

# ¿Cuán importante es $e^{-\beta E}$ ?

Raúl Toral

Instituto Mediterráneo de Estudios Avanzados (IMEDEA),

CSIC-Universitat de les Illes Balears

Ed. Mateu Orfila, Campus UIB, 07122 Palma de Mallorca

**Abstract** I discuss the possibility of using other factors than  $e^{-\beta E}$  to compute the equilibrium properties of physical systems. I will show that other choices can, in certain cases, lead to a simpler calculation of those properties.

**PACS** 05.20.-y, 05.20.Gg, 05.30.-d, 05.30.Ch

El alumno que termina la carrera y el profesional que no ha reflexionado más allá de lo que aprendió en la licenciatura de física suelen tener la vaga impresión de que la mecánica estadística descansa de manera esencial en el llamado factor de Boltzmann,  $e^{-\beta E}$ , que aparece en la colectividad canónica, de manera que vienen a creer que este factor constituye una especie de ley fundamental que describe a los sistemas en equilibrio y que toda suerte de desastres caerían sobre el aparato matemático creado por Gibbs si alguien osara cambiar de alguna manera este factor. Y, seamos honrados, ¿cuál sería la reacción de los que somos profesores de mecánica estadística si un alumno escribiera en el examen: “Calculemos la función de partición  $\mathcal{Z} = \sum e^{-\beta E^2}$ ”? Lo enviaríamos sin duda al limbo de septiembre, sin necesidad de leer el resto del examen y con un comentario en nuestras mentes: “no hay manera de que aprendan nada”.

Sin embargo, es fácil de entender que el factor  $e^{-\beta E}$  no es necesario si pensamos, por ejemplo, que se pueden obtener todas las propiedades termodinámicas del equilibrio trabajando exclusivamente en la colectividad microcanónica y sin usar, por tanto, en ningún momento el factor de Boltzmann.

En este artículo pretendo mostrar que uno puede desarrollar la mecánica estadística usando, en vez de  $e^{-\beta E}$ , otros muchos factores, incluyendo el  $e^{-\beta E^2}$  de nuestro incomprendido alumno avanzado, y reproducir sin embargo todos los resultados conocidos para sistemas en equilibrio.

Recordemos brevemente los fundamentos de la colectividad microcanónica. Consideramos un sistema de  $N$  partículas, con un total de  $f$  grados de libertad, descrito por pares de coordenadas canónicas  $(q, p) \equiv (q_i, p_i)_{i=1, \dots, f}$  y con un hamiltoniano

$\hat{\mathcal{H}}(q, p)$  conocido. A este hamiltoniano le corresponden una serie de estados propios<sup>1</sup>  $|m\rangle$  cada uno con energía  $E_m$ ,  $m = 0, 1, 2, 3, \dots$ . En la descripción estadística de la colectividad microcanónica, fijamos primero un valor de la energía  $U$  y asignamos a cada *microestado*  $|m\rangle$  una probabilidad  $p_m$  dada por

$$p_m = \begin{cases} 0 & \text{si } E_m \neq U \\ \Omega(U)^{-1} & \text{si } E_m = U \end{cases} \quad (1)$$

Siendo  $\Omega(U)$  la degeneración de  $U$ , es decir, el número de estados cuya energía es  $U$ . En general  $\Omega(U)$  dependerá de otras variables del sistema, tales como el volumen  $V$  o el número de constituyentes  $N$ ,  $\Omega(U, N, V)$ , pero esta dependencia sólo la haremos explícita en la notación cuando sea necesario. Es claro que estas probabilidades satisfacen  $\sum_m p_m = 1$ , como corresponde.

El gran genio de Boltzmann le hizo identificar la función de estado entropía con el logaritmo de  $\Omega(U)$ , en la famosa ecuación

$$S = k_B \ln \Omega \quad (2)$$

donde  $k_B$  es la llamada constante de Boltzmann. Esta fórmula, como todas las grandes fórmulas de la física, es a la vez sencilla y profunda. Se ha argumentado que toda persona culta, independientemente de su formación científica, debiera avergonzarse de no conocer algunas de las implicaciones de esta fórmula, así como los físicos debiéramos avergonzarnos de no haber leído las obras de Shakespeare o el Quijote.

