

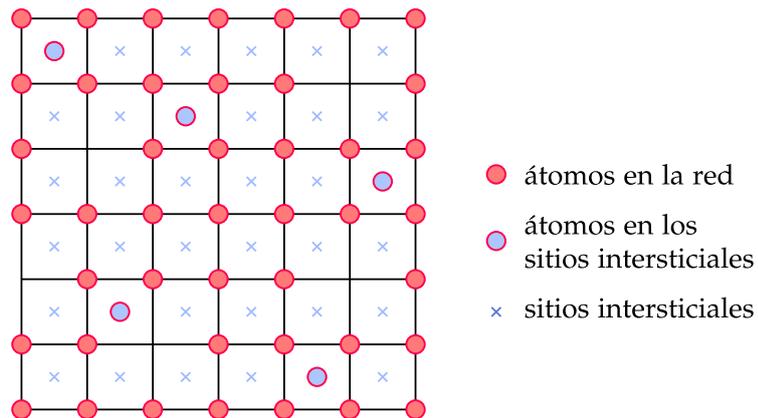
Física Teórica 3 – segundo cuatrimestre de 2024

GUÍA 5 – DEFECTOS DE FRENKEL*

■ **3. Defectos de Frenkel.** Una red cristalina está formada por N átomos de la misma especie. Si se extraen n átomos de sus lugares en la red y se los coloca en posiciones intersticiales, se obtienen n defectos de tipo Frenkel. El número N' de posiciones intersticiales en la red es del mismo orden de magnitud que N . Si $W > 0$ es la energía necesaria para producir un defecto, encuentre el número de defectos como función de la temperatura. En particular, muestre que si $e^{-\beta W} \ll 1$, entonces $n \simeq \sqrt{NN'}e^{-\beta W/2}$. Resuelva este problema en los tres ensambles. En el ensamble microcanónico, asuma que todos los números involucrados son mucho mayores que uno. En el ensamble canónico, asuma, además, que el logaritmo de la función de partición,

$$\log Z = \log \left[\sum_E \Omega(E) e^{-\beta E} \right], \quad (1)$$

puede aproximarse por el logaritmo del máximo término de la suma. En el ensamble gran canónico, note que el valor medio del número de átomos tiene que ser igual a N .



Ensamble microcanónico

El número de partículas está fijo y es igual a N . La energía del sistema, que también está fija, depende del número de defectos. Si hay n defectos, la energía es

$$E = nW. \quad (2)$$

Para calcular la entropía en el ensamble microcanónico, es necesario encontrar primero el número $\Omega(E)$ de microestados en donde hay n defectos. Luego,

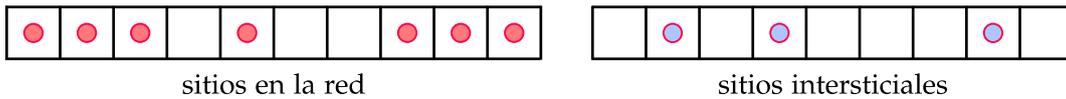
$$S(E) = k \log \Omega(E). \quad (3)$$

*zanellaj@df.uba.ar

Para encontrar $\Omega(E)$, se puede razonar así: hay $\binom{N}{n}$ maneras de elegir las n partículas que abandonan sus sitios en la red y, a su vez, hay otras $\binom{N'}{n}$ maneras en las que esas n partículas pueden distribuirse en los N' sitios intersticiales. Entonces, el número total de microestados es

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

Gráficamente, debemos contar de cuántas formas se pueden distribuir los n sitios vacíos en las N casillas de la figura de la izquierda y de cuántas formas se pueden distribuir los n sitios ocupados en las N' casillas de la figura de la derecha.



En definitiva, la entropía está dada por

$$S(E) = k \log \left[\binom{N}{n} \binom{N'}{n} \right], \quad (5)$$

donde $n = E/W$. Introduzcamos la variable intensiva

$$x = \frac{n}{N}, \quad (6)$$

que representa la fracción de defectos relativa al número de sitios de la red. Conviene considerar que la entropía es función de esta variable,

$$S(x) = k \log \left[\binom{N}{xN} \binom{N'}{xN} \right] = k \log \binom{N}{xN} + k \log \binom{N'}{xN}. \quad (7)$$

Si $N \gg 1$, los valores admisibles de x se distribuyen densamente en el intervalo $[0, 1]$. Podemos extender la definición de la entropía para cualquier x en ese intervalo. De esta manera, la entropía puede tratarse como una función continua y derivable. Esto es suficiente para poder calcular las derivadas de la entropía respecto de x .

Para obtener una expresión más manejable, supondremos que todos los números involucrados son mucho mayores que uno y usaremos la aproximación de Stirling:

$$\log n! \simeq n \log n - n. \quad (8)$$

Aplicada a los logaritmos de los números binomiales, esto implica

$$\begin{aligned} \log \binom{N}{xN} &= \log N! - \log(xN)! - \log[(1-x)N]! \\ &\simeq N \log N - N - xN \log xN + xN - (1-x)N \log(1-x)N + (1-x)N \\ &= -N [x \log x + (1-x) \log(1-x)]. \end{aligned} \quad (9)$$

En términos de la función

$$f(x) = -x \log x - (1 - x) \log(1 - x), \quad (10)$$

tenemos

$$\log \binom{N}{xN} \simeq Nf(x). \quad (11)$$

Introduzcamos el parámetro

$$\lambda = \frac{N}{N'}, \quad (12)$$

entonces, de manera similar,

$$\log \binom{N'}{n} = \log \binom{N'}{\lambda x N'} \simeq N' f(\lambda x) = \frac{N}{\lambda} f(\lambda x). \quad (13)$$

Finalmente, tratando estas aproximaciones con carácter de igualdad,

$$S(x) = Nk \left[f(x) + \frac{1}{\lambda} f(\lambda x) \right]. \quad (14)$$

Siempre es importante aislar la dependencia con el tamaño del sistema. Evidentemente, si uno mantiene fijas las fracciones x y λ , que son por naturaleza variables intensivas, la entropía es una función extensiva respecto de N .

