Capítulo 26

Principio de Mínima Acción

26.1 Introducción

¿Recuerdan como introdujimos la noción de principio integral?. Fue con el Principio de Fermat sobre la propagación de la luz, que establecía que al ir de un punto a otro, la luz lo hace por el camino más rápido. Con el Principio de Hamilton tal como lo hemos enunciado en el capítulo anterior, nunca podremos obtener un resultado como el postulado por Fermat, ya que el tiempo permanece inalterado al variar las trayectorias. Surge entonces la pregunta de si existirá y será posible construir en Mecánica algo parecido al Principio enunciado por Fermat para la óptica geométrica. Un ejemplo sencillo nos convence de que realmente es así.

Consideremos una partícula que experimenta un salto brusco de su energía potencial, pasando de una región del espacio de energía potencial constante V_1 a otra de energía V_2. Esto produce un cambio en la velocidad. Aplicando conservación de energía $mv_1^2/2 + V_1 = mv_2^2/2 + V_2$, obtenemos

$$v_2 = v_1 + \sqrt{v_1^2 + \frac{V_1 - V_2}{2}}$$

Esta variación es producida por una fuerza impulsiva perpendicular a la interfase, de manera tal que $v_{1\perp} \neq v_{2\perp}$, y la partícula cambia su dirección de movimiento, *i.e.* se difracta. Puesto que sobre la partícula no actúa ninguna fuerza tangencial a la interfase, se conserva dicha componente de la velocidad $v_{1\parallel} = v_{2\parallel}$, definiendo una *ley de Snell* dada por

$$v_1 \text{sen}\theta_1 = v_2 \text{sen}\theta_2$$

Pero ya hemos visto que podemos llegar a la misma ley aplicando el Principio de Fermat. Debe existir una versión mecánica del Principio de Fermat. Nos proponemos encontrarlo, junto con sus condiciones de validez.

26.2 Variación general de las variables

Consideremos entonces una variación infinitesimal arbitraria y sin restricciones de la trayectoria en el espacio de dimensión $3N-k+1$ determinado por las coordenadas generalizadas q_j $(j=1,...,3N-k)$ y el tiempo t. De hecho, estamos estudiando el caso más general posible, donde pueden variar tanto las coordenadas como el tiempo de los estados inicial y final del sistema. El valor del funcional $\mathcal{J}=\int_{t_1}^{t_2}\mathcal{L}dt$ en la trayectoria levemente variada es

$$\mathcal{J}+\mathrm{d}\mathcal{J}=\int_{t_1+\mathrm{d}t_1}^{t_2+\mathrm{d}t_2}\mathcal{L}(q_1+\delta q_1,...,q_n+\delta q_n;\dot{q}_1+\delta\dot{q}_1,...,\dot{q}_n+\delta\dot{q}_n;t)dt$$

Pero ya estudiamos esta situación en el capítulo dedicado al cálculo de variaciones. Operando a primer orden obtuvimos

$$\begin{aligned}\mathrm{d}\mathcal{J} &= \sum_{j=1}^{3N-k}\int_{t_i}^{t_f}\left(\frac{\partial\mathcal{L}}{\partial q_j}-\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right)\right)\delta q_j dt+\left(\mathcal{L}\,\mathrm{d}t\,+\sum_{j=1}^{3N-k}\left(\frac{\delta\mathcal{L}}{\delta\dot{q}_j}\right)\,\delta q_j\right)\Bigg|_{t_i}^{t_f}\\ &= \sum_{j=1}^{3N-k}\int_{t_i}^{t_f}\left(\frac{\partial\mathcal{L}}{\partial q_j}-\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right)\right)\delta q_j dt+\left(\mathcal{L}\,\mathrm{d}t\,+\sum_{j=1}^{3N-k}p_j\,\delta q_j\right)\Bigg|_{t_i}^{t_f}\end{aligned}$$

