente

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x_a} \right)^2 + V(x_a) + \frac{\partial S}{\partial t} = 0$$

Como vemos, es una ecuación no lineal, por lo que en principio puede resultar nada trivial resolverla.

Ejercicio:

Escriba la ecuación de Hamilton-Jacobi para una partícula libre relativista que se mueve en tres dimensiones.

La ecuación de Hamilton-Jacobi para un sistema con D coordenadas $\{q_i\}$ tiene $D + 1$ derivadas, con lo que su solución se escribirá en términos de $D + 1$ constantes de integración α_i, α . Nótese que, dado que en la ecuación sólo aparecen derivadas de S , una de las constantes de integración es aditiva. En efecto, si $S(q_i, \alpha_i, t)$ es solución de la ecuación, entonces $S(q_i, \alpha_i, t) + \alpha$ también lo será. En cuanto a las demás constantes de integración, sin pérdida de generalidad las podemos identificar con los valores constantes de las nuevas coordenadas $Q_i = \alpha_i$.

Dada una solución de la ecuación, podemos escribir la transformación canónica en la forma

$$p_i = \frac{\partial S}{\partial q_i}(q_i, \alpha_i, t) \quad P_i = \frac{\partial S}{\partial \alpha_i}(q_i, \alpha_i, t) = \beta_i$$

de donde se puede invertir para obtener una solución completa del problema

$$q_i = q_i(\alpha_i, \beta_i, t) \quad p_i = p_i(\alpha_i, \beta_i, t)$$

Es decir que obtuvimos una solución para el movimiento del sistema, en términos de las constantes de integración α_i, β_i , a partir de la primera función principal de Hamilton $S(q_i, \alpha_i, t)$ que es solución de la ecuación de Hamilton-Jacobi.

Ejemplo: relación con la mecánica cuántica

Es interesante notar que la ecuación de Hamilton-Jacobi para una partícula no relativista que se mueve en tres dimensiones se parece mucho a la ecuación del iconal que obtuvimos en la sección 13.3, cuando discutíamos el principio de acción estacionaria. Podemos hacer más precisa esta correspondencia escribiendo la ecuación de onda de la mecánica cuántica, es decir la ecuación de Schrödinger

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x_a^2} + V(x_a) \Psi = i\hbar \frac{\partial \Psi}{\partial t}$$

en la cual, y sin pérdida alguna de generalidad, podemos insertar una función de onda de la forma

$$\Psi(x_a, t) = e^{\frac{i}{\hbar} S(x_a, t)}$$

para obtener

$$-\frac{1}{2m} \left(i\hbar \frac{\partial^2 S}{\partial x_a^2} - \left(\frac{\partial S}{\partial x_a} \right)^2 \right) + V(x_a) + \frac{\partial S}{\partial t} = 0$$

Ahora bien, como sucedía en el caso del iconal antes discutido, cuando la longitud de onda es mucho menor que la escala característica del problema (que aquí viene dada por la distancia típica en la que el potencial experimenta cambios apreciables) podemos descartar el término que contiene el laplaciano. Con esto tenemos

$$\frac{1}{2m} \left(\frac{\partial S}{\partial x_a} \right)^2 + V(x_a) + \frac{\partial S}{\partial t} = 0$$

que no es otra cosa que la ecuación de Hamilton-Jacobi para una partícula no relativista que se mueve en tres dimensiones.

En otras palabras, la ecuación de Hamilton-Jacobi es la ecuación del iconal de la mecánica cuántica, lo que implica que la mecánica clásica es su límite de óptica geométrica.

Ahora bien, en la discusión sobre el principio de acción estacionaria, el análogo del iconal en el sistema mecánico era su acción ¿Significa esto que la primera función principal de Hamilton es la acción del problema mecánico?

Para tener una intuición sobre el significado físico de la primera función principal de Hamilton, escribamos su derivada temporal en la forma

$$\dot{S} = \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t}$$

El segundo término de esta expresión se puede obtener de la ecuación de Hamilton-Jacobi como menos el hamiltoniano. Esto resulta en la expresión

$$\dot{S} = \frac{\partial S}{\partial q_i} \dot{q}_i - H\left(\frac{\partial S}{\partial q_i}, q_i, t\right) = p_i \dot{q}_i - H(p_i, q_i, t)$$

La segunda igualdad pone de manifiesto que el valor numérico de esta expresión coincide con el lagrangiano del sistema. Sin embargo, los impulsos son aquí funciones de las coordenadas dadas por las derivadas de S , por lo que no se trata de la función lagrangiano propiamente dicha. Integrando esta expresión en el tiempo, entre un instante inicial cualquiera donde las coordenadas toman valores arbitrarios, hasta un instante final t donde las coordenadas valen q_i , vemos que la función principal de Hamilton coincide con la acción del problema, considerada como una función del punto final de la trayectoria. El punto inicial no resulta relevante, ya que su contribución se puede reabsorber en la constante aditiva α .

18.3 Separación de variables

Una técnica útil para encontrar soluciones de la ecuación de Hamilton-Jacobi es el método conocido como *separación de variables*. Se trata de proponer una forma tentativa o *Ansatz* para la solución, en términos de un conjunto de funciones desconocidas que dependen de variables diferentes. Si la forma es la correcta, al insertarla en la ecuación de Hamilton-Jacobi ésta se separa en un conjunto de ecuaciones distintas para cada función desconocida, donde cada ecuación depende de un conjunto diferente de variables.

