

Mecánica Estadística. Clase 6. Teoría cinética y mecánica estadística

Prof. Juan Mauricio Matera

31/03/2025

En las clases previas presentamos el formalismo de la termodinámica, y encontramos ciertas analogías entre la entropía termodinámica y la entropía estadística, asociada a una distribución de probabilidades. En ambos casos, se trata de cantidades que crecen linealmente con el número de *copias* del sistema descrito, y que aumentan cuando los sistemas se *mezclan* de forma *irreversible*. A partir de esta observación, y de la búsqueda explicaciones para las propiedades termodinámicas de los sistemas en términos de la dinámica *determinista* de sus constituyentes, Ludwig Boltzmann, James Clerk Maxwell y luego Josiah Willard Gibbs (quien le da nombre) comienzan a desarrollar el programa de la mecánica estadística. El ingrediente fundamental de este acercamiento es la renuncia a describir el estado instantáneo de los grados de libertad microscópicos en forma exhaustiva, para en su lugar hacerlo en términos de distribuciones de probabilidades.

En la clase de hoy vamos a explorar algunas de las ideas básicas que permiten construir las distribuciones de probabilidades de equilibrio para algunos sistemas simples, y a partir de estas recuperar la forma de una entropía estadística, que resulta formalmente equivalente a la entropía termodinámica. Para esto, vamos a considerar tres estrategias.

Un primer acercamiento a conectar las propiedades microscópicas con propiedades macroscópicas parte del llamado *teorema del virial*. Este teorema, deducido por primera vez por Rudolph Clausius, permite estimar promedios temporales para cantidades macroscópicas, como la energía interna, a partir de una descripción microscópica de las interacciones.

Luego analizaremos la denominada *teoría cinética*, que se basa en construir la distribución de probabilidades a partir de consideraciones de simetría, para luego conectar los parámetros libres con las propiedades macroscópicas que se esperan encontrar en estos modelos. Vía el teorema del virial y el uso de relaciones termodinámicas, este approach permite conectar los valores esperados de las magnitudes físicas consideradas en términos de promedios ergódicos.

A partir de estos modelos, se analizará la conexión entre la entropía termodinámica y la entropía estadística. Para esto, un aspecto a considerar tiene que ver con que la entropía estadística se define para distribuciones de probabilidad discretas, mientras que el espacio de estados de la mecánica es *continuo*. Luego, deberemos discutir un esquema de *discretización*, y cómo este puede o no afectar los resultados en el límite termodinámico.

Finalmente consideraremos el problema de la marcha al equilibrio, en términos de la llamada *ecuación de Boltzmann*. Esta ecuación busca describir la dinámica de una distribución de probabilidades, asumiendo que esta describe partículas que se mueven de acuerdo con las leyes de la mecánica. Veremos que vía el llamado *teorema H*, se puede mostrar que las soluciones de esta ecuación efectivamente *relajan* a soluciones estacionarias, y que dichas soluciones son consistentes con las que resultan de la teoría cinética. Más aún, del *teorema H* se sigue que la entropía estadística es una cantidad que nunca *decrece*.

Teoría Cinética

El espacio de fases de la Mecánica clásica

En un sistema mecánico, el estado de un sistema está completamente determinado por un conjunto de *coordenadas generalizadas* ξ , a partir de las cuales es posible calcular cualquier propiedad física de un sistema. Para una partícula libre, estas coordenadas pueden elegirse como sus las tres componentes de su vector posición \vec{r} en un sistema de coordenadas cartesianas, y las tres componentes de su *cantidad de movimiento* \vec{p} . Para un sistema de N partículas, las coordenadas generalizadas corresponden a la suma directa de las coordenadas generalizadas de cada partícula por

separado. Si el sistema es cerrado, las coordenadas generalizadas evolucionan de acuerdo a la dinámica Hamiltoniana

$$\frac{d\xi_i}{dt} = \mathcal{M} \frac{\partial H}{\partial \xi_i}$$

con $H(\xi)$ una función de las coordenadas generalizadas y \mathcal{M} una matriz que aplicada a $\xi = (\vec{x}, \vec{p})^T$, multiplica por 1 a las coordenadas x y por -1 a los momentos p . De esta manera, cualquier función de ξ y sus derivadas puede expresarse como una función de ξ .

Sobre el espacio de fases podemos definir una distribución de probabilidades $f(\xi)$, de manera que

$$\int f(\xi) d^{6N} \xi = 1$$

Podemos interpretar la integral de esta distribución sobre una región \mathcal{R} del espacio de fases como la *frecuencia relativa* $f_{\mathcal{R}}$ con la que el esperamos encontrar al sistema si lo observamos en instantes de tiempo al azar, o si observamos una copia del sistema sobre una gran cantidad de copias preparadas de la misma manera *en un instante dado*. Si en un sistema, bajo cierto contexto de observación, los resultados de ambas nociones coinciden, decimos que el sistema satisface la *hipótesis ergódica*.

Como cualquier cantidad física puede ser expresada como una función $W(\xi)$ que depende de las coordenadas, la correspondiente media estadística viene dada por

$$\langle W \rangle = \int W(\xi) f(\xi) d^{6N} \xi$$

En un sistema que satisface la hipótesis ergódica,

$$\langle W \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T W(\xi(\tau)) d\tau$$

de manera que $f(\xi)$ es proporcional al tiempo que pasa el sistema en el elemento de volumen $d^{6N} \xi$ del espacio de fases.

Notemos por otro lado que en un sistema *grande*, que evoluciona en forma cerrada, la convergencia del límite ergódico puede requerir tiempos ridículamente grandes. Por ejemplo, en un sistema de osciladores, el sistema pasa exactamente una vez por cada punto de la trayectoria en cada período. Si los osciladores se encuentran desacoplados, y tienen frecuencias $\omega_1 < \omega_2 < \dots < \omega_n$ conmensuradas ($\omega_i/\omega_n = p_i/q_i < 1$ para enteros p_i, q_i) el período del sistema resulta ser $T = Q \frac{2\pi}{\omega_n}$ con Q el mínimo común múltiplo de los q_i . Como ejemplo, supongamos que tenemos solo dos osciladores, uno con frecuencia ω_0 y otro con frecuencia $\omega_1 = \frac{99}{100} \omega_0$. El período de los osciladores por separado es prácticamente igual $T_0 = \frac{2\pi}{\omega_0}$, mientras que el período del conjunto es $T = 100 \times T_0$.

Por otro lado, si tomamos la interpretación de frecuencia sobre sistemas *preparados de forma equivalente*, debemos especificar qué características comunes tienen esos sistemas. Volveremos a esto la clase siguiente, cuando discutamos la noción de *ensamble estadístico*.