Ahora debemos hacer una aclaración. Las variaciones $\delta q_j(t_i)$ y $\delta q_j(t_f)$ en el último término están definidas a tiempo constante, *i.e.* son variaciones virtuales. Sin embargo, estamos permitiendo que los estados inicial y final varíen tanto en coordenada como el tiempo. De esta manera, la variación neta del estado inicial está dada por

$$\begin{aligned}\mathrm{d}q_j(t_i) &= [q_j(t_i+\mathrm{d}t_i)+\delta q_j(t_i+\mathrm{d}t_i)]-q_j(t_i)\\ &= \left[q_j(t_i)+\frac{dq_j}{dt}\bigg|_{t_i}\mathrm{d}t_i+\delta q_j(t_i)+\frac{d\delta q_j}{dt}\bigg|_{t_i}\mathrm{d}t_i\right]-q_j(t_i)\\ &= \dot{q}_j(t_i)\,\mathrm{d}t_i+\delta q_j(t_i)+\delta\dot{q}_j\,\mathrm{d}t_i\\ &= \dot{q}_j(t_i)\,\mathrm{d}t_i+\delta q_j(t_i)\end{aligned}$$

En el último paso hemos eliminado un término de segundo orden en la variación. Evidentemente, podemos obtener una ecuación similar para $\mathrm{d}q_j(t_f)$. Este resultado ños permite obtener una forma alternativa para la variación $\mathrm{d}\mathcal{J}$, reemplazamos δq_j por $\mathrm{d}q_j$,

$$\begin{aligned}\mathrm{d}\mathcal{J} &= \sum_{j=1}^{3N-k}\int_{t_i}^{t_f}\left(\frac{\partial\mathcal{L}}{\partial q_j}-\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right)\right)\delta q_j dt+\left(\mathcal{L}\,\mathrm{d}t\,+\sum_{j=1}^{3N-k}p_j\,\left(\mathrm{d}q_j-\dot{q}_j\,\mathrm{d}t\right)\right)\Bigg|_{t_i}^{t_f}\\ &= \sum_{j=1}^{3N-k}\int_{t_i}^{t_f}\left(\frac{\partial\mathcal{L}}{\partial q_j}-\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right)\right)\delta q_j dt+\left(\sum_{j=1}^{3N-k}p_j\,\mathrm{d}q_j-\mathcal{H}\,\mathrm{d}t\right)\Bigg|_{t_i}^{t_f}\end{aligned}$$

donde hemos utilizado la definición del Hamiltoniano $\mathcal{H}=\sum_{j=1}^{3N-k}p_j\,q_j-\mathcal{L}$

26.3 Principio de Mínima Acción

Supongamos ahora que establecemos las siguientes restricciones sobre la variación de las trayectorias en el resultado anterior

1. La variación neta de las coordenadas generalizadas $\mathrm{d}q_j$ $(j = 1, ..., 3N-k)$ en los puntos extremos de la trayectoria es nula (aunque se permite la variación del tiempo).

2. El Hamiltoniano del sistema a lo largo de la trayectoria real es constante y adopta el mismo valor a lo largo de la trayectoria variada

Aplicando estas condiciones tenemos que, para una trayectoria real

$$\begin{aligned} \mathrm{d}\left(\int_{t_i}^{t_f} \mathcal{H}\, dt\right) &= \sum_{j=1}^{3N-k} \int_{t_i}^{t_f} \left(\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right)\right) \delta q_j dt + \left(\sum_{j=1}^{3N-k} p_j \,\mathrm{d}q_j - \mathcal{H}\,\mathrm{d}t\right)\Bigg|_{t_i}^{t_f} \\ &= \sum_{j=1}^{3N-k} \int_{t_i}^{t_f} \left(\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right)\right) \delta q_j dt + 0 - \mathrm{d}\left(\int_{t_i}^{t_f} \mathcal{H}\, dt\right) \end{aligned}$$

O sea,

$$\mathrm{d}\int_{t_i}^{t_f} (\mathcal{L} + \mathcal{H})\; dt = \sum_{j=1}^{3N-k} \int_{t_i}^{t_f} \left(\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right)\right) \delta q_j dt$$

Ahora, si la trayectoria es real es necesario que se verifiquen las ecuaciones de Lagrange, con lo cual

$$\mathrm{d}\int_{t_i}^{t_f} (\mathcal{L} + \mathcal{H})\; dt = 0$$

y viceversa, si se verifica esta ecuación, también deben verificarse las ecuaciones de Lagrange en virtud de la independencia de las variaciones δq_j.