Por ejemplo, supongamos que el hamiltoniano no depende del tiempo. En ese caso la ecuación de Hamilton-Jacobi toma la forma

$$H\left(\frac{\partial S}{\partial q_i}, q_i\right) + \frac{\partial S}{\partial t} = 0$$

Si proponemos una solución tentativa para la separación de variables en términos de dos funciones $W(q_i)$ y $s(t)$ con la siguiente forma aditiva

$$S(q_i, t) = W(q_i) + s(t)$$

la ecuación se reescribe como

$$H\left(\frac{\partial W}{\partial q_i}, q_i\right) + \frac{\partial s}{\partial t} = 0$$

En esta ecuación, el primer término depende exclusivamente de las coordenadas y no del tiempo, mientras que el segundo depende del tiempo y no de las coordenadas. La única forma que puedan cancelarse es si ambos son constantes. Entonces podemos escribir

$$\frac{\partial s}{\partial t} = -E$$

$$H\left(\frac{\partial W}{\partial q_i}, q_i\right) = E$$

en términos de una *constante de separación* E . Vemos aquí que la primera ecuación sólo involucra a la función $s(t)$ y depende exclusivamente del tiempo, con lo que se resuelve inmediatamente según $s(t) = -Et + \alpha$, donde α es una constante de integración. Por otro lado la segunda ecuación se denomina *ecuación de Hamilton-Jacobi independiente del tiempo* y depende exclusivamente de las coordenadas. Determina la función $W(q_i)$ conocida como *función asociada de Hamilton*.

Dependiendo de la forma funcional del hamiltoniano, puede ser posible continuar con el proceso de separación. En efecto, supongamos que el hamiltoniano puede escribirse como

$$H(p_i, q_i) = H(p_j, q_j, h(p_k, q_k))$$

donde las variables canónicas $\{(p_i, q_i)\}$ con $i \in \{1, \dots, D\}$ entran separadas en dos subconjuntos diferentes $\{(p_k, q_k)\}$ con $k \in \{1, \dots, D'\}$ y $\{(p_j, q_j)\}$ con $j \in \{D' + 1, \dots, D\}$, el primero de ellos solamente dentro de una función $h(p_k, q_k)$. En tal caso, la ecuación de Hamilton-Jacobi independiente del tiempo se escribe

$$H\left(\frac{\partial W}{\partial q_j}, q_j, h\left(\frac{\partial W}{\partial q_k}, q_k\right)\right) = E$$

Proponemos una solución tentativa para la separación de variables con la forma aditiva

$$W(q_i) = W_1(q_j) + W_2(q_k)$$

donde cada uno de los términos depende de un subconjunto diferente de las variables. Insertando esta expresión en la ecuación, ésta toma la forma

$$H\left(\frac{\partial W_1}{\partial q_j}, q_j, h\left(\frac{\partial W_2}{\partial q_k}, q_k\right)\right) = E$$

Si ahora calculamos la derivada respecto de q_k obtenemos

$$\frac{\partial H}{\partial h} \frac{\partial h}{\partial q_k} = 0$$

Dado que por hipótesis $\partial H / \partial h \neq 0$, esta condición implica que $\partial h / \partial q_k = 0$, es decir que h debe tomar un valor constante

$$h\left(\frac{\partial W_2}{\partial q_k}, q_k\right) = \gamma$$

donde γ es la constante de separación. Por otro lado, también se debe satisfacer la ecuación original, que ahora se lee

$$H\left(\frac{\partial W_2}{\partial q_j}, q_j, \gamma\right) = E$$

En estas dos últimas ecuaciones, la primera determina completamente $W_2(q_k)$ y depende de las coordenadas $\{q_k\}$, mientras que la segunda ecuación determina $W_1(q_j)$ en términos de las $\{q_j\}$. Es decir que hemos separado la ecuación de Hamilton-Jacobi en dos ecuaciones que dependen de un número menor de variables.

En los casos en los que el proceso arriba descrito se pueda repetir un número suficiente de veces, podemos resolver analíticamente la ecuación de Hamilton-Jacobi para obtener la función principal de Hamilton. Si en algún momento el proceso se estanca porque la dependencia del hamiltoniano mezcla las variables de manera no separable, siempre se puede recurrir a métodos numéricos para obtener una solución.

Ejemplo:

Dado el hamiltoniano

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + V(q_2)$$

la ecuación de Hamilton-Jacobi se lee

$$\frac{1}{2m} \left(\frac{\partial S}{\partial q_1} \right)^2 + \frac{1}{2m} \left(\frac{\partial S}{\partial q_2} \right)^2 + V(q_2) + \frac{\partial S}{\partial t} = 0$$

Con lo que podemos separar el tiempo, usando

$$S(q_1, q_2, t) = W(q_1, q_2) + s(t)$$

Esto nos da una ecuación de Hamilton-Jacobi independiente del tiempo, con la forma

$$\frac{1}{2m} \left(\frac{\partial W}{\partial q_1} \right)^2 + \frac{1}{2m} \left(\frac{\partial W}{\partial q_2} \right)^2 + V(q_2) = E$$

mientras que la solución para la función $s(t)$ es, como explicamos $s(t) = -Et + \alpha$. En la ecuación independiente del tiempo podemos identificar un sector que depende exclusivamente de la coordenada q_2

$$\frac{1}{2m} \left(\frac{\partial W}{\partial q_1} \right)^2 + \underbrace{\frac{1}{2m} \left(\frac{\partial W}{\partial q_2} \right)^2 + V(q_2)}_{h(p_2, q_2)} = E$$

lo que nos permite utilizar la solución tentativa para la separación de variables

$$W(q_1, q_2) = W_1(q_1) + W_2(q_2)$$

Reemplazando en la ecuación, obtenemos

$$\frac{1}{2m} \left(\frac{\partial W_1}{\partial q_1} \right)^2 + \frac{1}{2m} \left(\frac{\partial W_2}{\partial q_2} \right)^2 + V(q_2) = E$$

