

Física Teórica 3 — segundo cuatrimestre de 2024

EL MÉTODO DEL TÉRMINO MÁXIMO*

Muchas veces, nos encontramos con sumas a primera vista intratables. Esto debe ser tomado con un valor subjetivo. Muy específicamente, su significado es que *nosotros* no tenemos la menor idea de cómo calcular ciertas sumas, salvo, acaso, por métodos numéricos. Supondremos que la suma que queremos calcular es de la forma

$$Z = \sum_{n=0}^N s(n), \quad (1)$$

con N finito y positivo. Usamos el símbolo Z por comodidad, pero esto puede ocurrir en cualquier ensamble. Introduciendo la variable

$$x = \frac{n}{N}, \quad (2)$$

queda

$$Z = \sum_x \sigma(x), \quad (3)$$

donde

$$\sigma(x) = s(xN). \quad (4)$$

El índice x toma los valores $0, \frac{1}{N}, \frac{2}{N}, \dots, 1$. Como los valores sucesivos de x están separados por una distancia $\Delta x = N^{-1}$, si $N \gg 1$, estos puntos cubrirán muy densamente el intervalo $[0, 1]$. De este modo, a todos los efectos prácticos, la función $\sigma(x)$ puede considerarse como una función continua y derivable, definida en todo el intervalo.

Lo primero que a uno suele ocurrírsele para obtener el valor aproximado de una suma es reemplazarla por una integral,

$$Z = \frac{1}{\Delta x} \sum_x \sigma(x) \Delta x \simeq N \int_0^1 dx \sigma(x). \quad (5)$$

La validez de esta aproximación sólo depende de que N sea lo suficientemente grande. El problema es que, incluso cuando tengamos más familiaridad con el cálculo de integrales que con el cálculo de sumas, la integral puede ser, al igual que la suma, intratable.

Resumido en pocas palabras, el método del término máximo afirma que, si $\sigma(x)$ es una función con forma de campana muy aguda en torno de cierto valor de x en donde alcanza su máximo, entonces, el logaritmo de Z puede aproximarse por el logaritmo del máximo término de la suma,

$$\log Z \simeq \log \max \sigma(x) = \max \log \sigma(x). \quad (6)$$

*zanellaj@df.uba.ar

Si nos hubieran pedido la peor aproximación que se nos viniera a la mente, es probable que no alcanzásemos nunca el grado de negligencia que trasunta la aproximación del término máximo. Es muy parecido a preguntarle a una persona cuánto vale la suma de todos los enteros entre uno y un millón y que nos contestara que esa suma es, muy aproximadamente, igual a un millón. Esta respuesta erra por mucho. Lo asombroso, lo que está más allá de toda intuición, es que, en mecánica estadística, el método del término máximo da aproximaciones de gran exactitud para N finito y exactas en el límite $N \rightarrow \infty$.

En primer lugar, debemos decir que el método del término máximo no es un método para calcular Z , sino su logaritmo,

$$\log Z = \log \sum_x \sigma(x). \quad (7)$$

El método del término máximo aproxima el logaritmo de Z por el logaritmo del máximo término de la suma. Aun así, si le preguntáramos a la misma persona de antes cuánto vale el logaritmo de la suma de todos los enteros entre uno y un millón, y nos contestara que esa suma está muy bien aproximada por el logaritmo de un millón, su respuesta no sería mucho más exacta que la anterior. El método es aplicable a muchas de las sumas que aparecen en mecánica estadística, no a cualquier suma. No es aplicable, por ejemplo, a la suma de todos los enteros entre uno y un millón.

El método asegura que, bajo ciertas condiciones,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Z = \lim_{N \rightarrow \infty} \max \frac{1}{N} \log \sigma(x). \quad (8)$$

Esto podría ser trivialmente cierto si ambos límites fueran cero. Asumiremos que no es el caso. En general, para N finito, valdrá la aproximación

$$\log Z \simeq \max \log \sigma(x). \quad (9)$$

Las condiciones necesarias para que sea aplicable el método del término máximo son dos. Primero, que el logaritmo de σ pueda desarrollarse alrededor de $N = \infty$ como

$$\log \sigma(x) = Ng(x) + o(N), \quad (10)$$

y, en segundo lugar, que la función $g(x)$ alcance su máximo global en un único punto x^* en el intervalo $[0, 1]$. En la práctica, es conveniente que ese punto esté en el interior del intervalo, de forma que pueda determinarse partir de la ecuación

$$g'(x^*) = 0. \quad (11)$$

En adelante, asumiremos que esto es cierto.

La demostración de la Ec. (8) se basa en una serie de desigualdades desconcertantes, porque no hay ningún esfuerzo en imponer cotas precisas. La demostración se limita a proponer las cotas menos precisas. Eso habla de la robustez del método.

Evidentemente, como la suma de la Ec. (7) tiene $N + 1$ términos,

$$\max \sigma(x) \leq \sum_x \sigma(x) \leq (N + 1) \max \sigma(x). \quad (12)$$

Con la Ec. (7) a la vista, si tomamos el logaritmo de la expresión anterior,

$$\log \max \sigma(x) \leq \log Z \leq \log(N + 1) + \log \max \sigma(x). \quad (13)$$

El logaritmo es una función creciente, por lo tanto podemos invertir el orden en el que calculamos el máximo y tomamos el logaritmo,

$$\max \log \sigma(x) \leq \log Z \leq \log(N + 1) + \max \log \sigma(x). \quad (14)$$

Dividiendo por N ,

$$\max \frac{1}{N} \log \sigma(x) \leq \frac{1}{N} \log Z \leq \frac{1}{N} \log(N + 1) + \max \frac{1}{N} \log \sigma(x). \quad (15)$$

Pero, por hipótesis,

$$\frac{1}{N} \log \sigma(x) = g(x) + \frac{o(N)}{N}. \quad (16)$$

Entonces,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \sigma(x) = g(x). \quad (17)$$

Luego,

$$\lim_{N \rightarrow \infty} \max \frac{1}{N} \log \sigma(x) = g(x^*). \quad (18)$$

Volviendo a la Ec. (15) y tomando el límite $N \rightarrow \infty$, resulta

$$g(x^*) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \log Z \leq g(x^*). \quad (19)$$

Hemos demostrado así que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Z = g(x^*). \quad (20)$$

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

$$\frac{1}{N} \log Z \simeq g(x^*). \quad (21)$$

Al mismo tiempo hemos demostrado que, cuando $N \rightarrow \infty$, $\log Z$ es una magnitud extensiva respecto de N . Si N es el número de partículas y se trata del ensamble microcanónico, habremos demostrado, por ejemplo, la extensividad de la entropía; si la suma es la función de partición en el ensamble canónico, habremos demostrado la extensividad de la energía libre de Helmholtz.