Puesto que lo que buscamos es una relación entre el número de defectos y la temperatura, debemos escribir la ecuación que relaciona la energía con la temperatura,

$$\frac{1}{kT} = \frac{1}{k} \frac{\partial S}{\partial E} = \frac{1}{NkW} \frac{\partial S}{\partial x}. \quad (15)$$

La última derivada es

$$\frac{1}{Nk} \frac{\partial S(x)}{\partial x} = f'(x) + f'(\lambda x), \quad (16)$$

donde

$$f'(x) = \log \left(\frac{1-x}{x} \right). \quad (17)$$

Así,

$$\frac{1}{Nk} \frac{\partial S(x)}{\partial x} = \log \left(\frac{1-x}{x} \right) + \log \left(\frac{1-\lambda x}{\lambda x} \right) = \log \left[\frac{(1-x)(1-\lambda x)}{\lambda x^2} \right]. \quad (18)$$

Volviendo a la Ec. (15), queda

$$\beta W = \log \left[\frac{(1-x)(1-\lambda x)}{\lambda x^2} \right]. \quad (19)$$

Equivalentemente,

$$\lambda x^2 = (1 - x)(1 - \lambda x)e^{-\beta W}. \quad (20)$$

Esta es una ecuación cuadrática para x . La solución nos permitirá escribir x como función de β . Si $e^{-\beta W} \ll 1$, es decir, en el régimen de bajas temperaturas, x será un número mucho menor que 1. En el límite $\beta W \rightarrow \infty$, la fracción de defectos será estrictamente cero. Podemos usar esto para calcular x de manera aproximada. Escribamos la ecuación anterior como una ecuación de punto fijo,

$$x = \sqrt{(1 - x)(1 - \lambda x)} \frac{e^{-\beta W/2}}{\sqrt{\lambda}}. \quad (21)$$

La solución de orden cero es $x_0 = 0$. Entonces, la solución de orden uno en la recursión es

$$x_1 = \sqrt{(1 - x_0)(1 - \lambda x_0)} \frac{e^{-\beta W/2}}{\sqrt{\lambda}} = \frac{e^{-\beta W/2}}{\sqrt{\lambda}}. \quad (22)$$

La solución de orden dos está dada por

$$\begin{aligned} x_2 &= \sqrt{(1 - x_1)(1 - \lambda x_1)} \frac{e^{-\beta W/2}}{\sqrt{\lambda}} = \sqrt{\left(1 - \frac{1}{\sqrt{\lambda}} e^{-\beta W/2}\right) \left(1 - \sqrt{\lambda} e^{-\beta W/2}\right)} \frac{e^{-\beta W/2}}{\sqrt{\lambda}} \\ &= \sqrt{1 + e^{-\beta W} - \frac{1 + \lambda}{\sqrt{\lambda}} e^{-\beta W/2}} \frac{e^{-\beta W/2}}{\sqrt{\lambda}}. \end{aligned} \quad (23)$$

Si expandimos en potencias de $y = e^{-\beta W}$, resulta,

$$x_2 \simeq \sqrt{\frac{y}{\lambda}} \left(1 - \frac{1 + \lambda}{2} \sqrt{\frac{y}{\lambda}}\right). \quad (24)$$

Los mismos resultados deberían obtenerse luego de resolver la cuadrática (20) y expandir en potencias de y . En efecto, escrita por extenso, la cuadrática (20) es

$$\lambda(1 - y)x^2 + (1 + \lambda)xy - y = 0. \quad (25)$$

Luego,

$$x = \frac{1}{2\lambda(1 - y)} \left[-(1 + \lambda)y \pm \sqrt{(1 - \lambda)^2 y^2 + 4\lambda y} \right]. \quad (26)$$

Debemos tomar la solución con signo positivo, debido a que x tiene que ser mayor o igual que cero. Es fácil ver que esa solución es, al mismo tiempo, menor o igual que el mínimo entre uno y λ^{-1} (¿por qué?). Expandamos la raíz cuadrada alrededor de $y = 0$,

$$\sqrt{(1 - \lambda)^2 y^2 + 4\lambda y} = 2\sqrt{\lambda y} \left[1 + \frac{(1 - \lambda)^2}{8\lambda} y \right] + \dots \quad (27)$$

Entonces,

$$x = \frac{1}{2\lambda(1-y)} \left[-(1+\lambda)y + 2\sqrt{\lambda y} + \frac{(1-\lambda)^2}{4\sqrt{\lambda}} y^{3/2} + \dots \right] \quad (28)$$

$$\simeq \sqrt{\frac{y}{\lambda}} \left(1 - \frac{1+\lambda}{2} \sqrt{\frac{y}{\lambda}} \right),$$

que es el resultado que obtuvimos antes.

Un tercer método, con mejor control del orden de la aproximación que el primero y más sencillo que el segundo, consiste en proponer una solución que sea una serie de potencias en la variable $y^{1/2}$,

$$x = c_0\sqrt{y} + c_1y + \dots \quad (29)$$

El siguiente término debería ser proporcional a $y^{3/2}$. Reemplazando en el miembro de la izquierda de la Ec. (20) y escribiendo explícitamente sólo los términos de hasta orden $y^{3/2}$,

$$\lambda (c_0\sqrt{y} + c_1y)^2 = \lambda (c_0^2y + 2c_0c_1y^{3/2} + \dots). \quad (30)$$

El mismo reemplazo en el miembro de la derecha de la Ec. (20) da

$$(1 - c_0\sqrt{y} - c_1y) \left(1 - \lambda c_0\sqrt{y} - \lambda c_1y \right) y = y - c_0(1+\lambda)y^{3/2} + \dots \quad (31)$$

La comparación de las Ecs. (30) y (31) implica

$$c_0 = \frac{1}{\sqrt{\lambda}}, \quad (32)$$

$$c_1 = -\frac{1+\lambda}{2\lambda},$$

y de aquí se obtiene el mismo resultado que antes,

$$x = \sqrt{\frac{y}{\lambda}} - \frac{1+\lambda}{2\lambda} y + \dots \simeq \sqrt{\frac{y}{\lambda}} \left(1 - \frac{1+\lambda}{2} \sqrt{\frac{y}{\lambda}} \right). \quad (33)$$

El primer método, que es una forma recursiva de obtener la solución, es el más fácil de implementar numéricamente, pero tiene la desventaja de que no está claro hasta qué orden es válida la aproximación. Por ejemplo, expandiendo la expresión (23) hasta orden $y^{3/2}$,

$$x_2 \simeq \sqrt{\frac{y}{\lambda}} - \frac{1+\lambda}{2} \frac{y}{\lambda} - \frac{1}{8}(1-\lambda)^2 \left(\frac{y}{\lambda} \right)^{3/2}. \quad (!) \quad (34)$$

El último término es incorrecto. La expansión de la solución exacta muestra que

$$x = \sqrt{\frac{y}{\lambda}} - \frac{1+\lambda}{2} \frac{y}{\lambda} + \frac{1}{8} (1+6\lambda+\lambda^2) \left(\frac{y}{\lambda} \right)^{3/2} + \mathcal{O}(y^2). \quad (35)$$

El segundo método es el más trabajoso. Además, si resolvemos exactamente la cuadrática no se ve la necesidad de hacer ninguna aproximación. Quisiéramos encontrar la expresión aproximada sin tener que resolver la cuadrática. El tercer método es sistemático en el orden de la aproximación. El valor de estos comentarios depende del contexto. Si tienen una computadora, no les cuesta nada resolver la cuadrática, graficar la solución y encontrar la expresión aproximada hasta un orden arbitrario. Pero aun si tienen una computadora, son pocas las ecuaciones que pueden resolverse exactamente, de manera que es fundamental saber que existen métodos que prescinden de tales soluciones.