Esto solo se puede satisfacer si se cumple independientemente que

$$\frac{1}{2m} \left(\frac{\partial W_1}{\partial q_1} \right)^2 + \gamma = E \quad \frac{1}{2m} \left(\frac{\partial W_2}{\partial q_2} \right)^2 + V(q_2) = \gamma$$

La ecuación para $W_1(q_1)$ se puede resolver inmediatamente como

$$W_1(q_1) = -\sqrt{2m(E - \gamma)} q_1 + \gamma_1$$

donde γ_1 es una constante de integración, y hemos hecho una elección particular del signo en la raíz cuadrada. Por otro lado, la ecuación para W_2 se puede reescribir en la forma

$$\frac{\partial W_2}{\partial q_2} = -\sqrt{2m(\gamma - V(q_2))}$$

donde nuevamente hemos hecho una elección para el signo de la raíz cuadrada. De aquí se obtiene la solución

$$W_2 = -\int dq_2 \sqrt{2m(\gamma - V(q_2))} + \gamma_2$$

con una nueva constante de integración γ_2 . Con esto, la solución general para la función principal de Hamilton toma la forma

$$S = -\sqrt{2m(E - \gamma)} q_1 - \int dq_2 \sqrt{2m(\gamma - V(q_2))} + \alpha + Et$$

donde hemos agrupado las constantes de integración aditivas en una nueva única constante $\alpha = \gamma_1 + \gamma_2$. La solución para las variables canónicas se escribe

$$p_1 = \frac{\partial S}{\partial q_1} = -\sqrt{2m(E - \gamma)} \quad p_2 = \frac{\partial S}{\partial q_2} = -\sqrt{2m(\gamma - V(q_2))}$$

$$\beta_1 = \frac{\partial S}{\partial E} = -q_1 \sqrt{\frac{m}{2(E - \gamma)}} + t$$

$$\beta_2 = \frac{\partial S}{\partial \gamma} = q_1 \sqrt{\frac{m}{2(E - \gamma)}} - \int dq_2 \sqrt{\frac{m}{2(\gamma - V(q_2))}}$$

Lo que se puede invertir parcialmente en la forma

$$p_1 = mv_0 \quad p_2 = -\sqrt{2m(\gamma - V(q_2))}$$

$$q_1 = v_1(t - t_0) + q_0 \quad t - t_0 = \int dq_2 \sqrt{\frac{2m}{\gamma - V(q_2)}}$$

donde hemos identificado las constantes de integración como $\beta_1 = t_0$, $\beta_2 = t_0 - q_0/v_0$, $v_0 = -\sqrt{2(E - \gamma)/m}$, siendo q_0 y v_0 los valores de la coordenada q_1 y su derivada temporal \dot{q}_1 en el instante inicial $t = t_0$. Esto por supuesto coincide con la solución que hubiéramos obtenido resolviendo simplemente las ecuaciones de Hamilton.

Nótese que elecciones distintas para los signos para las raíces cuadradas hubieran resultado en soluciones que difieren en los signos de sus velocidades iniciales. Esta aparición de diferentes *ramas* de soluciones es característica de las ecuaciones diferenciales no lineales.

18.4 Resumen

En esta clase, encontramos una descripción del movimiento en términos de una ecuación diferencial en derivadas parciales para una función sobre el espacio de configuración. De este modo, la Mecánica se parece al electromagnetismo o a la mecánica de fluidos, en el sentido de que sus soluciones se describen como excitaciones propagándose en algún tipo de sustrato que llena todo el espacio de configuración \mathcal{C} .

Como sucedía con el principio de acción estacionaria, la ecuación de Hamilton-Jacobi nos permite hacer contacto con el origen cuántico de las leyes de la Mecánica Analítica, interpretándolas como límite de óptica geométrica de la ecuación de Schrödinger.

19. Vínculos hamiltonianos

19.1 Objetivos

En el capítulo 11 aprendimos a tratar vínculos en la formulación lagrangiana, mediante el uso de coordenadas adaptadas o bien con multiplicadores de Lagrange. El primer método reduce el número de coordenadas generalizadas, mientras que el segundo agrega nuevas coordenadas para representar las fuerzas de vínculo.

En cualquier caso, aún cuando todos los vínculos del problema han sido incorporados a su lagrangiano, todavía pueden aparecer vínculos adicionales al pasar a la formulación hamiltoniana.

En efecto, al realizar la transformación de Legendre del lagrangiano para obtener el hamiltoniano, puede darse el caso de que no sea posible invertir algunos de los momentos generalizados para obtener las correspondientes velocidades como función de los impulsos. Esto establece vínculos entre impulsos y coordenadas, que debemos incorporar de algún modo a la formulación.



Paul Adrien Maurice Dirac

19.2 Vínculos hamiltonianos

Supongamos un sistema mecánico en cuya formulación lagrangiana ya hemos incorporado todos los vínculos holonómicos en términos de coordenadas adaptadas, y todos los no holonómicos utilizando multiplicadores de Lagrange. El espacio de configuración resultante \mathcal{C} estará parametri-

zados por coordenadas generalizadas $\{q_i\}$ con $i \in \{1, \dots, D\}$ (entre las cuales están incluidos los eventuales multiplicadores de Lagrange que hayamos introducido), y la dinámica estará descrita por un lagrangiano $L(\dot{q}_i, q_i, t)$.

