

Física Teórica 3 – 1er. cuatrimestre de 2019

Guía 3: El problema de los tres niveles

Problema 4. Un sistema está compuesto por N elementos distinguibles y no interactuantes, cada uno de los cuales puede tener tres valores de energía, $-\epsilon$, 0 y ϵ . Encuentre $S(T)$ y $U(T)$ por dos caminos alternativos: i) usando el ensamble microcanónico (¡dolor!), ii) usando el ensamble canónico.

Aquí nos concentraremos en el cálculo de la entropía en ambos ensambles.

En el ensamble microcanónico

A diferencia del problema de los dos niveles, ahora el valor de la energía no determina de manera única la población de cada nivel. Si las poblaciones de los niveles son A , B y C , donde A corresponde al nivel con energía ϵ , B al nivel con energía cero y C al nivel con energía $-\epsilon$, entonces debe cumplirse que

$$A + B + C = N, \quad \epsilon(A - C) = E, \quad (1)$$

donde E es la energía. Para cada elección posible de A , B y C habrá

$$\Omega(A, B, C) = \frac{N!}{A! B! C!} \quad (2)$$

microestados diferentes. El número total de microestados se obtiene considerando todas las ternas $\{A, B, C\}$ compatibles con las condiciones (1):

$$\Omega(E, N) = \sum' \Omega(A, B, C) = \sum' \frac{N!}{A! B! C!}, \quad (3)$$

donde la sumatoria primada indica que la suma está restringida por (1). Entonces, ¿cuáles son los valores posibles de A , B y C ? Podemos eliminar dos de los números de la terna en términos del restante y construir una suma ordinaria. Veremos que eso no simplifica mucho el problema de calcular la suma.

Si despejamos C y B en términos de A (que no parece ser lo más simétrico pero que lleva al resultado más simple), la suma triple se reducirá a una suma sobre A . Tenemos

$$B = N - 2A + M, \quad C = A - M, \quad (4)$$

donde

$$M = \frac{E}{\epsilon} \quad (5)$$

es el número de *cuantos* de energía. Con esta definición las condiciones (1) se leen como

$$A + B + C = N, \quad A - C = M. \quad (6)$$

Supongamos primero que $M \geq 0$. Entonces el mínimo valor que puede tomar A es M , con $B = N - M$ y $C = 0$. Cada vez que aumentamos A en una unidad, B disminuye en una unidad y C aumenta en la misma proporción. Estos números no pueden ser negativos, de modo que las relaciones (4) implican $2A \leq N + M$, y $M \leq A$. La segunda desigualdad era conocida. La primera se lee como

$$A \leq \left[\frac{1}{2} (N + M) \right], \quad (7)$$

donde los corchetes indican que debe tomarse la parte entera: si $N + M$ es par, no hay duda de que A puede llegar a tomar el valor entero $\frac{1}{2}(N + M)$, que coincide con su parte entera. En cambio, si $N + M = 2n + 1$, la desigualdad

$$A \leq \frac{N + M}{2} = n + \frac{1}{2} \quad (8)$$

implica que el máximo valor entero que puede tomar A es n , que es la parte entera de $\frac{1}{2}(N + M)$. En el primer caso el valor mínimo de B es cero y en el segundo es 1.

En resumen, luego de eliminar B y C , y si $M \geq 0$, la sumatoria (3) se lee como

$$\Omega(E, N) = \sum_{A=M}^{\left[\frac{1}{2}(N+M) \right]} \frac{N!}{A! (N - 2A + M)! (A - M)!}. \quad (9)$$

Si M fuera menor que cero no es necesario rehacer el cálculo. Los números de estados con M y $-M$ son iguales. Para verlo, supongamos $M < 0$, entonces $M = -|M|$, con lo que las ecuaciones (6) toman la forma

$$A + B + C = N, \quad C - A = |M|. \quad (10)$$

Alcanza con intercambiar los nombres A y C para reducir el problema de conteo al del caso con $M > 0$, después de todo la Ec. (3) es simétrica respecto de ese intercambio.

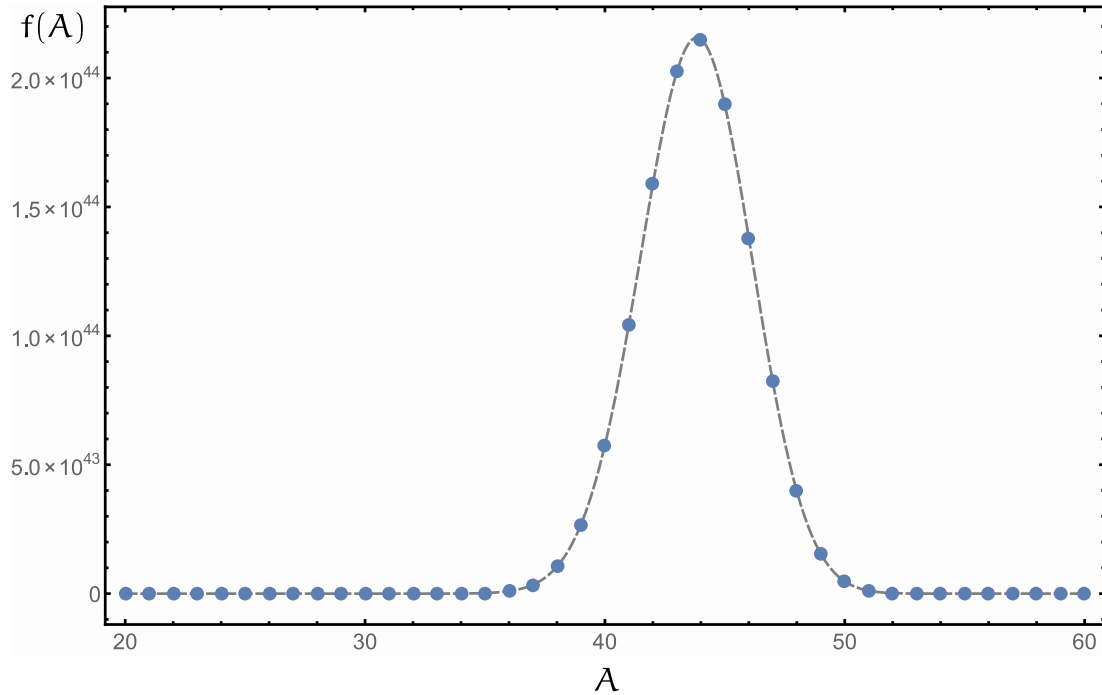
Podemos concluir entonces que el número de microestados con M cuantos de energía es, en general,