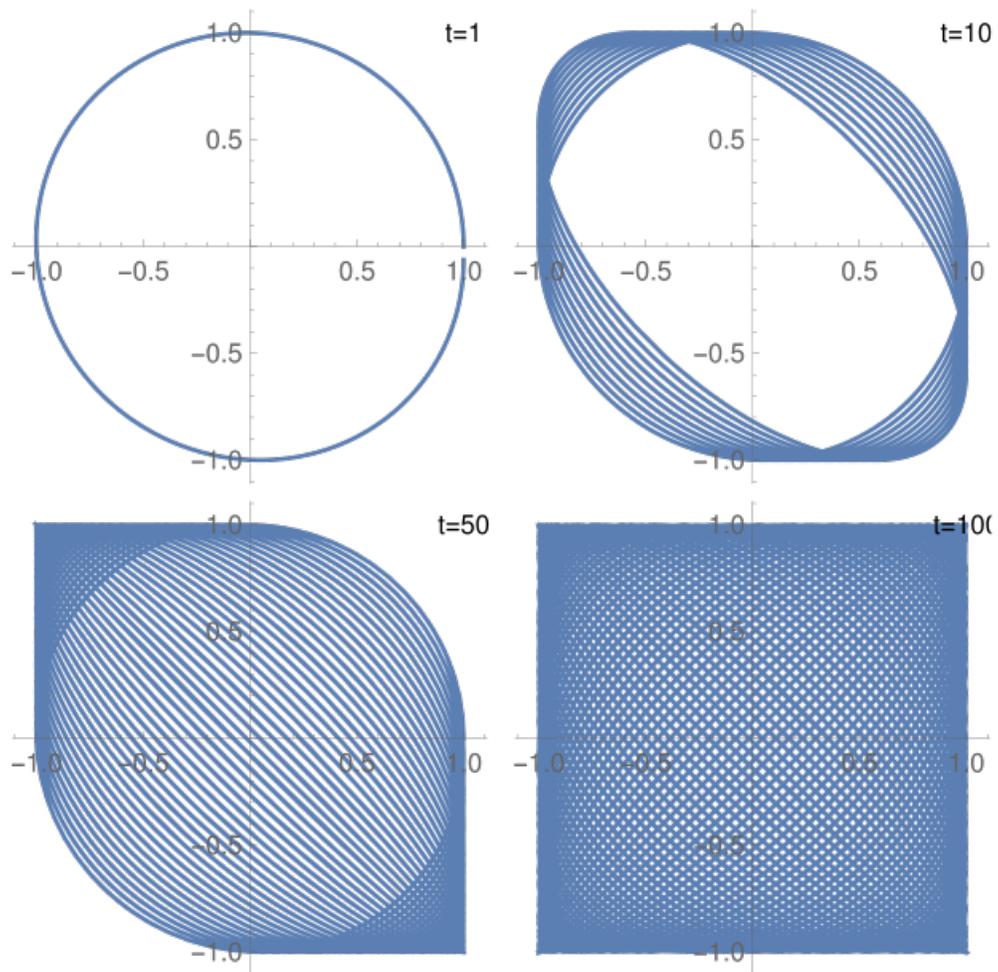


Figure 1: Trayectoria de dos osciladores acoplados con frecuencias próximas. La equivalencia entre el promedio temporal y el promedio sobre el ensamble sólo se realiza para tiempos del orden del período de recurrencia global, y no para el de un oscilador independiente.

Teorema del Virial



Rudolph Clausius
1822-1888

Para el cálculo de promedios temporales, un resultado útil deducido por primera vez por Rudolph Clausius es el llamado teorema del virial. Para formularlo, consideremos primero el *momento escalar de inercia*

$$I = \sum_i \frac{1}{2} m_i |\vec{r}_i|^2$$

con \vec{r}_i la posición relativa de la partícula i al centro de masas del sistema. Esencialmente, esta cantidad es proporcional a la traza del tensor de inercia

$$I = \frac{1}{4} \sum_{\mu} I_{\mu\mu}$$

con

$$I_{\mu\nu} = \sum_i m_i (\delta_{\mu\nu} |\vec{r}_i|^2 - r_{i\mu} r_{i\nu})$$

La derivada temporal de esta cantidad resulta ser

$$G = \sum_i m_i \vec{r}_i \cdot \vec{v}_i = \sum_i \vec{r}_i \cdot \vec{p}_i$$

Derivando nuevamente,

$$\frac{\partial G}{\partial t} = \sum_i \vec{v}_i \cdot \vec{p}_i + \sum_i \vec{r}_i \cdot \frac{d \cdot \vec{p}_i}{dt} = 2E_c + \sum_i \vec{r}_i \cdot \vec{F}_i$$

con E_c la energía cinética del sistema y \vec{F}_i la fuerza sobre la partícula i -ésima.

Promediando en el tiempo, e introduciendo la cantidad

$$\mathcal{V} = -\frac{1}{2} \langle \sum_i \vec{r}_i \cdot \vec{F}_i \rangle_T$$

denominada *virial del sistema*, se encuentra que si su escalar de inercia no crece de forma acelerada - esto es, el sistema no se expande en forma explosiva-, resulta que

$$\langle \frac{dG}{dt} \rangle_T = \lim_{T \rightarrow \infty} \frac{G(t)}{T} = 0$$

y por lo tanto

$$\langle E_c \rangle_T = \mathcal{V}$$

lo que constituye el resultado del *teorema del virial*.

Conviene ahora distinguir en el virial las contribuciones de fuerzas externas e internas:

$$\vec{F}_i = \vec{F}_i^{ext} + \sum_j \vec{F}_{ij}$$

con $\vec{F}_{ij} = -\vec{F}_{ji}$ la fuerza que ejerce la partícula j sobre la partícula i . El segundo miembro se expresa entonces

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

A partir de este resultado, analicemos algunos casos.

Gas de partículas no interactuantes

Si el sistema está compuesto por partículas no interactuantes en un contenedor de volumen V , el miembro de la derecha en la expresión del virial se expresa en términos de la presión como

$$-\left\langle \sum_i \vec{F}_i \cdot \vec{r}_i \right\rangle = - \int_S -P \vec{n} \cdot \vec{r} dS$$

que es una integral de flujo sobre una superficie del contenedor. Vía el teorema de la divergencia, y notando que

$$\vec{\nabla} \cdot \vec{r} = 3$$

podemos expresar

$$\int_S P \vec{n} \cdot \vec{r} dS = P \int \nabla \cdot \vec{r} dV = 3PV$$

y luego,

$$\langle E_c \rangle = \frac{3}{2} PV$$

o vía la ecuación de estado,

$$\langle E_c \rangle = \frac{3}{2} NRT$$

Sistemas acoplados por un potenciales

Otro caso importante a considerar es cuando el sistema es *cerrado*, y las partículas están acopladas por fuerzas conservativas. Este es el caso de un sistema de osciladores armónicos acoplados, o el de un sistema autogravitante. En tal caso, las fuerzas pueden expresarse en términos del gradiente de un potencial que se expresa como suma de potenciales de pares: $\varphi(\vec{r}_1, \dots, \vec{r}_N) = \sum_{i < j} \varphi_{ij}(\vec{r}_i - \vec{r}_j)$

de manera que

$$\vec{F}_{ij} = - \sum_j \nabla_i \varphi_{ij}(\vec{r}_i - \vec{r}_j)$$

Remplazando en la expresión de la derivada del virial obtenemos

$$\langle E_c \rangle = \frac{1}{2} \left\langle \sum_{i < j} (\vec{r}_i - \vec{r}_j) \cdot \nabla \varphi_{ij}(\vec{r}_i - \vec{r}_j) \right\rangle$$

que para el caso de un potencial $\varphi_{ij}(\vec{r}_i - \vec{r}_j) = \varphi_{ij}(|\vec{r}_i - \vec{r}_j|)$ central resulta

$$\langle E_c \rangle = \frac{1}{2} \left\langle \sum_{i < j} |\vec{r}_i - \vec{r}_j| \nabla \varphi'_{ij}(|\vec{r}_i - \vec{r}_j|) \right\rangle$$