En conclusión, cuando $e^{-\beta W} \ll 1$, la fracción de defectos es

$$x \simeq \sqrt{\frac{y}{\lambda}} = \sqrt{\frac{N'}{N}} e^{-\beta W/2}. \quad (36)$$

Esta fracción es más relevante que el número de defectos en sí mismo, que es

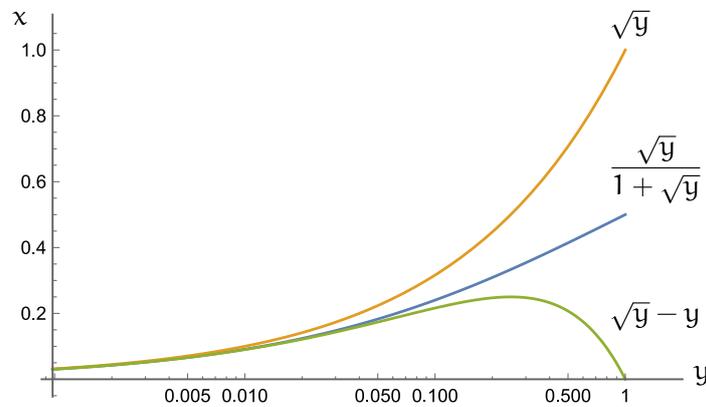
$$n = xN \simeq \sqrt{NN'} e^{-\beta W/2}. \quad (37)$$

Cuando la temperatura es tan baja que $n \sim 1$, la solución deja de ser aplicable.

Si $\lambda = 1$, la solución toma una forma muy simple. Esto se ve directamente de la Ec. (20),

$$x = \frac{\sqrt{y}}{1 + \sqrt{y}}. \quad (38)$$

La figura compara la solución exacta junto con las aproximaciones de orden $y^{1/2}$ y orden y .



Ensamble canónico

En este problema, la promesa de que el ensamble canónico simplifica los cálculos no se cumple. Aunque cambia el contexto, vamos a ver que los cálculos son los mismos.

No hay una forma obvia de organizar la suma sobre estados de modo que la función de partición se factorice. Recurriendo a su expresión en términos de la multiplicidad de estados, la función de partición es

$$Z = \sum_{\mathbf{E}} \Omega(\mathbf{E}) e^{-\beta \mathbf{E}} = \sum_{\mathbf{n}} \binom{N}{\mathbf{n}} \binom{N'}{\mathbf{n}} e^{-\beta W \mathbf{n}}. \quad (39)$$

La solución que han encontrado físicos y matemáticos para escribir el resultado de esta suma consiste en asignarle un nombre a la función así definida. Esta solución es insatisfactoria, pero los nombres son muy creativos. Aceptamos sin gran violencia que Z es cierta función hipergeométrica emparentada con los polinomios de Legendre,

$$Z = {}_2F_1(-N, -N'; 1; e^{-\beta W}). \tag{40}$$

Cuando $N = N'$,

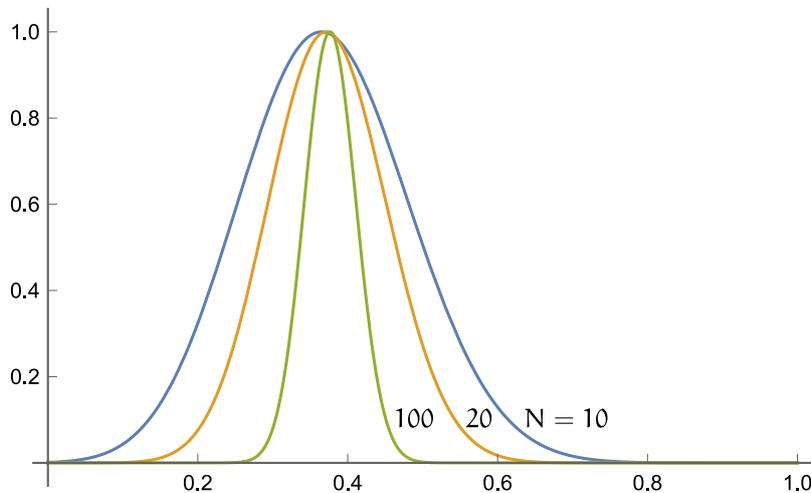
$$Z = (1 - e^{-\beta W})^N P_N\left(\frac{1 + e^{-\beta W}}{1 - e^{-\beta W}}\right). \tag{41}$$

Nos encontramos, así, por primera vez, con el caso de una suma intratable. Lo de intratable es una apreciación subjetiva. Los que sean diestros en la manipulación de funciones hipergeométricas no verán en la expresión (40) nada problemático.

En la Ec. (39), no hemos escrito los límites de la sumatoria. En principio, n podría barrer todos los enteros. Los números binomiales se van a anular cuando n sea negativo o mayor que N o N' . Esto se puede aprovechar muchas veces para manipular los índices de las sumatorias sin tener que preocuparnos por transformar los límites. Pero, si queremos evaluar la suma numéricamente en una computadora, esos límites deben escribirse. Para ser explícitos,

$$Z = \sum_{n=0}^{\min\{N, N'\}} \binom{N}{n} \binom{N'}{n} e^{-\beta W n}. \tag{42}$$

La siguiente figura muestra el gráfico del término general de la suma para el caso $N = N'$ y $\beta W = 1$. Se ha tomado $N = 10$, $N = 20$ y $N = 100$. Las curvas están normalizadas según los dos ejes, y n se ha tratado como una variable continua.



Claramente, hay un término dominante, o, mejor dicho, un intervalo de términos dominantes. Esto lleva a introducir la siguiente aproximación.