Definimos la *acción*

$$\mathcal{A} = \int_{t_i}^{t_f} (\mathcal{L} + \mathcal{H})\; dt = \int_{t_i}^{t_f} \sum_{j=1}^{3N-k} p_j dq_j$$

de manera tal que si realizamos una variación que mantenga constantes las coordenadas extremas de la trayectoria y el Hamiltoniano del sistema, entonces se verifica

$$\mathrm{d}\mathcal{A} = 0$$

En el capítulo 11 vimos que si los vínculos son esclerónomos, es decir si las ecuaciones de ligadura no incluyen explícitamente al tiempo, entonces la energía es una función cuadrática de las velocidades generalizadas y con ello el Hamiltoniano es igual a la energía total del sistema. En dicho caso tenemos que

$$\mathcal{A} = \int_{t_i}^{t_f} (\mathcal{L} + \mathcal{H})\; dt = \int_{t_i}^{t_f} ((T - V) + (T + V))\; dt = 2\int_{t_i}^{t_f} T\; dt$$

Bajo esta forma vemos claramente que $\mathcal{A}$ es dimensionalmente igual al producto de "energía por tiempo", igual que la constante fundamenta de la Mecánica Cuántica, la constante de Planck $h = 6.62517 \times 10^{-34}$ Joule.seg. Más adelante veremos que esta similitud no es casual.

26.4 Geodésicas

En el superficie plana definimos la distancia entre dos puntos como la integral del elemento de arco de longitud ds definido como

$$ds = \sqrt{dx^2 + dy^2}$$

y podemos generalizarla al espacio definido por las coordenadas generalizadas $q_1, ..., q_{3N-k}$ de un sistema como

$$ds = \sqrt{\sum_{j=1}^{3N-k} dq_j^2}$$

Esta definición puede parecer obvia, pero no es única y, hasta cierto punto es arbitraria. Por ejemplo, sobre una superficie esférica de radio R y eligiendo como coordenadas los ángulos azimutal θ y radial ϕ, nada nos impide usar la definición anterior y escribir

$$ds = \sqrt{d\theta^2 + d\phi^2}$$

Aunque esta expresión no necesariamente sea incorrecta, probablemente nos encontremos más a gusto con la expresión usual

$$ds = R\sqrt{d\theta^2 + \mathrm{sen}^2\theta d\phi^2}$$

En general, podemos definir la distancia en un dado espacio de dimensión $3N - k$ por medio de una forma diferencial cuadrática

$$ds = \sqrt{\sum_{j,\ell=1}^{3N-k} b_{j,\ell}(q_1, ..., q_{3N-k})\, dq_j\, dq_\ell} \tag{26.1}$$

Al escribir esta expresión decimos que hemos introducido una *métrica* en el espacio. Es evidente que dentro de la arbitrariedad con que podemos definir esta métrica, existen algunos requerimientos. Por ejemplo, debemos imponer el radicando sea positivo, es decir que $ds^2 \geq 0$ y sólo se anula si todos los coeficientes $b_{j,\ell}$ son nulos. Cuando esto se cumple, la forma cuadrática se denomina *definida positiva.*

En el capítulo 11 demostramos que si las ligaduras de un sistema son esclerónomas, entonces la energía cinética es una función cuadrática de las velocidades generalizadas

$$T = \sum_{j,\ell=1}^{3N-k} b_{j,\ell}(q_1, ..., q_{3N-k})\, \dot{q}_j\, \dot{q}_\ell$$