La primera objeción que a uno se le ocurre cuando le proponen aproximar el logaritmo de Z por el logaritmo del máximo término de la suma es que, si bien estamos de acuerdo en que habrá un máximo término en la suma, los términos vecinos no serán mucho menores que este máximo. Quizá haya 100 o 1000 o 100 000 términos que sean comparables al máximo término de la suma. Lo que demuestra la serie de desigualdades que hemos escrito más arriba es que, aun si los N términos de la suma fueran todos comparables al término máximo, seguiría siendo cierto que

$$\log[N \max \sigma(x)] \simeq \max \log \sigma(x), \quad (22)$$

porque $\max \sigma(x) \sim e^N$ y, entonces, el error relativo es del orden de $\frac{1}{N} \log N$. Por ejemplo,

$$\log(10^{23} e^{10^{23}}) = \log e^{23 \log 10 + 10^{23}} \simeq \log e^{10^{23}}. \quad (23)$$

No hay que mirar lo que pasa debajo, sino lo que pasa en los exponentes.

La aproximación (21) es válida para el logaritmo de Z , pero no para Z . No podemos asegurar que $Z \simeq e^{Ng(x^*)}$, del mismo modo que la relación (10) tampoco nos permite afirmar que $\sigma(x) \simeq e^{Ng(x)}$. Hemos sido tan poco estrictos en las cotas, que todo lo que podemos decir acerca de Z es que

$$\max \sigma(x) \leq Z \leq N \max \sigma(x). \quad (24)$$

No hay ninguna razón para afirmar, entonces, que $Z \simeq \max \sigma(x)$. Puede haber una discrepancia de orden N entre estas dos cantidades.

El hecho de que la aproximación del término máximo funcione para el logaritmo de la suma, pero no para la suma, puede compararse con el resultado análogo para la aproximación de Stirling. En su forma más elemental, la aproximación de Stirling afirma que

$$\log N! \simeq N \log N - N. \quad (25)$$

Esta aproximación es válida para $\log N!$, pero no puede usarse para obtener una aproximación de $N!$ de la forma

$$N! \simeq \left(\frac{N}{e}\right)^N. \quad (26)$$

La aproximación para Z puede corregirse del mismo modo en el que se corrige la aproximación de Stirling, mediante un factor multiplicativo, usualmente del orden de $\sqrt{N}$. En el caso de la aproximación de Stirling, la fórmula válida hasta orden $\sqrt{N}$ es

$$\log N! \simeq N \log N - N + \sqrt{2\pi N}. \quad (27)$$

El término siguiente es de orden N^{-1} . Entonces, sí es válida la aproximación

$$N! \simeq \sqrt{2\pi N} \left(\frac{N}{e}\right)^N. \quad (28)$$

La aproximación del término máximo funciona, en general, cuando Z está dada por una suma que tiene del orden de $N \gg 1$ términos y cuando el término general de la suma tiene un máximo global cuyo comportamiento está dominado por un término del orden de e^N . Esos son, en esencia, los dos resultados que nos permitieron llegar a la conclusión de que la aproximación del término máximo es exacta cuando $N \rightarrow \infty$.

Para convencernos, veámoslo con algunos números. Supongamos que $\max \sigma(x) \simeq e^{aN}$ y que $N = 10^{23}$, con a de orden uno. Entonces, podemos afirmar que

$$10^{23}a \leq \log Z \leq \log 10^{23} + 10^{23}a, \quad (29)$$

o bien,

$$a \leq \frac{1}{N} \log Z \leq 23 \times 10^{-23} \log 10 + a. \quad (30)$$

El ancho relativo de la cota para $\frac{1}{N} \log Z$ es del orden de 10^{-22} . El error es insignificante. Ese es el efecto nivelador del logaritmo. En cambio, si exponenciamos cada término de la desigualdad (29), obtenemos

$$e^{a10^{23}} \leq Z \leq 10^{23} e^{a10^{23}}. \quad (31)$$

Esta cota para Z es muy poco informativa. Es como si dijéramos que la altura de una persona estuviera entre $1 \text{ \AA}$ y el radio del sistema solar.

Derivadas de $\log Z$ bajo la aproximación del término máximo

Es importante saber cómo se calculan las derivadas de $\log Z$ bajo la aproximación del término máximo. El logaritmo de Z puede depender de varios parámetros, como la energía (si estamos en el ensamble microcanónico), la temperatura, la fugacidad, etc. Por simplicidad, supongamos que es $Z = Z(\lambda)$. El origen de esta dependencia sólo puede estar en el hecho de que la propia función $\sigma(x)$ dependa de λ . Por lo tanto, $g(x)$, definido por la Ec. (10), también dependerá de λ ,

$$g(x) = g(x, \lambda). \quad (32)$$

La aproximación del término máximo se lee como

$$\log Z(\lambda) \simeq Ng(x^*, \lambda), \quad (33)$$

donde x^* es el máximo global de $g(x, \lambda)$ respecto de x para λ fijo. Por hipótesis, $g(x, \lambda)$ alcanza su máximo en el interior del intervalo $[0, 1]$. Entonces,

$$\frac{\partial g}{\partial x}(x^*, \lambda) = 0. \quad (34)$$

Esta ecuación determina x^* y depende paramétricamente de λ . Luego, x^* será, en realidad, una función de λ ,