A partir de aquí, el formalismo de la colectividad microcanónica **define** la temperatura  $T$  a partir de la ecuación:

$$\frac{1}{T} = \left( \frac{\partial S}{\partial U} \right)_{N, V} \quad (3)$$

que es una fórmula previamente conocida en termodinámica si identificamos  $U$  con la energía interna. De esta manera, la colectividad microcanónica enlaza la descripción microscópica dada por el hamiltoniano con las propiedades termodinámicas observables. La fórmula anterior nos da la temperatura como función de  $U, N, V$ , pero es más usual despejar la energía interna  $U(T, N, V)$  y obtener así mismo la entropía como función de las mismas variables  $S(T, N, V) = S(U(T, N, V), N, V)$ . A partir de aquí podemos definir la energía libre del Helmholtz  $F(T, N, V) = U(T, N, V) - T S(T, N, V)$  y usar el aparato matemático de derivadas parciales de la termodinámica para obtener todas las propiedades del equilibrio a partir de cualquiera de los potenciales termodinámicos  $S, U, F$ .

Es bien sabido que aparte de la colectividad microcanónica, es posible usar otras colectividades, la canónica o la macrocanónica, por ejemplo, y obtener los mismos resultados en el caso de que el número de constituyentes  $N$  sea grande, entendiendo por

---

<sup>1</sup> Adoptamos el punto de vista cuántico, que es algo más natural, puesto que el conjunto de posibles estados es en general discreto numerable. En la versión clásica deberíamos cambiar las sumas sobre estados por integrales en el espacio fásico.

grande  $N \sim 10^{23}$  o, formalmente, en el límite  $N \rightarrow \infty$ . Veremos ahora que es posible ampliar grandemente el conjunto de colectividades para las cuales reproducimos los resultados de la microcanónica.

Una manera conveniente de deducir otras colectividades es la reformulación de Jaynes de la mecánica estadística. Según esta formulación, las probabilidades  $p_m$  de los microestados se obtienen mediante un proceso de maximización de la entropía definida mediante el funcional:

$$S = -k_B \sum_m p_m \ln p_m \quad (4)$$

La maximización de  $S$  se ha de conseguir siempre con las condiciones de que  $\sum_m p_m = 1$ ,  $p_m \geq 0$  y otras que dependen del tipo de colectividad escogida. La colectividad microcanónica escoge como condición adicional que  $E_m = U$ , un número prefijado. La colectividad canónica usa la condición adicional  $\sum_m p_m E_m = U$ . Nosotros escogemos como condicional adicional

$$\sum_m p_m \Phi(E_m) = \Phi(U) \quad (5)$$

siendo  $\Phi(E)$  una función dada y  $U$  un valor prefijado que haremos corresponder con la energía interna. Si  $\Phi(E) = E$  recobramos la colectividad canónica. Llamaremos a la nueva colectividad la “ $\Phi$ -canónica”.

La solución del problema de maximización de la entropía con las condiciones anteriores es[1]:

$$p_m = \frac{e^{-\beta\Phi(E_m)}}{\mathcal{Z}} \quad (6)$$

donde  $\mathcal{Z}(\beta)$  es la función de partición definida mediante:

$$\mathcal{Z} = \sum_m e^{-\beta\Phi(E_m)}. \quad (7)$$

En estas ecuaciones,  $\beta$  es un multiplicador de Lagrange que se obtiene mediante la ecuación (5):

$$\Phi(U) = \frac{\sum_m \Phi(E_m) e^{-\beta\Phi(E_m)}}{\sum_m e^{-\beta\Phi(E_m)}} \quad (8)$$

que es completamente equivalente a

$$\Phi(U) = - \left( \frac{\partial \ln \mathcal{Z}}{\partial \beta} \right)_{N,V}, \quad (9)$$

algo más sencilla de usar, en la práctica.

