

# Maximización de entropía e introducción a ensambles

Facundo Rost

Física Teórica 3 - 1er cuatrimestre de 2020

## Resumen

Resumimos aquí la teoría de ensambles genéricos, la conexión con la termodinámica, diversos tipos de ensambles y trucos muy importantes, con sus aplicaciones en mecánica estadística.

Recuadramos en cuadros grises las cosas más importantes sobre ensambles genéricos, y en cuadros azules las cosas más importantes sobre ensambles particulares en mecánica estadística (microcanónico, canónico, gran canónico e isobárico).

## Índice

<b>1. Ensambls genéricos</b>	<b>2</b>
1.1. Ensambls en continuo . . . . .	6
1.2. Ensambls en términos de copias del sistema . . . . .	7
1.3. Distribuciones de máxima entropía, ejemplo: Ejercicio 18b, guía 2 . . . . .	8
<b>2. Conexión con termodinámica</b>	<b>11</b>
<b>3. Ensambls particulares en mecánica estadística</b>	<b>13</b>
3.1. Ensamble microcanónico . . . . .	13
3.2. Ensamble canónico . . . . .	14
3.3. Ensamble gran canónico . . . . .	16
3.4. Ensamble isobárico . . . . .	17
<b>4. Trucos importantes y aplicaciones</b>	<b>19</b>
4.1. Indistinguibilidad y conteo correcto de Boltzmann . . . . .	19
4.2. Ensambls en términos de otros ensambles más chicos . . . . .	21
4.3. Trucos y suposiciones para simplificar el cálculo de $Z$ . . . . .	22
4.4. Truco para un ensamble con uno de sus vínculos: $\langle N \rangle = \bar{N}$ . . . . .	25

# 1. Ensembles genéricos

Un comentario previo: La resolución del ejercicio 18b de la guía 2 <sup>1</sup> es una versión un poco más resumida de la sección 1 de éste apunte. Así que quienes hayan leído dicha resolución pueden leer más rápido la sección 1, prestando atención a lo nuevo: partes nuevas de la sección 1 y el resto de las secciones. Y quienes no hayan leído la anterior resolución, pueden leer todo el apunte de una y listo (es autocontenido), sin necesariamente leer la anterior resolución (contenida en éste apunte). Ahora sí, desarrollemos la teoría de ensembles genéricos:

Consideremos un conjunto  $M$ , llamado espacio muestral, formado por posibles estados  $r \in M$  de nuestro sistema físico, donde cada estado  $r$  posee una probabilidad  $p_r$  de ser el estado del sistema. Por simplicidad, consideremos que  $M$  es discreto ( $\#(M) = \Omega \in \mathbb{N}$  es la multiplicidad del sistema y es un número natural), por lo que:

$$\sum_{r \in M} p_r = 1$$

Definimos la **entropía estadística o de Shannon del sistema** como:

$$S_{\text{est}} = - \sum_{r \in M} p_r \ln(p_r) = \langle -\ln(p_r) \rangle = \langle \ln \left( \frac{1}{p_r} \right) \rangle \quad (1)$$

Ver [https://en.wikipedia.org/wiki/Entropy\\_\(information\\_theory\)](https://en.wikipedia.org/wiki/Entropy_(information_theory)) para más detalle.

Notar que en general vale:  $S_{\text{est}} \geq 0$ , y ocurre que  $S_{\text{est}} = 0 \iff \exists r \in M / p_r = 1 \iff p_{r'} = \delta_{r'r}$ .

Si maximizamos  $S_{\text{est}}(p_1, \dots, p_r, \dots, p_\Omega)$  sujetos al vínculo  $\sum_{r \in M} p_r = 1$ , obtenemos (usando el método de los multiplicadores de Lagrange): <sup>2</sup>

$$W(p_r) = S_{\text{est}}(p_r) - \nu \left( \sum_{r \in M} p_r - 1 \right)$$

Luego, si maximizamos esta función respecto de las variables  $p_r$ , obtenemos (para todo  $r \in M$ ):

$$0 = \frac{\partial W}{\partial p_r} = -\ln(p_r) - 1 - \nu \iff p_r = e^{-1-\nu} \equiv A$$

<sup>1</sup>Ver [http://materias.df.uba.ar/ft3a2020c1/files/2020/04/Guia\\_2\\_Maximizaci%C3%B3n\\_de\\_entrop%C3%ADa\\_y\\_ejercicio\\_18b.pdf](http://materias.df.uba.ar/ft3a2020c1/files/2020/04/Guia_2_Maximizaci%C3%B3n_de_entrop%C3%ADa_y_ejercicio_18b.pdf)

<sup>2</sup>Probablemente hayan visto en Matemática 1 que para maximizar una función  $f(x_1, \dots, x_\Omega)$  sujeta a los  $n$  vínculos  $g_k(x_1, \dots, x_\Omega) = 0$  con  $1 \leq k \leq n$ , el método de los multiplicadores de Lagrange consiste en plantear que  $\nabla f(\vec{x}) = \sum_{k=1}^n \lambda_k \nabla g_k(\vec{x})$ , donde  $\lambda_k$  con  $1 \leq k \leq n$  son los multiplicadores de Lagrange. Es decir, que el gradiente de  $f$  es perpendicular a la superficie formada por los vínculos  $S = \{g_k(\vec{x}) = 0 \forall 1 \leq k \leq n\} \subseteq \mathbb{R}^\Omega$ . No es difícil ver que plantear ésto es equivalente a plantear:  $\nabla W(\vec{x}) \equiv \nabla \left( f(\vec{x}) - \sum_{k=1}^n \lambda_k g_k(\vec{x}) \right) = \vec{0}$  o equivalentemente que  $\frac{\partial W(\vec{x})}{\partial x_r} = \frac{\partial}{\partial x_r} \left( f(\vec{x}) - \sum_{k=1}^n \lambda_k g_k(\vec{x}) \right) = 0$  para todo  $1 \leq r \leq \Omega$ , donde se definió la función auxiliar  $W(\vec{x}) \equiv f(\vec{x}) - \sum_{k=1}^n \lambda_k g_k(\vec{x})$ .

Y como se debe verificar el vínculo:

$$\sum_{r \in M} p_r = A\Omega = p_r\Omega = 1 \implies p_r = \frac{1}{\Omega}$$

Vemos que el máximo de entropía se da cuando todos los estados son equiprobables, que es la distribución de probabilidad en la cual la incertidumbre antes de conocer el estado es máxima, y la información que se gana al conocer el estado es máxima.

En general, cuanto mayor sea  $S_{\text{est}}$ , mayor es la incertidumbre antes de conocer el estado, y mayor información se gana al conocer el estado.

Maximizar la entropía se corresponde luego con maximizar la incertidumbre antes de conocer el estado.

Ahora planteemos otra situación: Maximicemos la entropía dado el vínculo  $\sum_{r \in M} p_r = 1$ , y también dados los vínculos  $\langle A_k \rangle = a_k$  con  $1 \leq k \leq n$  (hay  $n$  valores medios fijos), donde  $a_k$  es un número real fijo y  $A_k$  es una cierta magnitud del sistema cuyo valor depende del estado del sistema: Si el estado es  $r$  entonces  $A_k$  vale  $A_{k,r}$  (O sea,  $A_k : M \rightarrow \mathbb{R}$ ). Cada uno de los  $n$  vínculos  $\langle A_k \rangle = a_k$  los podemos reescribir como:  $-a_k + \langle A_k \rangle = -a_k + \sum_{r \in M} p_r A_{k,r} = 0$ .

La distribución de probabilidad sobre el espacio muestral  $M$  que se obtenga a partir de maximizar la entropía  $S_{\text{est}}$  dados los anteriores vínculos de valores medios  $\langle A_k \rangle = a_k$  con  $1 \leq k \leq n$  (además del vínculo  $\sum_{r \in M} p_r = 1$ ), nos determina lo que denominamos **un ensamble estadístico con vínculos**  $\langle A_k \rangle = a_k$  con  $1 \leq k \leq n$ .