Aproximación del término máximo

Al principio esta aproximación les parecerá insensata. En primer lugar, no estamos directamente interesados en la función de partición, sino en su logaritmo,

$$\log Z = \log \left[\sum_{n=0}^{\min\{N, N'\}} \binom{N}{n} \binom{N'}{n} e^{-\beta W n} \right]. \quad (43)$$

O, mejor aún, en esta cantidad dividida por N . Debido a que tenemos pensado tomar el límite termodinámico, debemos construir cantidades que, eventualmente, sean intensivas. Lo que haremos será aproximar el logaritmo de la suma por el logaritmo del máximo término en la suma. En general, eso no tiene mucho sentido. La expresión

$$\log(1 + 2 + \dots + 100) \approx 8.5 \quad (44)$$

no está bien aproximada por $\log 100 \approx 4.6$.

Introduzcamos de nuevo la variable $x = n/N$ y el parámetro $\lambda = N/N'$, y escribamos

$$Z = \sum_x \sigma(x), \quad (45)$$

donde

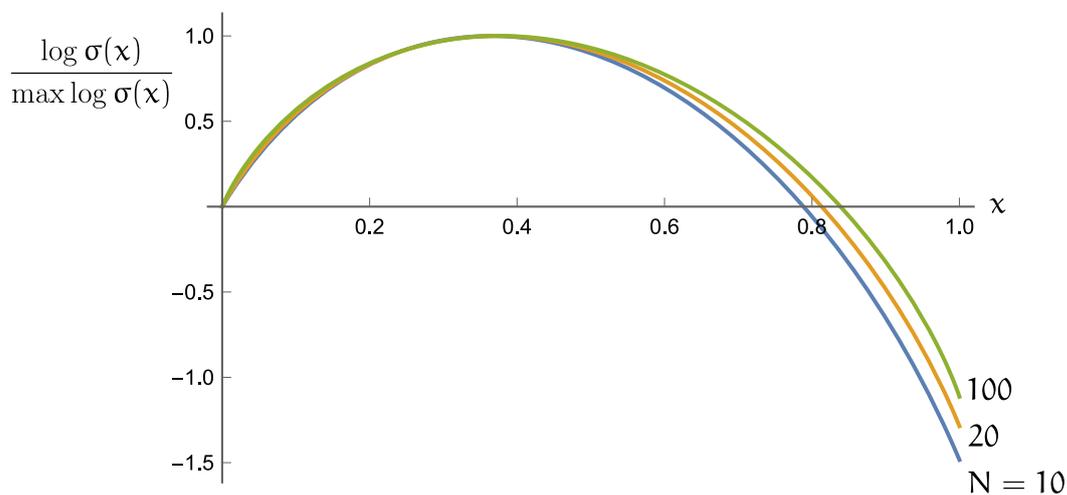
$$\sigma(x) = \binom{N}{xN} \binom{N'}{\lambda x N'} e^{-\beta W x N}. \quad (46)$$

La aproximación que estamos proponiendo es

$$\log Z \simeq \log \max \sigma(x) = \max \log \sigma(x). \quad (47)$$

Después daremos la justificación. Tienen derecho a declarar su incredulidad

La figura muestra la función $\log \sigma(x)$ para tres valores de N , con $N = N'$ y $\beta W = 1$. Notar que todas las curvas están normalizadas según el eje vertical.



Si $N \gg 1$, los valores admisibles de x estarán distribuidos densamente en el intervalo $[0, 1]$, de manera que tiene sentido considerar a $\sigma(x)$ como una función definida en todo ese intervalo, tal como hicimos antes con la función $S(x)$ en el ensamble microcanónico. Para buscar los puntos críticos de $\log \sigma(x)$, podremos usar entonces las herramientas habituales del cálculo diferencial. Asumiendo que todos los números involucrados son mucho mayores que uno, según la Ec. (11) de la sección anterior, la aproximación de Stirling aplicada a $\log \sigma(x)$ implica

$$\log \sigma(x) \simeq Ng(x), \quad (48)$$

con

$$g(x) = f(x) + \frac{1}{\lambda}f(\lambda x) - \beta Wx, \quad (49)$$

donde $f(x)$ es la misma función de la sección anterior,

$$f(x) = -x \log x - (1 - x) \log(1 - x). \quad (50)$$

Siempre es muy importante aislar la dependencia con el tamaño del sistema, como está hecho en la Ec. (48). La función g sólo contiene parámetros y variables intensivos.

A partir de ahora tomaremos el signo aproximado en la Ec. (48) con carácter de igualdad. Podemos cuantificar el error en la aproximación de Stirling, y sabemos que tiende a cero cuando los números involucrados tienden a infinito. El problema no es la aproximación de Stirling, sino la aproximación (47).

Para encontrar los puntos críticos, debemos estudiar las derivadas de $g(x)$:

$$\begin{aligned} g'(x) &= \log\left(\frac{1-x}{x}\right) + \log\left(\frac{1-\lambda x}{\lambda x}\right) - \beta W, \\ g''(x) &= -\frac{1}{x(1-x)} - \frac{1}{x(1-\lambda x)} = -\frac{1}{x} \left(\frac{1}{1-x} + \frac{1}{1-\lambda x} \right). \end{aligned} \quad (51)$$

Haber graficado antes la función $\log \sigma(x)$ es haber hecho un poco de trampa. En un parcial no van a tener una computadora para graficarla. Así que veamos todo lo que podemos extraer de las expresiones de g y de sus derivadas. Por definición, x es mayor o igual que cero y menor o igual que el mínimo entre 1 y λ^{-1} . Esto significa que $g''(x)$ es siempre negativa en ese intervalo. La función $g'(x)$ debe ser, entonces, monótona decreciente. Puesto que hay una completa simetría entre N y N' , podemos suponer que $\lambda \geq 1$, de modo que $x \leq \lambda^{-1} \leq 1$. Cuando $x \rightarrow 0^+$, $g'(x)$ tiende a más infinito. En cambio, cuando $x \rightarrow \lambda^{-1}$ por izquierda, $g'(x)$ tiende a menos infinito. No queda otra alternativa más que la función $g'(x)$ se anule en un único punto del intervalo $(0, \lambda^{-1})$. Ese es el punto crítico que estamos buscando. Además, ya que al atravesar ese punto crítico $g'(x)$ pasa de ser mayor que cero a ser menor que cero, el punto crítico es un máximo de $g(x)$.