Finalmente, si $\phi_{ij}(r) = \lambda r^a$, $r\phi'_{ij}(r) = a\phi_{ij}(r)$ y la expresión se reduce a

$$\langle E_c \rangle = \frac{a}{2} \langle \sum_{i < j} \nabla \phi_{ij}(|\vec{r}_i - \vec{r}_j|) \rangle = \frac{a}{2} \langle \varphi \rangle$$

De esta manera, para una interacción armónica, ($a = 2$),

$$\langle E_c \rangle = \langle \varphi \rangle$$

que significa que la energía mecánica total promedio del sistema se *reparte en partes iguales* entre la energía cinética y la energía potencial.

Por otro lado, si las partículas están acopladas por la interacción gravitatoria (no relativista),

$$\langle E_c \rangle = -\frac{1}{2} \langle \varphi \rangle$$

La distribución de Maxwell



James Clerk Maxwell
1831-1879

En cualquier caso, podemos suponer que en un sistema en equilibrio termodinámico, la distribución de probabilidades para el estado del sistema es *estacionaria*, y satisface ciertas simetrías. Partiendo de esta hipótesis, James Maxwell propuso que en un sistema de partículas puntuales, iguales y no interactuantes entre sí, confinadas a moverse en un cierto volumen, la distribución de probabilidades para el estado del sistema en equilibrio debería ser el producto de las distribuciones de probabilidad para cada partícula,

$$f(\xi) = \prod_{k=1}^N p_k(\xi_k)$$

con $p_k(\xi_k) \equiv p_1(\vec{r}_i, \vec{v}_i)$ una distribución tal que:

- sólo dependa de las velocidades y no de las posiciones.
- sea invariante ante rotaciones. Luego, $p_1(\vec{v}) = p_1(\vec{v})$
- las componentes cartesianas de la velocidad no *están correlacionadas*, de manera que $p_1(\vec{v}) = p_{1,x}(v_x)p_{1,y}(v_y)p_{1,z}(v_z)$

Bajo estas hipótesis, la distribución para una partícula debe tener la forma gaussiana

$$p_1(\xi_1) = \frac{e^{-A \frac{v^2}{2}}}{\sqrt{(2\pi/A)^3 V}}$$

con A una constante a determinar y V el volumen del contenedor.

Si el sistema de partículas presenta esta distribución, podemos estimar la presión media que ejerce sobre las paredes del contenedor calculando el intercambio de momento neto por unidad de tiempo y unidad de área debido a las

colisiones de las partículas. Para esto, notamos que si una partícula de masa m tiene velocidad \vec{v} , al colisionar con la pared en forma elástica, esta invierte la dirección de la componente de su velocidad normal a la pared. De esta manera, el momento absorbido por la pared en un choque individual viene dado por

$$\Delta\vec{p}(\vec{v}) = -2m\vec{n} \cdot \vec{v}\vec{n}$$

con \vec{n} el vector normal. Por otro lado, la presión sobre un elemento de área dS viene dada por

$$P = \int_{\vec{v} \cdot \vec{n} < 0} (-\vec{n} \Delta\vec{p}(\vec{v})) \times \frac{dN(\vec{v})}{dSdt} d^3\vec{v}$$

con $\frac{dN(\vec{v})}{dSdt} d^3\vec{v}$ el número de choques por unidad de tiempo y unidad de área que experimenta el elemento de área $d\vec{S}$ con partículas con velocidades en $d^3\vec{v}$. Este número puede estimarse como el producto del número total de partículas N , la distribución de probabilidades $p_1(\xi)$, y el volumen dV en el que se encuentran las partículas cuyas trayectorias intersectan el elemento de área $d\vec{S}$ dentro del intervalo de tiempo dt , dividido dt :

$$\frac{dN}{dSdt} = N p_1(\vec{v}) \frac{dV}{dSdt} = \frac{N}{V} \frac{e^{-A \frac{v^2}{2}}}{\sqrt{(2\pi/A)^3}} \frac{dV}{dSdt}$$

Como previo a la colisión, las partículas se mueven en línea recta, dV es el volumen de un cilindro oblicuo, de sección dS y alto $\vec{v}\vec{n}dt$, de manera que

$$\frac{dV}{dSdt} = -\vec{v}\vec{n}$$

Nótese además que la integral debe restringirse a las partículas cuyas velocidades están orientadas en dirección a la pared ($\vec{v} \cdot \vec{n} < 0$).

Poniendo todo junto, y eligiendo un sistema de coordenadas cartesianas tal que $-\vec{n}$ sea paralelo al eje z ,

$$P = \int_{v_z > 0} 2mv_z \times \frac{N}{V} \frac{e^{-A \frac{v^2}{2}}}{\sqrt{(2\pi/A)^3}} v_z dv_x dv_y dv_z = 2m \frac{N}{V} \int_0^\infty v_z^2 \frac{e^{-A \frac{v_z^2}{2}}}{\sqrt{2\pi/A}} dv_z$$

donde en el miembro de la derecha se evaluaron trivialmente las integrales en v_x y v_y usando la normalización. La integral restante se puede calcular de diferentes maneras. Una forma bastante directa consiste en observar que $x^2 e^{-\lambda x^2/2} = -2 \frac{\partial}{\partial \lambda} e^{-\lambda x^2/2}$, de manera que

$$\int_0^\infty v_z^2 e^{-A \frac{v_z^2}{2}} dv_z = -2 \frac{\partial}{\partial A} \int_0^\infty e^{-A \frac{v_z^2}{2}} dv_z \quad (1)$$

$$= -2 \frac{\partial}{\partial A} \frac{\sqrt{2\pi/A}}{2} = \frac{1}{2} \frac{\sqrt{2\pi/A}}{A} \quad (2)$$