Entonces, si maximizamos  $S_{\text{est}}$  sujeto a los anteriores vínculos usando multiplicadores de Lagrange, debemos maximizar la función:

$$W(p_r) = S_{\text{est}}(p_r) - \nu \left( \sum_{r \in M} p_r - 1 \right) - \sum_{k=1}^n \lambda_k \left( -a_k + \sum_{r \in M} p_r A_{k,r} \right)$$

Si maximizamos esta función respecto de las variables  $p_r$ , obtenemos (para todo  $r \in M$ ):

$$0 = \frac{\partial W}{\partial p_r} = -\ln(p_r) - 1 - \nu - \sum_{k=1}^n \lambda_k A_{k,r} \iff p_r = A \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right)$$

con  $A \equiv e^{-1-\nu}$ . Y como se debe verificar el vínculo:

$$1 = \sum_{r \in M} p_r = A \sum_{r \in M} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) \equiv AZ = 1 \implies A = \frac{1}{Z}$$

donde:

$$Z \equiv \sum_{r \in M} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) \quad (2)$$

se define como **la función de partición del sistema**. Luego, las probabilidades del

sistema que maximizan la entropía dados los vínculos es:

$$p_r = \frac{1}{Z} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) = \frac{\exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right)}{\sum_{r' \in M} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r'} \right)} \quad (3)$$

Si planteamos los otros vínculos:

$$a_k = \langle A_k \rangle = \sum_{r \in M} p_r A_{k,r} = \frac{1}{Z} \sum_{r \in M} A_{k,r} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) = - \frac{1}{Z} \frac{\partial Z}{\partial \lambda_k} \Big|_{\lambda_j, j \neq k} = - \frac{\partial (\ln Z)}{\partial \lambda_k} \Big|_{\lambda_j, j \neq k}$$

En definitiva obtenemos (recordar que por definición:  $d(f(x_1, \dots, x_n)) \equiv \sum_{k=1}^n \frac{\partial f}{\partial x_k} \Big|_{x_j, j \neq k} dx_k$ ):

$$a_k = - \frac{\partial (\ln Z)}{\partial \lambda_k} \Big|_{\lambda_j, j \neq k} \iff d(\ln Z) = - \sum_{k=1}^n a_k d\lambda_k \quad (4)$$

entendiendo a  $\ln Z(\vec{\lambda})$  como una función de los multiplicadores de Lagrange  $\vec{\lambda} = (\lambda_1, \dots, \lambda_n)$ .

Y si calculamos la entropía estadística que corresponde al estado que la maximiza (dados los vínculos):

$$\begin{aligned} S_{\text{est}}^{\text{max}} &= S_{\text{est}} \left( p_r = \frac{1}{Z} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) \right) = - \sum_{r \in M} \frac{1}{Z} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) \ln \left( \frac{1}{Z} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) \right) \\ &= \sum_{r \in M} p_r \ln Z + \sum_{r \in M} p_r \sum_{k=1}^n \lambda_k A_{k,r} = \left( \sum_{r \in M} p_r \right) \ln Z + \sum_{k=1}^n \lambda_k \left( \sum_{r \in M} p_r A_{k,r} \right) = \ln Z + \sum_{k=1}^n \lambda_k a_k \end{aligned}$$

En definitiva:

$$S_{\text{est}}^{\text{max}} = \ln Z + \sum_{k=1}^n \lambda_k a_k \quad (5)$$

y en consecuencia:

$$dS_{\text{est}}^{\text{max}} = d(\ln Z) + \sum_{k=1}^n (a_k d\lambda_k + \lambda_k da_k) = \sum_{k=1}^n (-a_k d\lambda_k + a_k d\lambda_k + \lambda_k da_k) = \sum_{k=1}^n \lambda_k da_k$$

o equivalentemente:

$$\lambda_k = \frac{\partial S_{\text{est}}^{\text{max}}}{\partial a_k} \Big|_{a_j, j \neq k} \iff dS_{\text{est}}^{\text{max}} = \sum_{k=1}^n \lambda_k da_k \quad (6)$$

Notemos que  $S_{\text{est}}^{\text{max}} = S_{\text{est}}^{\text{max}}(\vec{a})$  es la transformada total de Legendre de  $\ln Z$ , y viceversa.

Ahora bien, si calculamos:

$$\begin{aligned}
\frac{\partial^2(\ln Z)}{\partial\lambda_l\partial\lambda_k} &= -\frac{\partial\langle A_k\rangle}{\partial\lambda_l} = -\frac{\partial}{\partial\lambda_l}\left(\sum_{r\in M}p_rA_{k,r}\right) = -\frac{\partial}{\partial\lambda_l}\left(\frac{1}{Z}\sum_{r\in M}A_{k,r}\exp\left(-\sum_{k=1}^n\lambda_kA_{k,r}\right)\right) \\
&= \frac{1}{Z^2}\frac{\partial Z}{\partial\lambda_l}\sum_{r\in M}A_{k,r}\exp\left(-\sum_{k=1}^n\lambda_kA_{k,r}\right) + \frac{1}{Z}\sum_{r\in M}A_{k,r}A_{l,r}\exp\left(-\sum_{k=1}^n\lambda_kA_{k,r}\right) \\
&= \frac{\partial(\ln Z)}{\partial\lambda_l}\sum_{r\in M}A_{k,r}p_r + \sum_{r\in M}A_{k,r}A_{l,r}p_r = \langle A_kA_l\rangle - \langle A_l\rangle\langle A_k\rangle = \text{Cov}(A_k, A_l)
\end{aligned}$$

obtenemos los elementos de matriz de **la matriz de covarianza** (que es simétrica), **sólo con derivar dos veces a  $\ln Z$** . En definitiva:

$$\frac{\partial^2(\ln Z)}{\partial\lambda_l\partial\lambda_k} = -\frac{\partial\langle A_k\rangle}{\partial\lambda_l} = \langle A_kA_l\rangle - \langle A_l\rangle\langle A_k\rangle = \text{Cov}(A_k, A_l) \quad (7)$$

En particular, si  $k = l$ , obtenemos la varianza de  $A_k$ :

$$\frac{\partial^2(\ln Z)}{\partial\lambda_k^2} = \langle A_k^2\rangle - \langle A_k\rangle^2 = \text{Cov}(A_k, A_k) = \sigma_k^2$$

La matriz de covarianza  $K_{\vec{A}\vec{A}} = \langle(\vec{A} - \vec{a})(\vec{A} - \vec{a})^t\rangle = \frac{\partial^2(\ln Z)}{\partial\vec{\lambda}\partial\vec{\lambda}^t}$  es (semi-)positiva definida ya que (siendo  $w$  un vector constante):  $w^tK_{\vec{A}\vec{A}}w = w^t\langle(\vec{A} - \vec{a})(\vec{A} - \vec{a})^t\rangle w = \langle w^t(\vec{A} - \vec{a})(\vec{A} - \vec{a})^tw\rangle = \langle(w^t(\vec{A} - \vec{a}))^2\rangle \geq 0$  para todo vector  $w$  gracias a que  $w^t(\vec{A} - \vec{a}) = (\vec{A} - \vec{a})^tw$  es un escalar. En consecuencia la matriz de covarianza es semi-definida positiva.