$$x^* = x^*(\lambda), \quad (35)$$

y deberíamos escribir

$$\log Z(\lambda) \simeq Ng(x^*(\lambda), \lambda). \quad (36)$$

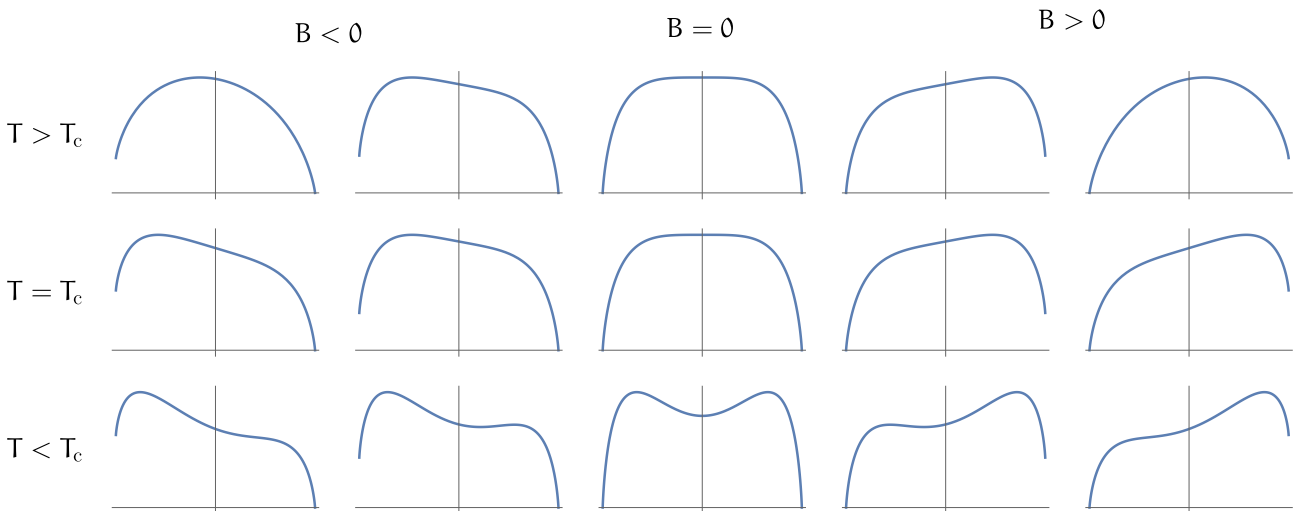
Finalmente, cuando calculemos la derivada de $\log Z$ respecto de λ , obtendremos

$$\frac{1}{N} \frac{\partial \log Z(\lambda)}{\partial \lambda} \simeq \frac{\partial g(x^*(\lambda), \lambda)}{\partial \lambda} = \frac{\partial g(x^*(\lambda), \lambda)}{\partial x} \frac{\partial x^*(\lambda)}{\partial \lambda} + \frac{\partial g(x^*(\lambda), \lambda)}{\partial \lambda}. \quad (37)$$

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

$$\frac{1}{N} \frac{\partial \log Z(\lambda)}{\partial \lambda} \simeq \frac{\partial g(x^*(\lambda), \lambda)}{\partial \lambda}. \quad (38)$$

Un punto esencial en todo este argumento es que $\sigma(x)$ tenga un sólo máximo global. Bien podría ocurrir que $\sigma(x)$ alcance su valor máximo en más de un punto. Es usual que $\sigma(x)$ dependa de varios parámetros. El paisaje de máximos y mínimos depende de esos parámetros. Así, no es imposible que, variando esos parámetros, uno consiga que la función $\sigma(x)$ tenga más de un máximo global. Eso suele indicar que hay una transición de fase. En ese caso, el método del término máximo requiere una generalización más o menos trivial. La figura muestra la función $\log \sigma(x)$ para un sistema ferromagnético que depende de dos parámetros, la temperatura y un campo externo B . Cuando $B = 0$, al atravesar la temperatura crítica, el sistema tiene una transición de fase: para $B = 0$ y $T > T_c$, $\sigma(x)$ tiene un máximo global en $x = 0$; para $B = 0$ y $T < T_c$, $\sigma(x)$ tiene dos máximos globales.



Por el momento, la existencia de múltiples máximos no será una de nuestras preocupaciones. Recién veremos este problema cuando estudiemos fenómenos críticos.

El sistema de los dos niveles en la aproximación del término máximo

Este es un buen ejemplo para demostrar la eficacia del método, porque conocemos la solución exacta. En el sistema de los dos niveles, hay N elementos distinguibles y no interactuantes, cada uno de los cuales puede estar en dos estados, con energías cero y ϵ . La función de partición canónica se factoriza trivialmente y resulta

$$\frac{1}{N} \log Z = \log(1 + e^{-\beta\epsilon}). \quad (39)$$

Ahora bien, uno puede estar en un mal día y no darse cuenta de que la función de partición se factoriza. Puede escribir que Z es

$$Z = \sum_E \Omega(E) e^{-\beta E} = \sum_{n=0}^N \binom{N}{n} e^{-\beta\epsilon n}. \quad (40)$$

Puede no darse cuenta de que esto es el desarrollo de un binomio y puede llegar al extremo de considerar que es una suma intratable, lo que, por definición, sería cierto en ese día en particular. Pero como uno conoce el método del término máximo, intenta aplicarlo.

Primero, introduciendo la variable $x = n/N$, el término general de la suma es

$$\sigma(x) = \binom{N}{xN} e^{-\beta\epsilon Nx}. \quad (41)$$

Mediante la aproximación de Stirling,

$$\log \sigma(x) \simeq N[f(x) - \beta\epsilon x] = Ng(x), \quad (42)$$

donde

$$f(x) = -x \log x - (1-x) \log(1-x). \quad (43)$$

Por otro lado,

$$g'(x) = f'(x) - \beta\epsilon = \log\left(\frac{1-x}{x}\right) - \beta\epsilon. \quad (44)$$

Esto se anula en x^* tal que

$$x^* = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}. \quad (45)$$

Es fácil verificar que, según el método del término máximo, resulta

$$\frac{1}{N} \log Z \simeq g(x^*) = \log(1 + e^{-\beta\epsilon}). \quad (46)$$

Para nuestra sorpresa, este es el resultado exacto. La desventaja de haberlo obtenido mediante el método del término máximo, más allá de que es mucho más trabajoso, es que uno está sujeto a asumir que el resultado es válido sólo cuando $N \gg 1$. En realidad sabemos que este resultado es válido aun cuando $N = 1$, y de manera exacta, no aproximada.

El problema de los defectos de Frenkel

No repetiremos aquí lo que está dicho en otras notas. Nos limitaremos a transcribir los resultados principales. Se trata del problema 3 de la Guía 5, tomando, por simplicidad, $N = N'$. Basta saber que la función de partición es

$$Z = \sum_{n=0}^N \binom{N}{n}^2 e^{-\beta W n}, \quad (47)$$

o, en términos de $x = n/N$ e $y = e^{-\beta W}$,

$$Z(y, N) = \sum_x \binom{N}{xN}^2 y^{xN}. \quad (48)$$

Aquí, n representa el número de defectos en una red de N átomos y x es la fracción de defectos. El término general de la suma es

$$\sigma(x) = \binom{N}{xN}^2 y^{xN}. \quad (49)$$

Usando la aproximación de Stirling, se encuentra

$$\sigma(x) \simeq N g(x) \quad (50)$$

con

$$g(x) = 2f(x) + x \log y, \quad (51)$$

donde $f(x)$ es la función definida en la Ec. (43). La función $g(x)$ alcanza su máximo en

$$x^* = \frac{\sqrt{y}}{1 + \sqrt{y}}. \quad (52)$$

Bajo esta aproximación, el valor medio de x es, justamente, x^* .

