

Física Teórica 3 – segundo cuatrimestre de 2024

GUÍA 5 – SISTEMA DE LOS DOS NIVELES*

- **1.** Un sistema está compuesto por N elementos distinguibles y no interactuantes. Cada elemento puede estar en dos estados, con energías 0 y $\epsilon > 0$.
- a) Encuentre el número $\Omega(E)$ de microestados con energía E .
 - b) En el ensamble microcanónico, asumiendo que las poblaciones de cada estado son mucho mayores que uno, calcule $S(E)$, $\beta(E)$ y $S(\beta)$. Escriba y grafique la entropía por elemento en función de la fracción x de elementos en el estado con energía ϵ .
 - c) Calcule la función de partición Z en el ensamble canónico mediante los siguientes dos métodos: i) por intermedio de $\Omega(E)$, sumando los estados del sistema según los valores de la energía; ii) organizando la suma sobre los estados del sistema como una suma sobre los estados de cada elemento.
 - d) En general, ¿bajo qué condiciones la función de partición de N elementos distinguibles, no necesariamente de la misma especie, puede escribirse como el producto de las funciones de partición de cada elemento?
 - e) A partir de Z , calcule $\Omega(E)$. Aunque en este sistema no hay una diferencia notable, a veces es más fácil ir desde el ensamble canónico al microcanónico que a la inversa. En el lenguaje de los problemas de combinatoria, ¿qué es $Z(\beta)$ en relación a $\Omega(E)$?
 - f) En el ensamble canónico, calcule la energía media y la entropía como funciones de β y escriba la entropía como función de la energía media. Compare con los resultados del ensamble microcanónico.
 - g) En el ensamble canónico, considere un elemento determinado del sistema. ¿Cuál es, en función de la temperatura, la probabilidad de encontrar a este elemento en cada uno de los estados posibles?
 - h) Misma pregunta pero en el ensamble microcanónico. ¿Cómo se comparan estas probabilidades con las calculadas en el ensamble canónico? ¿Cómo puede interpretar este resultado? ¿De qué condiciones depende?
 - i) Demuestre que, en general, la variancia de la energía en el ensamble canónico es

$$\sigma_E^2 = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = -\frac{\partial \langle E \rangle}{\partial \beta}. \quad (1)$$

- j) Calcule la fluctuación relativa $\sigma_E / \langle E \rangle$ para el sistema de los dos niveles. La equivalencia con el ensamble microcanónico requiere que $\sigma_E / \langle E \rangle \ll 1$. ¿Bajo qué condiciones esto es cierto?

*zanellaj@df.uba.ar

Sobre el sistema de los dos niveles

Este es uno de los sistemas más elementales que admiten una solución exacta. Su exposición bien podría ocupar una sola página. Ahora bien, como es la primera vez que tratamos con ensambles, es necesario explicar con cuidado la lógica de cada paso. Hay varias series de operaciones (como la aproximación de los números binomiales, la factorización de Z o el cálculo de valores medios a través de la función de partición) que, en un primer encuentro, deben explicarse con detalle, pero cuyo uso futuro deberá ocurrir de una manera más bien automática. Saber resolver este problema de principio a fin es muy importante. Después de completar la guía, este problema les tiene que resultar trivial. Al principio les va a parecer que son muchas cosas nuevas de golpe. Estas notas les tienen que servir de referencia general para el resto de los problemas.

La entropía en el ensamble microcanónico

Empecemos por el ensamble microcanónico. Para calcular la entropía, necesitamos encontrar el número $\Omega(E)$ de estados con una determinada energía E . Luego,

$$S(E) = k \log \Omega(E). \quad (2)$$

Para evitar complicaciones innecesarias, asumiremos que S está definida únicamente para los valores admisibles de la energía. El sistema está formado por N elementos y cada elemento puede estar en dos estados, con energías cero y $\epsilon > 0$, respectivamente. La energía total sólo depende de cuántos elementos hay en cada estado. Si hay n_0 elementos en el estado con energía cero y n_1 elementos en el estado con energía ϵ , entonces

$$E = n_1 \epsilon. \quad (3)$$

Debemos preguntarnos de cuántas maneras es posible realizar el estado macroscópico en el cual hay esta distribución de estados entre los elementos del sistema. El número de microestados es igual al número de maneras en que podemos elegir los n_0 elementos que hay en el estado con energía cero. Puesto que el número total de elementos es N ,

$$\Omega(E) = \binom{N}{n_0}. \quad (4)$$

Por simetría, esto tiene que ser igual al número de maneras en que podemos elegir los n_1 elementos en el estado con energía igual a ϵ . En efecto, debido a que $n_1 = N - n_0$, resulta

$$\binom{N}{n_1} = \binom{N}{N - n_0} = \frac{N!}{n_0!(N - n_0)!} = \binom{N}{n_0}. \quad (5)$$

En definitiva,

$$\Omega(E) = \binom{N}{n_0} = \binom{N}{n_1}. \quad (6)$$

Manteniendo la simetría entre n_0 y n_1 , también podemos escribir

$$\Omega(E) = \frac{N!}{n_0! n_1!}. \tag{7}$$

No hay que perder de vista que uno sólo de los números n_0 y n_1 es independiente. Como la energía es $n_1 \epsilon$, conviene quedarse con n_1 .

En conclusión, la entropía es

$$S(E) = k \log \binom{N}{n_1} = k \log \binom{N}{E/\epsilon}. \tag{8}$$

Por varias razones, es preferible definir una variable intensiva x tal que

$$n_1 = xN. \tag{9}$$

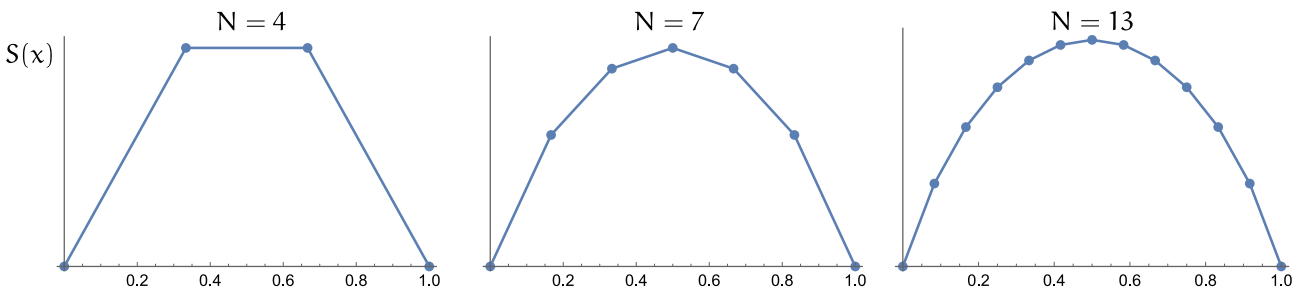
La variable x es la fracción de elementos en el estado con energía ϵ . Evidentemente, x pertenece al intervalo $[0, 1]$. Toma los valores, $0, \frac{1}{N}, \frac{2}{N}, \dots, 1$. La separación entre dos valores sucesivos de x es

$$\Delta x = \frac{\Delta n_1}{N} = \frac{1}{N}. \tag{10}$$

De modo que, a medida que aumentamos N , los valores permitidos de x en el intervalo $[0, 1]$ se agrupan cada vez más densamente.