Una vez tenemos las probabilidades  $p_m$ , la entropía se puede encontrar de (4) o de la fórmula equivalente

$$S = k_B \left( \frac{\partial(\beta^{-1} \ln \mathcal{Z})}{\partial(\beta^{-1})} \right)_{N,V} = k_B [\ln \mathcal{Z} + \beta\Phi(U)]. \quad (10)$$

Parecería que ya tenemos los potenciales termodinámicos y que, por tanto, ya podemos empezar a derivar para deducir las propiedades termodinámicas. Sin embargo, antes de ello, es esencial conocer la relación entre el parámetro  $\beta$  y la temperatura.

Para ello, adoptamos como válida (como no podía ser de otra manera dada su amplia verificación experimental) la relación (3). Derivando (10) con respecto a  $U$ , y después de un cuidadoso cálculo<sup>2</sup>, obtenemos

$$\left(\frac{\partial S}{\partial U}\right)_{N,V} = k_B \beta \Phi'(U) \quad (11)$$

que nos lleva a identificar:

$$\beta = \frac{1}{k_B T \Phi'(U)} \quad (12)$$

y permite expresar los potenciales  $S$  y  $U$  en función de las variables  $T, N, V$  y proceder entonces de la manera usual.

Veremos en los ejemplos siguientes como la aplicación de la colectividad  $\Phi$ -canónica produce los mismos resultados que la microcanónica para una gran variedad de funciones  $\Phi(E)$  y, en particular, obtenemos los mismos potenciales termodinámicos  $F, U, S$  como función de las variables  $T, N, V$ . Este resultado puede parecer sorprendente, ya que la manera de proceder y, en particular, las probabilidades asignadas a cada estado  $|m\rangle$  son muy diferentes en la colectividad microcanónica o en la nueva colectividad  $\Phi$ -canónica, compárese (1) con (6). La explicación, archirrepetida en los libros de texto en el caso de la colectividad canónica[2], es que lo que es relevante físicamente no es tanto la probabilidad del estado  $|m\rangle$ , que es inaccesible experimentalmente, sino la probabilidad  $P(E)$  de que el sistema tenga una energía  $E$ . En cualquiera de los casos, esto es igual a la probabilidad de que un estado  $|m\rangle$  tenga energía  $E$  multiplicado por el número de estados con energía igual a  $E$ . Es decir:  $P(E) = \Omega(E)p_m$  con  $E_m = E$ .

Por la definición de  $p_m$  en la colectividad microcanónica, es claro que

$$P_M(E) = \delta(E - U) \quad (13)$$

(el subíndice  $M$  hace referencia a la colectividad microcanónica) mientras que en la colectividad  $\Phi$ -canónica es

$$P_\Phi(E) = \Omega(E)e^{-\beta\Phi(E)} / \mathcal{Z} \quad (14)$$

(el subíndice  $\Phi$  hace referencia a la función  $\Phi(E)$  escogida). En principio,  $P_M(E)$  y  $P_\Phi(E)$  parecen ser dos funciones muy diferentes. Sin embargo, resultan ser la misma función en el límite de un sistema con muchos constituyentes,  $N \rightarrow \infty$ . Hay varias maneras de demostrar esto. Una es la de calcular las fluctuaciones de la energía en la colectividad  $\Phi$ :

$$\sigma^2[\Phi] = \langle \Phi^2 \rangle - \langle \Phi \rangle^2 = -\frac{\partial \Phi(U)}{\partial \beta} \quad (15)$$

---

<sup>2</sup>Recordemos que  $\beta$  es función de  $U$  a través de la ecuación (8).