Notemos que, si calculamos los elementos de matriz de covarianza relativos (una generalización trivial del cuadrado del error relativo  $\sigma_X^2/\langle X\rangle^2$ ):

$$\frac{1}{\langle A_l\rangle\langle A_k\rangle}\text{Cov}(A_k, A_l) = \frac{\langle A_kA_l\rangle}{\langle A_l\rangle\langle A_k\rangle} - 1 = -\frac{1}{\langle A_l\rangle\langle A_k\rangle}\frac{\partial\langle A_k\rangle}{\partial\lambda_l} = -\frac{1}{\langle A_l\rangle}\frac{\partial\ln(\langle A_k\rangle)}{\partial\lambda_l} = -\frac{1}{\langle A_k\rangle}\frac{\partial\ln(\langle A_l\rangle)}{\partial\lambda_k}$$

Notar que si  $k = l$ , se tiene el cuadrado del error relativo:

$$\frac{\sigma_k^2}{\langle A_k\rangle^2} = -\frac{1}{\langle A_k\rangle}\frac{\partial\ln(\langle A_k\rangle)}{\partial\lambda_k}$$

Si suponemos que las magnitudes  $A_k$  para todo  $1 \leq k \leq n$  son extensivas (es decir, son proporcionales al número de partículas  $N$  del sistema:  $A_k \propto N$ ), y que la entropía estadística es también una cantidad extensiva, entonces debe ocurrir trivialmente que  $\lambda_k = \left.\frac{\partial S_{\text{est}}^{\text{max}}}{\partial a_k}\right|_{a_j, j \neq k}$  deben ser cantidades intensivas (que no dependen de  $N$ ). Luego, al tomar el límite termodinámico  $N \rightarrow \infty$ , ocurre que los elementos de matriz de la covarianza relativos tienden a cero:

$$\frac{1}{\langle A_l\rangle\langle A_k\rangle}\text{Cov}(A_k, A_l) = -\frac{1}{\langle A_k\rangle}\frac{\partial\ln(\langle A_l\rangle)}{\partial\lambda_k} \propto \frac{1}{N} \rightarrow 0 \quad \text{si } N \rightarrow \infty \quad (8)$$

En particular, los errores relativos  $\sigma_k^2 \propto 1/N \rightarrow 0$  en el límite termodinámico  $N \rightarrow \infty$ .

Luego, en el límite termodinámico, las fluctuaciones de las variables  $A_k$  alrededor del valor medio  $a_k$  son despreciables. En consecuencia considerar un ensamble con uno de sus vínculos que sea que el valor medio de una magnitud  $\langle A_k\rangle$  fijo, es equivalente (en el límite termodinámico) a

considerar un ensamble cuyo espacio muestral sólo contenga sistemas cuyo valor de la magnitud  $A_k$  sea exactamente igual a  $a_k$ .

Por último, notar:

$$\frac{\partial^2 S_{\text{est}}^{\text{max}}}{\partial a_l \partial a_k} = \frac{\partial \lambda_k}{\partial a_l}$$

es el elemento  $l, k$  de la matriz  $\frac{\partial^2 S_{\text{est}}^{\text{max}}}{\partial \vec{a} \partial \vec{a}^t}$ , y no es difícil ver que es la matriz diferencial de la función  $\vec{\lambda}(\vec{a}) \equiv F(\vec{a})$ . Por lo tanto, por el teorema de la función inversa, ésta matriz es la inversa de la matriz diferencial de la función  $\vec{a}(\vec{\lambda}) \equiv F^{-1}(\vec{\lambda})$ , que como vimos, es igual a menos la matriz de covarianza  $K_{\vec{A}\vec{A}}$  (ya que sus elementos de matriz son  $\langle (A_l - a_l)(A_k - a_k) \rangle = \frac{\partial^2 (\ln Z)}{\partial \lambda_l \partial \lambda_k} = -\frac{\partial a_k}{\partial \lambda_l} = -(D(F^{-1}))_{lk}$ ):

$$\frac{\partial^2 S_{\text{est}}^{\text{max}}}{\partial \vec{a} \partial \vec{a}^t} = DF = (D(F^{-1}))^{-1} = (-K_{\vec{A}\vec{A}})^{-1} = -K_{\vec{A}\vec{A}}^{-1} = -\left(\frac{\partial^2 (\ln Z)}{\partial \vec{\lambda} \partial \vec{\lambda}^t}\right)^{-1} \quad (9)$$

donde, para invertir la matriz de covarianza se asumió que es inversible, o equivalentemente que es positiva definida (descartando el caso de que sea semi-positiva definida o equivalentemente que tenga algún autovalor nulo). Por lo tanto, el hessiano de  $S_{\text{est}}^{\text{max}}$  es negativo definido, y en consecuencia el extremo hallado de la entropía es un máximo.

## 1.1. Ensamblados en continuo

En caso de que  $M$  no sea discreto, la fórmula de la entropía estadística se obtiene reemplazando la sumatoria  $\sum_{r \in M}$  de la expresión 1 por una integral  $\int_M d^L R$ , si especificar el estado  $r \in M$  consiste en especificar el valor de  $L$  variables aleatorias  $\vec{R} = (R_1, \dots, R_L)$  que toman valores continuos:

$$S_{\text{est}} = - \int_M f(\vec{R}) \ln(f(\vec{R})) d^L R$$

y se denomina entropía diferencial o continua (para más detalles ver [https://en.wikipedia.org/wiki/Differential\\_entropy](https://en.wikipedia.org/wiki/Differential_entropy)).

Un pequeño comentario al respecto: Como se debe verificar  $\int_M f(\vec{R}) d^L R = 1$  entonces las unidades de la distribución de probabilidad son trivialmente:  $[f] = [d^L R]^{-1} = [R_1]^{-1}(\dots)[R_L]^{-1}$ . Ahora bien, puesto que en el cálculo de la entropía se está aplicando un logaritmo a  $f$ , entonces es necesario que  $f(\vec{R})$  sea adimensional (ver <http://math.ucr.edu/home/baez/physics/General/logs.html>), para lo cual el elemento de volumen  $d^L R$  debe ser adimensional. Si partimos de variables  $r_1, (\dots), r_L$  tales que  $d^L r$  no es adimensional, entonces podemos normalizar  $d\omega \equiv d^L r \rightarrow d^L R \equiv \frac{d\omega}{\omega_0} = \frac{d^L r}{\omega_0}$  donde  $\omega_0$  es una cantidad con unidades de volumen  $[\omega_0] = [d^L r]$  que va a servir para adimensionalizar el elemento de volumen  $d^L r$ .

En consecuencia se pasa de tener  $1 = \int_M \rho_N(\vec{r}) d^L r = \int_M \rho_N(\vec{r}) \omega_0 \frac{d^L r}{\omega_0} = \int_M f(\vec{R}) d^L R$ , donde  $\rho_N$  es una densidad normalizada. También se puede tener una densidad  $\rho$  no normalizada ( $V_0 \equiv \int_M \rho(\vec{r}) d^L r \neq 1$ ), por lo que podemos considerar que la densidad normalizada es  $\rho_N(\vec{r}) = \frac{\rho(\vec{r})}{V_0}$  y luego se tiene:  $f(\vec{R}) = \rho_N(\vec{r}) \omega_0 = \frac{1}{V_0} \rho(\vec{r}) \omega_0 = \frac{\rho(\vec{r})}{\int_M \rho(\vec{r}) d^L r} \omega_0 = \frac{\rho(\vec{r})}{\int_M \rho(\vec{r}) d^L r / \omega_0}$ .

Si **interpretamos a  $\omega_0$  como el volumen de un microestado**, entonces la cantidad de microestados en un espacio muestral  $M$  es igual a  $\Omega = \frac{\omega}{\omega_0} \equiv \frac{\text{Vol}(M)}{\omega_0} = \frac{1}{\omega_0} \int_M d^L R$ .

No es difícil ver, al maximizar  $S_{\text{est}}[f(\vec{R})]$  (que ahora es un funcional), que **todas las ecuaciones 2, 3, 4, 5, 6, 7, 8, 9** de la anterior sección **son válidas reemplazando**

$$\sum_{r \in M} \rightarrow \int_M d^L R = \int_M \frac{d^L r}{\omega_0} \quad (10)$$

## 1.2. Ensamblen en términos de copias del sistema

Otra forma de entender a los ensambles es en términos de copias del sistema: **Un ensamble es un conjunto de muchas copias de un mismo sistema en un mismo estado macroscópico de equilibrio en un mismo tiempo, pero posiblemente en distintos estados microscópicos.** Ésto nos da una distribución de probabilidad sobre los posibles estados microscópicos de un sistema dado que está en un cierto estado macroscópico en equilibrio.

Consideremos que hay  $\eta$  copias del sistema (con  $\eta$  un número gigante:  $\eta \rightarrow \infty$ ; por supuesto, las copias no interactúan entre sí) que se encuentran en el espacio muestral  $M$  distribuidas con la distribución de probabilidad  $\{p_r / r \in M\}$  que maximiza la entropía dados ciertos vínculos. Es decir, cada una de las  $\eta$  copias del sistema es un elemento de  $M$ , y van a haber  $\eta_r \equiv p_r \eta$  copias en el estado  $r \in M$  (gracias a que la probabilidad es frecuentista y que  $\eta \rightarrow \infty$ ). Por supuesto, se debe cumplir el vínculo trivial  $\sum_{r \in M} \eta_r = \eta$ , y si hay vínculos de la forma  $\langle A_k \rangle = a_k$

se traducen a:  $\sum_{r \in M} \eta_r A_{k,r} = \eta a_k$ .