Esto es fundamental si queremos pensar a la entropía con función de una variable continua. Cuando $N \gg 1$, podemos considerar que la entropía es función de la variable continua x .



Esto es importante, porque, en algún momento, tendremos que calcular derivadas de la entropía. Si la entropía es una función definida sólo en puntos discretos, sus derivadas no tienen un significado inmediato. Entonces, definamos

$$S(x) = k \log \binom{N}{xN}, \tag{11}$$

con $x \in [0, 1]$. Introduciendo la generalización natural de los números binomiales en términos de la función gamma, esto es suficiente para calcular las derivadas de S respecto de x .

La aproximación de Stirling

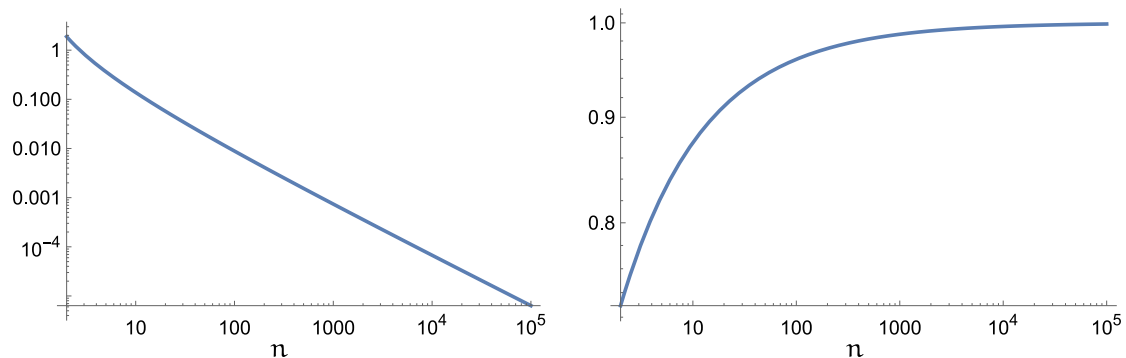
El análisis se simplifica si asumimos que tanto $n_1 = xN$ como $n_0 = (1 - x)N$ son números mucho mayores que uno. Dicho sea de paso, esto implica que N es mucho mayor que uno; después de todo, ya es una de las cosas que estábamos asumiendo. Cuando n es mucho mayor que uno, vale la siguiente aproximación, usualmente llamada aproximación de Stirling:

$$\log n! \simeq n \log n - n. \quad (12)$$

Es necesario hacer un comentario marginal en relación a esta fórmula. De manera ingenua, uno pensaría que una aproximación para $n!$ puede obtenerse directamente exponenciando la ecuación anterior,

$$n! \simeq \left(\frac{n}{e}\right)^n. \quad (!) \quad (13)$$

Lo cierto es que la Ec. (12) no implica la validez de la Ec. (13). La figura de la izquierda muestra el error relativo cuando se usa la aproximación (12) para el logaritmo de $n!$. La figura de la derecha muestra el error relativo cuando se usa la aproximación (13) para calcular $n!$,



Aquí se ve el peligro de usar la aproximación (12) para calcular $n!$. Lo que ocurre es que, en rigor, la aproximación de Stirling es

$$\log n! \simeq n \log n - n + \log \sqrt{2\pi n}. \quad (14)$$

Esto implica que la fórmula correcta es

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \quad (15)$$

Si se omite el último término de la Ec. (14), el error es despreciable para el logaritmo de $n!$, pero no para $n!$. Como regla general, si uno al final de todo va a tomar logaritmos, entonces la expresión (12) es suficiente. En el caso de la entropía, lo que nos interesa es el logaritmo del número binomial, de forma que podemos usar esta aproximación.

De acuerdo a la expresión (12), para el número binomial, tenemos

$$\begin{aligned} \log \binom{N}{xN} &= \log N! - \log(xN)! - \log[(1-x)N!] \\ &\simeq N \log N - N - xN \log xN + xN - (1-x)N \log(1-x)N + (1-x)N. \end{aligned} \tag{16}$$

Los términos lineales en N se simplifican y queda

$$\begin{aligned} \log \binom{N}{xN} &\simeq N \log N - xN \log xN - (1-x)N \log(1-x)N \\ &= N \left[\log N - x \log x - x \log N - (1-x) \log(1-x) - (1-x) \log N \right]. \end{aligned} \tag{17}$$

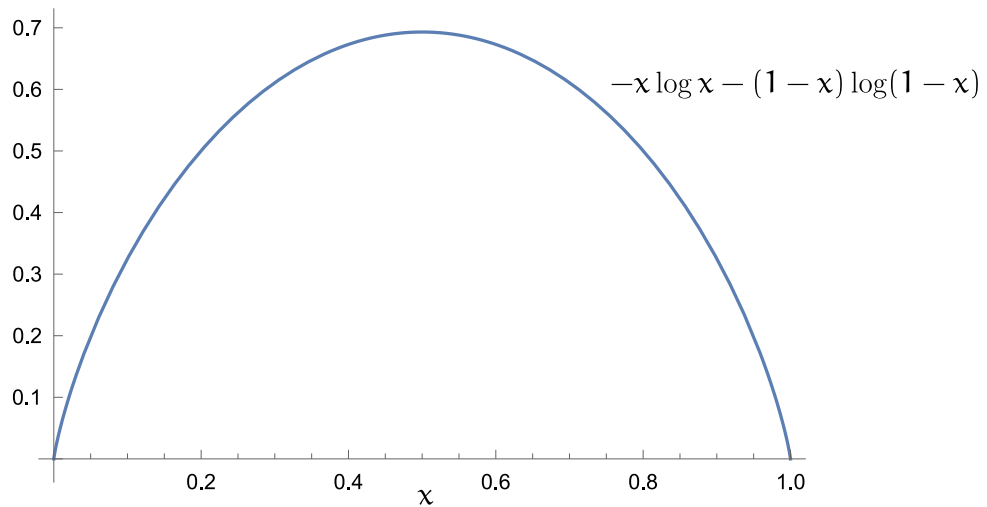
Ahora los términos proporcionales a log N se cancelan. Así, finalmente,