Además es definida positiva, en tanto que $T \geq 0$. Este resultado nos conduce a la idea de utilizar esta forma cuadrática para definir una métrica en el espacio de las coordenadas generalizadas $q_1, ..., q_{3N-k}$, según la ecuación (26.1). Ahora supondremos que se conserva la energía[1] y modificaremos levemente la métrica, escribiendo

$$ds = 2\, \left[(E - V(q_1, ..., q_{3N-k}))\, \sum_{j,\ell=1}^{3N-k} b_{j,\ell}(q_1, ..., q_{3N-k})\, dq_j\, dq_\ell \right]^{1/2}$$

Con esta definición, la acción se puede escribir como la distancia entre los estados inicial y final en el espacio de las coordenadas generalizadas $(q_1, ..., q_{3N-k})$

$$\begin{aligned} \mathcal{A} &= 2 \int_{t_i}^{t_f} T\, dt = 2 \int_{t_i}^{t_f} \sqrt{(E - V)\, T}\, dt \\ &= 2 \int_{t_i}^{t_f} \left[(E - V)\, \sum_{j,\ell=1}^{3N-k} b_{j,\ell}\, \dot{q}_j\, \dot{q}_\ell \right]^{1/2} dt = \int_i^f ds \end{aligned}$$

Con este resultado vemos que el principio de Mínima Acción puede traducirse en términos geométricos como la búsqueda de la distancia más corta entre dos puntos en un dado espacio particular. Por motivos históricos relacionados con la navegación en alta mar, se denomina *geodésica* a la curva que subtiende esta distancia mínima. En un cilindro, la geodésica está dada por una hélice y en una esfera por el arco del círculo máximo que pasa por ambos puntos. Cualquier otra curva por ambos puntos tendrá una longitud mayor.

En este punto es interesante destacar que los principios variacionales sólo dicen que el funcional es un extremo sobre la trayectoria real, pero no que este deba ser un mínimo. Sin embargo, se puede probar que si el estado final se elige suficientemente cercano al inicial (en el sentido de la métrica definida en el espacio), la trayectoria real es aquella que minimiza al funcional. Pero si la distancia es suficientemente grande, puede ocurrir que existan dos trayectorias reales que pasen por ambos estados, ó inclusive un número infinito de ellas. En el caso de una superficie esférica de radio R, esto ocurre cuando la distancia entre ambos puntos es igual a $s = \pi R$. En ese caso hay infinitos círculos máximos que pasan por ambos puntos, y con ellos infinitas geodésicas. En tal caso se dice que el

[1] Recordar que para vínculos esclerónomos (independientes del tiempo), la constancia del Hamiltoniano implica la conservación de la energía.

punto f es el *foco cinético conjugado* del punto i. Consideremos una trayectoria entre dos estados i y f sobre la cual se hace extrema la acción. En dicho caso se puede demostrar en forma absolutamente general que esa acción es menor que la correspondiente a cualquier otra trayectoria, si sobre ella no hay ningún foco conjugado f' del estado i.

26.5 Pierre-Louis Moreau de Maupertuis

26.6 Un problema de denominación

En esta sección estamos entrando en un área de la Mecánica donde no hay un acuerdo preciso sobre la denominación de los distintas cantidades y leyes. Mencionaremos aquí algunas denominaciones en uso.

Algunos autores (Landau, por ejemplo) prefieren la denominación de Principio de Maupertuis en lugar de Principio de Mínima Acción, reservando este último nombre para nuestro Principio de Acción. Por otra parte, se suele llamar *Acción abreviada* a nuestra *Acción*, y *Acción* al funcional $\mathcal{J}$ del principio de Hamilton, que también recibe el nombre de *Función Principal de Hamilton*.