$$\Omega(E, N) = \sum_{A=|M|}^{\left[\frac{1}{2}(N+|M|) \right]} \frac{N!}{A! (N - 2A + |M|)! (A - |M|)!}. \quad (11)$$

La forma de evaluar este tipo de sumas, a primera vista intratables, consiste en aprovecharse de los grandes números involucrados, por lo que vamos a asumir que N, A, B y C son todos mucho mayores que uno. Ustedes recordarán que los números binomiales $\binom{N}{n}$, graficados como funciones de la variable n , tienden a ser campanas muy agudas, con máximos de orden e^N y dispersiones de orden N . Del mismo modo

$$f(A) = \frac{N!}{A! (N - 2A + |M|)! (A - |M|)!} \quad (12)$$

dibuja una campana que decrece rápidamente a ambos lados del máximo. La figura muestra como ejemplo el caso $N = 100, M = 20$.



En general, ustedes pueden construir la gaussiana aproximada y verificar que tiene un máximo en

$$A^* = \frac{N}{6} \left[4 + 3 \frac{M}{N} - \sqrt{4 - 3 \left(\frac{M}{N} \right)^2} \right], \tag{13}$$

que en el máximo alcanza un valor de orden e^N y que la dispersión es de orden N . En ese sentido no se trata de nada muy distinto de los números binomiales usuales. Los parámetros de la gaussiana son más complicados, pero el resultado cualitativo es el mismo.

Tratemos ahora de aproximar la suma (11). Es trivial darse cuenta de que la suma estará acotada por

$$f(A^*) \leq \Omega(E, N) \leq \mathcal{O}(N) f(A^*). \tag{14}$$

Esto se entiende así: si el único sumando existente fuera el término correspondiente al máximo, entonces la suma sería igual a $f(A^*)$. De ahí la cota inferior en la desigualdad (14). Pero el término máximo no es el único de la suma. En el peor de los casos puede haber hasta $\mathcal{O}(N)$ términos, todos iguales al valor máximo. De ahí la cota superior en la desigualdad (14). (No es tan importante si son N o $(N - M)/2$ o $2N$ sumandos, lo que importa es el orden de magnitud, que no se verá afectado siempre que $|M|$ no sea demasiado cercano a N). Ahora bien, la cantidad relevante en el cálculo de la entropía no es Ω por sí misma, sino su logaritmo. La desigualdad anterior se traslada al logaritmo como

$$\log f(A^*) \leq \log \Omega(E, N) \leq \mathcal{O}(\log N) + \log f(A^*). \tag{15}$$

Pero, según dijimos, $f(A^*)$ es de orden e^N , de modo que $\log f(A^*)$ es de orden N . La desigualdad (15) acota el logaritmo de Ω entre dos números de orden N que difieren en una cantidad de orden $\log N$. Puesto que estamos asumiendo que $N \gg 1$, $\log N$ es mucho menor que N . A todos los fines prácticos, la desigualdad (15) se puede reemplazar por

$$\log f(A^*) \leq \log \Omega(E, N) \leq \log f(A^*). \quad (16)$$

En otras palabras, el logaritmo de Ω es, a todos los fines prácticos, igual al logaritmo del término más grande en la suma (11).

Notar que estas afirmaciones tan extraordinarias son válidas para el logaritmo, pero no para Ω . Para Ω sólo logramos una cota que está entre $f(A^*)$ y $\mathcal{O}(N)f(A^*)$. El margen de error relativo es de orden $N \gg 1$. En cambio, el error relativo en la aproximación $\log \Omega \approx \log f(A^*)$ es de orden $(\log N)/N \ll 1$.

Para conjurar el misterio tomemos $N = 10^6$, un número mucho mayor que 1, pero todavía órdenes de magnitud por debajo del número de Avogadro. La desigualdad para Ω se leería como

$$e^{10^6} \leq \Omega \leq 10^6 e^{10^6}. \quad (17)$$

El error relativo puede ser tan alto como 10^6 . En cambio, para los logaritmos tenemos

$$10^6 \leq \log \Omega \leq 6 \log 10 + 10^6,$$

$$1\,000\,000 \leq \log \Omega \leq 1\,000\,014. \quad (18)$$

El error relativo es de orden 10^{-5} . Pasamos de un error relativo de 10^6 , para la función, a otro de 10^{-5} , para su logaritmo.