$$\log \binom{N}{xN} \simeq -N [x \log x + (1-x) \log(1-x)]. \tag{18}$$

Esta es una fórmula para recordar. La función

$$f(x) = -x \log x - (1-x) \log(1-x) \tag{19}$$

aparece con mucha frecuencia. Es fácil verificar que es simétrica respecto de $x = \frac{1}{2}$, que se anula en $x = 0$ y en $x = 1$ y que tiene un máximo en $x = \frac{1}{2}$, donde alcanza el valor $\log 2$.



La función $f(x)$ puede definirse en términos de

$$\phi(x) = -x \log x, \tag{20}$$

como

$$f(x) = \phi(x) + \phi(1-x). \tag{21}$$

La función $\phi(x)$ es particularmente relevante para establecer una relación entre entropía e información. Miren al respecto el capítulo 17 del libro de Callen.

Reemplazando el resultado (18) en la expresión (11) para la entropía, obtenemos

$$\frac{S(x)}{k} \simeq Nf(x) = -N[x \log x + (1-x) \log(1-x)]. \quad (22)$$

A partir de ahora tomaremos el signo aproximado con carácter de igualdad.

Extensividad en el ensamble microcanónico

Un aspecto importante de la expresión (22) es que la entropía es extensiva. La lógica de haber introducido la variable intensiva x es que en el resultado final la extensividad de la entropía es evidente. Recordemos que un sistema con entropía extensiva es más simple de tratar, pero la extensividad por sí misma no es algo que esté garantizado. En cualquier caso, es una buena práctica introducir variables intensivas e intentar aislar la dependencia con el número de elementos, sean estos espines o partículas o lo que fueren.

La definición de la entropía en el ensamble microcanónico depende de que N sea mucho mayor que uno. La extensividad también depende de eso. Si conserváramos más términos en la fórmula de Stirling, encontraríamos que la entropía no es exactamente extensiva.

La temperatura como función de la energía. Temperaturas negativas.

La entropía (22) es extensiva, pero, como puede verse en la figura de la página anterior, no es una función monótona creciente de la energía, que era otra de las propiedades contingentes de la entropía. Debido a que

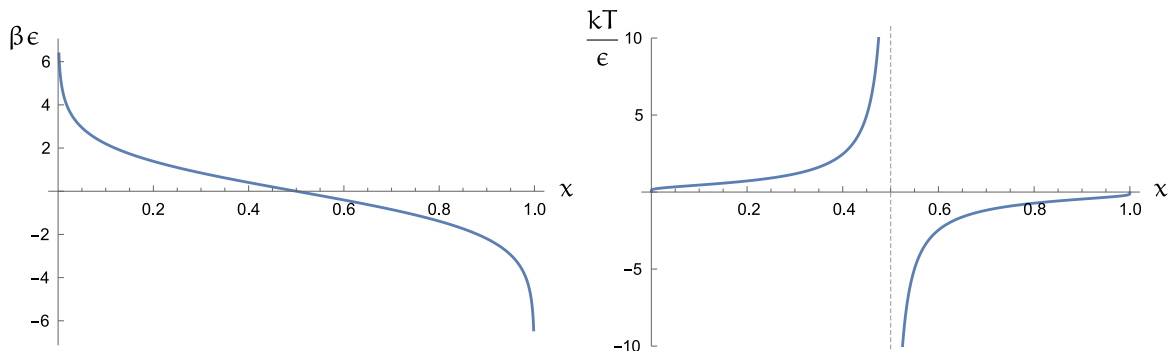
$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{1}{\epsilon N} \frac{\partial S}{\partial x} = \frac{k}{\epsilon} f'(x) = \frac{k}{\epsilon} \log\left(\frac{1-x}{x}\right), \quad (23)$$

cuando la energía es mayor que $\frac{1}{2}N\epsilon$, la temperatura es negativa. No vamos a discutir aquí el tema de las temperaturas negativas. Pueden leer el libro de Pathria o buscar papers recientes. El asunto no está del todo cerrado.

La Ec. (23) puede escribirse en una forma más conveniente como

$$\frac{\epsilon}{kT} = \beta\epsilon = \log\left(\frac{1-x}{x}\right). \quad (24)$$

La figura muestra los gráficos de las dos funciones, $\beta\epsilon$ y kT/ϵ .



A diferencia de la función kT , la función $\beta\epsilon$ es continua. Esta función se anula cuando $x = \frac{1}{2}$. Si la energía tiende a $\frac{1}{2}N\epsilon$ por izquierda, la temperatura tiende a más infinito y, cuando la energía tiende a $\frac{1}{2}N\epsilon$ por derecha, la temperatura tiende a menos infinito. Por otro lado, $\beta\epsilon$ tiende a más infinito cuando la energía tiende a cero. Esto significa que la temperatura tiende a 0^+ cuando la energía tiende a cero. En el otro extremo del intervalo, $\beta\epsilon$ tiende a menos infinito cuando E tiende a $N\epsilon$, lo que significa que la temperatura tiende a 0^- . La energía es máxima cuando T tiende a 0^- .

Siempre resulta un poco desconcertante que, cuando la temperatura es infinita, la energía no toma su valor máximo $N\epsilon$, sino el valor $E_\infty = \frac{1}{2}N\epsilon$. Si pensamos a la temperatura como una medida de la agitación, es natural que, al agitar más al sistema, más homogénea será la distribución de los elementos del sistema respecto a los estados. En el estado de máxima agitación, las poblaciones de cada estado tienden a igualarse. Esto resultará más fácil de entender una vez que hayamos visto el ensamble canónico.

La energía y la entropía como funciones de la temperatura

A partir de la Ec. (24), podemos escribir la energía como función de la temperatura,

$$x = \frac{E}{\epsilon N} = \frac{1}{1 + e^{\beta\epsilon}} = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (25)$$

Más adelante quedará claro por qué escribimos la energía en esta forma, en lugar de detenemos en la expresión intermedia, que es más simple. La razón está en la interpretación probabilística de la última expresión.