Reemplazando,

$$P = \frac{N m}{V A}$$

Finalmente, comparando esta expresión con la ecuación de estado de un gas ideal, se sigue que

$$\frac{Nm}{A} = \nu RT$$

El miembro de la derecha no depende de m , por lo que A debe ser proporcional a m/T . Por otro lado, el cociente entre el número de partículas N que forman el sistema, y el número de moles ν

$$N_A = \frac{N}{\nu}$$

no dependen ni del tipo de partículas, ni de la temperatura. Por lo tanto tiene que valer lo mismo para cualquier sistema, y depende sólo de cómo definimos el *mol*. Vía la observación del efecto Browniano, fue posible estimar este

número en una cantidad del orden de $N_A \approx 6 \times 10^{23}$. Medidas más precisas, basadas en la determinación de la carga del electrón, permitieron establecer el valor de esta cantidad en

$$N_A = 6,02214076 \times 10^{23}$$

que en la actualidad se considera un valor exacto (por definición). A partir de esta constante, podemos expresar A como

$$A = \frac{mN_A}{RT} = \frac{m}{kT}$$

con $k = R/N_A = 1.380649 \times 10^{-23}$ Joule/Kelvin la *constante de Boltzmann*.

De esta manera, la distribución de Maxwell para un gas se expresa como

$$p_1(\vec{r}, \vec{v}) = \frac{e^{-\frac{mv^2}{2kT}}}{\sqrt{(2\pi kT/m)^3 V}}$$

Identificando $\frac{mv^2}{2}$ con la *energía cinética* de una partícula, podemos expresar la distribución para el sistema completo de N partículas no interactuantes como

$$f(\xi) = \frac{e^{-\frac{H(\xi)}{kT}}}{V^N \sqrt{2\pi mkT}^{3N}}$$

donde implementamos el cambio de variables $\vec{v}_i = \vec{p}_i/m$ y el hecho de que el Hamiltoniano $H(\xi) = \sum_i \frac{p_i^2}{2m} = \sum_i E_{c,i}$ es igual a la suma de las energías cinéticas de cada partícula del sistema.

A partir de esta distribución, es posible estimar la velocidad media cuadrática de las partículas

$$\langle v^2 \rangle = \int p_1(\vec{v}) v^2 d^3v = 3 \int p_1(\vec{v}) v_z^2 d^3v = \frac{3kT}{m}$$

de manera que la energía cinética media por partícula resulta

$$\langle E_c \rangle = \frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} kT$$

que para un mol de gas, resulta ser

$$U = N_A \langle E_c \rangle = \frac{3}{2} N_A kT = \frac{3}{2} RT$$

lo que está de acuerdo con las observaciones para el calor específico de gases diluídos.

Nota histórica: La propuesta de que a una temperatura y presión fija, el volumen de un gas es proporcional a la cantidad de moléculas que lo forman fue propuesta por primera vez en 1811 por el químico italiano Amadeo Avogadro. Claro, ya dijimos que en esa época la hipótesis atómica no era considerada demasiado seriamente por la comunidad científica. 50 años más tarde, Johann Loschmidt estima la constante de proporcionalidad a partir de la estimación del tamaño de las moléculas y del camino libre medio.

La determinación precisa de la constante de proporcionalidad fue realizada por el físico francés Jean Perrín cien años más tarde, lo que lo llevó a ganar el Premio Nobel de Química en 1926. Perrín propone entonces nombrar N_A en honor a Avogadro.



Amedeo Avogadro
1776-1856



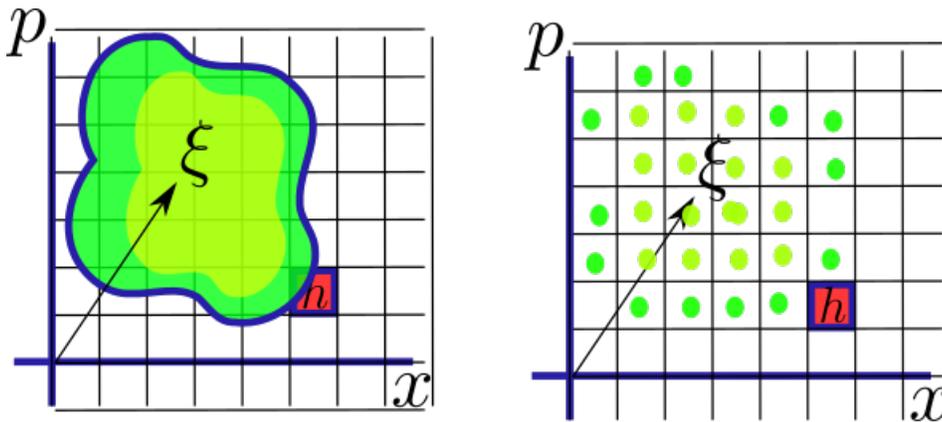
Jean Perrin
1870-1942



Josef Loschmidt
1821-1895

Entropía de la distribución de Maxwell

A partir de la distribución de Maxwell, nos interesa ahora construir su entropía estadística. Sin embargo, la entropía estadística está definida para distribuciones discretas y, como el espacio de fases es continuo, las distribuciones definidas sobre este espacio también lo son. Para calcular una entropía debemos discretizar el espacio y la distribución correspondiente. La manera de hacerlo es dividir el espacio de fases en regiones disjuntas de tamaño fijo, y calcular la probabilidad de encontrar al sistema en cada una de esas regiones.



$$f(\xi) \rightarrow f_\xi = \int_\xi f(\xi) d\xi$$

Una elección simple consiste en dividir el espacio en una grilla uniforme, de tamaño h^{3N} , con h una constante. Si h es suficientemente pequeño, la energía dentro del elemento de volumen puede considerarse constante, y podemos expresar la probabilidad p_ξ de encontrar al sistema en el elemento de volumen que contiene a ξ como

$$f_\xi = \int_\xi f(\xi) d^{6N}\xi \approx f(\xi) h^{3N}$$

Observamos que

$$\sum_\xi f_\xi = \sum_\xi f(\xi) \int_\xi d^{6N}\xi = \int f(\xi) d^{6N}\xi = 1$$

donde usamos que $h^{3N} = \int_\xi d^{6N}\xi$, la integral dentro del elemento de volumen que contiene a ξ , y que para cualquier

función $w(\xi)$,

$$\sum_{\xi} w(\xi) f_{\xi} = \int w(\xi) f(\xi) d^{6N} \xi$$

siempre que $w(\xi) \approx \int_{\xi} w(\xi) d^{6N} \xi$.