El valor relativo de las fluctuaciones se mide como  $\sigma[\Phi]/\langle\Phi\rangle$ . Una sencilla estimación de órdenes de magnitud, en el caso que la energía interna  $U$  sea proporcional al número de constituyentes  $N$ , lleva a que cuando  $N \rightarrow \infty$  es  $\sigma[\Phi]/\langle\Phi\rangle \sim N^{-1/2} \rightarrow 0$ . De manera que la anchura de la distribución  $P(E)$  es completamente despreciable frente al valor medio para sistemas macroscópicos para los que  $N^{-1/2} \sim 10^{-12}$ . Si dibujáramos, en un papel de unos pocos centímetros,  $P(E)$  como función de la energía  $E$  de manera que el valor medio  $U$  estuviera en el centro de la gráfica, la anchura de la distribución sería menor que el tamaño del átomo de hidrógeno, es decir, sería a todos los efectos prácticos una función delta de Dirac<sup>3</sup>:  $P_\Phi(E) \rightarrow \delta(E - U)$ . En estas condiciones los resultados que obtenemos de la distribución  $\Phi$ -canónica coinciden con los de la microcanónica y, por tanto, con los de la canónica.

¿Todavía no lo creemos? Veamos un ejemplo. Consideremos el gas ideal no relativista en dimensión  $d = 3$ . Como los microestados tienen energías muy cercanas entre sí, las sumas sobre estados de una función arbitraria  $f(E)$  se pueden substituir por una integral[3]:

$$\sum_m f(E_m) \longrightarrow \int_0^\infty dE g(E) f(E) = C(N, V) \int_0^\infty dE E^{3N/2-1} f(E) \quad (16)$$

siendo  $g(E) = C(N, V)E^{3N/2-1}$ , con  $C(N, V) = \frac{[2\pi m h^{-2} V^{2/3}]^{3N/2}}{\Gamma(\frac{3N}{2})N!}$ , la densidad de estados (hemos incluido el factor  $N!$  que da cuenta de manera aproximada de la indistinguibilidad de las partículas según la cuenta correcta de Boltzmann).

Elegimos la función  $\Phi(E) = E^q$  con  $q > 0$  arbitrario. Es claro que el caso  $q = 1$  se reduce al usual de la colectividad canónica. La función de partición se calcula fácilmente:

$$\mathcal{Z}(\beta) = C(N, V) \int_0^\infty dE E^{3N/2-1} e^{-\beta E^q} = \frac{C(N, V)}{q} \Gamma\left(\frac{3N}{2q}\right) \beta^{-3N/2q} \quad (17)$$

La relación entre la energía interna  $U$  y el parámetro  $\beta$  se obtiene de (9):

$$U^q = \frac{3N}{2q\beta} \quad (18)$$

y la relación entre  $\beta$  y la temperatura de (12):

$$\beta = \frac{1}{k_B T q U^{q-1}} \quad (19)$$

que substituida en (18) lleva a

$$\frac{U}{N} = \frac{3}{2} k_B T \quad (20)$$

un resultado bien conocido.

---

<sup>3</sup>Estrictamente hablando, la distribución delta de Dirac es válida para la variable intensiva  $E/N$ .

Calculemos ahora la entropía usando (10)

$$\frac{S}{k_B} = \ln \mathcal{Z} + \beta U^q \quad (21)$$

Si usamos ahora (17) y (18), y utilizamos la aproximación de Stirling  $\ln \Gamma(x) \approx x \ln x - x$ , válida para  $x \rightarrow \infty$ , podemos simplificar esta expresión a

$$\frac{S}{Nk_B} = \frac{\ln C(N, V)}{N} + \frac{3}{2} \ln U - \frac{\ln q}{N} \quad (22)$$

que coincide con el resultado microcanónico  $S = k_B \ln g(U)$  salvo términos (como el último de la ecuación anterior) que son completamente despreciables en el límite  $N \rightarrow \infty$ .

Como comprobación de la equivalencia, podemos calcular las fluctuaciones de la energía usando (15):

$$\sigma^2[U^q] = -\frac{\partial U^q}{\partial \beta} = \frac{3N}{2q\beta^2} \quad (23)$$

y las fluctuaciones relativas

$$\frac{\sigma[U^q]}{U^q} = \sqrt{\frac{2q}{3N}} \quad (24)$$

que, efectivamente, tienden a 0 cuando  $N \rightarrow \infty$ , demostrando la equivalencia entre colectividades.