Para terminar, escribamos la entropía como función de la temperatura, reemplazando la expresión (25) en la Ec. (22):

$$\begin{aligned} \frac{S(\beta)}{Nk} &= f(x(\beta)) = -\frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} \log\left(\frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}\right) + \frac{1}{1 + e^{-\beta\epsilon}} \log(1 + e^{-\beta\epsilon}) \\ &= \beta\epsilon \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} + \log(1 + e^{-\beta\epsilon}). \end{aligned} \quad (26)$$

En la terminología de Callen, esta no es una ecuación fundamental. A partir de ella no podemos, sin otra información, deducir la Ec. (22), que sí es una ecuación fundamental.

Ensamble canónico

En el ensamble canónico, el objeto principal es la función de partición $Z(\beta)$. El número de elementos está fijo, pero la energía puede tomar cualquier valor admisible. La función de partición canónica está definida como

$$Z = \sum_{\text{estados}} e^{-\beta E(\text{estado})}. \quad (27)$$

La expresión anterior no deja lugar a interpretaciones acerca de qué es lo que se está sumando y cuál es el recorrido de la suma. Sin embargo, la experiencia indica que es un punto que se presta a confusiones nebulosas. De modo que hay que enfatizar esto: la suma es sobre los estados, no sobre los valores de la energía.

Para abreviar, usaremos el símbolo \mathcal{C} en lugar de "estado". La "C" es de configuración. Si agrupamos los estados de acuerdo a las energías,

$$\{\mathcal{C}\} = \{\mathcal{C}|E(\mathcal{C}) = E_0\} \cup \{\mathcal{C}|E(\mathcal{C}) = E_1\} \cup \dots \cup \{\mathcal{C}|E(\mathcal{C}) = E_N\}. \quad (28)$$

Así,

$$\sum_{\mathcal{C}} = \sum_{\{\mathcal{C}|E(\mathcal{C})=E_0\}} + \sum_{\{\mathcal{C}|E(\mathcal{C})=E_1\}} + \dots + \sum_{\{\mathcal{C}|E(\mathcal{C})=E_N\}} = \sum_E \sum_{\{\mathcal{C}|E(\mathcal{C})=E\}}. \quad (29)$$

Entonces,

$$Z = \sum_E \sum_{\{\mathcal{C}|E(\mathcal{C})=E\}} e^{-\beta E} = \sum_E \Omega(E) e^{-\beta E}. \quad (30)$$

Aquí hemos usado el hecho trivial de que la suma del factor de Boltzmann sobre los estados que tienen energía E es igual al número de estados que tienen energía E multiplicado por el factor de Boltzmann correspondiente a esa energía.

Presten atención a la diferencia entre las dos expresiones que tenemos para Z ,

$$Z = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} = \sum_E \Omega(E) e^{-\beta E}. \quad (31)$$

Este resultado es general y difícilmente requiera una demostración tan detallada. Pero esa impresión corre el riesgo de ser retrospectiva, así que no está mal por una vez ser en extremo formales. La Ec. (31) dice que, si conocemos $\Omega(E)$, podemos calcular Z , al menos en principio. Pero, si conocemos $\Omega(E)$, entonces podríamos trabajar directamente en el ensamble microcanónico. No pareciera haber motivos suficientes para abandonar este ensamble. El punto principal es que, a veces, es mucho más sencillo calcular Z que $\Omega(E)$, a partir de la definición (27). Ahí radica la superioridad del ensamble canónico. Aunque la Ec. (31) indica cómo calcular Z conocido $\Omega(E)$, en la práctica suele darse el camino inverso: uno usa la función de partición en el ensamble canónico para calcular $\Omega(E)$.

El cálculo directo de $\Omega(E)$ muchas veces involucra un problema de combinatoria complicado. La definición de la función de partición canónica es análoga a la definición de la función generatriz que vimos al considerar problemas de combinatoria y probabilidades. Un problema combinatorio que es difícil de resolver de manera directa puede simplificarse si se lo expresa en términos de la función generatriz. La función de partición Z es la función generatriz de los números $\Omega(E)$. La única diferencia con la definición habitual es que en lugar de ser una serie de potencias en la variable z es una serie de potencias en la variable $e^{-\beta}$.

Como regla general, cuando se trata de contar estados, el proceso es más complicado cuantos más vínculos existan. En el ensamble microcanónico, N y E están fijos. En el ensamble canónico, se elimina el vínculo sobre E . Al pasar al ensamble gran canónico, se elimina el vínculo sobre N . En especial, cuando estudiemos estadística cuántica, veremos que calcular funciones de partición en el ensamble canónico es muy complicado, mientras que en el ensamble gran canónico el cálculo es casi trivial.

El sistema de los dos niveles es tan sencillo que es posible calcular Z siguiendo ambos métodos: a partir del camino constructivo que va desde el ensamble microcanónico al canónico a través de la función $\Omega(E)$, o siguiendo un camino que no depende del cálculo de $\Omega(E)$. En el primer caso, escribimos

$$Z = \sum_E \Omega(E) e^{-\beta E} = \sum_{n=0}^N \binom{N}{n} e^{-\beta \epsilon n}. \quad (32)$$

Esta suma es el desarrollo del binomio, de modo que

$$Z = (1 + e^{-\beta \epsilon})^N. \quad (33)$$

Veamos ahora el otro camino. Para no depender del cálculo $\Omega(E)$, podemos escribir directamente

$$Z = \sum_c e^{-\beta E(c)}. \quad (34)$$

El estado microscópico del sistema queda definido por el estado de cada elemento. Así, la suma sobre los estados es una suma sobre los estados de todos los elementos. Denotemos con el símbolo s_i al estado del elemento i -ésimo. Lo más cómodo es tomar $s_i = 0$ si el elemento está en el estado con energía cero y $s_i = 1$ si el elemento está en el estado con energía ϵ . Entonces, resulta

$$Z = \sum_{s_1=0}^1 \sum_{s_2=0}^1 \dots \sum_{s_N=0}^1 e^{-\beta E(s_1, \dots, s_N)}. \quad (35)$$

Pero la energía total es la suma de las energías de cada elemento, que, por construcción, son iguales a $s_i \epsilon$,

$$E(s_1, \dots, s_N) = (s_1 + \dots + s_N) \epsilon. \quad (36)$$

Por lo tanto, la expresión (35) se factoriza en el producto de N sumas, y queda

$$Z = \left(\sum_{s_1=0}^1 e^{-\beta \epsilon s_1} \right) \dots \left(\sum_{s_N=0}^1 e^{-\beta \epsilon s_N} \right) = \left(\sum_{s=0}^1 e^{-\beta \epsilon s} \right)^N = (1 + e^{-\beta \epsilon})^N, \quad (37)$$

que es el mismo resultado que obtuvimos antes.