Consideremos que el espacio muestral  $M = \{1, \dots, \Omega\}$  es discreto y posee  $\Omega$  elementos (etiquetados por números entre 1 y  $\Omega \in \mathbb{N}$ ). Entonces la multiplicidad de ésta distribución de estados, es decir la cantidad de formas que hay de elegir  $\eta$  copias tales que hayan  $\eta_r$  copias en el estado etiquetado por  $r$  para todo  $r \in M$  es igual a: La cantidad de formas  $\binom{\eta}{\eta_1}$  de elegir  $\eta_1$  copias de las  $\eta$  copias totales en el estado 1, multiplicadas por la cantidad de formas  $\binom{\eta - \eta_1}{\eta_2}$  de elegir  $\eta_2$  copias de las  $\eta - \eta_1$  restantes en el estado 2, multiplicadas por la cantidad de formas  $\binom{\eta - \eta_1 - \eta_2}{\eta_3}$  de elegir  $\eta_3$  copias de las  $\eta - \eta_1 - \eta_2$  restantes en el estado 3, y así sucesivamente hasta multiplicar por la cantidad de formas  $\binom{\eta - \eta_1 - \eta_2 - \dots - \eta_{\Omega-1}}{\eta_\Omega}$  de elegir  $\eta_\Omega$  copias de las  $\eta - \eta_1 - \eta_2 - \dots - \eta_{\Omega-1}$

restantes en el (último) estado  $\Omega$ :

$$\begin{aligned}\Gamma_D &= \binom{\eta}{\eta_1} \binom{\eta - \eta_1}{\eta_2} \binom{\eta - \eta_1 - \eta_2}{\eta_3} (\dots) \binom{\eta - \eta_1 - (\dots) - \eta_{\Omega-1}}{\eta_\Omega} \\ &= \frac{\eta!}{\eta_1! (\eta - \eta_1)!} \frac{(\eta - \eta_1)!}{\eta_2! (\eta - \eta_1 - \eta_2)!} \frac{(\eta - \eta_1 - \eta_2)!}{\eta_3! (\eta - \eta_1 - \eta_2 - \eta_3)!} (\dots) \frac{(\eta - \eta_1 - (\dots) - \eta_{\Omega-1})!}{\eta_\Omega! (\eta - \eta_1 - (\dots) - \eta_{\Omega-1} - \eta_\Omega)!} \\ &= \frac{\eta!}{\eta_1! \eta_2! \eta_3! (\dots) \eta_\Omega!} = \frac{\eta!}{\prod_{r \in M} \eta_r!}\end{aligned}$$

en caso de que las  $\eta$  copias del sistema sean distinguibles. Se utilizó  $\eta - \sum_{r \in M} \eta_r = \eta - \eta_1 - (\dots) - \eta_{\Omega-1} - \eta_\Omega = 0$  y  $0! = 1$ .

Pero si las  $\eta$  copias del sistema son indistinguibles, podemos aplicar el conteo correcto de Boltzmann dividiendo lo anterior por  $\eta!$ :

$$\Gamma_I = \frac{1}{\eta!} \Gamma_D = \frac{1}{\prod_{r \in M} \eta_r!}$$

Y si calculamos (utilizando la aproximación de Sterling):

$$\begin{aligned}\ln \Gamma_D &= \ln(\eta!) - \sum_{r \in M} \ln(\eta_r!) \simeq \eta(\ln \eta - 1) - \sum_{r \in M} \eta_r(\ln \eta_r - 1) \\ &= \eta(\ln \eta - 1) + \sum_{r \in M} \eta_r - \eta \sum_{r \in M} \frac{\eta_r}{\eta} \ln \left( \frac{\eta_r}{\eta} \right) = \eta \left( \ln \eta - 1 + 1 - \sum_{r \in M} \frac{\eta_r}{\eta} \ln \left( \frac{\eta_r}{\eta} \right) - \ln \eta \sum_{r \in M} \frac{\eta_r}{\eta} \right) \\ &= \eta \left( \ln \eta - 1 + 1 - \ln \eta - \sum_{r \in M} \frac{\eta_r}{\eta} \ln \left( \frac{\eta_r}{\eta} \right) \right) = \eta \left( - \sum_{r \in M} \frac{\eta_r}{\eta} \ln \left( \frac{\eta_r}{\eta} \right) \right) = \eta \left( - \sum_{r \in M} p_r \ln(p_r) \right) \\ &= \eta S_{\text{est}}\end{aligned}$$

Por otro lado, si aplicamos el conteo correcto de Boltzmann:

$$\ln \Gamma_I = \ln \Gamma_D - \ln(\eta!) \simeq \eta(S_{\text{est}} - \ln \eta + 1)$$

Entonces, notamos que maximizar la entropía estadística  $S_{\text{est}}$  es equivalente a maximizar  $\ln \Gamma_D = \eta S_{\text{est}}$  (ya que  $\eta$  es un número fijo), lo cual también es equivalente a maximizar  $\ln \Gamma_I = \ln \Gamma_D - \ln(\eta!) \simeq \eta(S_{\text{est}} - \ln \eta + 1)$  y a su vez es equivalente trivialmente a maximizar  $\Gamma_D$  y a maximizar  $\Gamma_I$ .

En definitiva, vemos que **maximizar la entropía es equivalente a maximizar la multiplicidad (independientemente de si las copias son distinguibles o indistinguibles), luego el equilibrio físico se corresponde con éste máximo.**

### 1.3. Distribuciones de máxima entropía, ejemplo: Ejercicio 18b, guía 2

La distribución de probabilidad que maximice la entropía va a depender fuertemente de los vínculos  $\langle A_k(\vec{R}) \rangle = a_k$  en cuestión, y del dominio de las variables aleatorias. De hecho,



podemos ver en la tabla de la página [https://en.wikipedia.org/wiki/Maximum\\_entropy\\_probability\\_distribution#Other\\_examples](https://en.wikipedia.org/wiki/Maximum_entropy_probability_distribution#Other_examples) que se pueden obtener distribuciones de probabilidad que maximicen la entropía muy diversas, que dependen en esencia de los vínculos elegidos y del dominio de las variables aleatorias.

En particular, mostraremos como es la distribución de probabilidad de máxima entropía en un sistema con una variable aleatoria  $x$  que pueden tomar cualquier valor real (continua), con los vínculos  $\langle x \rangle = \mu$  (con multiplicador de Lagrange asociado  $\lambda$ ) y  $\langle \Delta x^2 \rangle = \langle (x - \mu)^2 \rangle = \sigma^2$  (con multiplicador de Lagrange asociado  $\xi$ ). Entonces la función de partición del sistema queda de la forma:

$$\begin{aligned} Z &= \int_{\mathbb{R}} dx e^{-\lambda x - \xi(x-\mu)^2} = \int_{\mathbb{R}} dx e^{-\lambda(x-\mu+\mu) - \xi(x-\mu)^2} = e^{-\lambda\mu} \int_{\mathbb{R}} dx e^{-\lambda(x-\mu) - \xi(x-\mu)^2} \\ &= e^{-\lambda\mu} \int_{\mathbb{R}} dx e^{-\lambda\Delta x - \xi\Delta x^2} \end{aligned}$$

donde definimos  $\Delta x \equiv x - \mu$ . Luego, para poder integrar, completamos cuadrados en la expresión dentro de la exponencial del integrando:

$$\begin{aligned} -\lambda\Delta x - \xi\Delta x^2 &= -\xi \left( \Delta x^2 + \frac{\lambda}{\xi} \Delta x \right) = -\xi \left( \Delta x^2 + 2 \left( \frac{\lambda}{2\xi} \right) \Delta x \right) \\ &= -\xi \underbrace{\left( \Delta x^2 + 2 \left( \frac{\lambda}{2\xi} \right) \Delta x + \frac{\lambda^2}{4\xi^2} \right)}_{\left( \Delta x + \frac{\lambda}{2\xi} \right)^2} + \frac{\lambda^2}{4\xi} = -\xi \left( \Delta x + \frac{\lambda}{2\xi} \right)^2 + \frac{\lambda^2}{4\xi} \end{aligned} \quad (11)$$

donde se asumió que  $\xi \neq 0$ . Entonces la cuenta que hay que hacer es la siguiente:

$$\begin{aligned} Z(\lambda, \xi) &= e^{-\lambda\mu} \int_{\mathbb{R}} dx e^{-\lambda\Delta x - \xi\Delta x^2} \\ &= e^{-\lambda\mu} \int_{\mathbb{R}} dx e^{-\xi \left( \Delta x + \frac{\lambda}{2\xi} \right)^2 + \frac{\lambda^2}{4\xi}} \\ &= e^{-\lambda\mu + \frac{\lambda^2}{4\xi}} \int_{\mathbb{R}} dx e^{-\xi \left( \Delta x + \frac{\lambda}{2\xi} \right)^2} \\ &= e^{-\lambda\mu + \frac{\lambda^2}{4\xi}} \int_{\mathbb{R}} dy e^{-\xi y^2} \\ &= e^{-\lambda\mu + \frac{\lambda^2}{4\xi}} \sqrt{\frac{\pi}{\xi}} \end{aligned} \quad (12)$$

donde en el segundo renglón se utilizó la ecuación 11, en el tercer renglón se extrajo de la integral el factor constante  $e^{\frac{\lambda^2}{4\xi}}$ , en el cuarto renglón se realizó el cambio de variable  $x \rightarrow y \equiv x - \mu + \frac{\lambda}{2\xi} = \Delta x + \frac{\lambda}{2\xi}$  con jacobiano igual a 1, y en el quinto y último renglón se utilizó que la integral de una gaussiana en los reales es precisamente  $\int_{\mathbb{R}} dy e^{-\xi y^2} = \sqrt{\frac{\pi}{\xi}}$ .<sup>3</sup>

---

<sup>3</sup>Podemos demostrar ésto si lo elevamos al cuadrado, lo vemos como una integral en dos dimensiones y pasamos a coordenadas polares:  $\left( \int_{\mathbb{R}} dy e^{-\xi y^2} \right)^2 = \left( \int_{\mathbb{R}} dx e^{-\xi x^2} \right) \left( \int_{\mathbb{R}} dy e^{-\xi y^2} \right) = \iint_{\mathbb{R}^2} dx dy e^{-\xi(x^2+y^2)} =$

Y luego:

$$\ln Z(\lambda, \xi) = \frac{1}{2} \ln \pi - \lambda \mu + \frac{\lambda^2}{4\xi} - \frac{1}{2} \ln \xi \quad (13)$$

Entonces, si se plantean los vínculos se obtiene:

$$\mu = -\frac{\partial}{\partial \lambda} \ln Z = \mu - \frac{\lambda}{2\xi} \iff \frac{\lambda}{2\xi} = 0 \iff \lambda = 0$$

Donde se utilizó que  $\xi \neq 0$ .

Y si calculamos ahora la derivada segunda de  $\ln Z$  respecto a  $\lambda$ , deberíamos obtener la matriz de covarianza de  $x$  (o sea, como hay una sola variable, deberíamos obtener la varianza  $\sigma^2$ ):

$$\sigma^2 = \langle (x - \mu)^2 \rangle = \frac{\partial^2 (\ln Z)}{\partial \lambda^2} = -\frac{\partial}{\partial \lambda} \left( -\frac{\partial}{\partial \lambda} \ln Z \right) = -\frac{\partial}{\partial \lambda} \left( \mu - \frac{\lambda}{2\xi} \right) = \frac{1}{2} \xi^{-1} \iff \xi = \frac{1}{2\sigma^2}$$

También podemos obtener la varianza  $\sigma^2$  derivando a  $\ln Z$  respecto al multiplicador de Lagrange  $\xi$  asociado al vínculo que fija la varianza:

$$\sigma^2 = -\frac{\partial}{\partial \xi} \ln Z = -\frac{\partial}{\partial \xi} \left( \frac{\lambda^2}{4\xi} - \frac{1}{2} \ln \xi \right) = -\frac{\lambda^2}{4} \frac{\partial}{\partial \xi} (\xi^{-1}) + \frac{1}{2} \frac{\partial}{\partial \xi} (\ln \xi) = \frac{\lambda^2}{4\xi^2} + \frac{1}{2\xi} = \frac{1}{2} \xi^{-1}$$

donde se utilizó en el último paso que  $\lambda = 0$  por el anterior vínculo. De ésta forma se reobtuvo que  $\sigma^2 = (2\xi)^{-1}$ .

Ésto nos permite escribir la distribución de probabilidades de la forma:

$$\begin{aligned} f(x) &= \frac{1}{Z} e^{-\lambda x - \xi(x-\mu)^2} = \frac{1}{Z} e^{-\lambda \mu} e^{-\lambda \Delta x - \xi \Delta x^2} = \frac{1}{Z} e^{-\lambda \mu + \frac{\lambda^2}{4\xi}} e^{-\xi \left( \Delta x + \frac{\lambda}{2\xi} \right)^2} \\ &= \sqrt{\frac{\xi}{\pi}} e^{-\xi \left( x - \mu + \frac{\lambda}{2\xi} \right)^2} = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (x-\mu)^2} \end{aligned}$$

donde se utilizó la ecuación 11 en el tercer paso, se reemplazó  $Z(\lambda, \xi)$  en el cuarto paso, y en el quinto y último paso se utilizó lo obtenido a partir de los vínculos:  $\lambda = 0$  y  $\xi = (2\sigma^2)^{-1}$ . En definitiva, la función de distribución de probabilidad de una variable aleatoria  $x$  con dominio en  $\mathbb{R}$ , que maximiza la entropía sujeto a los vínculos  $\langle x \rangle = \mu$  y  $\langle (x - \mu)^2 \rangle = \sigma^2$ , es una distribución gaussiana de una variable.

También se puede calcular la entropía estadística de ésta distribución simplemente haciendo:

$$\begin{aligned} S_{\text{est}}(f(\vec{R})) &= S_{\text{est}}^{\text{max}} = \ln Z + \lambda \mu + \xi \sigma^2 = \frac{1}{2} \ln \pi + \frac{\lambda^2}{4\xi} - \frac{1}{2} \ln \xi + \xi \sigma^2 \\ &= \frac{1}{2} \ln \pi - \frac{1}{2} \ln (2\sigma^2)^{-1} + (2\sigma^2)^{-1} \sigma^2 = \frac{1}{2} (\ln \pi + \ln (2\sigma^2) + 1) = \frac{1}{2} \ln (2\pi e \sigma^2) \end{aligned}$$

donde en el tercer paso se reemplazó la expresión de  $\ln Z$ , en el cuarto paso se reemplazaron las expresiones  $\lambda = 0$  y  $\xi = (2\sigma^2)^{-1}$  obtenidas con los vínculos, y luego se hicieron cuentas.

Nótese de hecho que derivando a  $S_{\text{est}}^{\text{max}}(\mu, \sigma^2)$  podemos reobtener:  $\lambda = \frac{\partial S_{\text{est}}^{\text{max}}}{\partial \mu} = 0$  y también

$$\xi = \frac{\partial S_{\text{est}}^{\text{max}}}{\partial (\sigma^2)} = \frac{1}{2\sigma^2}, \text{ es decir: Recuperamos lo obtenido con los vínculos.}$$

---


$$\int_0^{2\pi} d\theta \int_0^{+\infty} dr r e^{-\xi r^2} = 2\pi \frac{1}{2\xi} \int_0^{+\infty} dr e^{-u} = \frac{\pi}{\xi} (-0 + 1) = \frac{\pi}{\xi}, \text{ donde se utilizó que el jacobiano al pasar de coordenadas cartesianas a polares en } \mathbb{R}^2 \text{ es igual a } r, \text{ y se realizó el cambio de variable } r \rightarrow u \equiv \xi r^2 \text{ (luego } r dr = \frac{1}{2\xi} du \text{). Por lo tanto, tomando raíz cuadrada de lo anterior: } \int_{\mathbb{R}} dy e^{-\xi y^2} = \sqrt{\frac{\pi}{\xi}}.$$

## 2. Conexión con termodinámica

En termodinámica, sabemos por la Segunda Ley que en un sistema aislado en equilibrio, la entropía termodinámica  $S_{\text{termo}}$  es máxima.