Ahora calcularemos el error cometido al aplicar el método del término máximo para aproximar Z , $\log Z$ y $\langle x \rangle$. En lo que sigue escribiremos directamente x en lugar de $\langle x \rangle$.

Existe para Z una fórmula explícita en términos de los polinomios de Legendre,

$$Z = (1 - e^{-\beta W})^N P_N \left(\frac{1 + e^{-\beta W}}{1 - e^{-\beta W}} \right), \quad (53)$$

o, usando la variable y ,

$$Z(y, N) = (1 - y)^N P_N \left(\frac{1 + y}{1 - y} \right). \quad (54)$$

Si uno está lo suficientemente familiarizado con los polinomios de Legendre, no verá en la Ec. (47) ninguna suma intratable, confirmando que el calificativo de intratable tiene un carácter relativo.

La fracción media de defectos, obtenida a partir de la expresión exacta (54), es

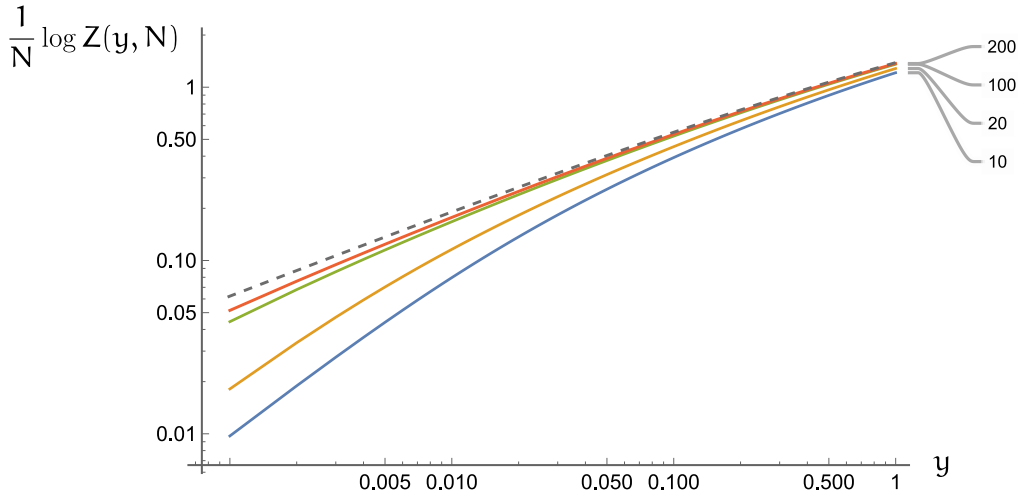
$$x(y, N) = \frac{y}{N} \frac{\partial \log Z}{\partial y} = \frac{1}{2N} \left[(N+1) \frac{P_{N+1} \left(\frac{1+y}{1-y} \right)}{P_N \left(\frac{1+y}{1-y} \right)} - \frac{N+1+y+3Ny}{1-y} \right]. \quad (55)$$

Esta expresión, en sí misma, no tiene mucho interés. Sólo la vamos a usar para hacer gráficos y comparar con las aproximaciones. Desde un punto de vista computacional, puede ser más eficiente el cálculo directo del promedio que el uso de la función anterior.

Por otro lado, bajo la aproximación del término máximo, la fracción media de defectos está dada por la Ec. (52), y el logaritmo de la función de partición es

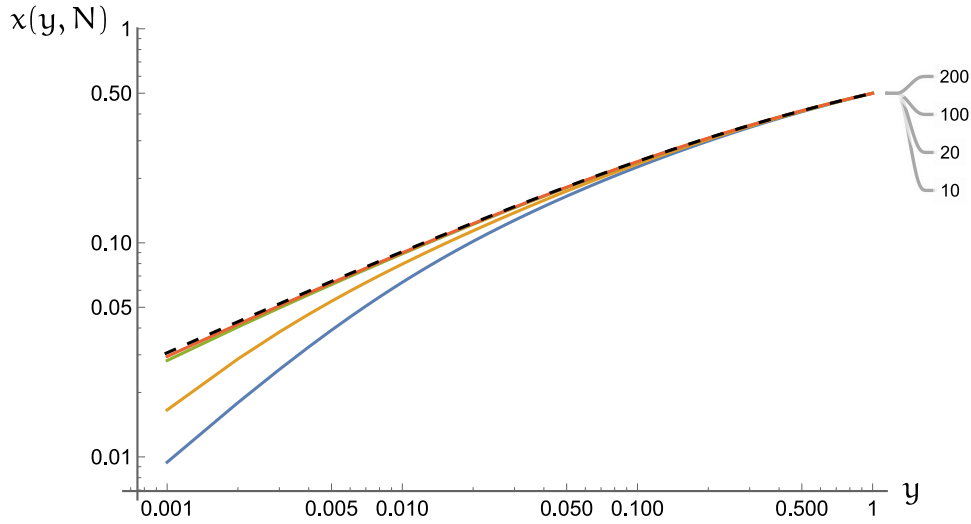
$$\frac{1}{N} \log Z \simeq g(x^*(y), y) = 2 \log(1 + \sqrt{y}). \quad (56)$$

La siguiente figura muestra la función $\frac{1}{N} \log Z(y, N)$ calculada mediante la expresión exacta para valores crecientes de N . En línea de trazos está graficada la aproximación (56). Notar que las escalas son logarítmicas. Puesto que $y = e^{-\beta W}$, el intervalo de variación de y es $[0, 1]$. El extremo izquierdo del intervalo corresponde a $T = 0$, mientras que el extremo derecho corresponde a $T \rightarrow \infty$.



Al aumentar N , las curvas se acercan a la función (56), pero persiste una característica común: el error es mayor cuanto menor es la temperatura. Esto es fácil de entender: al disminuir la temperatura, el número de defectos disminuye. Una de las limitaciones de la aproximación del término máximo es haber usado la aproximación de Stirling. Entonces, es razonable que, cuando el número medio de defectos no sea mucho mayor que uno, la aproximación del término máximo falle. En cristales macroscópicos, el número de defectos siempre será mucho mayor uno, de modo que estos escrúpulos sólo tienen un valor ilustrativo. Un indicio de que la aproximación del término máximo es inadecuada para valores pequeños de y lo da el hecho de que la función de la Ec. (56) no es derivable en $y = 0$, mientras que la expresión (54) no tiene ninguna patología.