Entonces, supongamos que x^* es la solución con sentido físico de la ecuación $g'(x) = 0$, que, en vista de la primera Ec. (51), puede escribirse como

$$\lambda x^2 = (1 - x)(1 - \lambda x)e^{-\beta W}. \quad (52)$$

El análisis del párrafo anterior nos exige de tener que demostrar que x^* existe y que corresponde a un máximo, lo que sería bastante laborioso de hacer a través de la solución explícita de la cuadrática. Sin proponérselo, estamos encontrando los mismos cálculos que tuvimos que hacer en el ensamble microcanónico. Primero, la aproximación de Stirling de los números binomiales, y ahora, la misma cuadrática de la Ec. (20).

En resumen, la aproximación para el logaritmo de Z es

$$\log Z \simeq \max \log \sigma(x) = Ng(x^*). \quad (53)$$

Notemos que, debido a que el término general de la suma en la Ec. (43) es proporcional a la probabilidad, el significado de x^* es el de ser el valor más probable de la fracción de defectos.

Conviene ser más explícitos en cuanto a las dependencias con β . La función σ depende de β a través del término $-N\beta Wx$ en la Ec. (49). Y, a su vez, x^* depende de β , porque es la solución de una ecuación cuadrática que depende de β paramétricamente. Entonces, sería más apropiado escribir

$$\log Z(\beta) \simeq Ng(x^*(\beta), \beta). \quad (54)$$

Para calcular el número medio de defectos, debemos calcular el valor medio de x , lo que es equivalente a calcular el valor medio de la energía, ya que $E = xNW$. Tenemos

$$\langle x \rangle = \frac{1}{NW} \langle E \rangle = -\frac{1}{NW} \frac{\partial \log Z(\beta)}{\partial \beta}. \quad (55)$$

Tienen que prestar atención a esto. La derivada de Z respecto de β es

$$\frac{1}{N} \frac{\partial \log Z(\beta)}{\partial \beta} \simeq \frac{\partial g(x^*(\beta), \beta)}{\partial \beta} = \frac{\partial g}{\partial x}(x^*(\beta), \beta) \frac{\partial x^*(\beta)}{\partial \beta} + \frac{\partial g}{\partial \beta}(x^*(\beta), \beta). \quad (56)$$

Pero, por construcción, la derivada de $g(x, \beta)$ respecto de x se anula en x^* , de modo que

$$\frac{1}{N} \frac{\partial \log Z(\beta)}{\partial \beta} \simeq \frac{\partial g}{\partial \beta}(x^*(\beta), \beta) = -Wx^*(\beta). \quad (57)$$

Esto nos evita calcular la derivada de x^* respecto de β . En definitiva,

$$\langle x \rangle \simeq x^*(\beta). \quad (58)$$

Es natural que, en esta aproximación, el valor medio de x sea igual al valor más probable, porque es el único valor que hemos tenido en cuenta para calcular el logaritmo de la función de partición. Cualquier otro resultado hubiera sido sospechoso.

No necesitamos repetir el análisis que hicimos en el ensamble microcanónico para calcular x^* , porque x^* es solución de la misma cuadrática que apareció al tratar el problema en ese ensamble. Lo que sí requiere una justificación es la aproximación que nos llevó a escribir que el logaritmo de Z puede reemplazarse por el logaritmo del máximo término de la suma.

La justificación del método del término máximo[†]

Por un lado, tenemos que

$$\log Z = \log \sum_x \sigma(x). \quad (59)$$

El número de términos en la suma es igual al mínimo entre N y N' (más uno). Por hipótesis, N y N' son del mismo orden de magnitud. Por otro lado,

$$\log \sigma(x) = Ng(x), \quad (60)$$

donde

$$g(x) = f(x) + \frac{1}{\lambda} f(\lambda x) - \beta Wx. \quad (61)$$

En la Ec. (60), hemos escrito el signo de igualdad, aunque esta expresión se deduce a partir de la aproximación de Stirling. La aproximación de Stirling es asintóticamente válida cuando $N, N' \rightarrow \infty$, de modo que podemos tomarnos la licencia de reemplazar el signo aproximado por el igual, con las mismas limitaciones que en la aproximación de Stirling. Es decir, no podemos afirmar que $\sigma(x) \simeq e^{Ng(x)}$.

La observación crucial es que $g(x)$, restringida al intervalo $[0, 1]$, es una función continua, está acotada y el orden de magnitud de los valores que toma es, a lo sumo, igual al orden de magnitud de βW . Esto quiere decir que, si βW no toma valores extremos, el orden de magnitud de $\log \sigma(x)$ será igual aproximadamente igual al orden de magnitud de N . Y lo mismo será válido para el máximo valor de $\log \sigma$. Es decir,

$$\log \sigma(x^*) = Ng(x^*) \sim N. \quad (62)$$

Evidentemente, como la suma (59) tiene al menos N términos (más uno) y hemos demostrado que el término general de la suma alcanza un máximo en $x = x^*$,

$$\sigma(x^*) \leq \sum_x \sigma(x) \leq N\sigma(x^*). \quad (63)$$

Con la Ec. (59) a la vista, si tomamos el logaritmo de la expresión anterior,

$$\log \sigma(x^*) \leq \log Z \leq \log N + \log \sigma(x^*). \quad (64)$$

[†]Aquí pueden encontrar una explicación más detallada.

Dividiendo por N ,

$$\frac{1}{N} \log \sigma(x^*) \leq \frac{1}{N} \log Z \leq \frac{1}{N} \log N + \frac{1}{N} \log \sigma(x^*). \quad (65)$$

Pero

$$\frac{1}{N} \log \sigma(x^*) = g(x^*). \quad (66)$$

Entonces,

$$g(x^*) \leq \frac{1}{N} \log Z \leq \frac{1}{N} \log N + g(x^*). \quad (67)$$

Si consideramos el límite $N \rightarrow \infty$, obtenemos una igualdad:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Z = g(x^*). \quad (68)$$

Esto es lo que justifica que, para $N \gg 1$, escribamos

$$\log Z \simeq Ng(x^*). \quad (69)$$

Al mismo tiempo hemos demostrado que, cuando $N \rightarrow \infty$, $\log Z$ es una magnitud extensiva y que el límite termodinámico está bien definido.