Muchos autores denominan a las ecuaciones de Lagrange como *ecuaciones de Lagrange de segunda especie* ó *ecuaciones de Lagrange en coordenadas independientes*, para distinguirlas de aquellas que se pueden despejar del Principio de d'Alembert en coordenadas cartesianas, por medio de la inclusión de multiplicadores de Lagrange. A estas últimas ecuaciones se las suele denominar *ecuaciones de Lagrange de primera especie*.

También hay una serie de distintas denominaciones en la clasificación de las ligaduras. Pero no me parece oportuno que nos detengamos en esta taxonomía.

26.7 Analogía entre la mecánica y la óptica geométrica

Volvamos a la pregunta con la que iniciamos este capítulo. Nos habíamos preguntado si sería posible encontrar en la mecánica algo parecido al Principio de Fermat sobre la propagación de la luz, que establecía que al ir de un punto a otro, la luz lo hace por el camino más rápido. Con el Principio de Hamilton tal como lo habíamos enunciado en el capítulo anterior, nunca podríamos haber obtenido un resultado como el postulado por Fermat, ya que el tiempo permanece inalterado al variar las trayectorias. Con el principio de Mínima Acción, en cambio, hemos introducido una variación del tiempo en el principio variacional, acercándonos a nuestro objetivo. Sin embargo, el proverbial lector sagaz ya habrá advertido que se mantiene una diferencia esencial entre el Principio de Mínima Acción y el de

Fermat. La ley de Snell para la luz es

$$\frac{1}{v_1}\mathrm{sen}\theta_1 = \frac{1}{v_2}\mathrm{sen}\theta_2$$

y para una partícula

$$v_1\mathrm{sen}\theta_1 = v_2\mathrm{sen}\theta_2$$

Veamos esta diferencia en la situación más general posible. Consideremos una única partícula moviéndose en un campo de fuerzas. La acción se escribe

$$\mathcal{A} = \int_{t_i}^{t_f} \sum_{j=1}^{3N-k} p_j dq_j = \int_{t_i}^{t_f} \sum_{j=1}^{3} p_j dr_j = \int_{t_i}^{t_f} \mathbf{p}.d\mathbf{r} = \int_{t_i}^{t_f} p\, ds$$

En el último paso hemos aprovechado que los vectores $\mathbf{p}$ y $d\mathbf{r}$ tienen la misma dirección. Luego el principio de mínima acción se escribe

$$\mathrm{d}\int_{i}^{f} p\, ds = 0$$

o equivalentemente

$$\mathrm{d}\int_{i}^{f} v\, ds = 0$$

Es cierto que esta expresión guarda una notable analogía con el *principio de Fermat*

$$\mathrm{d}\int_{i}^{f} \frac{1}{v}\, ds = 0$$

Pero también es cierto que es esencialmente diferente a él. Ahora podemos comprender la confusión de Descartes, corregida por Fermat, al suponer que su ley de refracción implicaba que la luz se movía más rápido en el medio más denso.

La analogía entre la óptica y la mecánica requiere definir el índice de refracción de un medio como el cociente entre las velocidades de la luz en el vacío y en el medio, $n = c/v$. De esta manera podemos escribir el Principio de Fermat como

$$\mathrm{d}\int_{i}^{f} n\, ds = 0$$

Ahora si, esta expresión tiene la misma estructura que el Principio de Mínima Acción, siempre que asociemos la velocidad de la partícula con el índice de refracción de la luz.

La cantidad de movimiento está ligada con la energía potencial del sistema por la ley de conservación de la energía $mv^2/2 + V = E$ y en consecuencia podemos asociar el índice de refracción con la energía potencial en una dada región del espacio

$$v = \sqrt{\frac{E - V}{m/2}}$$

Vemos que las variaciones del índice de refracción de la luz en un medio no homogéneo son análogas a las variaciones de energía potencial para el movimiento de una partícula. Es decir que a un incremento del índice de refracción le corresponde un incremento de la energía potencial.

26.8 Comentario final sobre los Principios Integrales

26.9 Gottfried Wilhelm Leibniz