Otro ejemplo más exagerado es considerar la diferencia entre los logaritmos de Ne^N y e^N cuando $N = 10^{23}$. Resulta

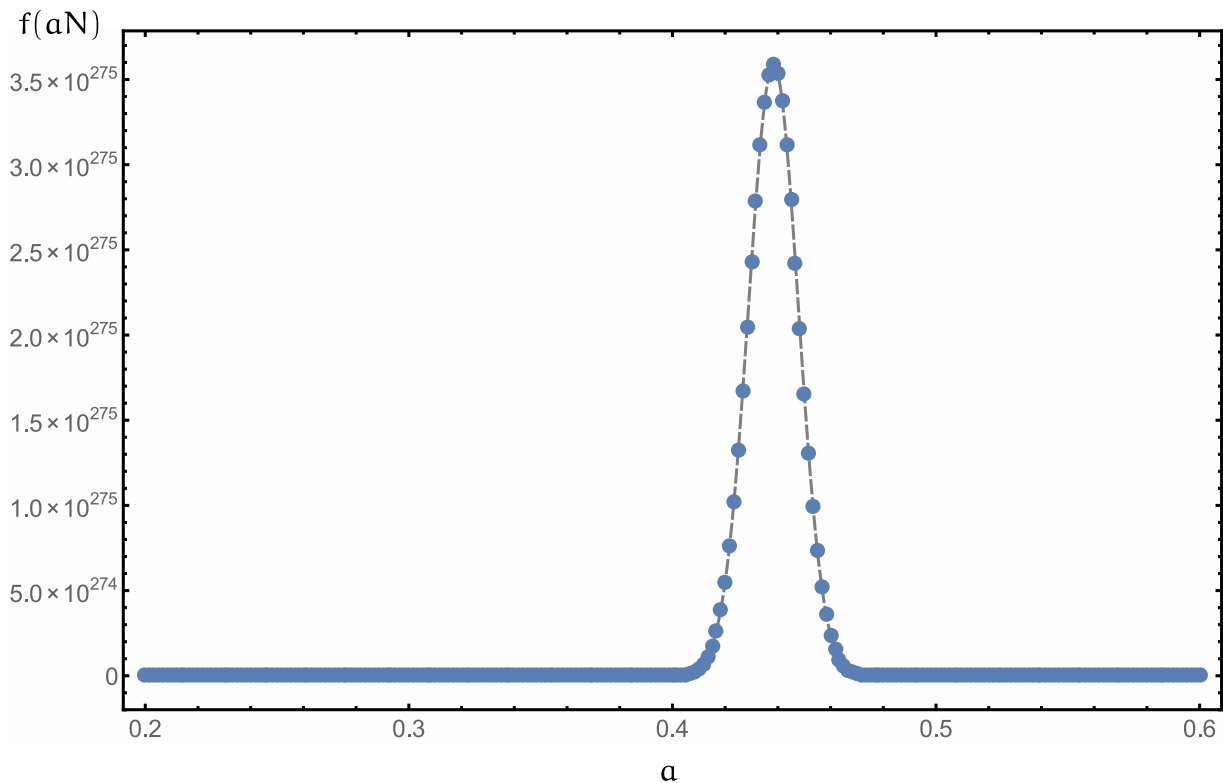
$$\log(Ne^N) = \log N + N = 23 \log 10 + 10^{23} \simeq 53 + 10^{23}, \quad (19)$$

que debe compararse con $\log(e^N) = 10^{23}$. La diferencia es de apenas 53 partes en 10^{23} . Para la función logaritmo da lo mismo $10^{23} e^{10^{23}}$ que $e^{10^{23}}$. Noten cómo $10^{23} e^{10^{23}} \gg e^{10^{23}}$, pero $\log(10^{23} e^{10^{23}}) \simeq \log(e^{10^{23}})$.

Resumiendo: hemos argumentado que $\log \Omega(E, N)$ puede aproximarse por $\log f(A^*)$ con un error despreciable, donde A^* es el valor de A que maximiza el sumando en la Ec. (11), es decir, que maximiza la función $f(A)$ de la Ec. (14).

Como es habitual, tratamos este problema de maximización como si A fuera una variable continua. Es más sencillo convencerse de que este paso al continuo está justificado si reescribimos todo en términos de las fracciones de las poblaciones de cada nivel relativas a N . Esto normaliza el intervalo de definición de las variables, independizándolo de N . Las fracciones A/N , B/N y C/N varían entre 0 y 1 y cada par de valores consecutivos está separado por intervalos de longitud $1/N$. Cuanto más grande sea N , más densamente

ocupado estará el intervalo de definición de cada variable. El gráfico de cualquier función de los números $a = A/N$, $b = B/N$ y $c = C/N$ estará formado por un conjunto discreto de puntos, tan estrechamente juntos, que podremos ubicarlos sobre la gráfica de una función suave. El paso al continuo consiste en trabajar con esta función. Así podemos derivar e integrar y resolver problemas de extremos sin mayor dificultad. El carácter discreto de los puntos se pierde cuando $N \gg 1$. La figura siguiente muestra el gráfico discreto de los puntos y la curva continua de $f(aN)$ con el intervalo escaleado entre 0 y 1. Se ha tomado $N = 600$ y $m = 120$.