Podemos entonces expresar la entropía asociada a f_{ξ} como

$$S(f_{\xi}) = \sum -k f_{\xi} \ln(f_{\xi}) = \int -k f(\xi) (\ln(f(\xi) h^{3N}) d^{6N} \xi$$

donde usamos k como escala para la entropía. En particular, si la distribución es la de Maxwell,

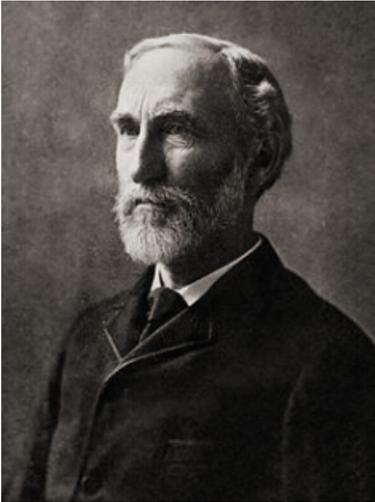
$$S(\{f_{\xi}\}) = \int f(\xi) \left(\frac{E(p)}{T} + Nk \ln((\sqrt{2\pi mkT}/h)^3 V) d^{3N} p d^{3N} r = \frac{\langle E \rangle}{yT} + Nk \ln((\sqrt{2\pi mkT}/h)^3 V)$$

pero $\langle E \rangle = \frac{3}{2} NkT$, y por lo tanto,

$$S(\{f_{\xi}\}) = \frac{3}{2} Nk + Nk \ln \left(\frac{(E/N)^{3/2}}{(h^2/(4/3\pi m))^{3/2} V} \right)$$

que tiene la forma de la entropía termodinámica, salvo por un factor $1/N$ faltante dentro del logaritmo, que es el que garantiza la extensividad.

La paradoja de Gibbs



Josiah Willard Gibbs 1839-1903

Del análisis anterior observamos que la entropía estadística parece fallar en la propiedad de extensividad. La solución *ad-hoc* para este problema es agregar un término extra a la entropía $S_g = -k \log(N!)$ que en el límite de $N \gg 1$ puede expresarse vía la aproximación de Stirling como

$$S_g = -Nk \log(N)$$

Pero este término aparece naturalmente si asumimos que los elementos del espacio de fases que se obtienen por permutaciones de las coordenadas son *equivalentes*:

$$f(\xi) \rightarrow f_{\text{indistinguishable}}(\xi) = \frac{f(\xi)}{N!}$$

La equivalencia de los elementos del espacio de fase puede argumentarse planteando la *indistinguibilidad* de las partículas, esto es, que no existe proceso físico capaz de distinguir dos partículas. En mecánica clásica, como las trayectorias de las partículas están perfectamente definidas, si etiquetamos a las partículas en un cierto instante a partir de sus posiciones iniciales, podemos identificar cuál es cuál en términos de la trayectoria que siguen, independientemente de cuán cerca pasen estas trayectorias.

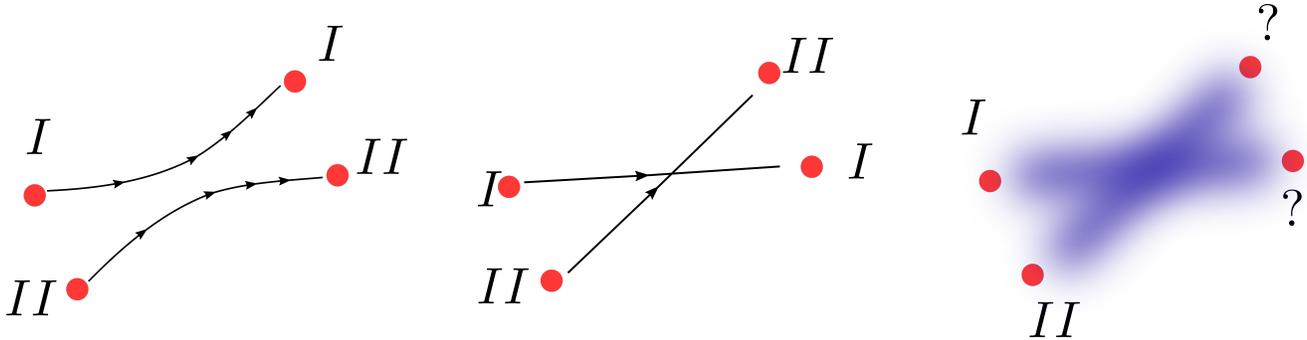


Figure 2: Partículas distinguibles por trayectoria. En mecánica clásica, las partículas *I* y *II* mantienen su identidad debido a que están conectadas por su trayectoria. Si las trayectorias no están bien definidas, esa indistinguibilidad de principio se pierde.

Por otro lado, en mecánica cuántica, el principio de incerteza implica que no existen trayectorias definidas. Luego, esta identificación de la identidad de las partículas no es posible, de manera que estas se vuelven indistinguibles.

La ecuación de Boltzmann



Ludwig Boltzmann
1844-1906

Al principio del curso cuando consideramos la evolución de una distribución de probabilidades en un sistema mecánico cerrado, vimos que para una evolución *determinista* se satisface

$$f(\xi(\xi_0, t), t) = f(\xi_0, t_0)$$

donde $\xi(\xi_0, t) = (x(t), p(t))$ representa la posición del sistema en el espacio de fases al tiempo t como función de las condiciones iniciales $\xi_0 = (x(t_0), p(t_0))$ a t_0 . Derivando esta relación respecto de t se sigue que

$$\frac{d}{dt} f(\xi(\xi_0, t), t) = \frac{\partial f}{\partial t} + \frac{d\xi}{dt} \cdot \nabla_{\xi} f = 0$$

donde el segundo miembro se anula ya que no depende de t . Conviene reescribir el segundo miembro usando las ecuaciones de Hamilton:

$$\frac{\partial \xi}{\partial t} = \begin{pmatrix} \nabla_{\vec{p}} H \\ -\nabla_{\vec{x}} H \end{pmatrix} = \begin{pmatrix} \vec{v} \\ \vec{F} \end{pmatrix}$$

con $\vec{v} = \nabla_{\vec{p}}H$ el vector de velocidades generalizadas y con $\vec{F} = -\nabla_{\vec{x}}H$ el vector de fuerzas generalizadas. Reemplazando obtenemos

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla_{\vec{x}}f + \vec{F} \cdot \nabla_{\vec{p}}f = 0$$

En la práctica, no interesa conocer la distribución f asociada a todas las coordenadas, sino el marginal asociado a un subsistema pequeño que involucra una o unas pocas partículas. Integrando sobre las coordenadas del resto de las partículas obtenemos

$$\frac{\partial f_S}{\partial t} + \vec{v}_S \cdot \nabla_{\vec{x}_S}f_S + \vec{F}_S \cdot \nabla_{\vec{p}_S}f_S = - \int d\Omega_{\bar{S}}\vec{v}_{\bar{S}} \cdot \nabla_{\vec{x}_{\bar{S}}}f - \int d\Omega_{\bar{S}}\vec{F}_{\bar{S}} \cdot \nabla_{\vec{p}_{\bar{S}}}f - \int d\Omega_{\bar{S}}\vec{F}_{S,\bar{S}} \cdot \nabla_{\vec{p}_S}f - \int d\Omega_{\bar{S}}\vec{F}_{\bar{S},S} \cdot \nabla_{\vec{p}_{\bar{S}}}f$$

donde $f_S = \int f d\Omega_{\bar{S}}$ es la distribución sobre el sistema de interés S , \vec{v}_S y \vec{F}_S son funciones de las coordenadas del sistema de interés, y $\vec{v}_{\bar{S}}$ y $\vec{F}_{\bar{S}}$ son funciones de las coordenadas del entorno \bar{S} . $\vec{F}_{S,\bar{S}}$ y $\vec{F}_{\bar{S},S}$ por otro lado son las fuerzas que el sistema \bar{S} ejerce sobre S y viceversa, y por lo tanto son funciones de las coordenadas de ambos sistemas.