La aproximación (69) es válida para el logaritmo de Z , pero no para Z . No podemos asegurar que $Z \simeq e^{Ng(x^*)}$. Hemos sido tan poco restrictivos en las cotas, que lo máximo que podemos decir acerca de Z es que

$$\sigma(x^*) \leq Z \leq N\sigma(x^*). \quad (70)$$

No hay ninguna razón para afirmar, entonces, que $Z \simeq \sigma(x^*)$. Si $N = 10^{23}$, el argumento sólo nos permite decir que la incerteza relativa en el valor de Z será igual a 10^{23} . Esto no significa que, en la práctica, la aproximación falle en un factor de este orden, sino que las cotas que encontramos no nos permiten decir nada mejor que eso.

El otro punto importante para recordar es cómo se calcula el valor medio de la energía bajo la aproximación del término máximo. Repitamos el argumento de un modo un poco más general. La aproximación nos permite concluir que

$$\log Z \simeq \log \sigma(x^*) = Ng(x^*), \quad (71)$$

donde x^* es el máximo global de $g(x)$. Que g alcance su máximo global en x^* implica

$$g'(x^*) = 0, \quad (72)$$

salvo casos excepcionales en los que el máximo sea alcanzado en algún extremo del intervalo. En general, esta ecuación depende paraméricamente de β , porque g depende

paramétricamente de β . Por lo tanto, x^* es, en realidad, una función de β ,

$$x^* = x^*(\beta), \quad (73)$$

y deberíamos escribir

$$\log Z(\beta) \simeq Ng(x^*(\beta), \beta). \quad (74)$$

El valor medio de la energía está dado por

$$\langle E \rangle = -\frac{\partial \log Z(\beta)}{\partial \beta} \simeq N \frac{\partial g(x^*(\beta), \beta)}{\partial \beta} = N \left[\frac{\partial g}{\partial x}(x^*(\beta), \beta) \frac{\partial x^*(\beta)}{\partial \beta} + \frac{\partial g}{\partial \beta}(x^*(\beta), \beta) \right]. \quad (75)$$

Por construcción, la derivada de g respecto de x evaluada en x^* es cero, de modo que

$$\langle E \rangle = N \frac{\partial g}{\partial \beta}(x^*(\beta), \beta). \quad (76)$$

Ensamble gran canónico

Hay al menos tres maneras de pensar este sistema en el ensamble gran canónico:

- Como un sistema en sí mismo, formado por los N sitios de la red y los N' sitios intersticiales.
- Como un sistema compuesto, formado por dos subsistemas. Por un lado, los N sitios de la red; por el otro, los N' sitios intersticiales. Los dos subsistemas están en equilibrio termodinámico y en equilibrio químico.
- Como un sistema compuesto, formado por $N + N'$ subsistemas. Los N sitios de la red y los N' sitios intersticiales considerados individualmente.

Evidentemente, hay otras alternativas. Por ejemplo, el sistema puede tratarse como un sistema compuesto por dos subsistemas, los sitios de la red y los sitios intersticiales, y, a su vez, cada uno de estos subsistemas puede pensarse como la unión de N o N' sistemas elementales. Es difícil mantenerse siempre dentro de una misma interpretación, y tampoco hay ninguna obligación de hacerlo.

El sistema considerado como un todo

El sistema está formado por los N sitios de la red y por los N' sitios intersticiales, ocupados por un número variable de partículas. En promedio, tiene que haber una partícula por cada sitio de la red. De modo que el número medio de partículas debe ser N . La función de partición es

$$\mathcal{Z}(\beta, z) = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} = \sum_m \sum_{\mathcal{C}_m} e^{-\beta E(\mathcal{C}_m)} z^m = \sum_{m=0}^{N+N'} Z_m(\beta) z^m. \quad (77)$$

Aquí, $\{\mathcal{C}\}$ es el conjunto de todos los estados; $E(\mathcal{C})$ y $m(\mathcal{C})$ son, respectivamente, la energía y el número de partículas del estado \mathcal{C} . El conjunto $\{\mathcal{C}_m\}$ es el conjunto de estados de m partículas, Z_m es la función de partición canónica de m partículas y z es la fugacidad. Es cierto que hay una relación entre la fugacidad y el potencial químico,

$$z = e^{\beta\mu}, \quad (78)$$

pero la variable de la que depende \mathcal{Z} es la fugacidad. La fugacidad no debe considerarse una función de la temperatura. Toda la dependencia con β que tiene la función de partición \mathcal{Z} viene a través del factor de Boltzmann. Al tomar la derivada parcial de \mathcal{Z} respecto de β , no hay que derivar la fugacidad. Ese es uno de los errores más comunes.

Si la única manera de calcular la función de partición en el ensamble gran canónico fuera a través de la función de partición canónica, entonces no habría realmente una ventaja en pasar al ensamble gran canónico. Lo que mostraremos es que el cálculo de \mathcal{Z} puede prescindir del conocimiento de Z_m . Hay que notar que Z_m no es la función de partición que hemos calculado en la sección anterior. En la sección anterior había tantas partículas como sitios en la red. Ahora puede haber un número arbitrario de partículas, siempre que sea menor que $N + N'$. En la sección anterior, sólo calculamos Z_N .

Los estados con un número arbitrario de partículas pueden recorrerse sumando independientemente sobre el número de partículas en los sitios de la red y en los sitios intersticiales,

$$\mathcal{Z} = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} = \sum_{i=0}^N \sum_{j=0}^{N'} \binom{N}{i} \binom{N'}{j} e^{-\beta W_j} z^{i+j} = (1+z)^N (1+ze^{-\beta W})^{N'}. \quad (79)$$

Los coeficientes binomiales tienen en cuenta que, si hay i partículas en los sitios de la red, existen $\binom{N}{i}$ maneras de distribuirlas entre los N sitios; y análogamente para los sitios intersticiales. Para este sistema, la función de partición en el ensamble gran canónico se factoriza. Esta simplificación es análoga a la que ocurre en el sistema de los dos niveles cuando se pasa del ensamble microcanónico al canónico. El problema de los defectos es mucho más simple en el ensamble gran canónico.