Observen que, no obstante guardar la misma relación $N/m = 5$ que la figura anterior, el hecho de haber aumentado el número de elementos de 100 a 600 no sólo ha vuelto más denso el conjunto de puntos del gráfico discreto, sino que ha reducido el ancho relativo de la campana.

Volviendo al problema de maximizar el sumando en la Ec. (11). Hay al menos dos alternativas para resolver el problema de extremos. A una de las alternativas ya la exploramos, aunque sin dar los detalles: consistió en extremar directamente la función de una variable $f(A)$. Este camino puede dar una sensación de seguridad en las cuentas, al precio de ser bastante tedioso. Más que tedioso, en realidad. Y una cosa es calcular el A^* de la Ec. (13), algo relativamente sencillo, y otra muy distinta es evaluar $f(A^*)$, cosa que nos abstuvimos de hacer. Dijimos que había más de una alternativa para resolver el problema de extremos. Una de las otras alternativas consiste en considerar un problema de extremos condicionados para la función original

$$F(A, B, C) = \frac{N!}{A! B! C!} \tag{20}$$

La idea es resolver el problema de extremos para F , sujeto a las condiciones (6), mediante el método de los multiplicadores de Lagrange. Este camino resulta mucho más práctico.

Debido a que hay factoriales involucrados, el problema se simplifica si en lugar de F trabajamos con $\log F$. Usando la aproximación de Stirling, es

$$\log F(A, B, C) = N \log N - (A \log A + B \log B + C \log C) + (A + B + C - N). \quad (21)$$

Pueden verificar que si definimos $a = A/N$, $b = B/N$ y $c = C/N$, la función que hay que extremar es

$$\begin{aligned} g(a, b, c) &= \frac{1}{N} \log F(aN, bN, cN) \\ &= (a + b + c - 1)(1 - \log N) - (a \log a + b \log b + c \log c) \\ &= (\log N - 1) - (a + b + c) \log N - (a \log a - a + b \log b - b + c \log c - c) \end{aligned} \quad (22)$$

Noten que hemos preservado la simetría entre las tres variables, que es lo primero que se perdería si eliminásemos dos de ellas en términos de la restante. Noten, además, que no hemos utilizado ninguna de las ecuaciones de vínculo. Hasta aquí las variables a , b y c son tratadas como independientes. El problema de extremos para $g(a, b, c)$ está condicionado por las ecuaciones

$$a + b + c = 1, \quad a - c = m, \quad (23)$$

donde hemos introducido la variable reducida

$$m = \frac{M}{N} = \frac{E}{\epsilon N}. \quad (24)$$

Con el propósito de usar el método de los multiplicadores de Lagrange, escribimos primero las ecuaciones que resultan de variar g y las condiciones de vínculo:

$$\begin{aligned} (\log aN) \delta a + (\log bN) \delta b + (\log cN) \delta c &= 0, \\ \delta a + \delta b + \delta c &= 0, \\ \delta a - \delta c &= 0. \end{aligned} \quad (25)$$

Introduciendo multiplicadores de Lagrange λ y α , las tres ecuaciones que resultan son

$$\lambda + \alpha + \log aN = 0, \quad \lambda + \log bN = 0, \quad \lambda - \alpha + \log cN = 0. \quad (26)$$

Ya que $\log N$ es una constante, podemos redefinir $\lambda + \log N \rightarrow \lambda$ y quedarnos con el sistema de ecuaciones

$$\lambda + \alpha + \log a = 0, \quad \lambda + \log b = 0, \quad \lambda - \alpha + \log c = 0, \quad (27)$$

sujeto a los vínculos (23). De aquí resultan

$$a = e^{-\lambda-\alpha}, \quad b = e^{-\lambda}, \quad c = e^{-\lambda+\alpha}. \quad (28)$$

Es útil definir $x = e^{-\lambda}$ y $y = e^{-\alpha}$, de modo que las ecuaciones de arriba se lean como

$$a = xy, \quad b = x, \quad c = \frac{x}{y}. \quad (29)$$

Cuando volvemos a escribir la función $g(a, b, c)$ evaluada en el extremo, queda

$$\text{máx}\{g(a, b, c)\} = -(a \log a + b \log b + c \log c) = (\lambda + \alpha)a + \lambda b + (\lambda - \alpha)c = \lambda + m\alpha. \quad (30)$$

En la última igualdad hemos aplicado las condiciones $a + b + c = 1$ y $a - c = m$. Finalmente,

$$\text{máx}\{g(a, b, c)\} = -\log x - m \log y. \quad (31)$$

El sistema de ecuaciones (23), ahora reescrito en términos de x y y es

$$x \left(y + 1 + \frac{1}{y} \right) = 1, \quad x \left(y - \frac{1}{y} \right) = m. \quad (32)$$

Los valores de x e y que aparecen en la Ec. (31) son las soluciones mayores que cero del sistema (32). Dividiendo entre sí las dos ecuaciones, resulta una ecuación cuadrática:

$$(1 - m)y^2 - my - (1 + m) = 0, \quad (33)$$

de donde se obtiene

$$\begin{aligned} x &= \frac{-1 + \sqrt{4 - 3m^2}}{3}, & \frac{1}{x} &= \frac{\sqrt{4 - 3m^2} + 1}{1 - m^2}, \\ y &= \frac{m + \sqrt{4 - 3m^2}}{2(1 - m)}, & \frac{1}{y} &= \frac{\sqrt{4 - 3m^2} - 1}{2(m + 1)}, \end{aligned} \quad (34)$$

y pueden verificar que son resultados válidos independientemente del signo de m .