El primer paso para obtener la formulación hamiltoniana del sistema consiste en calcular sus momentos generalizados según la fórmula

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \equiv p_i(\dot{q}_j, q_j, t)$$

A continuación deberíamos invertir esta expresión para obtener las velocidades generalizadas como función de los impulsos y las coordenadas canónicas

$$\dot{q}_i = v_i(p_i, q_i, t)$$

para con esto escribir el hamiltoniano como la transformada de Legendre del lagrangiano

$$H(p_i, q_i, t) = p_i v_i(p_i, q_i, t) - L(v_i(p_i, q_i, t), q_i, t)$$

Sin embargo, podría darse el caso que no fuera posible invertir las ecuaciones que definen los impulsos canónicos para obtener la totalidad de las velocidades generalizadas $\{\dot{q}_i\}$ con $i \in \{1, \dots, D\}$. En el caso general, sólo podremos invertir para un subconjunto de velocidades, digamos D_1 de ellas, y restarán un cierto número de velocidades que no se podrán despejar en función de los impulsos. Es decir que las ecuaciones

$$p_l - p_l(\dot{q}_i, q_i, t) \equiv \phi_l(p_i, q_i, t) = 0 \quad \text{para } l \in \{D_1 + 1, \dots, D\}$$

no se podrán invertir, dando origen a un conjunto de vínculos entre las variables canónicas

$$\phi_l(p_i, q_i, t) = 0 \quad \text{para } l \in \{D_1 + 1, \dots, D\}$$

Estos vínculos que aparecen como un obstáculo al intentar formular el formalismo hamiltoniano se conocen como vínculos *primarios*.

Ejemplo: efecto Hall

Supongamos que tenemos una partícula cargada sometida a un campo magnético uniforme de valor B en la dirección \hat{k} . Como hemos demostrado antes, el lagrangiano que describe esta dinámica se lee

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e(\Phi(x, y, z) - Bx\dot{y})$$

nótese que dado que la velocidad en x es finita, siempre es posible poner un campo magnético B tal que $m\dot{y}/2e \ll Bx$ con lo cual

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) + eBx\dot{y} + e\Phi(x, y, z)$$

En este lagrangiano, el momento generalizado en la dirección y está dado por

$$p_y = eBx$$

con lo que tenemos la ecuación

$$p_y - eBx = \phi(p_y, x) = 0$$

Pero aquí no entra la velocidad \dot{y} por lo que es imposible escribir $v_y(p_y)$ y tenemos un vínculo primario.

Ejemplo: lagrangiano lineal en las velocidades

Si tuviéramos un lagrangiano lineal en las velocidades generalizadas

$$L = \dot{q}_i g_i(q_j, t)$$

donde $g_i(q_j, t)$ son funciones arbitrarias de las coordenadas y el tiempo, entonces tendríamos

$$p_i = g_i(q_j, t)$$

Por lo tanto en la ecuación

$$p_i - g_i(q_j, t) = \phi_i(p_j, q_j, t) = 0$$

no se puede despejar \dot{q}_i como función de las variables canónicas (q_i, p_i) . Esto implica que tendremos un vínculo primario por cada coordenada q_i .

Si llamamos \mathcal{E}_{can} al espacio parametrizado por las variables canónicas $\{(p_i, q_i)\}$, vemos que en presencia de vínculos hamiltonianos este espacio no es accesible en su totalidad. Esto sucede porque el movimiento tiene lugar en la hipersuperficie $\mathcal{E}_{\text{vinc}} \subset \mathcal{E}_{\text{can}}$ determinada por los vínculos. Como veremos más adelante, el espacio de estados físicos del sistema estará contenido dentro de tal hipersuperficie $\mathcal{E} \subset \mathcal{E}_{\text{vinc}}$.

19.2.1 Vínculos primarios y hamiltoniano total

Supongamos que escribimos un hamiltoniano tentativo para describir la evolución temporal de un sistema con vínculos primarios como los que acabamos de definir, en la forma

$$H = p_i \dot{q}_i - L(\dot{q}_i, q_i, t)$$

donde no nos hemos preocupado por invertir explícitamente las velocidades en función de los impulsos, es decir que es por ahora una función de $\{q_i, \dot{q}_i, p_i\}$ y eventualmente el tiempo. Calculando la diferencial de esta función tendremos que

$$dH = p_i d\dot{q}_i + dp_i \dot{q}_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial q_i} dq_i$$

aquí podemos reemplazar las derivadas del lagrangiano, usando la definición de los momentos generalizados en el tercer término y las ecuaciones de Lagrange en el último, para obtener

$$dH = (p_i - \dot{p}_i(\dot{q}_j, q_j, t)) d\dot{q}_i + dp_i \dot{q}_i - \dot{p}_i(\dot{q}_j, q_j, t) dq_i$$

El primer término se anula dada la definición de los impulsos canónicos, incluso cuando alguno de los momentos $p_i(\dot{q}_j, q_j, t)$ no sea invertible. Esto implica que el hamiltoniano no es una función de las velocidades, incluso cuando hayamos fallado en resolverlas en términos de los impulsos. Tomando las derivadas del hamiltoniano para reescribir la diferencial del lado izquierdo, tenemos

$$\frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i = dp_i \dot{q}_i - \dot{p}_i dq_i$$

o bien, reordenando

$$\left(\frac{\partial H}{\partial p_i} - \dot{q}_i \right) dp_i + \left(\frac{\partial H}{\partial q_i} + \dot{p}_i \right) dq_i = 0$$