Habíamos anticipado que este problema es demasiado sencillo como para que existiera una diferencia notable en el grado de dificultad en el cálculo de Z por uno u otro camino. A veces, la diferencia existe y es importante. En otras ocasiones no queda más alternativa que calcular $\Omega(E)$.

Una observación importante es que podemos escribir

$$Z = Z_1^N, \quad (38)$$

donde

$$Z_1 = \sum_{\epsilon_1} e^{-\beta \epsilon(\epsilon_1)} = 1 + e^{-\beta \epsilon} \quad (39)$$

es la función de partición de un elemento. Este resultado es general. Si un sistema está formado por elementos distinguibles y no interactuantes, la función de partición del sistema es igual al producto de las funciones de partición de cada elemento. Es fundamental que los elementos sean no interactuantes, porque, de otra forma, la energía no podría escribirse como la suma de las energías de cada elemento considerado independientemente. Imagínense que cada uno de los elementos del sistema tuviera estados con energías 0 y ϵ_i , donde i indica al elemento i -ésimo. En el ensamble microcanónico sería muy complicado calcular $\Omega(E)$ cuando las energías ϵ_i son diferentes, porque el problema de sumar un dado número a partir de un conjunto determinado de sumandos no es un problema simple. Pero la función de partición canónica se calcula inmediatamente,

$$Z = (1 + e^{-\beta \epsilon_1}) (1 + e^{-\beta \epsilon_2}) \dots (1 + e^{-\beta \epsilon_N}). \quad (40)$$

En general, para elementos distinguibles no interactuantes, de la misma especie o no,

$$Z = Z_1^{(1)} Z_1^{(2)} \dots Z_1^{(N)}, \quad (41)$$

donde cada $Z_1^{(i)}$ es la función de partición canónica del elemento i -ésimo, que puede, a su vez, ser en sí mismo un sistema compuesto por otros elementos.

Sabiendo esto, podríamos haber encontrado Z de una forma más directa: los elementos que forman el sistema son todos de la misma especie, distinguibles y no interactuantes. Como cada elemento puede estar en dos estados, la función de partición de un elemento involucra sólo dos términos:

$$Z_1 = 1 + e^{-\beta \epsilon}. \quad (42)$$

Luego,

$$Z = Z_1^N = (1 + e^{-\beta \epsilon})^N. \quad (43)$$

En una vasta mayoría de problemas, la función de partición no se factoriza. Es importante entender las condiciones de las cuales depende el resultado (41).

$\Omega(E)$ a partir de $Z(\beta)$

Resta ver cómo ir desde el ensamble canónico hacia el microcanónico. La función de partición está dada por la Ec. (43). Para obtener $\Omega(E)$, debemos desarrollar Z como una serie de potencias en $e^{-\beta}$. Usando el desarrollo del binomio,

$$Z = \sum_{n=0}^N \binom{N}{n} e^{-\beta \epsilon n}. \quad (44)$$

Si esto ha de ser igual a

$$Z = \sum_E \Omega(E) e^{-\beta E}, \quad (45)$$

entonces el espectro de energías debe ser $E_n = n\epsilon$, con $n = 0, 1, \dots, N$, y, por otro lado,

$$\Omega(E_n) = \binom{N}{n}. \quad (46)$$

Este ejercicio parece trivial, porque, en verdad, $\Omega(E_n)$ es la solución de un problema combinatorio muy sencillo. En otros modelos, el problema combinatorio puede ser muy complicado, mientras que la solución a través Z puede ser, en cambio, relativamente sencilla.

Extensividad en el ensamble canónico

Antes de continuar, notemos que la energía libre de Helmholtz está dada por

$$F(T, N) = -kT \log Z(\beta, N) = -kT \log Z_1(\beta)^N = -NkT \log Z_1(\beta). \quad (47)$$

Esto muestra que el sistema es extensivo. Si multiplicamos el número de elementos por un número $\lambda > 0$, la energía libre de Helmholtz se multiplica por el mismo factor. Esta propiedad de extensividad también apareció en el ensamble microcanónico, pero expresada a través de la entropía en la Ec. (22).

La energía media

Calculemos ahora la energía media como función de la temperatura,

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta} = N\epsilon \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}. \quad (48)$$

Es decir,

$$\frac{\langle E \rangle}{N\epsilon} = \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}. \quad (49)$$

Esto coincide con el resultado (25) del ensamble microcanónico si identificamos $\langle E \rangle$ con E .

Es importante entender la lógica de la relación

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta}. \quad (50)$$

Puede deducirse de dos modos. De una manera probabilística, que suele ser la más intuitiva, y de una manera termodinámica. Es esencial ver que ambos caminos conducen al mismo resultado. Empecemos por la deducción probabilística. Por definición,

$$Z = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}. \quad (51)$$

La probabilidad de un determinado estado en el ensamble canónico es

$$P(\mathcal{C}) = \frac{e^{-\beta E(\mathcal{C})}}{Z}. \quad (52)$$

Entonces, el valor medio de la energía es

$$\langle E \rangle = \sum_{\mathcal{C}} E(\mathcal{C})P(\mathcal{C}) = \frac{1}{Z} \sum_{\mathcal{C}} E(\mathcal{C})e^{-\beta E(\mathcal{C})} = -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} = -\frac{\partial \log Z}{\partial \beta}. \quad (53)$$

Esto muestra que tomar la derivada del logaritmo de Z respecto de β , con signo menos, no hace otra cosa que construir el valor medio de la energía, según su definición primitiva. Muchos valores medios pueden calcularse en esta forma, sin una conexión aparente con la termodinámica. Pero es fácil ver que tal conexión existe.