Ahora estudiaremos la **conexión de la entropía estadística con la entropía termodinámica de un sistema aislado en equilibrio (por lo tanto posee entropía máxima)**:

$$S_{\text{termo}} = k_B S_{\text{est}}^{\text{max}} \quad (14)$$

lo cual nos permite conectar (usando la ecuación 6):

$$k_B dS_{\text{est}}^{\text{max}} = \sum_{k=1}^n k_B \lambda_k da_k = dS_{\text{termo}} = \left( \frac{1}{T} dU - \sum_{i=1}^m \frac{Y_i}{T} d\bar{X}_i \right) \Big|_M \quad (15)$$

Por supuesto, se utilizó para la expresión de  $dS_{\text{termo}}$  el primer principio de la termodinámica:  $\delta Q = dU - \delta W_{\text{ext}}^{\text{tot}} = dU - \sum_{i=1}^m Y_i d\bar{X}_i$ ; y se utilizó el hecho de que, en procesos reversibles ocurre que  $\delta Q_R = T dS$ .

Donde  $-\delta W^{(i)} = \delta W_{\text{ext}}^{(i)} = Y_i d\bar{X}_i$  es un trabajo externo genérico realizado sobre el sistema, y  $\bar{X}_i$  es en general una variable extensiva (en general es el valor medio de una cierta magnitud extensiva:  $\bar{X}_i = \langle X_i \rangle$ ) e  $Y_i$  es una variable intensiva, que puede ser de la forma típica  $Y_i d\bar{X}_i = -pdV$  para un sistema con presión  $p$  y volumen variable (con  $Y_i = -p$  y  $\bar{X}_i = V$ ), y puede ser de la forma típica  $Y_i d\bar{X}_i = \mu dN$  para un sistema con potencial químico  $\mu$  y número de partículas variable (con  $Y_i = \mu$  y  $\bar{X}_i = N$ ). También podemos considerar que se tienen varios componentes químicos en el sistema con número de partículas variable:  $Y_i d\bar{X}_i = \mu_i dN_i$ . Pero también pueden haber trabajos más atípicos, por ejemplo de la forma  $Y_i d\bar{X}_i = T dL$  para una soga con tensión  $T$  y longitud  $L$ , de la forma  $Y_i d\bar{X}_i = s dA$  para una lámina superficial con área  $A$  y tensión superficial  $s$ , de la forma  $Y_i d\bar{X}_i = E dP_E$  para un sólido dieléctrico con campo eléctrico externo  $E$  y polarización total  $P_E$ , de la forma  $Y_i d\bar{X}_i = \mu_0 H dM$  para un sólido magnético con campo magnético externo  $H$  y momento magnético total  $M$ <sup>4</sup>, de la forma  $Y_i d\bar{X}_i = V_0 dq$  para una pila reversible.

Podemos considerar también  $Y_i d\bar{X}_i = -dU$  (con  $Y_i = -1$  y  $\bar{X}_i = U$ ), que no es un trabajo externo y además ya lo consideramos en la expresión anterior, pero igual posee esa forma.

Por último, observamos que la anterior expresión está sujeta al espacio muestral  $M$ . Por ejemplo, si en el espacio muestral  $M$  sólo se consideran los sistemas cuyo volumen es exactamente igual a un cierto  $V$ , entonces se debe considerar que  $dV|_M = 0$ . De la misma forma, si en el espacio muestral  $M$  sólo se consideran los sistemas con  $X_i = X_{i,0}$  entonces:  $d\bar{X}_i|_M = 0$ .

<sup>4</sup>En sistema de unidades gaussianas como cgs se toma  $\mu_0 = 1$ ; ver apunte [http://materias.df.uba.ar/ft3a2020c1/files/2020/04/adicional\\_g1p12.pdf](http://materias.df.uba.ar/ft3a2020c1/files/2020/04/adicional_g1p12.pdf) para más detalles acerca de éste tipo de trabajo.

Consideremos un ensamble, definido por un espacio muestral  $M = \{X_i = X_{i,0} = \text{cte} \forall n+1 \leq i \leq m\}$  (con  $0 \leq n \leq m$ ) cuyos valores medios fijos (vínculos) son:  $a_0 = \langle X_0 \rangle = \bar{X}_0 = U$  la energía media del sistema, y las cantidades  $a_k = \langle X_k \rangle = \bar{X}_k$  para  $1 \leq k \leq n$ . En éste caso, la anterior ecuación se traduce a (como  $d\bar{X}_i|_M = 0$  para  $n+1 \leq i \leq m$  y  $a_0 = U$  y  $a_k = \bar{X}_k$  para  $1 \leq k \leq n$ ):

$$k_B dS_{\text{est}}^{\text{max}} = \sum_{k=0}^n k_B \lambda_k da_k = k_B \lambda_U dU + \sum_{k=1}^n k_B \lambda_k d\bar{X}_k = dS_{\text{termo}} = \frac{1}{T} dU - \sum_{k=1}^n \frac{Y_k}{T} d\bar{X}_k$$

**La anterior ecuación nos permite conectar los multiplicadores de lagrange  $\lambda_k$  con las variables intensivas  $Y_i$  con significado físico en el problema:** Si se toma  $a_k = \bar{X}_k$  para un cierto  $k$ , entonces a partir de la anterior igualdad podemos deducir:  $k_B \lambda_k da_k = -\frac{Y_k}{T} d\bar{X}_k \implies \lambda_k = -\frac{Y_k}{k_B T}$ .

Notar **en particular** que para  $a_0 = \bar{X}_0 = U$  se tiene que el multiplicador de lagrange correspondiente es:

$$\beta \equiv \lambda_U = \frac{1}{k_B T} \quad (16)$$

En particular para  $a_k = \bar{X}_k = N$  se tiene que el multiplicador de lagrange correspondiente es:

$$\gamma \equiv \lambda_N = -\frac{\mu}{k_B T} = -\beta \mu \quad (17)$$

En éste caso particular se suele definir una cantidad llamada fugacidad de la forma:  $z \equiv e^{-\gamma} = e^{\beta \mu} = e^{\mu/(k_B T)}$ , gracias a que como  $N$  es adimensional por ser un número,  $\gamma$  también lo es y  $z$  también.

En particular para  $a_k = \bar{X}_k = V$  se tiene que el multiplicador de lagrange correspondiente es:

$$\delta \equiv \lambda_V = -\frac{(-p)}{k_B T} = \beta p \quad (18)$$

Por último, en general se tiene que el multiplicador de Lagrange para  $a_k = \bar{X}_k$  es:

$$\lambda_k = \lambda_{\bar{X}_k} = -\frac{Y_k}{k_B T} = -\beta Y_k \quad (19)$$

Luego, la expresión de la función de partición en el anterior ensamble es la siguiente:

$$Z = \sum_{r \in M} \exp \left( -\sum_{k=0}^n \lambda_k A_{k,r} \right) = \sum_{r \in M} \exp \left( -\beta \left( E_r - \sum_{k=1}^n Y_k X_{k,r} \right) \right) \quad (20)$$

con distribución de probabilidad asociada:

$$p_r = \frac{1}{Z} \exp \left( -\beta \left( E_r - \sum_{k=1}^n Y_k X_{k,r} \right) \right) \quad (21)$$

Si en este caso calculamos:

$$-k_B T \ln(Z) = -k_B T \left( S_{\text{est}}^{\text{max}} - \sum_{k=0}^n \lambda_k a_k \right) = -T S_{\text{termo}} + k_B T \beta \left( U - \sum_{k=1}^n Y_k \bar{X}_k \right)$$

Luego, como  $k_B T \beta = 1$ , se tiene:

$$G_Z \equiv -k_B T \ln(Z) = U - T S_{\text{termo}} - \sum_{k=1}^n Y_k \bar{X}_k \quad (22)$$

Como el primer principio de la termodinámica implica (en procesos reversibles)  $dU = TdS + \sum_{i=1}^m Y_i d\bar{X}_i$  (con  $m \geq n$ ), entonces si calculamos el diferencial de  $G_Z$ , se tiene:

$$dG_Z = -SdT - \sum_{k=1}^n \bar{X}_k dY_k + \sum_{i=n+1}^m Y_i d\bar{X}_i \quad (23)$$

Es decir, **la elección del ensamble** (las magnitudes cuyos valores medios están fijos en el ensamble y las restantes magnitudes que están fijas en el espacio muestral) **determina un potencial termodinámico**  $G_Z(T, Y_1, \dots, Y_n, \bar{X}_{n+1}, \dots, \bar{X}_m)$ , **que es una transformada de Legendre parcial de**  $U(S, \bar{X}_1, \dots, \bar{X}_m)$  **en las variables cuyos valores medios están fijos en el ensamble** (es decir:  $S, \bar{X}_1, \dots, \bar{X}_n$ ).