Los primeros dos términos del segundo miembro son de la forma

$$\iint \vec{r}(u)\nabla_v f(u, v, w) dudv = \iint \nabla_v \cdot (\vec{r}(u)f(u, v, w)) dudv$$

que se reducen a una integral sobre el borde en una de las variables. En el caso del primer término, podemos asumir que la velocidad sobre el borde se anula, y por lo tanto lo hace la integral. El segundo término es una integral de borde sobre las variables de momento, que no están acotadas. Pero si la energía media del sistema es finita, f tiene que anularse para valores grandes de momento, por lo que también se anula.

El tercer término puede reescribirse como

$$- \int d\Omega_{\bar{S}}\vec{F}_{S,\bar{S}} \cdot \nabla_{\vec{p}_S}f = -\nabla_{\vec{p}_S} \cdot \int d\Omega_{\bar{S}}\vec{F}_{S,\bar{S}}f + \int (\nabla_{\vec{p}_S} \cdot \vec{F}_{S,\bar{S}})f d\Omega_{\bar{S}} = -\nabla_{\vec{p}_S} \cdot \int d\Omega_{\bar{S}}\vec{F}_{S,\bar{S}}f$$

donde el segundo término se anula si asumimos que las fuerzas entre los sistemas no dependen de las velocidades. De manera análoga, el cuarto término también se anula:

$$- \int d\Omega_{\bar{S}}\vec{F}_{\bar{S},S} \cdot \nabla_{\vec{p}_{\bar{S}}}f = -\nabla_{\vec{p}_{\bar{S}}} \cdot \int d\Omega_{\bar{S}}\vec{F}_{\bar{S},S}f + \int (\nabla_{\vec{p}_{\bar{S}}} \cdot \vec{F}_{\bar{S},S})f d\Omega_{\bar{S}} = 0$$

Aquí, el primer término del segundo miembro se anula por ser una integral sobre el *borde* en la coordenada de momento, y el segundo lo hace bajo la hipótesis de que las fuerzas no dependen de los momentos. Finalmente,

$$\frac{\partial f_S}{\partial t} + \vec{v}_S \cdot \nabla_{\vec{x}_S}f_S + \vec{F}_S \cdot \nabla_{\vec{p}_S}f_S = -\nabla_{\vec{p}_S} \cdot \int \vec{F}_{S,\bar{S}}f d\Omega_{\bar{S}} = I_{col}[f]$$

El problema con esta expresión es que el segundo miembro involucra una integral sobre la distribución del sistema completo, por lo que para evaluarla, debemos imponer alguna hipótesis extra sobre como es f .

La integral de colisiones

Para el caso de un gas, Boltzmann propuso asumir que $f(x_1, p_1, x_2, p_2, \dots, t) = f_1(x_1, p_1, t)f_1(x_2, p_2, t)\dots$ de manera que

$$\frac{\partial f_1}{\partial t} + \vec{v} \cdot \nabla_{\vec{x}}f_1 + \vec{F} \cdot \nabla_{\vec{p}}f_1 = I_{col}[f_1]$$

con I_{col} una integral que a cuenta de las *colisiones* entre pares de partículas *descorrelacionadas*, todas siguiendo la *misma distribución* f_1 . Esta hipótesis, conocida como *hipótesis de caos molecular* (o en alemán, *Stosszahlansatz*, que quiere decir literalmente *hipótesis de tasa de impacto*) permite escribir $I_{col}[f_1]$ como una funcional de la forma

$$I_{col}[f_1] = \int N\Sigma(\vec{p}, \vec{p}', \vec{p}'', \vec{p}''')(f_1(\vec{x}, \vec{p}'')f_1(\vec{x}, \vec{p}''') - f_1(\vec{x}, \vec{p}')f_1(\vec{x}, \vec{p}))d^3p'd^3p''d^3p'''$$

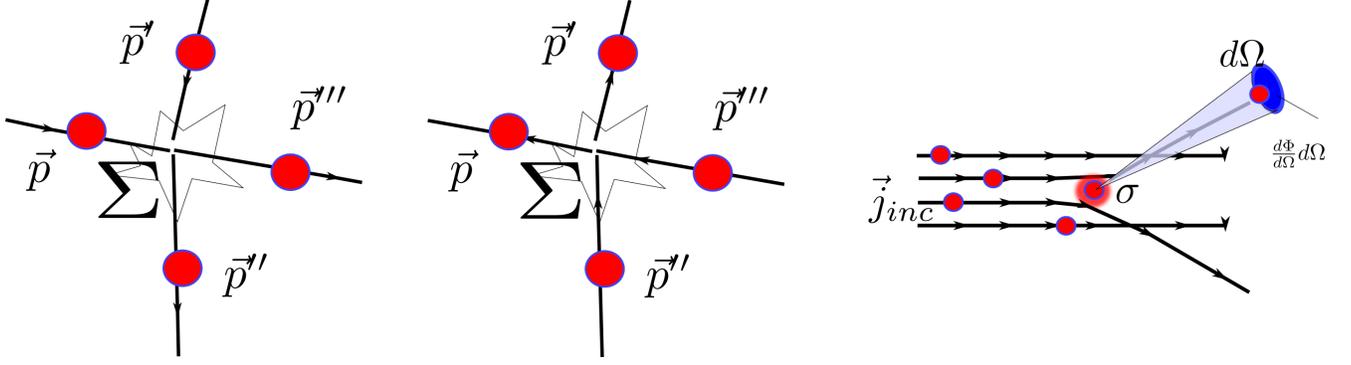


Figure 3: Colisión de dos partículas y sección eficaz. Izquierda/centro: el proceso en el que dos partículas con momentos \vec{p} y \vec{p}' coinciden en una misma región del espacio de fases y se dispersan con momentos \vec{p}'' y \vec{p}''' tiene la misma probabilidad por unidad de tiempo y volumen que el proceso inverso. Derecha: descripción del proceso en términos de una sección eficaz de dispersión. La probabilidad por unidad de tiempo de que ocurra una dispersión se expresa en términos de la sección eficaz asociada a la interacción, y la densidad de flujo incidente del resto de las partículas sobre la partícula blanco, en el sistema de referencia en que esta se encuentra en reposo.

con N el número total de partículas en el sistema, y $\Sigma(\vec{p}, \vec{p}', \vec{p}'', \vec{p}''')$ la probabilidad por unidad de tiempo de que dos partículas que se encuentran en el mismo elemento de volumen con momentos \vec{p}'' y \vec{p}''' colisionen, y terminen con momentos \vec{p} y \vec{p}' . Esta forma del término de colisiones asume que las interacciones entre las partículas son *puntuales, instantáneas, y reversibles*, en el sentido de que el proceso inverso, en el que dos partículas comienzan con momentos \vec{p} y \vec{p}' y se dispersan con momentos \vec{p}'' y \vec{p}''' tiene la misma probabilidad por unidad de tiempo.