A pesar de que esta ecuación es sugerente, no podemos deducir de aquí que se cumplen las ecuaciones de Hamilton. En efecto, dado que las variables canónicas deben satisfacer los vínculos $\phi_l(p_i, q_i, t) = 0$ con $l \in \{D_1 + 1, \dots, D\}$, tenemos que sus diferenciales $\{dp_i, dq_i\}$ deben cumplir la restricción

$$d\phi_l = \frac{\partial \phi_l}{\partial q_i} dq_i + \frac{\partial \phi_l}{\partial p_i} dp_i = 0 \quad \text{con } l \in \{D_1 + 1, \dots, D\}$$

Esto nos impide poner $dp_i = 0$ o bien $dq_i = 0$ independientemente en la ecuación de más arriba, para obtener de allí las ecuaciones de Hamilton. Esta restricción sólo deja D_1 coordenadas e impulsos canónicos cuyas diferenciales son independientes, por lo que solamente podríamos extraer D_1 ecuaciones de Hamilton.

Para resolver el problema, multiplicamos la última ecuación por una función arbitraria u_l de las variables canónicas y el tiempo

$$u_l d\phi_l = u_l \frac{\partial \phi_l}{\partial q_i} dq_i + u_l \frac{\partial \phi_l}{\partial p_i} dp_i = 0$$

Ahora podemos sumar esta expresión a la diferencial de más arriba, para escribir

$$\left(\frac{\partial H}{\partial p_i} - \dot{q}_i + u_l \frac{\partial \phi_l}{\partial p_i} \right) dp_i + \left(\frac{\partial H}{\partial q_i} + \dot{p}_i + u_l \frac{\partial \phi_l}{\partial q_i} \right) dq_i = 0$$

Aquí tenemos D_1 coordenadas e impulsos canónicos para los cuales podemos escribir $dp_k = 0$ o $dq_k = 0$ independientemente, con $k \in \{1, \dots, D_1\}$. Por otro lado, tenemos $D - D_1$ funciones arbitrarias u_l que podemos ajustar de modo de anular los coeficientes de los diferenciales que restan. Esto implica que la relación de más arriba se puede resolver, con lo que obtenemos las ecuaciones de movimiento

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} + u_l \frac{\partial \phi_l}{\partial p_i} & \dot{p}_i &= -\frac{\partial H}{\partial q_i} - u_l \frac{\partial \phi_l}{\partial q_i} \\ \phi_l(p_i, q_i, t) &= 0 \end{aligned}$$

Donde en la segunda línea hemos agregado los vínculos primarios, que deben resolverse junto con las ecuaciones. Es importante notar que en este punto todavía tenemos D_1 funciones desconocidas u_l , por lo que el problema no está aún bien definido.

Podemos reescribir las ecuaciones en la primera línea usando paréntesis de Poisson, para obtener

$$\dot{q}_i = \{q_i, H\} + u_l \{q_i, \phi_l\} \quad \dot{p}_i = \{p_i, H\} + u_l \{p_i, \phi_l\}$$

$$\phi_l(p_i, q_i, t) = 0$$

Lo que se puede reacomodar usando la regla de Leibniz y la bilinealidad de los paréntesis de Poisson, según

$$\dot{q}_i = \{q_i, H + u_l \phi_l\} - \{q_i, u_l\} \phi_l \quad \dot{p}_i = \{p_i, H + u_l \phi_l\} - \{p_i, u_l\} \phi_l$$

$$\phi_l(p_i, q_i, t) = 0$$

El segundo término en las ecuaciones de la primera línea se anula al usar los vínculos, cumpliendo una propiedad a veces se denomina *nulidad débil*. El conjunto de ecuaciones de movimiento puede entonces escribirse como paréntesis de Poisson con una nueva magnitud, en la forma

$$\dot{q}_i = \{q_i, H_{\text{tot}}\} \quad \dot{p}_i = \{p_i, H_{\text{tot}}\}$$

$$\phi_l(p_i, q_i, t) = 0$$

donde hemos definido un *hamiltoniano total* según la fórmula

$$H_{\text{tot}}(p_i, q_i, t) = H(p_i, q_i, t) + u_l \phi_l(p_i, q_i, t)$$

Con estas ecuaciones, podemos calcular la derivada total de cualquier función F de las variables canónicas

$$\dot{F} = \frac{\partial F}{\partial p_i} \dot{p}_i + \frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial t} = \frac{\partial F}{\partial p_i} \{p_i, H_{\text{tot}}\} + \frac{\partial F}{\partial q_i} \{q_i, H_{\text{tot}}\} + \frac{\partial F}{\partial t} = \{F, H_{\text{tot}}\} \frac{\partial F}{\partial t}$$

En conclusión, la evolución temporal puede ahora escribirse en términos de un hamiltoniano total H_{tot} que involucra $D - D_1$ funciones desconocidas. Como veremos más adelante, estas funciones no son completamente arbitrarias, sino que deben satisfacer requisitos de consistencia.

Nótese que en ningún punto en los cálculos de arriba hemos impuesto los vínculos dentro de los paréntesis de Poisson. Esto se debe a que los mismos no involucran el valor de la función $\phi_l(q_i, p_i, t)$ sino de sus derivadas.