La siguiente figura muestra la fracción de defectos calculada con la fórmula exacta para valores crecientes de N . En línea de trazos está la aproximación (52).



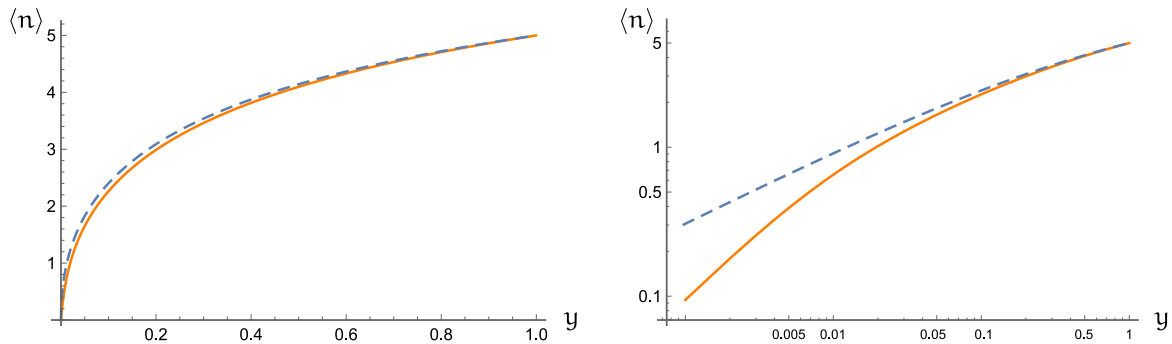
Nuevamente, el acuerdo es excelente para $N \gtrsim 10$, pero el error es grande para y suficientemente pequeño.

Consideremos como sujeto de estudio el caso $N = 10$. El máximo número medio de defectos está por debajo de la aplicabilidad de la aproximación de Stirling,

$$\log 5! \approx 4.8, \quad (57)$$

$$5 \log 5 - 5 \approx 3.0.$$

Se trata, entonces, de una situación en la que no esperaríamos que la aproximación del término máximo diera resultados válidos. Sin embargo, si graficamos el número medio de defectos en función de y , usando escalas lineales en ambos ejes, queda lo que muestra la figura de la izquierda. La curva en línea de trazos corresponde a la aproximación del término máximo. El acuerdo parece excelente, pero la figura de la derecha, que usa escalas logarítmicas, lo desmiente.



Tal como habíamos previsto, la aproximación falla cuando el número de defectos es del orden de uno. Lo asombroso es que sea válida aun para un número de defectos tal que la aproximación de Stirling es completamente inválida.

A primera vista es irrazonable que el acuerdo sea tan bueno, siendo que las hipótesis bajo las cuales se obtuvo la aproximación del término máximo no se cumplen. La explicación de este hecho es mucho menos sorprendente que el hecho en sí mismo.

El argumento consta de dos partes. En primer lugar, para encontrar x^* , hemos usado la fórmula de Stirling para aproximar los coeficientes binomiales. Esa aproximación podrá ser incorrecta cuando los números involucrados no son mucho mayores que uno, pero respeta la simetría de los coeficientes binomiales. En principio,

$$B(x) = \log \binom{N}{xN} = \log N! - \log(xN)! - \log[(1-x)N]!. \quad (58)$$

Evidentemente,

$$B(x) = B(1-x). \quad (59)$$

Esto significa que B alcanza su máximo en $x = \frac{1}{2}$, puesto que la ecuación

$$B'(x) = -B'(1-x) \quad (60)$$

evaluada en $x = \frac{1}{2}$ implica

$$B'\left(\frac{1}{2}\right) = 0. \quad (61)$$

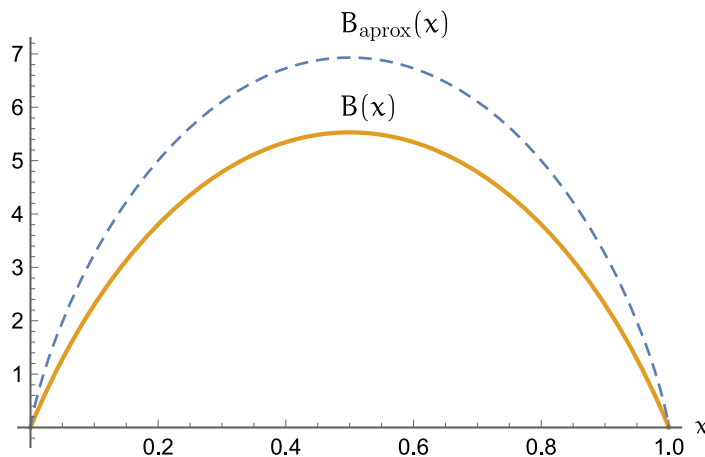
Por otro lado, al aplicar la aproximación de Stirling, obtuvimos

$$B(x) \simeq B_{\text{aprox}}(x) = N[\phi(x) + \phi(1-x)], \quad (62)$$

donde

$$\phi(x) = -x \log x. \quad (63)$$

Entonces, $B_{\text{aprox}}(x)$ también alcanza su máximo en $x = \frac{1}{2}$, porque, al igual que $B(x)$, es una función simétrica respecto de $x = \frac{1}{2}$. Para $N = 10$, la función B_{aprox} no es una buena aproximación de $B(x)$, pero alcanza su máximo en el mismo punto.



La segunda parte del argumento que explica la irrazonable validez de la aproximación del término máximo, aun cuando para $N = 10$ no están dadas las condiciones para su aplicabilidad, consiste en notar que la condición que determina x^* ,

$$g'(x) = 0, \quad (64)$$

y que se lee como

$$\frac{d}{dx} [2B_{\text{aprox}}(x) - \beta Wx] = 0, \quad (65)$$

cuando $\beta W = 0$ se reduce a la condición del máximo para $B_{\text{aprox}}(x)$. Aunque la aproximación sea incorrecta, la posición de máximo está correctamente determinada. Entonces, cuando $\beta W = 0$, x^* coincide con el valor exacto de x . De modo que la aproximación da el valor exacto de x cuando $y = 1$. Esto explica que, en las figuras de la página 10, la diferencia entre las dos curvas sea cero cuando $y = 1$. Por continuidad, la diferencia debe ser pequeña en todo un entorno de $y = 1$; las dos curvas deben convergir de manera suave.