Podemos conectar la cantidad Σ con la *sección eficaz de dispersión* para la interacción entre las partículas. Para esto, notamos que por la conservación del impulso lineal, y la simetría ante intercambio de partículas,

$$\Sigma(\vec{p}, \vec{p}', \vec{p}'', \vec{p}''') = \delta^3(\vec{p} + \vec{p}' - \vec{p}'' - \vec{p}''') \Sigma(\vec{p} - \vec{p}', \vec{p}'' - \vec{p}''')$$

con lo que es una cantidad que resulta invariante ante cambios entre sistemas de referencia inerciales. Si elegimos un sistema de referencia en el que, por ejemplo, $\vec{p}''' = 0$, podemos entender a $N f(\vec{x}, \vec{p}''') \Sigma$ como la probabilidad por unidad de tiempo de que una partícula con momento \vec{p}'' sea dispersada en la dirección de \vec{p} . Pero esto se escribe en términos de la sección eficaz diferencial como

$$\frac{dP}{dt} = |\vec{j}_{inc}| \frac{d\sigma}{d\Omega} \frac{\delta(|\Delta\vec{p}| - |\Delta\vec{p}'|)}{|\Delta\vec{p}|^2} d\Omega$$

con $\vec{j}_{inc} = N f(\vec{x}, \vec{p}''') \frac{\vec{p}'' - \vec{p}'''}{m}$ la *densidad de flujo* asociada a las partículas incidentes, y el factor en la delta de $|\Delta\vec{p}|$ imponiendo la conservación de la energía en el proceso.

Finalmente,

$$\Sigma(\Delta\vec{p}, \Delta\vec{p}') = \frac{|\vec{p}'|}{m} \frac{d\sigma}{d\Omega} \frac{\delta(|\Delta\vec{p}| - |\Delta\vec{p}'|)}{|\Delta p|^2}$$

De esta manera, la ecuación para la evolución de $f_1(\vec{x}, \vec{p})$ toma la forma de Boltzmann

$$\frac{\partial f_1}{\partial t} = -\vec{v} \cdot \nabla_{\vec{x}} f_1 + -\vec{F} \cdot \nabla_{\vec{p}} f_1 + \iint N \frac{|\vec{p}' - \vec{p}|}{m} \frac{d\sigma}{d\Omega} (f_1''' f_1'' - f_1' f_1) \frac{d\sigma}{d\omega} d\Omega d^3 p$$

con los momentos satisfaciendo las relaciones de conservación para el momento

$$\vec{p} + \vec{p}' = \vec{p}'' + \vec{p}'''$$

y la energía

$$|\vec{p} - \vec{p}'| = |\vec{p}'' - \vec{p}'''|$$

Los primeros dos términos de la ecuación, que involucran gradientes, corresponden a los llamados *términos difusivos*, y que determinan la evolución para la distribución de probabilidades de una partícula aislada, que sigue una dinámica de Liouville.

Soluciones estacionarias

Resolver la ecuación de Boltzmann en forma explícita es una tarea desafiante, ya que es una ecuación no lineal, que involucra derivadas parciales en seis coordenadas, y la evaluación de un término integral con restricciones no triviales sobre las variables. Sin embargo, sí es posible encontrar en forma más o menos directa soluciones estacionarias. Para ello, vemos que el término de colisiones se anula si en cada punto del espacio de coordenadas

$$f(\vec{x}, \vec{p})f(\vec{x}, \vec{p}') = f(\vec{x}, \vec{p}'')f(\vec{x}, \vec{p}''')$$

para cualquier elección de \vec{p} , \vec{p}' , \vec{p}'' y \vec{p}''' que cumplan las ecuaciones de conservación.

Tomando logaritmo a ambos miembros, podemos reescribir la condición sobre productos como una condición sobre sumas de cantidades:

$$\ln f(\vec{x}, \vec{p}) + \ln f(\vec{x}, \vec{p}') = \ln f(\vec{x}, \vec{p}'') + \ln f(\vec{x}, \vec{p}''')$$

que se cumple si expresamos a $\ln f$ como combinación lineal de términos escalares formados con las cantidades conservadas *aditivas*

$$\ln f_1(\vec{x}, \vec{p}) = -\beta(-\vec{V}_0 \cdot \vec{p} + \beta E(\vec{p}, \vec{x})) - \ln Z$$

con $E(\vec{x}, \vec{p}) = \frac{|\vec{p}|^2}{2m} + U(\vec{x})$ la energía mecánica de la partícula y \vec{V}_0 un vector constante y Z una constante de normalización. El factor global β permite asociar \vec{V}_0 con la *velocidad de centro de masa* del sistema, de manera que ese término puede absorberse en la energía mecánica vía un cambio de sistema de referencia inercial. Por otro lado, el signo $-$ global garantiza que si $\beta > 0$, $f(\vec{x}, \vec{p})$ se mantenga acotado. Finalmente,

$$f_1^*(\vec{x}, \vec{p}) = \frac{e^{-\beta E(\vec{x}, \vec{p})}}{Z(\beta)}$$

Como E es una *constante de movimiento*, esta elección para f_1 anula automáticamente el primer término, ya que

$$\frac{\partial f_1^*}{\partial t} + \vec{v} \cdot \nabla f_1 + \vec{F} \cdot \nabla f_1^* = 0 - \beta f_1^*(\vec{v} \cdot (\nabla_{\vec{x}} E) + \vec{F} \cdot (\nabla_{\vec{p}} E)) = -\beta f_1^*(-\vec{v} \cdot \vec{F} + \vec{F} \cdot \vec{v}) = 0$$

En ausencia de interacciones, la distribución de equilibrio f^* resulta ser la misma que encontramos usando los métodos de la teoría cinética de Maxwell, lo que motivó a llamar a esta distribución *distribución de Maxwell-Boltzmann*.

Teorema H

Otro resultado importante que podemos probar para la ecuación de Boltzmann es el llamado *Teorema H de Boltzmann*. Este teorema establece que la cantidad

$$H[f_1] = \int f_1(\xi) \ln f_1(\xi) d^6 \xi$$

es una funcional que no crece en el tiempo:

$$\frac{d}{dt} H[f_1] \leq 0$$

Notemos la semejanza de esta cantidad con la entropía estadística asociada a f_1 (salvo por un signo menos y la normalización), de manera que el enunciado del teorema es equivalente a decir que la entropía estadística de la distribución para una partícula nunca decrece, y sólo es constante para distribuciones estacionarias (esto es, en equilibrio).