Como último paso, reuniendo todos los resultados intermedios, escribimos

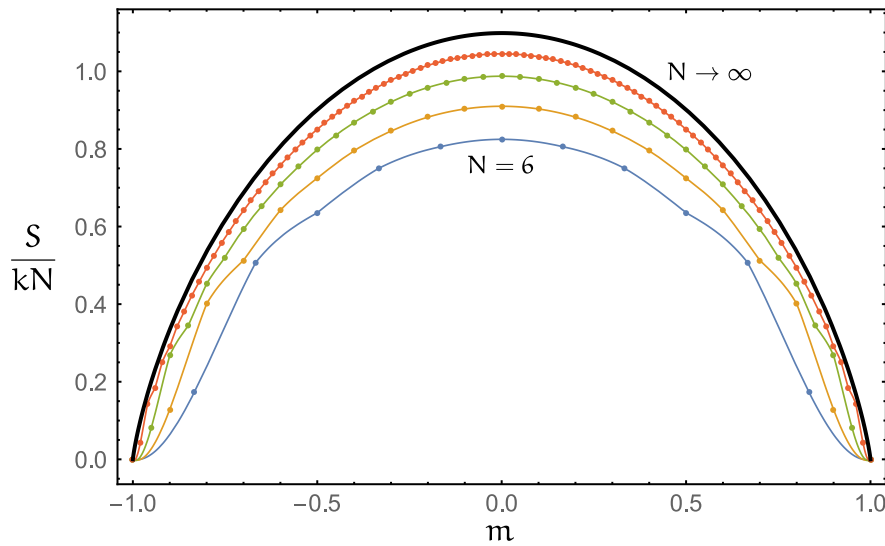
$$\begin{aligned} \frac{1}{N} \log \Omega(E, N) &\approx \frac{1}{N} \log \text{máx}\{F(A, B, C)\} = \\ &\log \text{máx}\{g(a, b, c)\} = \log \left(\frac{\sqrt{4 - 3m^2} + 1}{1 - m^2} \right) + m \log \left(\frac{\sqrt{4 - 3m^2} - m}{2(m + 1)} \right). \end{aligned} \quad (35)$$

La entropía está dada entonces por

$$\frac{S(m, N)}{kN} = \frac{1}{N} \log \Omega(E, N) = \log \left(\frac{\sqrt{4 - 3m^2} + 1}{1 - m^2} \right) + m \log \left(\frac{\sqrt{4 - 3m^2} - m}{2(m + 1)} \right), \quad (36)$$

donde el signo aproximado ha sido reemplazado por el igual, de acuerdo a lo que ya discutimos.

Entre otras cosas, es fácil verificar que S en la Ec. (36) es una función par, como debe ser, ya que el cambio $E \rightarrow -E$ o $m \rightarrow -m$ equivale a cambiar ϵ por $-\epsilon$, que a su vez equivale al intercambio de las etiquetas A y C en las poblaciones de los niveles. Obviamente un simple cambio de nombres no puede modificar el resultado. El máximo de S/kN se alcanza en $m = 0$ (ese es el macroestado con mayor multiplicidad) y vale $\log 3$, lo que no es casualidad. El gráfico de la función $S(m, N)/kN$ se muestra más abajo, junto con los resultados de haber calculado S mediante la suma exacta (11) para $N = 6, 10, 20$ y 50 . La curva en trazo negro es la solución (36), que representa el límite termodinámico del sistema. Las curvas que unen los puntos de los casos discretos son interpolaciones sin mayor significado y sólo están para facilitar la lectura de la figura.



Para concluir con esta sección, un punto para notar, ya que es de carácter muy general, es que podemos escribir

$$\Omega(E, N) = e^{NI(m)}. \quad (37)$$

En primer lugar, tal como anticipamos, $\Omega(E, N)$ es de orden e^N . En segundo lugar, la función $I(m)$ es lo que llamaríamos una cantidad intensiva. Sólo depende del cociente $m = E/\epsilon N$. Esto hace que $S = NI(m)$ sea extensiva. Un comportamiento asintótico para $N \gg 1$ del tipo del de la Ec. (37) es típico de la Teoría de grandes desviaciones. La función $I(m)$, con signo menos, recibe el nombre de *rate function*.

En el ensamble canónico

La función de partición en el ensamble canónico es trivial si se la escribe adecuadamente. Contemplemos varias alternativas:

$$Z(\beta, N) = \sum_{\text{estados}} e^{-\beta E(\text{estado})} = \sum_{E=-\epsilon N}^{\epsilon N} e^{-\beta E} \Omega(E, N). \quad (38)$$

La primera igualdad es la definición formal: una suma sobre todos los estados del sistema pesada por el factor de Boltzmann. En esa suma se pueden agrupar todos los estados que

tienen la misma energía. Para una dada energía, existen $\Omega(E, N)$ estados posibles. La suma sobre los estados se reemplaza por una suma sobre los valores de la energía, pesada por el número $\Omega(E, N)$ y por el factor de Boltzmann.