Por tanto, hemos visto que podemos usar la distribución  $e^{-\beta E^q}$  con  $q \neq 1$  y reproducir los resultados estándar de  $e^{-\beta E}$ . La siguiente pregunta es ¿por qué usamos siempre  $e^{-\beta E}$ ? Hay una muy buena razón para ello ya esgrimida por Gibbs y es, nada más y nada menos, que **la sencillez en el cálculo**. Para sistemas ideales, el hamiltoniano se escribe como suma de hamiltonianos individuales  $\mathcal{H} = \sum_i \mathcal{H}_i$  y para estos sistemas el factor de Boltzmann se desacopla en producto de términos independientes y se simplifica grandemente el cálculo de la función de partición.

Sin embargo, si no tenemos esta propiedad aditiva, no tiene por qué ser necesariamente mejor el factor de Boltzmann que algún otro factor. Veamos otro ejemplo: el modelo de Ising con interacciones de largo alcance:

$$\mathcal{H} = -\frac{2J}{N} \sum_{1 \leq i < j \leq N} S_i S_j \quad (25)$$

Aquí,  $S_i$ ,  $i = 1, \dots, N$  son variables de Ising que pueden tomar los valores  $S_i = \pm 1$  y localidades en los nudos  $\vec{r}_i$  de una red regular. En este hamiltoniano todas las variables interactúan con todas las demás independientemente de su posición en la red. Se puede escribir también como:

$$\mathcal{H} = J - \frac{J}{N} \left( \sum_{i=1}^N S_i \right)^2 \quad (26)$$

Es posible utilizar tanto la colectividad microcanónica como la canónica para obtener las propiedades del equilibrio de este sistema. En particular, el cálculo en la colectividad canónica es complicado porque utiliza la transformación de Hubbard–Stratonovich y una integración por punto de silla[4]. Es mucho más sencillo utilizar nuestra colectividad  $\Phi$ -canónica con  $\Phi(E) = (J - E)^{1/2}$ . La función de partición es

$$\mathcal{Z} = \sum_{S_1=\pm 1, \dots, S_N=\pm 1} e^{\beta \sqrt{\frac{J}{N}} \sum_{i=1}^N S_i} \quad (27)$$

Como las sumas se desacoplan, se pueden hacer de manera trivial

$$\mathcal{Z} = \left[ 2 \cosh \left( \beta \sqrt{\frac{J}{N}} \right) \right]^N \quad (28)$$

La energía interna se obtiene de (9)

$$(J - U)^{1/2} = -N \sqrt{\frac{J}{N}} \tanh \left( \beta \sqrt{\frac{J}{N}} \right) \quad (29)$$

Si ahora obtenemos la relación entre  $\beta$  y la temperatura  $T$  usando (12)

$$\beta = \frac{-2(J - U)^{1/2}}{k_B T} \quad (30)$$

obtenemos la energía en función de la temperatura como la solución de la ecuación trascendente:

$$(J - U)^{1/2} = \sqrt{JN} \tanh \left( \frac{2}{k_B T} \sqrt{\frac{J}{N}} (J - U) \right) \quad (31)$$

que es el resultado ya conocido por otros métodos. Si definimos  $u = \frac{(J-U)}{JN}$  y  $T_c = 2J/k_B$ , esta ecuación se escribe:

$$u^{1/2} = \tanh \left( \frac{T_c}{T} u^{1/2} \right) \quad (32)$$

Fijémonos que esta ecuación tiene soluciones no nulas solamente para  $T < T_c$ , la temperatura crítica de este modelo. Por último, podemos calcular la entropía de (10)

$$\frac{S}{Nk_B} = \ln \left[ 2 \cosh \left( \frac{T_c}{T} u^{1/2} \right) \right] - \frac{T_c}{T} u \quad (33)$$

Así pues, hemos visto cómo es posible reobtener las propiedades del equilibrio de un sistema físico usando el arbitrario  $e^{-\beta\Phi(E)}$  en vez del factor de Boltzmann,  $e^{-\beta E}$ . Mientras en la gran mayoría de ocasiones será más sencillo usar el conocido  $e^{-\beta E}$ , no es menos cierto que hay algunos casos en los que es más fácil utilizar  $e^{-\beta\Phi(E)}$  con una función  $\Phi(E)$  *ad hoc*.