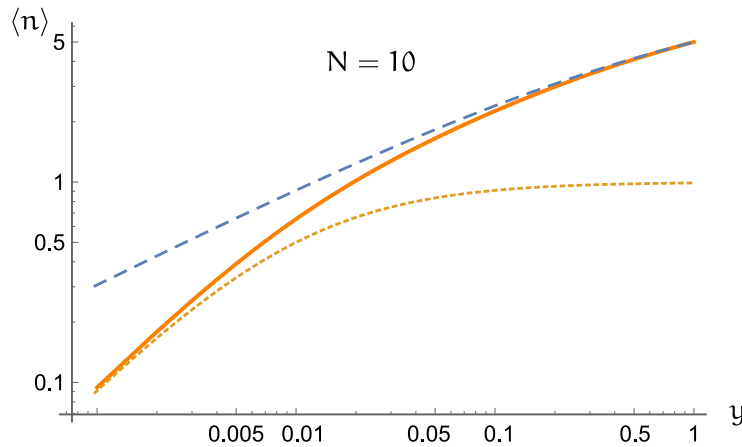
Para N finito, la aproximación del término máximo siempre falla para temperaturas suficientemente bajas, porque en algún momento el número medio de defectos se vuelve de orden uno. Ahora bien, en ese régimen, es posible obtener una nueva aproximación para la función de partición, basada en el hecho de que $y \ll 1$. En efecto,

$$Z = \sum_{n=0}^N \binom{N}{n}^2 e^{-\beta W n} = \sum_{n=0}^N \binom{N}{n}^2 y^n \simeq 1 + N^2 y. \quad (66)$$

Entonces,

$$\langle n \rangle = y \frac{\partial \log Z}{\partial y} \simeq \frac{y N^2}{1 + y N^2}. \quad (67)$$

La siguiente figura muestra, en línea llena, el número medio de defectos calculado exactamente para $N = 10$; en línea de trazos, la curva que corresponde a la aproximación del término máximo y , en línea de puntos, la aproximación de bajas temperaturas.



La Ec. (52) predice que, a bajas temperaturas, $\langle n \rangle \simeq n\sqrt{y}$. En cambio, la aproximación (67) da un comportamiento lineal con y .

Más allá del método del término máximo: el método de Laplace

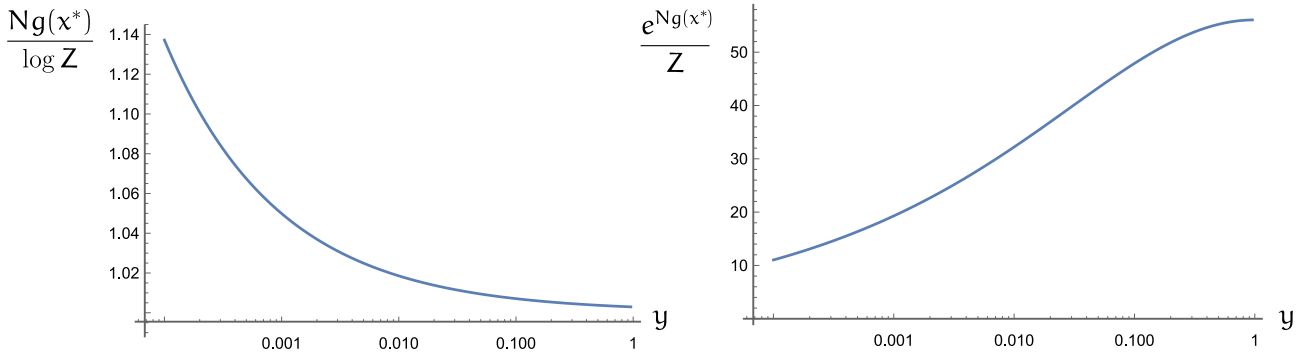
Por todo lo que vimos hasta aquí, está claro que la aproximación del término máximo funciona muy bien para el logaritmo de Z ,

$$\log Z \simeq Ng(x^*). \quad (68)$$

En principio, las cotas que usamos para encontrar esta aproximación no nos dicen mucho acerca de qué tan cierto sea que

$$Z \simeq e^{Ng(x^*)}. \quad (69)$$

La figura de la izquierda muestra el cociente $Ng(x^*)/\log Z$ en el caso de los defectos de Frenkel. Hemos tomado $N = 1000$, lo que es más que suficiente para asegurar la validez de la aproximación (68). La figura de la derecha muestra $e^{Ng(x^*)}/Z$.



Es evidente que la expresión (69) no es una buena aproximación para Z . Tal como anticipamos, la discrepancia es de orden $\sqrt{N}$. Es necesario calcular con mayor precisión la suma

$$Z = \sum_x \sigma(x). \quad (70)$$

No basta con decir que hay un término dominante. Son necesarias dos cosas. Primero, como estamos calculando $\sigma(x)$ a partir de $\log \sigma(x)$, debemos partir de una mejor aproximación de $\log \sigma(x)$. Esta aproximación debe incluir términos de orden $\log N$. Si el término dominante en el desarrollo de $\log \sigma(x)$ es de orden N , los términos de orden $\log N$ no son importantes para el cálculo aproximado de $\log \sigma(x)$, pero aportan factores multiplicativos que no pueden despreciarse cuando se quiere aproximar $\sigma(x)$. Es suficiente asumir que

$$\log \sigma(x) = Ng(x, N) + \log h(x, N) + o(1). \quad (71)$$

donde tanto $g(x, N)$ como $\log h(x, N)$ son $o(N)$. Esto quiere decir que, cuando $N \gg 1$,

$$\sigma(x) \simeq h(x, N)e^{Ng(x, N)}, \quad (72)$$

donde la función $h(x, N)$ crece con N más lentamente que la función exponencial y $g(x, N)$, más lentamente que N .

En segundo lugar, debemos reemplazar la aproximación del término máximo por una aproximación que sume de manera más precisa todos los términos que contribuyen apreciablemente a la suma de la Ec. (70). Primero, la idea es aproximar la suma (70) por una integral, para luego aplicar el método de Laplace. Así, en principio,

$$Z = \sum_x \sigma(x) = \frac{1}{\Delta x} \sum_x \sigma(x) \Delta x \simeq \frac{1}{\Delta x} \int_0^1 dx \sigma(x) = \frac{1}{\Delta x} \int_0^1 dx e^{\log \sigma(x)}. \quad (73)$$