19.2.2 Vínculos secundarios y hamiltoniano extendido

El procedimiento de la sección anterior permitió extraer las ecuaciones de movimiento para las variables canónicas, a partir de un hamiltoniano total que contiene funciones desconocidas. Sin embargo, esta construcción no garantiza que, dada una condición inicial que satisface los vínculos, la evolución temporal mantenga tal restricción. Necesitamos asegurar las condiciones para que la solución de las ecuaciones de Hamilton respete los vínculos a medida que transcurre el tiempo. En otras palabras, tiene que cumplirse que la derivada temporal de los vínculos se anule

$$\dot{\phi}_l = \{\phi_l, H_{\text{tot}}\} = 0$$

Usando la forma explícita del hamiltoniano total, esto se puede escribir como una ecuación de la forma

$$\{\phi_l, H\} + u_k \{\phi_l, \phi_k\} = 0$$

Cuando esta restricción se aplica a los diferentes valores de $l \in \{D_1 + 1, \dots, D\}$, puede dar origen a cuatro tipos de situaciones

- **Inconsistencias:** es decir ecuaciones que no pueden satisfacerse, del estilo $1 = 0$. Esto significa que el lagrangiano original es inconsistente. Por ejemplo, si definimos un sistema unidimensional cuyo lagrangiano tiene la forma $L = q$, sus ecuaciones de Lagrange resultan en $1 = 0$ y son inconsistentes. Esto significa que no podemos empezar con cualquier lagrangiano, sino solo con aquéllos que den origen a ecuaciones de Lagrange consistentes. No se trata en realidad de ninguna restricción novedosa, ya que la venimos aplicando tácitamente desde el comienzo del curso.
- **Consistencias:** es decir ecuaciones que se satisfacen automáticamente, de la forma $0 = 0$. Esto sucede cuando ambos términos de la ecuación se anulan independientemente $\{\phi_l, H\} = 0$ y $\forall k : \{\phi_l, \phi_k\} = 0$. Es decir que las ecuaciones de Hamilton calculadas con el hamiltoniano total respetan automáticamente el vínculo $\phi_l(p_i, q_i, t) = 0$ y lo mantienen a lo largo del tiempo.
- **Nuevos vínculos:** se trata de relaciones que no involucran a los u_l , sino solamente a las variables canónicas. Aparecen cuando el primer término de la ecuación no se anula $\{\phi_l, H\} \neq 0$, pero el segundo sí $\forall k : \{\phi_l, \phi_k\} = 0$. Podemos escribirlas en una forma similar a la de los vínculos primarios

$$\phi_l(p_i, q_i, t) = 0 \quad \text{donde ahora } l \in \{D_2, \dots, D_1\}$$

Debemos asegurar que estos *vínculos secundarios* también se satisfagan, y que se mantengan a lo largo del movimiento. Esto implica que sus derivadas temporales también deben anularse, lo que a su vez puede dar origen a nuevos vínculos secundarios, y así sucesivamente hasta agotar las relaciones de consistencia.

- **Restricciones sobre las u_l :** finalmente, para algunos de los vínculos ϕ_l obtendremos condiciones que involucran a las variables u_l , de la forma

$$\{\phi_l, H\} + u_k \{\phi_l, \phi_k\} = 0$$

Aquí el índice k recorre los valores $k \in \{D_1 + 1, \dots, D\}$, mientras que el índice l corre sobre sólo algunos valores dentro del conjunto $\{D_2 + 1, \dots, D\}$, aquéllos que tengan $\{\phi_l, \phi_k\} \neq 0$ para algún valor de k . Esto establece un conjunto de relaciones lineales entre los u_k , que puede resolverse como

$$u_l = U_l(q_i, p_i, t) + v_a V_{al}(q_i, p_i, t)$$

Donde $U_l(q_i, p_i)$ son soluciones particulares de la ecuaciones no-homogéneas, mientras que $V_{al}(q_i, p_i)$ son soluciones de las ecuaciones homogéneas, multiplicados por coeficientes arbitrarios v_a que son funciones del tiempo.

Con esto, el hamiltoniano total puede reescribirse como

$$\begin{aligned} H_{\text{tot}}(q_i, p_i, t) &= \underbrace{H(q_i, p_i, t) + U_l(q_i, p_i, t)\phi_l(q_i, p_i)}_{H'(q_i, p_i, t)} + v_a \underbrace{V_{al}(q_i, p_i, t)\phi_l(q_i, p_i, t)}_{\phi_a(p_i, q_i, t)} \\ &= H'(q_i, p_i, t) + v_a \phi_a(q_i, p_i, t) \end{aligned}$$

Donde tanto $H'(q_i, p_i, t)$ como $\phi_a(q_i, p_i, t)$ son funciones completamente determinadas de las variables canónicas $\{q_i, p_i\}$ y el tiempo, pero aún tenemos funciones indeterminadas del tiempo v_a . Estas funciones afectarán la evolución temporal, dado que las ecuaciones de movimiento se leen

$$\dot{q}_i = \{q_i, H_{\text{tot}}\} = \{q_i, H'\} + v_a \{q_i, \phi_a\} \quad \dot{p}_i = \{p_i, H_{\text{tot}}\} = \{p_i, H'\} + v_a \{p_i, \phi_a\}$$

Nótese que estas nuevas ecuaciones de movimiento no necesitan ser complementadas por los vínculos, ya que nuestra construcción asegura que si los mismos se satisfacen en el instante inicial, lo harán a lo largo de todo el movimiento. En otras palabras, el sistema evoluciona dentro de $\mathcal{E}_{\text{vinc}}$.

Sin embargo, la presencia de funciones arbitrarias en las ecuaciones de Hamilton implica que las condiciones iniciales no determinan completamente la evolución ulterior. Al transcurrir el tiempo, distintas elecciones de las funciones v_a resultarán en diferentes valores para las variables canónicas. Esto implica que el espacio de estados \mathcal{E} , que está definido por el conjunto de todas las condiciones iniciales posibles, está contenido dentro de la hipersuperficie de vínculos $\mathcal{E} \subset \mathcal{E}_{\text{vinc}}$. Cualquier valor para las variables canónicas que satisfaga los vínculos puede ser considerado como una condición inicial para el movimiento, por lo que éstas determinan completamente el estado del sistema. Sin embargo, el estado del sistema no determina un valor único para las coordenadas canónicas, ya que el sistema puede haber arribado a ese estado con cualquier elección de las funciones v_a .