Queda una última objeción. ¿No es verdad que se ha medido experimentalmente el factor de Boltzmann? Por ejemplo, ¿no se ha medido la distribución de velocidades de las moléculas de un gas y verificado muy precisamente que las velocidades están distribuidas según  $e^{-\frac{m}{2k_B T} \vec{v}^2}$ ? Es decir, ¿no se ha verificado experimentalmente la ley  $e^{-\mathcal{E}/k_B T}$  cuando  $\mathcal{E}$  es la energía de una molécula? Sin duda. Pero este resultado no implica que la distribución de energía del **conjunto** de las moléculas del gas sea también  $e^{-E/k_B T}$ . Hay que distinguir entre la distribución de energías de todo el gas y la de una molécula. La única distribución medible de energía del conjunto de las moléculas del gas es  $\delta(E - U)$  siendo  $U$  la energía interna, y los experimentos nos dicen que la distribución de energías para una molécula es  $e^{-\mathcal{E}/k_B T}$ . ¿Podemos compaginar ambos resultados?

No hay ninguna contradicción. Imaginemos que el sistema global esté descrito por la distribución de probabilidad  $e^{-\beta\Phi(E)}$ . Aislemos ahora una pequeña parte del sistema global (por ejemplo, una molécula en el caso de un gas) y calculemos la probabilidad de  $\mathcal{E}$ , la energía de esa pequeña parte. Para ello separamos la energía total como  $E = E' + \mathcal{E}$  siendo  $\mathcal{E}$  la energía de la molécula y  $E'$  la de todo el sistema sin la molécula en cuestión. Como  $\mathcal{E} \ll E'$  desarrollamos  $\Phi(E) = \Phi(E') + \Phi'(E')\mathcal{E} + \dots$  y reemplazamos en la distribución  $e^{-\beta\Phi(E)} = e^{-\beta\Phi(E')}e^{-\beta\Phi'(E')\mathcal{E}}$ . Recordemos ahora la ecuación (12) donde podemos substituir la energía total  $U$  por  $E'$ . Esto nos lleva a la distribución  $e^{-\beta\Phi(E')}e^{-\mathcal{E}/k_B T}$ . Integrando las variables que no dependen de la molécula, nos queda que la distribución de  $\mathcal{E}$  es  $e^{-\mathcal{E}/k_B T}$  **independientemente de la forma de la función  $\Phi(E)$**  y en acuerdo con los resultados experimentales. Así que sea cual sea la distribución  $e^{-\beta\Phi(E)}$  válida para el sistema global, la distribución de energía para una pequeña parte del sistema será siempre  $e^{-\mathcal{E}/k_B T}$ . Un resultado a recordar cuando se sugieren distribuciones alternativas para encontrar distribuciones en forma de leyes de potencias para magnitudes que dependen de un solo grado de libertad.

**Agradecimientos** Este trabajo ha sido financiado por los proyectos BFM2001-0341-C02-01 y BFM2000-1108 del Ministerio de Ciencia y Tecnología y FEDER. Agradezco también a E. Hernández-García, J. Marro y J. M. Sancho una lectura del manuscrito y muy interesantes sugerencias.

## Referencias

- [1] La formulación de Jaynes de la mecánica estadística se puede encontrar en W.T. Grandy, Jr. *Foundations of statistical mechanics*, D. Reidel Pub. Co., Dordrecht (1987). En el capítulo 2 de este libro se demuestran muchas de las fórmulas de la teoría de Jaynes que se utilizan aquí.
- [2] R. K. Pathria, *Statistical mechanics*, 2<sup>nd</sup> edition, Butterworth-Heinemann, Oxford (1996).



- [3] R. Balescu, *Equilibrium and nonequilibrium statistical mechanics*, John Wiley and Sons, New York (1975).
- [4] Este resultado fue obtenido por M. Kac y su trabajo está muy bien explicado en H.E. Stanley, *Introduction to phase transitions and critical phenomena*, Oxford University Press, Oxford (1971).