Aquí, $\Delta x = N^{-1}$, porque x toma $N + 1$ valores equiespaciados en el intervalo $[0, 1]$. Usando la expresión (72), cuya validez debe demostrarse en cada problema en particular,

$$Z \simeq N \int_0^1 dx h(x, N) e^{Ng(x, N)}. \quad (74)$$

El siguiente paso es desarrollar $g(x)$ alrededor de su máximo,

$$g(x, N) \simeq g(x^*, N) + \frac{1}{2} g''(x^*, N) (x - x^*)^2. \quad (75)$$

N hace las veces de un parámetro. Se entiende que todas las derivadas están tomadas respecto de x . La función $e^{Ng(x, N)}$ decae rápidamente cuando x se aleja de x^* . Entonces, podemos reemplazar los límites en las integrales de la Ec. (73) por menos y más infinito,

$$Z \simeq N e^{Ng(x^*, N)} \int_{-\infty}^{\infty} dx h(x, N) e^{\frac{1}{2} Ng''(x^*, N) (x - x^*)^2}. \quad (76)$$

Asumiremos que la función $h(x, N)$ es una función suave, de manera que tome un valor prácticamente constante en todo el ancho de la campana. Podemos sacar h fuera de la integral, con el siguiente resultado:

$$Z \simeq N h(x^*, N) e^{Ng(x^*, N)} \int_{-\infty}^{\infty} dx e^{\frac{1}{2} Ng''(x^*, N) (x - x^*)^2}. \quad (77)$$

La integral puede resolverse explícitamente. Luego,

$$Z \simeq h(x^*, N) \sqrt{\frac{2\pi N}{|g''(x^*, N)|}} e^{Ng(x^*, N)}. \quad (78)$$

Esta es la aproximación buscada para Z . No tiene mucho valor acordarse de esta fórmula. Lo importante es saber deducirla. El procedimiento que lleva de la Ec. (74) a la Ec. (78) se conoce como *método de Laplace*.

Resumido en pocas palabras, el método de Laplace aproxima la integral de una función cuyo gráfico tiene forma de campana aguda, por la integral de una gaussiana. La gaussiana se construye a partir de la aproximación cuadrática del logaritmo de la función original alrededor de su máximo.

Aplicación al problema de los defectos de Frenkel

En el problema de los defectos de Frenkel, habíamos llegado a escribir

$$\log \sigma(x) \simeq Ng(x). \quad (79)$$

Ahora tenemos que encontrar una mejor aproximación para $\log \sigma(x)$ e identificar la función h de la Ec.(71). Para eso usaremos la aproximación de Stirling correcta hasta términos de orden $\log n$,

$$\log n! \simeq n \log n - n + \log \sqrt{2\pi n} + \mathcal{O}(n^{-1}). \quad (80)$$

Con esto, para el coeficiente binomial, obtenemos

$$\log \binom{N}{xN} = Nf(x) - \frac{1}{2} \log [2\pi x(1-x)N] + \mathcal{O}(N^{-1}). \quad (81)$$

La función f está definida en la Ec. (43). Entonces,

$$\log \sigma(x) = \log \left[\binom{N}{xN}^2 y^{xN} \right] = 2Nf(x) - \log [2\pi x(1-x)N] + Nx \log y + \mathcal{O}(N^{-1}), \quad (82)$$

o, de manera más compacta,

$$\log \sigma(x) = Ng(x) + \log h(x, N) + \mathcal{O}(N^{-1}), \quad (83)$$

donde

$$\begin{aligned} g(x) &= 2f(x) + x \log y, \\ h(x, N) &= \frac{1}{2\pi Nx(1-x)}. \end{aligned} \quad (84)$$

Vemos que se satisfacen las condiciones de la Ec. (71). La función g alcanza su máximo en

$$x^* = \frac{\sqrt{y}}{1 + \sqrt{y}}, \quad (85)$$

y en este punto

$$g(x^*) = 2 \log(1 + \sqrt{y}). \quad (86)$$

Su derivada segunda es

$$g''(x) = -\frac{2}{x(1-x)}, \quad (87)$$

de modo que

$$g''(x^*) = -\frac{2(1 + \sqrt{y})^2}{\sqrt{y}}. \quad (88)$$

Lo único que nos falta calcular es $h(x^*, N)$,

$$h(x^*, N) = \frac{1}{2\pi N x^* (1 - x^*)} = \frac{(1 + \sqrt{y})^2}{2\pi N \sqrt{y}}. \quad (89)$$

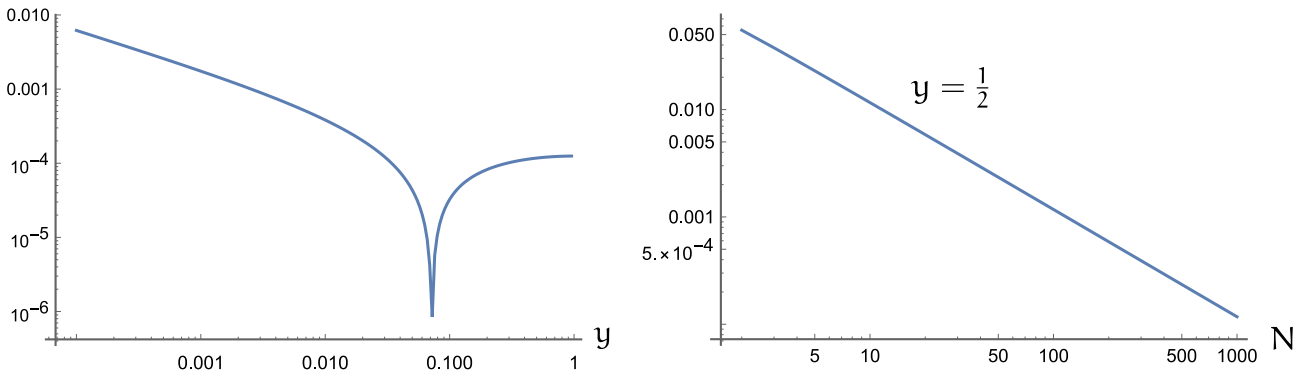
Tenemos todo lo necesario para evaluar la expresión (78),

$$Z \simeq \frac{(1 + \sqrt{y})^{2N+1}}{2\sqrt{\pi N \sqrt{y}}}. \quad (90)$$

Esto debe compararse con la estimación más rudimentaria, por no decir incorrecta,

$$Z \simeq e^{Ng(x^*)} = (1 + \sqrt{y})^{2N}. \quad : (\quad (91)$$

La figura de la izquierda muestra el error relativo para $N = 100$ como función de y si se usa para Z la aproximación (90). Para ver cómo depende el error con el valor de N , la figura de la derecha muestra el error relativo para un valor de y fijo y N entre 2 y 1000.