Si sólo considerásemos un sitio de la red, su función de partición sería

$$\mathcal{Z}_1 = \sum_{n=0}^1 z^n = 1 + z. \quad (80)$$

Del mismo modo, si sólo considerásemos un sitio intersticial, su función de partición sería

$$\mathcal{Z}'_1 = \sum_{n=0}^1 e^{-\beta W_n} z^n = 1 + ze^{-\beta W}. \quad (81)$$

De manera que el resultado (79) puede escribirse como

$$\mathcal{Z} = \mathcal{Z}_1^N \mathcal{Z}'_1^{N'}. \quad (82)$$

Hay todavía una forma más simple de ver que en este problema la función de partición debe factorizarse. Los estados con un número arbitrario de partículas pueden recorrerse especificando el número de partículas en cada sitio de la red y en cada sitio intersticial,

$$\mathcal{Z} = \left[\sum_{r_1=0}^1 \dots \sum_{r_N=0}^1 \right] \left[\sum_{i_1=0}^1 \dots \sum_{i_{N'}=0}^1 \right] z^{r_1+\dots+r_N+i_1+\dots+i_{N'}} e^{-\beta W(i_1+\dots+i_{N'})}. \quad (83)$$

Los índices r_j corresponden a los sitios de la red y los índices i_j a los sitios intersticiales. Como las exponenciales se factorizan, queda

$$\mathcal{Z} = \left(\sum_{j=0}^1 z^j \right)^N \left(\sum_{j=0}^1 e^{-\beta W_j} z^j \right)^{N'} = \mathcal{Z}_1^N \mathcal{Z}'_1^{N'}. \quad (84)$$

Es el mismo resultado que obtuvimos antes.

En resumen,

$$\mathcal{Z} = \mathcal{Z}_1^N \mathcal{Z}'_1^{N'} = (1+z)^N (1+ze^{-\beta W})^{N'}. \quad (85)$$

Nadie está muy interesado en la fugacidad. La fugacidad es un mero expediente para calcular la fracción de defectos. Ahora veremos cómo. El número medio de partículas es

$$\langle m \rangle = z \frac{\partial \log \mathcal{Z}}{\partial z} = N \left(z \frac{\partial \log \mathcal{Z}_1}{\partial z} \right) + N' \left(z \frac{\partial \log \mathcal{Z}'_1}{\partial z} \right). \quad (86)$$

Esta ecuación puede leerse del siguiente modo: $\langle m \rangle$ es igual a N veces el número medio de ocupación de cada sitio de la red más N' veces el número medio de ocupación de cada sitio intersticial. O, si se prefiere, diciendo que el número medio de partículas es igual al número medio de partículas en los sitios de la red más el número medio de partículas en los sitios intersticiales. Explícitamente,

$$\langle m \rangle = \langle m_r \rangle + \langle m_i \rangle = \frac{Nz}{1+z} + \frac{N'ze^{-\beta W}}{1+ze^{-\beta W}}. \quad (87)$$

El número medio de partículas debe ser igual al número de sitios de la red. Luego,

$$N = \frac{Nz}{1+z} + \frac{N'ze^{-\beta W}}{1+ze^{-\beta W}}. \quad (88)$$

Como antes, definamos $y = e^{-\beta W}$ y $\lambda = N/N'$. Entonces,

$$1 = \frac{z}{1+z} + \frac{yz}{\lambda(1+yz)}. \quad (89)$$

Esta es una ecuación cuadrática para z . Para encontrar el número medio de defectos, debemos encontrar la solución \bar{z} de esta ecuación, y luego evaluar el número medio de defectos en ese valor de z . La fracción media de defectos será

$$\bar{x} = \frac{y\bar{z}}{\lambda(1 + y\bar{z})}. \quad (90)$$

Pero como lo que nos interesa es comparar \bar{x} con las fracciones de defectos obtenidas en los otros ensambles, conviene reescribir la cuadrática (89) en términos de x , donde

$$x = \frac{yz}{\lambda(1 + yz)} \quad (91)$$

es la fracción media de defectos para un z cualquiera. De aquí se obtiene

$$yz = \frac{\lambda x}{1 - \lambda x}. \quad (92)$$

Reemplazando en la Ec. (89),

$$1 = \frac{\lambda x}{y - y\lambda x + \lambda x} + x. \quad (93)$$

Entonces,

$$\lambda(1 - y)x^2 + y(1 + \lambda)x - y = 0. \quad (94)$$

La solución de esta ecuación es \bar{x} . Esta es la misma ecuación cuadrática que apareció en los otros ensambles, por ejemplo, en la Ec. (25). Por construcción, la solución es la fracción media de defectos cuando el número medio de partículas es N . Aunque en el ensamble canónico también obtuvimos una fracción media de defectos, la interpretación es levemente diferente. En el ensamble canónico, el número de partículas estaba fijo. En el ensamble gran canónico, ese número fluctúa. En el ensamble gran canónico, \bar{x} fluctúa porque los sitios de la red intercambian partículas con los sitios intersticiales y porque los sitios intersticiales intercambian partículas con el reservorio. Esta diferencia entre los ensambles no se manifiesta en los valores medios de x , pero debería aparecer en el cálculo de las fluctuaciones. Esto queda como ejercicio.

El sistema considerado como un sistema compuesto

Sin alterar ninguno de los resultados anteriores ni afectar los cálculos en forma alguna, el problema puede pensarse de una manera alternativa. Es posible tratar al sistema de los sitios de la red y de los sitios intersticiales como dos sistemas independientes. Los dos sistemas están en equilibrio y pueden intercambiar partículas entre sí y con un reservorio.

Entonces, uno puede escribir la función de partición para cada sistema por separado,

$$\begin{aligned} \mathcal{Z}_r &= \sum_{m=0}^N \binom{N}{m} z^m = (1+z)^N, \\ \mathcal{Z}_i &= \sum_{m=0}^{N'} \binom{N'}{m} e^{-\beta W m} z^m = (1 + ze^{-\beta W})^{N'}. \end{aligned} \quad (95)$$

Aquí, \mathcal{Z}_r es la función de partición gran canónica de los sitios de la red y \mathcal{Z}_i es la función de partición de los sitios intersticiales. El hecho de que los dos sistemas estén en equilibrio y de que puedan intercambiar partículas implica que comparten los mismos valores de β y z . En otras palabras, están a la misma temperatura y tienen el mismo potencial químico. Alternativamente, podríamos haber escrito, de manera más directa,

$$\begin{aligned} \mathcal{Z}_r &= \mathcal{Z}_1^N, \\ \mathcal{Z}_i &= \mathcal{Z}'_1^{N'}, \end{aligned} \quad (96)$$

donde \mathcal{Z}_1 y \mathcal{Z}'_1 son las funciones de partición gran canónicas de un sitio de la red y de un sitio intersticial, respectivamente; Ecs. (80) y (81).