### 3. Ensamblares particulares en mecánica estadística

Aquí sólo consideraremos los trabajos externos típicos de la forma  $-pdV$  y  $\mu dN$ , sin considerar ningún otro trabajo externo  $Yd\bar{X}$ . Es decir, el primer principio es:

$$dU = TdS - pdV + \mu dN$$

#### 3.1. Ensamble microcanónico

Consideremos un **ensamble**, llamado **microcanónico**, formado por un espacio muestral  $M = \{(E, V, N) = \text{ctes}\}$  que consiste en todos los estados termodinámicos que posean  $E = U = \text{cte}$ ,  $V = \text{cte}$  y  $N = \text{cte}$ ; que **no posea vínculos más allá del vínculo trivial**  $\sum_{r \in M} p_r = 1$ .

Entonces, **la función del partición** (ver ecuación 2) **del ensamble microcanónico es igual a la multiplicidad del espacio muestral**  $\Omega = \#(M) \in \mathbb{N}$  (considerando que éste es discreto y finito):

$$Z_{\text{microcanónico}} = \sum_{r \in M} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \sum_{r \in M} 1 = \Omega \quad (24)$$

Ésto es ya que no hay vínculos  $\langle A_k \rangle = a_k$  más allá del vínculo trivial, entonces  $n = 0$  y en consecuencia cualquier sumatoria  $\sum_{k=1}^n \lambda_k A_{k,r} = 0$  es siempre nula.

En consecuencia, la distribución de probabilidad que maximiza la entropía en éste ensemble (que es la que le corresponde al ensemble), es la distribución equiprobable (ver ecuación 3):

$$p_r = \frac{1}{\Omega} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \frac{1}{\Omega} \quad (25)$$

Luego, como no hay vínculos más allá del vínculo trivial, no tiene sentido plantear las ecuaciones 4 ni 6.

Sí tiene sentido plantear la ecuación 5:

$$S_{\text{termo}}^{\text{microcanónico}} = k_B S_{\text{est}}^{\text{max}} = k_B (\ln Z_{\text{microcanónico}} + \sum_{k=1}^n \lambda_k a_k) = k_B \ln \Omega \quad (26)$$

Puesto que

$$dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} dN$$

podemos deducir la temperatura  $T^{-1} = \left. \frac{\partial S}{\partial U} \right|_{V,N}$ , la presión  $p = T \left. \frac{\partial S}{\partial V} \right|_{U,N}$  y el

potencial químico  $\mu = -T \left. \frac{\partial S}{\partial N} \right|_{U,V}$  a partir de derivadas primeras de la entropía  $S(U, V, N) = k_B \ln \Omega$ .

### 3.2. Ensemble canónico

Consideremos un **ensemble**, llamado **canónico**, formado por un espacio muestral  $M = \{(V, N) = \text{ctes}\}$  que consiste en todos los estados termodinámicos que posean  $V = \text{cte}$  y  $N = \text{cte}$ ; que **posea sólo el vínculo**  $\langle E \rangle = U$ , además del vínculo trivial  $\sum_{r \in M} p_r = 1$ .

Entonces, la función del partición (ver ecuación 2) del ensemble canónico es igual a:

$$Q \equiv Z_{\text{canónico}} = \sum_{r \in M} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \sum_{r \in M} e^{-\beta E_r} \quad (27)$$

En consecuencia, la distribución de probabilidad que maximiza la entropía en éste ensemble (que es la que le corresponde al ensemble), es la siguiente (ver ecuación 3):

$$p_r = \frac{1}{Q} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \frac{1}{Q} e^{-\beta E_r} \quad (28)$$

Si planteamos la ecuación de vínculo 4, se obtiene:

$$U = - \frac{\partial (\ln Q(\beta))}{\partial \beta} \quad (29)$$

Sí planteamos la ecuación 5, se obtiene:

$$S_{\text{termo}}^{\text{canónico}} = k_B S_{\text{est}}^{\text{max}} = k_B (\ln Q + \beta U) \quad (30)$$

Si planteamos la ecuación 6, obtenemos:

$$\beta = \frac{\partial S_{\text{est}}^{\text{max}}}{\partial U} = \frac{1}{k_B} \frac{\partial S_{\text{termo}}^{\text{canónico}}}{\partial U} = \frac{1}{k_B T} \quad (31)$$

donde se utilizó el primer principio  $dS = \frac{1}{T}dU + \frac{p}{T}dV - \frac{\mu}{T}d\bar{N}$ , y luego  $\left. \frac{\partial S}{\partial U} \right|_{V, \bar{N}} = \frac{1}{T}$ . Y se utilizó que  $(V, N) = \text{ctes}$  ya que en el espacio muestral  $M$  ocurre que  $(V, N) = \text{ctes}$ .

Además, si multiplicamos a  $S_{\text{termo}}^{\text{canónico}}$  por  $T$  en la anteúltima ecuación y despejamos  $-k_B T \ln Q$ , obtenemos (considerando  $k_B T \beta = 1$ ) la energía libre de Helmholtz:

$$-k_B T \ln Q = U - T S_{\text{termo}} = F \quad (32)$$

Es decir, **la energía libre de Helmholtz  $F$  es el potencial termodinámico** determinado por la elección **del ensamble canónico**.

Por último, si empezamos a calcular derivadas segundas, considerando la ecuación 7:

$$\sigma_E^2 = \frac{\partial^2 (\ln Q)}{\partial \beta^2} = -\frac{\partial U}{\partial \beta} = -\frac{\partial T}{\partial \beta} \frac{\partial U}{\partial T} = \frac{1}{k_B \beta^2} C_V = k_B T^2 C_V \quad (33)$$

donde se utilizó que  $U = -\frac{\partial (\ln Q)}{\partial \beta}$ , que  $T(\beta) = (k_B \beta)^{-1}$ , y que la capacidad calorífica a volumen constante es  $C_V = \left. \frac{\partial U}{\partial T} \right|_{V, N} = \frac{\partial U}{\partial T}$  ya que en el ensamble canónico,  $V$  y  $N$  ya son de por sí constantes en el espacio muestral  $M$ .

Además, si calculamos el error relativo al cuadrado de  $E$ , obtenemos:

$$\frac{\sigma_E^2}{U^2} = k_B T^2 \frac{C_V}{U^2} \propto \frac{1}{N} \rightarrow 0 \text{ si } N \rightarrow +\infty \quad (34)$$

donde se recordó que  $T$  es una variable intensiva, y que  $C_V$  y  $U$  son variables extensivas, y por lo tanto proporcionales al número de partículas  $N$ . Y se aplicó al final el límite termodinámico  $N \rightarrow +\infty$ .

Por lo tanto, en el límite termodinámico  $N \rightarrow +\infty$ , da igual considerar el ensamble microcanónico (con  $E = \text{cte}$ ) que el ensamble canónico (con  $\langle E \rangle = U = \text{cte}$ ): Ambos son equivalentes.