La suma sobre estados también puede desarrollarse como una suma sobre los valores de las poblaciones de cada estado:

$$\begin{aligned} Z(\beta, N) &= \sum_{A,B,C=0}^N \frac{N!}{A! B! C!} e^{-\beta(A-C)\epsilon} \delta_{A+B+C,N} \\ &= \sum_{A,B,C=-\infty}^{\infty} \frac{N!}{A! B! C!} e^{-\beta(A-C)\epsilon} \delta_{A+B+C,N}. \end{aligned} \quad (39)$$

La primera igualdad es la forma más cruda de escribirlo. Si las poblaciones son A , B y C , la forma en que pueden elegirse los elementos del sistema respetando esos números es el número de permutaciones con repetición,

$$\frac{N!}{A! B! C!}. \quad (40)$$

Poco empeño se pone en acotar el intervalo de variación de cada población. De ese trabajo nos libera la delta de Kronecker. Aún no es tan fácil ver el resultado de la suma. La segunda igualdad en la Ec. (39) es una forma aún más despojada, pero resulta práctica en ocasiones: los límites de la suma se han liberado, extendiéndolos entre menos y más infinito. En primer lugar, mover el extremo inferior hasta menos infinito no tiene ningún efecto, ya que los factoriales de los enteros negativos valen $\pm\infty$ y esos términos, por lo tanto, se anulan. Por otro lado, mover los límites de las sumas hasta infinito tampoco tiene ninguna consecuencia, porque la delta se encarga de que ninguna de las poblaciones supere N . Esta forma es útil como paso intermedio para hacer cambios de las variables de sumación, ya que los límites se arreglan solos. Hay autores (por ejemplo Knuth) que escribirían simplemente

$$Z(\beta, N) = \sum_{A,B,C} \frac{N!}{A! B! C!} e^{-\beta(A-C)\epsilon} \delta_{A+B+C,N}. \quad (41)$$

La siguiente manera de escribir la función de partición es más amable:

$$Z(\beta, N) = \sum_{A=0}^N \sum_{B=0}^{N-A} \binom{N}{A} \binom{N-A}{B} e^{-\beta(2A+B-N)\epsilon}. \quad (42)$$

Se ha usado la condición $A + B + C = N$ para eliminar C , y la condición $C \geq 0$ se ha usado para fijar el límite superior de la suma sobre B . Esta forma ya podría evaluarse de manera inmediata, pues es el desarrollo multinomial de

$$(e^{-\beta\epsilon} + 1 + e^{\beta\epsilon})^N. \quad (43)$$

En efecto,

$$(e^{-\beta\epsilon} + 1 + e^{\beta\epsilon})^N = \sum_{A=0}^N \sum_{B=0}^{N-A} \binom{N}{A} \binom{N-A}{B} e^{-\beta A\epsilon} \times 1^B \times e^{\beta(N-A-B)\epsilon}, \quad (44)$$

que se reduce a la Ec. (42). Con el resultado a la vista nos damos cuenta de que podríamos haber usado ya desde la Ec. (39) el desarrollo multinomial en su forma más general:

$$\sum_{n_1, \dots, n_N} \frac{N!}{n_1! \dots n_N!} x_1^{n_1} \dots x_N^{n_N} \delta_{\sum n_i, N} = (x_1 + \dots + x_N)^N. \quad (45)$$

Finalmente, una forma por completo diferente de organizar la suma sobre estados consiste en ir variando por separado los estados de cada elemento. El primer elemento puede estar en tres estados, $n_1 = -1, 0, 1$, cada uno con energía ϵn_1 . Lo mismo para cada elemento. La suma sobre los estados del sistema es la suma sobre todas las posibles combinaciones de los n_i . Es decir,

$$\begin{aligned} Z(\beta, N) &= \sum_{n_1=-1}^1 \sum_{n_2=-1}^1 \dots \sum_{n_N=-1}^1 e^{-\beta(n_1+n_2+\dots+n_N)\epsilon} \\ &= \left(\sum_{n_1=-1}^1 e^{-\beta n_1 \epsilon} \right) \left(\sum_{n_2=-1}^1 e^{-\beta n_2 \epsilon} \right) \dots \left(\sum_{n_N=-1}^1 e^{-\beta n_N \epsilon} \right) \\ &= \left(\sum_{n=-1}^1 e^{-n\beta\epsilon} \right)^N = (e^{-\beta\epsilon} + 1 + e^{\beta\epsilon})^N. \end{aligned} \quad (46)$$

Los números n_i no están ligados por ninguna condición. Eso hace que la suma se factorice: los elementos son independientes. La suma sobre el estado de cada elemento es igual para todos los elementos. Eso hace que todos los factores en la segunda línea de la Ec. (46) sean idénticos. El resultado final está a la vista. Ahora haremos la conexión con la termodinámica.