Veamos ahora el camino que lleva desde la termodinámica a la Ec. (50). La relación entre la termodinámica y la mecánica estadística en el ensamble canónico es a través de la energía libre de Helmholtz,

$$-\beta F = \log Z. \quad (54)$$

A su vez, F es la transformada de Legendre de U con respecto a S ,

$$F = U - TS, \quad (55)$$

que es lo mismo que decir que

$$U = F + TS. \quad (56)$$

Pero el primer principio de la termodinámica, escrito en términos de F , se lee como

$$dF = -SdT + \mu dN. \quad (57)$$

Notemos que en el sistema de este problema no hay ningún mecanismo por el cual realizar trabajo, de modo que el primer principio se limita sólo a esos dos términos en el segundo

miembro de la ecuación anterior. Lo importante es que

$$S = -\frac{\partial F}{\partial T} = \frac{1}{kT^2} \frac{\partial F}{\partial \beta} = k\beta^2 \frac{\partial F}{\partial \beta}. \quad (58)$$

Entonces, reemplazando en la Ec. (56), obtenemos

$$U = -\frac{1}{\beta} \log Z - \beta \frac{\partial}{\partial \beta} \left(\frac{1}{\beta} \log Z \right) = -\frac{\partial \log Z}{\partial \beta}, \quad (59)$$

que es el mismo resultado que obtuvimos siguiendo el camino meramente probabilístico.

Otra observación importante es que la energía media de todo el sistema es igual a la suma de las energías medias de cada elemento. Cada elemento considerado por separado tiene una función de partición

$$Z_1 = 1 + e^{-\beta \epsilon}. \quad (60)$$

La energía media de cada elemento es

$$\langle E_1 \rangle = -\frac{\partial \log Z_1}{\partial \beta} = \frac{\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}. \quad (61)$$

Entonces se verifica que $\langle E \rangle = N \langle E_1 \rangle$.

La entropía en el ensamble canónico

Para terminar de comparar los ensambles microcanónico y canónico de este sistema, calculemos la entropía en el ensamble canónico. En principio, tenemos dos alternativas. Una consiste en escribir

$$S = -\frac{\partial F(T, N)}{\partial T}. \quad (62)$$

La dificultad con este método (que es una dificultad mínima) es que la Ec. (54) da F como función de β , o, de una manera híbrida, en la forma

$$F = -kT \log Z(\beta). \quad (63)$$

Entonces hay que derivar un producto y también derivar β respecto de T . El segundo camino consiste en usar la definición de F como transformada de Legendre de la energía, Ec. (55), lo que implica

$$S = \frac{U - F}{T}. \quad (64)$$

Pero, por otro lado,

$$U = -\frac{\partial \log Z}{\partial \beta}. \quad (65)$$

Como Z es una función de β , esta derivada es muy simple de calcular. Reuniendo todos estos resultados, queda

$$\frac{S}{k} = -\beta \frac{\partial \log Z}{\partial \beta} + \log Z. \quad (66)$$

Explícitamente, usando la Ec. (33),

$$\frac{S}{Nk} = \beta \epsilon \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} + \log(1 + e^{-\beta \epsilon}) = -\frac{\beta \epsilon}{1 + e^{-\beta \epsilon}} - \log\left(\frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}\right). \quad (67)$$

Queda como ejercicio que verifiquen que el mismo resultado se obtiene a partir de la Ec. (62).

Noten que ya hemos encontrado antes la Ec. (67) a través del ensamble microcanónico, Ec. (26). De forma que con esto queda demostrado que, mediante uno u otro ensamble, llegamos a los mismos resultados. Lo que sigue ahora será entonces un poco redundante, porque la equivalencia ya está demostrada. De todos modos, el ejercicio es ilustrativo. Aunque ya sepamos que las expresiones son equivalentes, es bueno asegurarse de que somos capaces de demostrar esto en uno y en otro sentido de la equivalencia.

El objetivo es escribir S como función de la energía a partir de la Ec. (67), para recuperar la ecuación fundamental. Sabemos que la energía media es

$$\langle E \rangle = \frac{N\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}, \quad (68)$$

y, además, hemos definido x tal que

$$\langle x \rangle = \frac{\langle E \rangle}{N\epsilon}. \quad (69)$$

Luego,

$$\beta \epsilon = \log\left(\frac{N\epsilon}{\langle E \rangle} - 1\right) = \log\left(1 - \frac{\langle E \rangle}{N\epsilon}\right) - \log\frac{\langle E \rangle}{N\epsilon} = \log(1 - \langle x \rangle) - \log \langle x \rangle, \quad (70)$$

y

$$\frac{1}{1 + e^{-\beta \epsilon}} = 1 - \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = 1 - \langle x \rangle. \quad (71)$$

Entonces, según la Ec. (67),

$$\begin{aligned} \frac{S}{Nk} &= -(1 - \langle x \rangle) \left[\log(1 - \langle x \rangle) - \log \langle x \rangle \right] - \log \langle x \rangle \\ &= -(1 - \langle x \rangle) \log(1 - \langle x \rangle) - \langle x \rangle \log \langle x \rangle. \end{aligned} \quad (72)$$

Si identificamos $\langle x \rangle$ con x , esto coincide con la expresión (22).

Probabilidades

Teníamos pendiente una explicación. Luego de la Ec. (25), dejamos escrito que había motivos para preferir la expresión

$$x = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}, \quad (73)$$

en lugar de la forma equivalente y algo más simple

$$x = \frac{1}{1 + e^{\beta\epsilon}}. \quad (74)$$

Son los mismos motivos por los que escribimos la Ec. (61) como

$$\langle E_1 \rangle = \frac{\epsilon e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (75)$$

y no como

$$\langle E_1 \rangle = \frac{\epsilon}{1 + e^{\beta\epsilon}}. \quad (76)$$

Recién ahora podemos explicar la razón de esa preferencia. La cuestión pasa por la lectura probabilística que puede hacerse de la Ec. (75). La energía media de un elemento es

$$\langle E_1 \rangle = 0 \times p(0) + \epsilon \times p(\epsilon), \quad (77)$$

donde $p(E_1)$ es la probabilidad de que el elemento esté en el estado con energía E_1 . Estas probabilidades están determinadas por el factor de Boltzmann y por la función de partición, de acuerdo al principio general

$$p(\mathcal{C}) = \frac{e^{-\beta E(\mathcal{C})}}{Z}. \quad (78)$$

Como hay una relación uno a uno entre estados y energías, no hay ambigüedad si escribimos $p(E_1)$ en lugar de $p(\mathcal{C})$, de modo que

$$p(0) = \frac{1}{1 + e^{-\beta\epsilon}}, \quad p(\epsilon) = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (79)$$

Finalmente,

$$\langle E_1 \rangle = 0 \times \frac{1}{1 + e^{-\beta\epsilon}} + \epsilon \times \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} = \frac{\epsilon e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (80)$$

Vemos que la Ec. (75), escrita en esa forma, y no como en la Ec. (76), se lee directamente como la definición del valor medio de E_1 en términos de las probabilidades. Si uno ve escrita la Ec. (75), comprende de inmediato que es correcta. En cambio, no puede decirse a primera vista si la Ec. (76) es cierta o no.