Aun para valores tan pequeños como $N = 3$, el error es menor al 5%. Esto es bastante asombroso. El asombro debe trasladarse a la aproximación de Stirling de la Ec. (80) que se ha usado para aproximar $\sigma(x)$. Aun para números de orden uno, la aproximación (80) tiene un error muy pequeño. Por ejemplo, si $n = 2$,

$$2! \approx \sqrt{4\pi} \left(\frac{2}{e}\right)^2 \approx 1.92. \quad (92)$$

Contrariamente a lo que esperaríamos, no es necesario que sea $n \gg 1$ para obtener aproximaciones razonables. Para $n = 10$, el error es del 0.8%. En general, el error relativo es del orden de $(10n)^{-1}$.

Resulta un poco inquietante que la fórmula de Stirling pueda deducirse, como mostraremos a continuación, a partir del método de Laplace. En algún sentido, esto quiere decir que, para obtener la Ec. (90), hemos iterado el método de Laplace.

La fórmula de Stirling

El método de Laplace puede aplicarse para deducir la fórmula de Stirling. En general, la función factorial está definida en términos de la función gamma,

$$N! = \Gamma(N + 1), \quad (93)$$

donde, si $n > 0$, la función gamma está dada por

$$\Gamma(n) = \int_0^{\infty} dx \, x^{n-1} e^{-x}. \quad (94)$$

Que los límites de integración sean cero e infinito no introduce ninguna diferencia en el argumento que nos permitió llegar a la Ec. (78). Elijamos $n = N + 1$ y reescribamos el integrando como

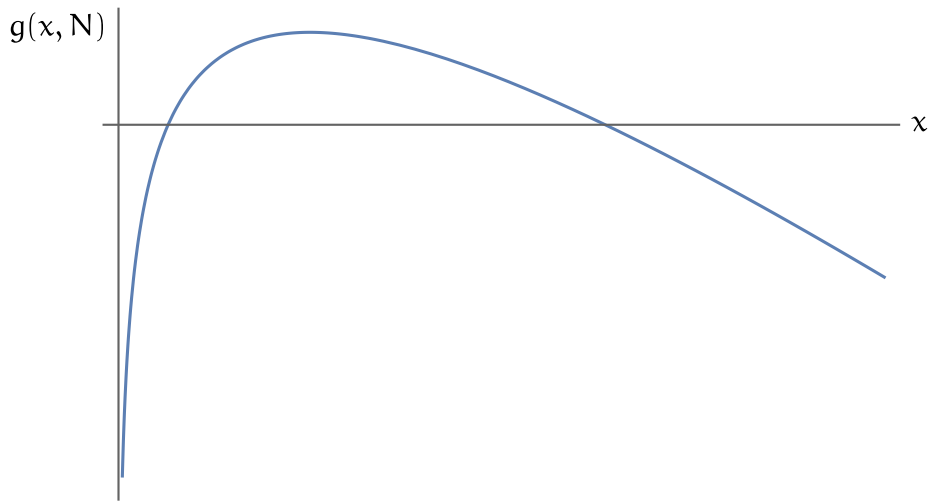
$$x^N e^{-x} = \exp(N \log x - x) = \exp\left[N \left(\log x - \frac{x}{N}\right)\right]. \quad (95)$$

Las funciones g y h que definimos a través de la Ec. (71) son

$$g(x, N) = \log x - \frac{x}{N}, \quad (96)$$

$$h(x, N) = 1.$$

Esta no es la única descomposición posible. Es fácil convencerse de que el gráfico de la función $g(x, N)$ es como muestra la figura.



La función g alcanza su máximo cuando $x = x^* = N$. En ese punto,

$$g(x^*, N) = \log N - 1, \quad g''(x^*, N) = -\frac{1}{N^2}. \quad (97)$$

Luego,

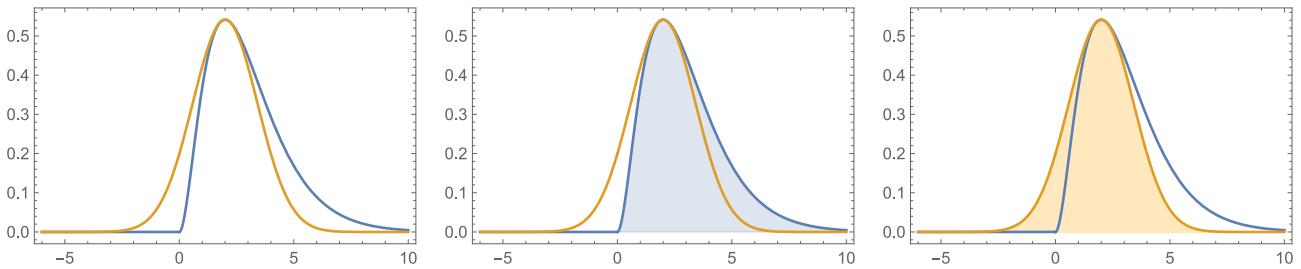
$$N! \simeq e^{N(\log N - 1)} \int_{-\infty}^{\infty} dx \, e^{-\frac{1}{2N}(x-N)^2} = \sqrt{2\pi N} \left(\frac{N}{e}\right)^N. \quad (98)$$

Como señalamos antes, el método de Laplace reemplaza la integral de una función que tiene un máximo muy agudo, por la integral de la gaussiana que aproxima a la función en un entorno del máximo. La figura siguiente muestra el integrando de la Ec. (94) con $n = N + 1$,

$$x^N e^{-x}, \quad (99)$$

junto con la gaussiana que lo aproxima cuando $N = 2$,

$$\left(\frac{N}{e}\right)^N e^{-\frac{1}{2N}(x-N)^2}. \quad (100)$$



Según la Ec. (92), las áreas bajo las dos curvas difieren en menos del 5%. A primera vista, uno nunca diría que la diferencia es tan pequeña. Lo que sucede es que a la izquierda del máximo el error es positivo, mientras que a la derecha es negativo. Estos errores tienden a cancelarse. La razón última de esta cancelación escapa a mi erudición.

Un llamado a la cordura

El método del término máximo es una herramienta para aproximar sumas, pero no es la única herramienta ni es aplicable a cualquier suma. En mecánica estadística, son comunes otros dos métodos para aproximar sumas, la fórmula de Euler–Maclaurin y la fórmula de sumación de Poisson. Es probable que vean el primero en la clase teórica, cuando analicen el rotor cuántico, para estudiar la contribución de los grados de libertad de rotación al calor específico de un gas ideal. Quizá veamos la fórmula de sumación de Poisson en las primeras clases de estadística cuántica, cuando aproximemos sumas sobre estados cuánticos por integrales en el espacio de fase. Ya conocen el dicho del martillo y los clavos, así que estén prevenidos contra ustedes mismos.