La relación entre la mecánica estadística y la termodinámica viene dada a través de

$$F(T, N) = -kT \log Z(\beta, N). \quad (47)$$

Para simplificar la notación definiremos $z = e^{-\beta\epsilon}$, con lo que resulta

$$Z(\beta, N) = (z + 1 + 1/z)^N. \quad (48)$$

y

$$F(T, N) = -NkT \log(z + 1 + 1/z). \quad (49)$$

La entropía es la derivada parcial de F respecto de T , con signo menos,

$$\begin{aligned} \frac{S_C(T, N)}{kN} &= \log(z + 1 + 1/z) + T \frac{1 - 1/z^2}{z + 1 + 1/z} \frac{\partial z}{\partial T} \\ &= \log(z + 1 + 1/z) + \frac{z - 1/z}{z + 1 + 1/z} \frac{\epsilon}{kT} \\ &= \log(z + 1 + 1/z) - \frac{z - 1/z}{z + 1 + 1/z} \log z. \end{aligned} \quad (50)$$

El subíndice "C" es de canónico. El último término en esta ecuación puede relacionarse con la energía media por elemento, debido a que

$$\frac{\langle E \rangle}{N} = -\frac{1}{N} \frac{\partial \log Z(\beta, N)}{\partial \beta} = \frac{z - 1/z}{z + 1 + 1/z} \epsilon = -kT \frac{z - 1/z}{z + 1 + 1/z} \log z. \quad (51)$$

Entonces,

$$\frac{S_C(T, N)}{kN} = \log(z + 1 + 1/z) + \frac{\langle E \rangle}{kNT}. \quad (52)$$

Antes de seguir, conviene notar una cosa: en una segunda mirada, este tipo de resultados tiene que resultar evidente. Estamos calculando S a partir de F . Por definición $F = U - TS$, de manera que $S = (U - F)/T$. Debido a que $F/kNT = -\log(z + 1 + 1/z)$, la Ec. (52) no está diciendo otra cosa que

$$S_C(\beta, N) = \frac{-F + \langle E \rangle}{T}. \quad (53)$$

Volviendo al cálculo principal. Si definimos, análogamente a la Ec. (24),

$$\langle m \rangle = \frac{\langle E \rangle}{\epsilon N} = \frac{z - 1/z}{z + 1 + 1/z}, \quad (54)$$

que es el número medio por elemento de cuantos de energía en el sistema, y reemplazamos en la Ec. (50), queda

$$\frac{S_C(T, N)}{kN} = \log\left(z + 1 + \frac{1}{z}\right) - \langle m \rangle \log z. \quad (55)$$

Si queremos relacionar esta entropía con el resultado del microcanónico tenemos que dejar escrito todo en las mismas variables. Puesto que en el ensamble microcanónico calculamos $S(E, N)$, una opción es eliminar ahora T en favor de la energía media y dejar escrito $S_C(\langle E \rangle, N)$, y comparar entonces con $S(E, N)$. El valor medio $\langle m \rangle$ es igual a $\langle E \rangle / \epsilon N$, así que no necesitamos tocarlo, ya es una función de $\langle E \rangle$. Dejando de lado $\langle m \rangle$, lo único que depende de T es z , así que debemos despejar z en términos de $\langle E \rangle$. Eso se hace a través de la ecuación (54): z debe ser la solución positiva de la cuadrática

$$(1 - \langle m \rangle)z^2 - \langle m \rangle z - (1 + \langle m \rangle). \quad (56)$$

Salvo un detalle, es la misma ecuación cuadrática que encontramos en el microcanónico, Ec. (33). La reescribimos aquí abajo resaltando el hecho de que en cada ensamble el símbolo m significa cosas distintas. En el ensamble microcanónico nos encontramos con la ecuación

$$(1 - m_\mu)y^2 - m_\mu y - (1 + m_\mu) = 0. \quad (57)$$

El subíndice “ μ ” es por microcanónico. Reservamos el símbolo m para la variable aleatoria asociada al ensamble canónico. La Ec. (31), que daba S/kN en el microcanónico, puede ser reescrita usando la primera Ec. (32) como

$$\frac{S(m_\mu, N)}{kN} = \text{máx}\{g(a, b, c)\} = \log\left(y + 1 + \frac{1}{y}\right) - m_\mu \log y, \quad (58)$$

donde y es la solución positiva de la cuadrática (57). Comparando con la Ec. (55), vemos que la entropía en el ensamble canónico, como función de la energía, se escribe de la misma forma que la entropía en el ensamble microcanónico. Lo que cambia es que z es la solución de la cuadrática en donde aparece $\langle m \rangle$, mientras que y es la solución de la cuadrática en donde aparece m_μ . Dicho en otras palabras, las entropías en el microcanónico y en el canónico vienen dadas exactamente por la misma función, sólo que están evaluadas la primera en m_μ (que está exactamente definida) y la otra en $\langle m \rangle$ (donde m es una variable aleatoria). Es decir, en tanto que funciones de su primera variable,

$$S_C(x, N) = S(x, N). \quad (59)$$

Si la temperatura del sistema en el ensamble canónico se ajusta de tal modo que $m_\mu = \langle m \rangle$, ambos ensambles tendrán la misma entropía.

El microcanónico a partir del canónico

Está claro cómo pasar del ensamble microcanónico al canónico. Conocida la función $\Omega(E, N)$ en el microcanónico, la función de partición canónica se obtiene mediante la fórmula

$$Z(\beta, N) = \sum_E \Omega(E, N) e^{-\beta E}. \quad (60)$$

Ahora veremos cómo recorrer el camino inverso.