También habíamos prometido que en el ensamble canónico iba a ser más fácil entender por qué cuando la temperatura tiende a infinito la energía no tiende a su valor máximo, sino

a $\frac{1}{2}N\epsilon$. Si la temperatura es muy alta comparada con ϵ , es decir, si $\beta\epsilon \ll 1$, la separación entre las energías de los dos estados de cada elemento es despreciable frente a kT . A todos los fines prácticos, los dos estados son equiprobables. Esto puede verse directamente a partir de las Ecs. (79):

$$\frac{p(\epsilon)}{p(0)} = e^{-\beta\epsilon} \rightarrow 1. \quad (81)$$

Notar que aquí no importa el signo de ϵ . Si los dos estados son equiprobables, la energía media debe tender a N veces el promedio de las energías de cada estado, esto es, $\frac{1}{2}\epsilon$.

En el ensamble microcanónico, la energía del sistema está fija, pero la energía de cada elemento puede fluctuar. En el ensamble microcanónico, no tiene mucho interés preguntarse por la probabilidad de que la energía del sistema sea E o por las fluctuaciones en la energía total. La probabilidad de que la energía sea E es o bien cero o bien uno, y las fluctuaciones son cero. Pero sí tiene sentido preguntarse por el valor medio y por las fluctuaciones de la energía de un elemento determinado del sistema. Intuitivamente, el resto de los elementos deberían actuar como un reservorio de energía para el elemento considerado, tal como si se tratara de un sistema considerado en el ensamble canónico. Claro que, para que esto sea cierto, el sistema debe ser, en algún sentido, lo suficientemente grande.

Si la energía del sistema es $E = x\epsilon N$, el número de elementos en el estado con energía ϵ es xN . Entonces, la probabilidad de que un elemento determinado esté en el estado con energía ϵ es

$$p_{\mu C}(\epsilon) = x. \quad (82)$$

Usamos el subíndice μC para indicar que estas probabilidades están calculadas en el ensamble microcanónico. Por complementariedad, la probabilidad de que el elemento considerado esté en el estado con energía cero es

$$p_{\mu C}(0) = 1 - p_{\mu C}(\epsilon) = 1 - x. \quad (83)$$

Por lo tanto, según la Ec. (25), que era válida cuando $N \gg 1$,

$$p_{\mu C}(0) = \frac{1}{1 + e^{-\beta\epsilon}}, \quad p_{\mu C}(\epsilon) = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (84)$$

Comparando con las probabilidades (79), esto significa que cada elemento del sistema en el ensamble microcanónico tiene una distribución de probabilidad canónica respecto a la energía. El resto del sistema se comporta efectivamente como el reservorio que uno introduce para definir el ensamble canónico. Los $N-1$ elementos restantes son el reservorio "ideal" con el que está en contacto el elemento considerado. Esto muestra de una manera muy clara cómo es el paso de un ensamble al otro.

Fluctuaciones I

Según hemos dicho antes, en el ensamble canónico la probabilidad de un estado es

$$P(\mathcal{C}) = \frac{e^{-\beta E(\mathcal{C})}}{Z}. \quad (85)$$

Por definición, la variancia de la energía es

$$\sigma_E^2 = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 - 2E \langle E \rangle + \langle E \rangle^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2. \quad (86)$$

Ahora bien, el valor medio de E^2 está dado por

$$\langle E^2 \rangle = \sum_{\mathcal{C}} E(\mathcal{C})^2 P(\mathcal{C}) = \frac{1}{Z} \sum_{\mathcal{C}} E(\mathcal{C})^2 e^{-\beta E(\mathcal{C})}. \quad (87)$$

Podemos construir esta expresión tomando derivadas respecto de β ,

$$\langle E^2 \rangle = \frac{1}{Z} \frac{\partial^2}{\partial \beta^2} \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}. \quad (88)$$

Pero, por otra parte, también sabemos que

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}. \quad (89)$$

Esto significa que

$$\sigma_E^2 = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 = \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right) = -\frac{\partial \langle E \rangle}{\partial \beta}. \quad (90)$$

Para el problema de los dos niveles, teníamos

$$\langle E \rangle = N\epsilon \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} = \frac{N\epsilon}{1 + e^{\beta\epsilon}}. \quad (91)$$

Entonces,

$$\sigma_E^2 = \frac{N\epsilon^2 e^{\beta\epsilon}}{(1 + e^{\beta\epsilon})^2} = \frac{e^{\beta\epsilon}}{N} \langle E \rangle^2. \quad (92)$$

Una medida de la dispersión en la energía es

$$\frac{\sigma_E}{\langle E \rangle} = \frac{e^{\beta\epsilon/2}}{\sqrt{N}}. \quad (93)$$

Cuando $N \rightarrow \infty$, la expresión anterior tiende a cero. En ese límite la energía del sistema está perfectamente definida, tal como en el ensamble microcanónico. Una condición necesaria para que los resultados de cada ensamble sean aproximadamente equivalentes es que

$$\frac{e^{\beta\epsilon/2}}{\sqrt{N}} \ll 1. \quad (94)$$

Esto no es tan simple como decir que \sqrt{N} debe ser mucho mayor que 1. Las fluctuaciones son más importantes cuanto menor es la temperatura. Para temperaturas lo suficientemente bajas, $e^{\beta\epsilon/2}$ puede ser del mismo orden de magnitud que \sqrt{N} . En ese régimen, los ensambles no serán equivalentes.

La Ec. (90) puede escribirse en términos de la capacidad calorífica,

$$\frac{\partial \langle E \rangle}{\partial \beta} = \frac{\partial T}{\partial \beta} \frac{\partial \langle E \rangle}{\partial T} = -kT^2 C, \quad (95)$$

donde

$$C = \frac{\partial \langle E \rangle}{\partial T}. \quad (96)$$

Entonces,

$$\sigma_E^2 = kT^2 C. \quad (97)$$

Como comentario marginal, las fluctuaciones en la energía de un sólo elemento están dadas por la misma Ec. (90) con E reemplazada por E_1 . Es fácil ver que

$$\frac{\sigma_{E_1}}{E_1} = e^{\beta\epsilon/2}. \quad (98)$$

Estas fluctuaciones no dependen del tamaño del sistema, y no tendrían por qué depender. Cada elemento es independiente de los otros. Las fluctuaciones sólo tienden a cero cuando $N \rightarrow \infty$ si se considera la energía total.