Las funciones arbitrarias v_a también afectan la evolución temporal de cualquier función de las coordenadas canónicas

$$\dot{F} = \{F, H_{\text{tot}}\} = \{F, H'\} + v_a \{F, \phi_a\}$$

Esto significa que si la función es un observable físico, entonces no debería depender de la elección de las v_a . Con esto concluimos que los observables físicos están dados por aquellas funciones de las variables canónicas que satisfacen las restricciones

$$\{F, \phi_a\} = 0$$

Una consecuencia interesante de esta definición se obtiene haciendo uso de la identidad de Jacobi, para escribir

$$\{\{\phi_a, \phi_b\}, F\} + \{\{\phi_b, F\}, \phi_a\} + \{\{F, \phi_a\}, \phi_b\} = \{\{\phi_a, \phi_b\}, F\} = 0$$

De aquí vemos que el observable también queda invariante si sumamos al hamiltoniano términos de la forma $v_{ab}\{\phi_a, \phi_b\}$, con v_{ab} nuevas funciones arbitrarias del tiempo. Esto nos permite definir un *hamiltoniano extendido*

$$H_{\text{ext}}(p_i, q_i, t) = H'(p_i, q_i, t) + v_a \phi_a(p_i, q_i, t) + v_{ab} \{\phi_a(p_i, q_i, t), \phi_b(p_i, q_i, t)\}$$

a partir del cual podemos definir nuevas ecuaciones de movimiento que tendrán el mismo contenido físico.

$$\dot{q}_i = \{q_i, H_{\text{ext}}\} \quad \dot{p}_i = \{p_i, H_{\text{ext}}\}$$

Estas ecuaciones definen la evolución más general posible para un sistema hamiltoniano con vínculos. Los valores de las funciones arbitrarias v_a y v_{ab} no afectan los observables.

Un punto importante es que las funciones $\{\phi_a, \phi_b\}$ podrían ser combinaciones lineales de los vínculos primarios. En tal caso, no es necesario agregarlas en el hamiltoniano extendido dado que ya se encuentran presentes en el hamiltoniano total. Otra opción es que al calcular $\{\phi_a, \phi_b\}$ obtengamos una combinación lineal que incluye algún vínculo secundario. En ese caso la función sí contribuye con un término nuevo al hamiltoniano extendido. Una última opción, que será de interés más adelante, es cuando las funciones $\{\phi_a, \phi_b\}$ no son una combinación lineal de vínculos ni primarios ni secundarios. En ese caso, también agregan nuevos términos al hamiltoniano extendido.

19.2.3 Transformaciones de gauge

Según las definiciones de la sección anterior, las funciones arbitrarias en nuestro hamiltoniano extendido no afectan el estado físico. Ahora bien ¿cómo se relacionan los valores de las coordenadas canónicas cuando los hacemos evolucionar a partir de la misma condición inicial con diferentes elecciones para las funciones arbitrarias? Para ver esto, supongamos que hacemos una evolución infinitesimal con un tiempo ε para una elección posible de dichas funciones, digamos v_a y v_{ab}

$$q_i(t + \varepsilon) = q_i(t) + \varepsilon (\{q_i, H'\} + v_a \{q_i, \phi_a\} + v_{ab} \{q_i, \{\phi_a, \phi_b\}\})$$

$$p_i(t + \varepsilon) = p_i(t) + \varepsilon (\{p_i, H'\} + v_a \{p_i, \phi_a\} + v_{ab} \{p_i, \{\phi_a, \phi_b\}\})$$

El resultado dependerá de la dependencia temporal de las funciones indeterminadas v_a y v_{ab} . Si las cambiáramos por funciones diferentes \tilde{v}_a y \tilde{v}_{ab} , tendríamos

$$\tilde{q}_i(t + \varepsilon) = q_i(t) + \varepsilon (\{q_i, H'\} + \tilde{v}_a \{q_i, \phi_a\} + \tilde{v}_{ab} \{q_i, \{\phi_a, \phi_b\}\})$$

$$\tilde{p}_i(t + \varepsilon) = p_i(t) + \varepsilon (\{p_i, H'\} + \tilde{v}_a \{p_i, \phi_a\} + \tilde{v}_{ab} \{p_i, \{\phi_a, \phi_b\}\})$$

Con lo cual la diferencia entre evolucionar con v_a y v_{ab} respecto de hacerlo con \tilde{v}_a y \tilde{v}_{ab} se puede obtener restando ambos casos, como

$$\tilde{q}_i(t + \varepsilon) = q_i(t + \varepsilon) + \varepsilon ((\tilde{v}_a - v_a) \{q_i, \phi_a\} + (\tilde{v}_{ab} - v_{ab}) \{q_i, \{\phi_a, \phi_b\}\})$$

$$\tilde{p}_i(t + \varepsilon) = p_i(t + \varepsilon) + \varepsilon ((\tilde{v}_a - v_a) \{p_i, \phi_a\} + (\tilde{v}_{ab} - v_{ab}) \{p_i, \{\phi_a, \phi_b\}\})$$