Para probarlo, escribimos explícitamente la derivada temporal:

$$\frac{d}{dt}H[f_1] = \int (1 + \ln f_1(\xi)) \frac{\partial f_1(\xi)}{\partial t} d^6\xi$$

y escribimos la derivada en el integrando usando la ecuación de Boltzmann:

$$\frac{d}{dt}H[f_1] = \int (1 + \ln f_1(\xi)) \left. \frac{\partial f_1(\xi)}{\partial t} \right|_{dif} d^6\xi + \int (1 + \ln f_1(\xi)) \left. \frac{\partial f_1(\xi)}{\partial t} \right|_{col} d^6\xi$$

con

$$\left. \frac{\partial f_1(\xi)}{\partial t} \right|_{dif} = -\frac{\vec{p}}{m} \cdot \nabla_{\vec{x}} f_1 - \vec{F} \cdot \nabla_{\vec{p}} f_1$$

la contribución a la derivada de f de los términos de *difusión* y

$$\left. \frac{\partial f_1(\xi)}{\partial t} \right|_{col} = \iiint N \Sigma (f'' f''' - f f') d^3 p' d^3 p'' d^3 p'''$$

la contribución a $\partial f / \partial t$ que viene de los términos de colisión.

Analicemos primero la contribución del término difusivo. Para esto, notamos que

$$\int (1 + \ln f_1(\xi)) \left. \frac{\partial f_1(\xi)}{\partial t} \right|_{dif} d^6\xi = - \int (1 + \ln f_1(\xi)) \xi \nabla f_1 d^6\xi = - \int \xi \cdot \nabla (f_1 \ln f_1) d^6\xi$$

Pero, integrando por partes,

$$\int \xi \cdot \nabla (f_1 \ln f_1) d^6\xi = \int \nabla \cdot \xi (f_1 \ln f_1) d^6\xi - \int (f_1 \ln f_1) (\nabla \cdot \xi) d^6\xi$$

Pero en una evolución Liuvilliana, $\nabla \cdot \xi = 0$:

$$\nabla \cdot \xi = \nabla_{\vec{x}} \nabla_{\vec{p}} H - \nabla_{\vec{p}} \nabla_{\vec{x}} H = 0$$

por lo que en una evolución exclusivamente difusiva, $H[f]$ es una constante de movimiento.

Por otro lado, la contribución del término de colisiones resulta ser

$$\left. \frac{dH}{dt} \right|_{col} = \int (1 + \ln f) \Sigma (f'' f''' - f f') d^3 p d^3 p' d^3 p'' d^3 p'''$$

Pero salvo por el factor $(1 + f)$, el integrando es una función simétrica ante el intercambio $f \leftrightarrow f'$ y antisimétrica ante los intercambios $f \leftrightarrow f''$ y $f \leftrightarrow f'''$, de manera que

$$\begin{aligned} \left. \frac{dH}{dt} \right|_{col} &= \frac{1}{4} \int ((1 + \ln f) + (1 + \ln f') - (1 + \ln f'') - (1 + \ln f''')) \Sigma (f'' f''' - f f') d^3 p d^3 p' d^3 p'' d^3 p'''$$

$$= \frac{1}{4} \int \Sigma \ln \left(\frac{f f'}{f'' f'''} \right) (f'' f''' - f f') d^3 p d^3 p' d^3 p'' d^3 p'''$$

La derivada resulta entonces numéricamente igual a la integral del producto de Σ por una cantidad de la forma $\ln(a/b)(b-a)$, que no puede ser positiva. Pero Σ representa una probabilidad por unidad de tiempo, y por lo tanto es una cantidad positiva. Luego el integrando siempre es menor o igual a cero, y por lo tanto la integral también lo es, lo que prueba el teorema.

Estados asintóticos. Principio de Máxima entropía.



Edwin Jaynes

1922-1998

El teorema H establece que debido a las interacciones de una partícula con el resto del sistema, la funcional H sólo puede decrecer. Por otro lado, como f es una cantidad positiva, H sólo puede decrecer disminuyendo el valor de $\ln f$. Pero como además la integral de f está normalizada, si f decrece en una región, tiene que crecer en otra. Como resultado, a medida que H decrece, f se *desparrama* más y más. Pero esto eventualmente conlleva un aumento del volumen y la energía mecánica media del sistema. Luego, una consecuencia del teorema H es que el sistema eventualmente alcanzará un estado asintótico estacionario, en el que H toma el valor mínimo posible, sujeto a las restricciones impuestas por los valores de la energía total y el volumen disponible.

Veamos que efectivamente la distribución de Maxwell-Boltzmann satura $H[f]$. Para eso, vamos a buscar una distribución f que cumpla con las siguientes restricciones:

$$\int f d^3\xi = 1 \quad \text{normalización}$$
$$\int_V f(\xi) f d^3\xi = 1 \quad \text{restricción sobre el volumen}$$
$$\int_V f(\xi) E(\xi) d^3\xi = \frac{E_{total}}{N}$$

Utilizando la técnica de los multiplicadores de Lagrange, podemos buscar la distribución que minimiza $H[f]$ sujeta a las restricciones pidiendo que la variación de la funcional

$$\frac{\delta}{\delta f} \left\{ H[f] + \lambda_1 \int f(\xi) d^3\xi + \lambda_2 \int_V f(\xi) d^3\xi + \beta \int_V f(\xi) E(\xi) d^3\xi \right\}$$

se anule. Fuera de V , la condición de extremo implica

$$1 + \ln f(\xi) = \lambda_1$$

que se satisface si $f(\xi)$ es *constante*. Pero esa constante tiene que ser necesariamente 0, ya que de otro modo $H[f]$ no sería finita. Luego, podemos eliminar el constraint ($\lambda_1 = 0$) y seguir adelante considerando sólo funciones con soporte en V .

Dentro de V , la condición estacionaria implica que

$$1 + \ln f(\xi) + \lambda_2 + \beta E(\xi) = 0$$

lo que implica que $f(\xi) = \exp(-\beta E(\xi) - (1 + \lambda_2)) = \frac{e^{-\beta E}}{Z}$ que no es otra cosa que la distribución de Maxwell-Boltzmann.

Inspirado en estos resultados, y en la identificación que hizo Gibbs entre la entropía termodinámica y la entropía estadística, Edwin Janes propuso en 1957 que independientemente de la dinámica que genera una distribución estadística, es natural asignar a un sistema aquella distribución que maximice la entropía estadística sujeta a las restricciones impuestas por la información disponible, en forma de valores medios de cantidades. Este enunciado, establecido al nivel de principio de la estadística, se conoce como *principio de Jaynes*, *principio de Máxima Entropía* o *principio Max-Ent*.