Fluctuaciones II

Existe otro método para llegar a la expresión (92). La idea es encontrar una densidad de probabilidad aproximada que describa la distribución de x cuando x se considera una variable continua. El método tiene interés en sí mismo, porque muestra cómo aproximar mediante una distribución gaussiana una probabilidad fuertemente concentrada.

La probabilidad de que el sistema esté en un estado con energía E es igual a la probabilidad de que esté en cualquier estado con energía E multiplicada por el número de estados con energía E . Definiendo nuevamente la energía como $E = xN\epsilon$, en términos de x ,

$$P(x) = \frac{1}{Z} \binom{N}{xN} e^{-\beta x N \epsilon}. \quad (99)$$

Consideremos la función

$$g(x) = \binom{N}{xN} e^{-\beta x N \epsilon}. \quad (100)$$

Conviene expresarla en términos de su logaritmo,

$$g(x) = e^{h(x)}, \quad (101)$$

donde

$$h(x) = \log \binom{N}{xN} - \beta x N \epsilon. \quad (102)$$

Es fácil verificar que, si se usa la aproximación de Stirling *correcta* de la Ec. (14),

$$\log n! \simeq n \log n - n + \log \sqrt{2\pi n}, \quad (103)$$

resulta

$$\binom{N}{xN} \simeq N f(x) - \log \sqrt{2\pi x(1-x)N}, \quad (104)$$

donde $f(x)$ es la función que definimos en la Ec. (19),

$$f(x) = -x \log x - (1-x) \log x. \quad (105)$$

Puesto que $f(x)$ tiene, a grandes rasgos, la forma de una cuadrática, entonces

$$h(x) \approx N [f(x) - \beta \epsilon x], \quad (106)$$

tendrá también, a grandes rasgos, la forma de una parábola invertida, debido a que una función lineal más una cuadrática es también una cuadrática. Al calcular $e^{h(x)}$, ya que $h(x)$ es aproximadamente proporcional a N y $N \gg 1$, la función exponencial tenderá a producir una campana muy aguda centrada en el máximo de $h(x)$. La idea, entonces, es desarrollar $h(x)$ alrededor del punto x_0 en donde alcanza su máximo. Si escribimos

$$h(x) \simeq h(x_0) + \frac{1}{2} h''(x_0) (x - x_0)^2, \quad (107)$$

al reemplazar en la definición de $P(x)$, obtendremos, como primera aproximación, una gaussiana centrada en x_0 , aunque todavía no será una densidad propiamente dicha.

Tomando el signo aproximado en la Ec. (104) con carácter de igualdad, el máximo de h se alcanza en x_0 tal que

$$h'(x_0) = N \left[\log \left(\frac{1-x_0}{x_0} \right) - \beta \epsilon \right] - \frac{1-2x_0}{2x_0(1-x_0)} = 0. \quad (108)$$

Cuando $N \gg 1$, podemos despreciar el segundo término. Luego,

$$x_0 \simeq \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}}. \quad (109)$$

Es decir, según lo que obtuvimos antes, Ec. (69),

$$x_0 \simeq \langle x \rangle. \quad (110)$$

Por otro lado,

$$h''(x_0) \simeq -N \left(\frac{1}{1-x_0} + \frac{1}{x_0} \right) = -\frac{N}{x_0(1-x_0)} \quad (111)$$

Además, usando el resultado de la Ec. (26) para evaluar $f(x_0)$,

$$\begin{aligned} h(x_0) &\simeq N [f(x_0) - \beta \epsilon x_0] - \log \sqrt{2\pi x_0(1-x_0)N} \\ &= N \log(1 + e^{-\beta \epsilon_0}) - \log \sqrt{2\pi x_0(1-x_0)N}. \end{aligned} \quad (112)$$

Finalmente,

$$P(x) \simeq \frac{1}{Z} e^{h(x_0) + \frac{1}{2} h''(x_0)(x-x_0)^2} \simeq \frac{1}{\sqrt{2\pi N x_0(1-x_0)}} \exp \left[-\frac{N(x-x_0)^2}{2x_0(1-x_0)} \right]. \quad (113)$$

Pero, por otra parte,

$$x_0(1-x_0) = \frac{e^{-\beta \epsilon}}{(1 + e^{-\beta \epsilon})^2} = x_0^2 e^{\beta \epsilon}. \quad (114)$$

En definitiva,

$$P(x) \simeq \frac{1}{\sqrt{2\pi N x_0^2 e^{\beta \epsilon}}} \exp \left[-\frac{N(x-x_0)^2}{2x_0^2 e^{\beta \epsilon}} \right]. \quad (115)$$

Cuando $N \gg 1$, los valores admisibles de x se distribuyen densamente. Por lo tanto, en lugar de trabajar con la probabilidad de x , conviene introducir una densidad de probabilidad. La probabilidad de que x esté en el intervalo $[y, y + \Delta x]$ es aproximadamente igual a $P(y)$ por el número de valores admisibles de x en el intervalo Δx . Ese número es $N\Delta x$, ya que los valores sucesivos de x están separados por una distancia $1/N$. Entonces,

$$P(y \leq x \leq y + \Delta x) \simeq P(y)N\Delta x. \quad (116)$$

La densidad de probabilidad $p(x)$ se define de modo tal que

$$P(y \leq x \leq y + \Delta x) \simeq p(y)\Delta x. \quad (117)$$

Entonces,

$$p(x) \simeq NP(x) \simeq \sqrt{\frac{N}{2\pi x_0^2 e^{\beta \epsilon}}} \exp \left[-\frac{N(x-x_0)^2}{2x_0^2 e^{\beta \epsilon}} \right] \quad (118)$$

Esta es una distribución gaussiana, centrada en x_0 y con variancia

$$\sigma_x^2 = \frac{x_0^2 e^{\beta \epsilon}}{N}. \quad (119)$$

Puesto que $E = N\epsilon x$, la variancia de E es

$$\sigma_E^2 = N^2 \epsilon^2 \sigma_x^2 = \frac{\langle E \rangle^2 e^{\beta \epsilon}}{N}. \quad (120)$$

Así, obtenemos nuevamente el resultado (93).