El número medio de partículas en los sitios de la red es

$$\langle m_r \rangle = z \frac{\partial \log \mathcal{Z}_r}{\partial z} = \frac{Nz}{1+z}, \quad (97)$$

y en los sitios intersticiales,

$$\langle m_i \rangle = z \frac{\partial \log \mathcal{Z}_i}{\partial z} = \frac{N'ze^{-\beta W}}{1 + ze^{-\beta W}}. \quad (98)$$

Además, existe la condición de que el número medio total de partículas sea igual a N ,

$$\langle m_r \rangle + \langle m_i \rangle = N. \quad (99)$$

Es fácil ver que, a partir de aquí, se obtienen los mismos resultados de antes. Esta forma de pensar el problema en términos de un sistema compuesto resulta útil en muchos otros casos.

Otras relaciones en el ensamble gran canónico

Antes escribimos que el número medio de partículas en los sitios intersticiales es

$$\langle m_i \rangle = z \frac{\partial \log \mathcal{Z}_i}{\partial z} = N'z \frac{\partial \log \mathcal{Z}'_1}{\partial z} = \frac{N'ze^{-\beta W}}{1 + ze^{-\beta W}}. \quad (100)$$

Debido a que hay una relación unívoca entre la energía y el número de defectos, deberíamos poder obtener $\langle m_i \rangle$ a partir de $\langle E \rangle$. Debería ser

$$\langle m_i \rangle = \frac{1}{W} \langle E \rangle. \quad (101)$$

En el ensamble gran canónico vale la misma fórmula que en el canónico,

$$\langle E \rangle = -\frac{\partial \log \mathcal{Z}}{\partial \beta}. \quad (102)$$

Teniendo en cuenta la Ec. (85),

$$\langle E \rangle = -N' \frac{\partial \log \mathcal{Z}'_1}{\partial \beta} = N' W \frac{ze^{-\beta W}}{1 + ze^{-\beta W}}. \quad (103)$$

Entonces, tal como habíamos previsto,

$$\langle m_i \rangle = \frac{N' ze^{-\beta W}}{1 + ze^{-\beta W}}. \quad (104)$$

Esto no es un gran avance, porque de todas maneras tendremos que resolver la Ec. (89) o, equivalentemente, la Ec. (94). Sólo muestra la consistencia del formalismo.

El cálculo anterior sirve como excusa para formular la siguiente observación: cuando se calcula la energía media derivando respecto de β , la función que se deriva es $\mathcal{Z}(\beta, z)$, no $\mathcal{Z}(\beta, e^{\beta\mu})$. La confusión que produce esta regla es motivo de muchos errores. Desde una perspectiva probabilística, no hay lugar a dudas. La distribución de probabilidad en el ensamble gran canónico es

$$P(\mathcal{C}) = \frac{1}{\mathcal{Z}} e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})}. \quad (105)$$

El valor medio de la energía está dado por

$$\langle E \rangle = \sum_{\mathcal{C}} E(\mathcal{C}) P(\mathcal{C}) = \frac{1}{\mathcal{Z}} \sum_{\mathcal{C}} E(\mathcal{C}) e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} = -\frac{1}{\mathcal{Z}} \frac{\partial}{\partial \beta} \left[\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} \right]. \quad (106)$$

Finalmente,

$$\langle E \rangle = -\frac{\partial \log \mathcal{Z}(\beta, z)}{\partial \beta}. \quad (107)$$

Esto mismo puede demostrarse a partir de la conexión entre la mecánica estadística y la termodinámica. En el ensamble gran canónico,

$$\Omega = -kT \log \mathcal{Z}, \quad (108)$$

donde Ω es el gran potencial. Lo importante es que Ω es la transformada de Legendre de la energía respecto a la entropía y al número de partículas,

$$\Omega = U - TS - \mu m. \quad (109)$$

Además, el primer principio expresado en términos de Ω se escribe como

$$d\Omega = -SdT - m d\mu. \quad (110)$$

El significado de esta ecuación es que

$$\frac{\partial \Omega(T, \mu)}{\partial T} = -S, \quad \frac{\partial \Omega(T, \mu)}{\partial \mu} = -m. \quad (111)$$

Entonces, si queremos calcular la energía a partir de Ω , tendremos

$$U = \Omega + TS + \mu m = \Omega - T \frac{\partial \Omega}{\partial T} - \mu \frac{\partial \Omega}{\partial \mu}. \quad (112)$$

Las derivadas que aparecen aquí son las derivadas de Ω considerado como una función de T y μ . La relación (108) debe escribirse con atención a este hecho,

$$\Omega(T, \mu) = -kT \log \mathcal{Z}(\beta(T), z(\beta(T), \mu)). \quad (113)$$

Luego,

$$U = -kT \log \mathcal{Z} - T \left[-k \log \mathcal{Z} - kT \left(\frac{\partial \log \mathcal{Z}}{\partial \beta} + \frac{\partial \log \mathcal{Z}}{\partial z} \frac{\partial z}{\partial \beta} \right) \frac{\partial \beta}{\partial T} \right] + \mu kT \frac{\partial \log \mathcal{Z}}{\partial z} \frac{\partial z}{\partial \mu}. \quad (114)$$

Luego de una sucesión de cancelaciones fortuitas, y con gran peligro de cometer algún error, llegamos al mismo resultado de antes,

$$U = -\frac{\partial \log \mathcal{Z}(\beta, z)}{\partial \beta}. \quad (115)$$

El argumento probabilístico es mucho más directo.

Del mismo modo, hay al menos dos caminos para demostrar que el número medio de partículas está dado por

$$\langle m \rangle = z \frac{\partial \log \mathcal{Z}(\beta, z)}{\partial z}. \quad (116)$$

En términos de las probabilidades,

$$\langle m \rangle = \sum_{\mathcal{C}} m(\mathcal{C}) P(\mathcal{C}) = \frac{1}{\mathcal{Z}} \sum_{\mathcal{C}} m(\mathcal{C}) e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} = \frac{z}{\mathcal{Z}} \frac{\partial}{\partial z} \left[\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} z^{m(\mathcal{C})} \right] = z \frac{\partial \log \mathcal{Z}(\beta, z)}{\partial z}. \quad (117)$$

Partiendo desde la termodinámica, deberían calcular

$$m = -\frac{\partial \Omega(T, \mu)}{\partial \mu} = kT \frac{\partial}{\partial \mu} \left[\log \mathcal{Z}(\beta(T), z(\beta(T), \mu)) \right] \quad (118)$$

y verificar que se obtiene la misma expresión. Esta verificación queda como ejercicio. Tengan presente la siguiente salvedad: lo que en termodinámica es una cantidad perfectamente definida, en mecánica estadística suele ser un valor medio.