Primero recordemos que, cuando calculamos las cosas en el ensamble microcanónico, al escribir la Ec. (4) se mencionó, sin mayores explicaciones, que la eliminación de B y C en términos de A no parecía ser la elección más simétrica, pero que era la más simple. La elección más simétrica hubiera sido despejar A y C en términos de B . De hecho, en una primera versión de estas notas, las variables eliminadas eran A y C , y la cuenta sobre estados se llevaba a través de la variable B . Además de que quedaban sumas cuyos términos (no sólo sus límites) dependían de si $N \pm M$ era par o impar, la variable de sumación B daba pasos de 2 en 2 (¿pueden ver por qué?). En resumen: un verdadero lío.

En la versión final de estas notas se optó por despejar B y C en términos de A, en lugar del despeje simétrico de A y C en términos de B. El indicio de que así se llegaría a un resultado más simple para $\Omega(E, N)$ surgió luego de resolver el problema en el ensamble canónico. La misma observación puede resultar útil en otros problemas. Además puede llevarse un paso más arriba en la familia de ensambles: en ocasiones resultará más fácil obtener la función de partición canónica a partir de la función de partición gran canónica.

Veamos cómo obtener $\Omega(E, N)$ a partir de la función de partición $Z(\beta, N)$. En principio, la relación que tenemos es

$$Z(\beta, N) = \sum_{E=-\epsilon N}^{\epsilon N} e^{-\beta E} \Omega(E, N) = \sum_{M=-N}^N e^{-\beta M \epsilon} \Omega(E, N) = \sum_{M=-N}^N z^M \Omega(M, N), \quad (61)$$

donde $E = \epsilon M$ y $z = e^{-\beta \epsilon}$, y donde escribimos indistintamente $\Omega(E, N)$ u $\Omega(M, N)$. De la última igualdad se ve que $\Omega(M, N)$ es el coeficiente que acompaña a z^M en la expansión en potencias de z de la función de partición $Z(\beta, N)$.* Puesto que $Z(\beta, N)$ se calcula de manera muy sencilla, no se pierde nada en tratar de escribir su expansión en potencias de z . Si resulta fácil de hacer, entonces obtendremos $\Omega(M, N)$ con muy poco trabajo.

Vimos que $Z = (z + 1 + 1/z)^N$. La expresión más simple para su expansión en potencias de z es

$$Z = \sum_{i=0}^N \sum_{j=0}^{N-i} \binom{N}{i} \binom{N-i}{j} z^{i-j}. \quad (62)$$

Como señalamos al comienzo de la sección anterior, es indistinto que escribamos o no los límites de las sumatorias, debido a que los números binomiales se anulan automáticamente fuera de los intervalos que aparecen en la Ec. (62),

$$Z = \sum_i \sum_j \binom{N}{i} \binom{N-i}{j} z^{i-j}. \quad (63)$$

La ventaja de proceder así es que un cambio en los índices de sumación ya no está entorpecido por los límites de las sumatorias. La combinación de índices que da la potencia de z es $i - j$. Entonces convendrá elegir como nuevo índice de sumación la combinación $M = i - j$. Si i y j varían entre menos y más infinito, M también variará entre los mismos límites. Al cambiar de índices $\{i, j\} \rightarrow \{i, M\}$, la Ec. (63) se transforma en

$$Z = \sum_M \left[\sum_i \binom{N}{i} \binom{N-i}{i-M} \right] z^M. \quad (64)$$

El término entre corchetes es el coeficiente que acompaña la potencia M -ésima de z . Explícitamente,

$$Z = \sum_M \left[\sum_i \frac{N!}{i! (N - 2i + M)! (i - M)!} \right] z^M. \quad (65)$$

*Matemáticamente, $Z(z, N)$ es la función generatriz de los números $\Omega(M, N)$

Puesto que el coeficiente que acompaña a z^M en la expansión de Z en potencias de z es $\Omega(M, N)$, a partir de la expresión anterior se deduce que

$$\Omega(M, N) = \sum_i \frac{N!}{i! (N - 2i + M)! (i - M)!}, \quad (66)$$

que coincide con la Ec. (9) si identificamos i con la población A de elementos con energía ϵ . Y así obtenemos $\Omega(M, N)$ y vemos que la manera eficaz de eliminar variables mediante las condiciones (1), para reescribir la Ec. (3), es despejar C y B en términos de A .

Hay que hacer una salvedad, pero a nuestro favor. La Ec. (9) había sido deducida para $M \geq 0$. Pero nada hubo que decir acerca de M al escribir la Ec. (66). Entonces descubrimos que (9) es igualmente válida para $M < 0$, siempre que dejemos libres los límites de la sumatoria.

Resumiendo: en una primera versión de estas notas, al tratar el ensamble microcanónico se despejaban A y C en términos de B , en la inteligencia de que era la forma más simétrica de manejar las variables de sumación. Se llegaba por ese camino a expresiones para $\Omega(M, N)$ que necesitaban adaptarse según fuera $N \pm M$ par o impar. Luego de redactar la sección sobre el ensamble canónico, al intentar la reconstrucción de $\Omega(M, N)$ a partir de la expansión de Z en potencias de z , inferimos que hubiera sido más práctico al principio haber despejado B y C en términos de A . La primera versión de la solución en el microcanónico fue reemplazada entonces por la que figura actualmente y se agregó esta sección final sobre el cálculo de $\Omega(M, N)$ a partir de $Z(\beta, N)$. (Un claro ejemplo de cómo el futuro puede influir en el pasado).