Esto tiene la forma de una transformación canónica generada por las funciones ϕ_a y ϕ_{ab} , lo que se ve más claramente si definimos los parámetros infinitesimales de transformación como $\varepsilon_a = \varepsilon(\tilde{v}_a - v_a)$ y $\varepsilon_{ab} = \varepsilon(\tilde{v}_{ab} - v_{ab})$, lo que nos permite reescribir la formula de arriba como

$$\tilde{q}_i = q_i + \varepsilon_a \{q_i, \phi_a\} + \varepsilon_{ab} \{q_i, \{\phi_a, \phi_b\}\}$$

$$\tilde{p}_i = p_i + \varepsilon_a \{p_i, \phi_a\} + \varepsilon_{ab} \{p_i, \{\phi_a, \phi_b\}\}$$

Pero sabemos que la evolución con v_a y v_{ab} y la evolución con \tilde{v}_a y \tilde{v}_{ab} corresponden al mismo estado físico. Esto quiere decir que las funciones ϕ_a y $\{\phi_a, \phi_b\}$ generan las transformaciones canónicas que no cambian el estado físico, que se denominan las *transformaciones de gauge* del sistema.

Ejemplo: paréntesis de Dirac

Las transformaciones de gauge generadas por las funciones ϕ_a son muy similares a las simetrías. Si bien no dejan invariantes las ecuaciones de movimiento, si lo hacen con el estado físico que resulta de la evolución temporal. Por otro lado, sus generadores ϕ_a son vínculos que, como vimos, se conservan a lo largo del movimiento.

Sin embargo, aparece una diferencia importante y es que el paréntesis de Poisson de dos de tales cantidades conservadas $\{\phi_a, \phi_b\}$ no necesariamente resulta en una cantidad conservada. Esto sólo sucede cuando el resultado es una función de los vínculos primarios o secundarios, pero como vimos más arriba también puede darse el caso de que sea una función enteramente nueva. Por esta razón, los generadores ϕ_a de transformaciones de gauge no constituyen una estructura de álgebra cuando se los combina utilizando el paréntesis de Poisson, como sí sucede con los generadores de cualquier simetría.

Podemos sin embargo idear una modificación del paréntesis de Poisson, conocida como *paréntesis de Dirac* y definida por la fórmula

$$\{F, G\}_D = \{F, G\} + \{F, \phi_{\bar{k}}\} [\{\phi_{\bar{k}}, \phi_{\bar{l}}\}]^{-1} \{\phi_{\bar{l}}, G\}$$

donde estamos tomando la inversa de la matriz $\{\phi_{\bar{k}}, \phi_{\bar{l}}\}$. En esta expresión, los índices \bar{k}, \bar{l} corren sobre los vínculos *de segunda clase*, que son aquéllos cuyo paréntesis de Poisson $\{\phi_{\bar{k}}, \phi_{\bar{l}}\}$ no es una combinación lineal de los vínculos. Por contraposición los vínculos *de primera clase* son aquéllos donde eso sí sucede.

Utilizando el paréntesis de Dirac, tenemos que $\{\phi_a, \phi_b\}_D$ sí es una combinación lineal de los vínculos, con lo que recuperamos una estructura de álgebra. Esto resulta particularmente importante a la hora de cuantizar un sistema mecánico con vínculos hamiltonianos.

19.3 Resumen

En esta clase, estudiamos la aparición de vínculos entre las variables canónicas, como una obstrucción para la construcción del formalismo hamiltoniano a partir de la transformada de Legendre del hamiltoniano. Fuimos capaces de encontrar las ecuaciones de evolución temporal, en función de un hamiltoniano extendido que contiene funciones arbitrarias del tiempo.

Entendimos que los estados físicos son insensibles a la elección de tales funciones arbitrarias, y que diferentes elecciones se relacionan entre sí por transformaciones de gauge.

Bibliografía

- Landau, Lev D.; Lifshitz, Evgeny M. (1976). *Mechanics*. Vol. 1 (3rd ed.). Butterworth-Heinemann. ISBN 978-0-7506-2896-9.
- Goldstein, Herbert; Poole, C. P.; Safko, J. L. (2001). *Classical Mechanics* (3rd ed.). Addison-Wesley. ISBN 978-0-201-65702-9.
- Arnold, Vladimir I. (1989). *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics. Vol. 60. Springer-Verlag. ISBN 978-0-387-96890-2. OCLC 18681352.
- Dirac, P. (1964). *Lectures on Quantum Mechanics*. Yeshiva University. ISBN 978-0-486-41713-4
- Vucetich, H. (2008). *Introducción a la Mecánica Analítica*. Eudeba. ISBN 978-9-502-31588-1

Grandi, Nicolás

Notas de mecánica analítica : preparadas para el curso de la Licenciatura en Física / Nicolás Grandi. - 1a ed. - La Plata : Universidad Nacional de La Plata ; La Plata : EDULP, 2025.

Libro digital, PDF - (Libros de cátedra)

Archivo Digital: descarga

ISBN 978-950-34-2539-8

1. Física. 2. Mecánica. 3. Matemática. I. Título.
CDD 530.1

Diseño de tapa: Dirección de Comunicación Visual de la UNLP

Universidad Nacional de La Plata – Editorial de la Universidad de La Plata

48 N.º 551-599 / La Plata B1900AMX / Buenos Aires, Argentina

+54 221 644 7150

edulp.editorial@gmail.com

www.editorial.unlp.edu.ar

EduLP integra la Red de Editoriales Universitarias Nacionales (REUN)

Primera edición, 2025

ISBN 978-950-34-2539-8

© 2025 - EduLP

e
exactas


EDITORIAL DE LA UNLP



UNIVERSIDAD
NACIONAL
DE LA PLATA