### 3.3. Ensemble gran canónico

Consideremos un **ensamble**, llamado **gran canónico**, formado por un espacio muestral  $M = \{V = \text{cte}\}$  que consiste en todos los estados termodinámicos que posean  $V = \text{cte}$ ; que **posean los vínculos**  $\langle E \rangle = U$  y  $\langle N \rangle = \bar{N}$ , además del vínculo trivial  $\sum_{r \in M} p_r = 1$ .

Entonces, la función del partición (ver ecuación 2) del ensemble gran canónico es igual a:

$$Z_{GC} \equiv Z_{\text{gran canónico}} = \sum_{r \in M} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \sum_{r \in M} e^{-\beta E_r - \gamma N_r} = \sum_{r \in M} z^{N_r} e^{-\beta E_r} \quad (35)$$

donde se utilizó la definición de la fugacidad  $z \equiv e^{-\gamma}$ , que puede reemplazar al multiplicador de Lagrange  $\gamma$  asociado al vínculo  $\langle N \rangle = \bar{N}$ .

En consecuencia, la distribución de probabilidad que maximiza la entropía en éste ensemble (que es la que le corresponde al ensemble), es la siguiente (ver ecuación 3):

$$p_r = \frac{1}{Z_{GC}} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \frac{1}{Z_{GC}} e^{-\beta E_r - \gamma N_r} \quad (36)$$

Si planteamos la ecuación de vínculo 4, se obtiene:

$$U = - \left. \frac{\partial (\ln Z_{GC}(\beta, \gamma))}{\partial \beta} \right|_{\gamma} \quad (37)$$

equivalentemente, la derivada puede ser tomada a  $z$  constante.

y también se obtiene

$$\bar{N} = - \left. \frac{\partial (\ln Z_{GC}(\beta, \gamma))}{\partial \gamma} \right|_{\beta} = - \left. \frac{\partial z}{\partial \gamma} \frac{\partial (\ln Z_{GC}(\beta, z))}{\partial z} \right|_{\beta} = z \left. \frac{\partial (\ln Z_{GC}(\beta, z))}{\partial z} \right|_{\beta} \quad (38)$$

donde se expresó la derivada en términos de  $\gamma$  y de la fugacidad  $z$ .

Sí planteamos la ecuación 5, se obtiene:

$$S_{\text{termo}}^{GC} = k_B S_{\text{est}}^{\text{max}} = k_B (\ln Z_{GC} + \beta U + \gamma \bar{N}) \quad (39)$$

Si planteamos la ecuación 6, obtenemos:

$$\beta = \left. \frac{\partial S_{\text{est}}^{\text{max}}}{\partial U} \right|_{\bar{N}} = \frac{1}{k_B} \left. \frac{\partial S_{\text{termo}}^{GC}}{\partial U} \right|_{\bar{N}} = \frac{1}{k_B T} \quad (40)$$

y además

$$\gamma = \left. \frac{\partial S_{\text{est}}^{\text{max}}}{\partial \bar{N}} \right|_U = \frac{1}{k_B} \left. \frac{\partial S_{\text{termo}}^{GC}}{\partial \bar{N}} \right|_U = - \frac{\mu}{k_B T} = -\beta \mu \quad (41)$$

donde en las dos anteriores ecuaciones se utilizó el primer principio  $dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} d\bar{N}$ , y luego  $\left. \frac{\partial S}{\partial U} \right|_{V, \bar{N}} = \frac{1}{T}$  y  $\left. \frac{\partial S}{\partial \bar{N}} \right|_{U, V} = -\frac{\mu}{T}$ . Y se utilizó que en ambas ocurre que  $V = \text{cte}$  ya que en el espacio muestral  $M$  ocurre que  $V = \text{cte}$ .



Además, si multiplicamos a  $S_{\text{termo}}^{GC}$  por  $T$  en la ecuación 39 y despejamos  $-k_B T \ln Z_{GC}$ , obtenemos (considerando  $k_B T \beta = 1$  y  $k_B T \gamma = -\mu$ ) el gran potencial:

$$-k_B T \ln Z_{GC} = U - TS_{\text{termo}} - \mu \bar{N} = \Omega = -pV \quad (42)$$

Es decir, **el gran potencial  $\Omega$  es el potencial termodinámico** determinado por la elección **del ensamble gran canónico**.

Respecto a la última igualdad, si consideramos que  $U(S, V, N)$  es una función homogénea de primer orden  $U(bS, bV, bN) = bU(S, V, N) \forall b \in \mathbb{R}$  (ver ejercicio 5 de la guía 1), entonces es válido que  $dU = TdS - pdV + \mu dN \implies U = TS - pV + \mu N$ , y luego:  $\Omega = U - TS - \mu N = -pV$ . O equivalentemente:  $pV = k_B T \ln Z_{GC}$ .

Por último, si empezamos a calcular derivadas segundas, considerando la ecuación 7, podemos obtener nuevamente  $\sigma_E^2/U^2 = k_B T^2 C_V/U^2 \propto 1/\bar{N} \rightarrow 0$  si  $\bar{N} \rightarrow \infty$ . Y también podemos obtener:

$$\sigma_N^2 = \left. \frac{\partial^2(\ln Z_{GC})}{\partial \gamma^2} \right|_{\beta} = - \left. \frac{\partial \bar{N}}{\partial \gamma} \right|_T = - \left. \frac{\partial \mu}{\partial \gamma} \right|_T \left. \frac{\partial \bar{N}}{\partial \mu} \right|_T = -k_B T V \left. \frac{\partial(1/v)}{\partial \mu} \right|_T = k_B T \frac{V}{v^2} \left. \frac{\partial v}{\partial \mu} \right|_T = k_B T \frac{\bar{N}}{v} \left. \frac{\partial v}{\partial \mu} \right|_T$$

donde se utilizó que  $\bar{N} = - \left. \frac{\partial(\ln Z_{GC})}{\partial \gamma} \right|_{\beta}$ , que  $\beta = \text{cte} \iff T = \text{cte}$ , que  $\gamma = -\mu/(k_B T)$ , que  $V = v\bar{N}$  y que en el ensamble gran canónico,  $V$  ya es de por sí constante en el espacio muestral  $M$ .

Además, si calculamos el error relativo al cuadrado de  $N$ , obtenemos:

$$\frac{\sigma_N^2}{\bar{N}^2} = k_B T \frac{1}{V} \left. \frac{\partial v}{\partial \mu} \right|_T \propto \frac{1}{\bar{N}} \rightarrow 0 \text{ si } \bar{N} \rightarrow +\infty \quad (43)$$

donde se recordó que  $T, v, \mu$  son variables intensivas, y que  $V = v\bar{N}$  es una variable extensiva. Y se aplicó al final el límite termodinámico  $\bar{N} \rightarrow +\infty$ .

Por lo tanto, **en el límite termodinámico  $\bar{N} \rightarrow +\infty$ , da igual considerar el ensamble canónico (con  $N = \text{cte}$ ) que el ensamble gran canónico (con  $\langle N \rangle = \bar{N} = \text{cte}$ ):** Ambos son equivalentes entre sí, **y a su vez son equivalentes al ensamble microcanónico.**

### 3.4. Ensamble isobárico

Consideremos un **ensamble**, llamado **isobárico**, formado por un espacio muestral  $M = \{N = \text{cte}\}$  que consiste en todos los estados termodinámicos que posean  $N = \text{cte}$ ; que **posean los vínculos  $\langle E \rangle = U$  y  $\langle V \rangle = \bar{V}$** , además del vínculo trivial  $\sum_{r \in M} p_r = 1$ . Entonces, la función del partición (ver ecuación 2) del ensamble isobárico es igual a:

$$Z_{IB} \equiv Z_{\text{isobárico}} = \sum_{r \in M} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \sum_{r \in M} e^{-\beta E_r - \delta V_r} \quad (44)$$

En consecuencia, la distribución de probabilidad que maximiza la entropía en éste ensemble (que es la que le corresponde al ensemble), es la siguiente (ver ecuación 3):

$$p_r = \frac{1}{Z_{IB}} e^{-\sum_{k=1}^n \lambda_k A_{k,r}} = \frac{1}{Z_{IB}} e^{-\beta E_r - \delta V_r} \quad (45)$$

Si planteamos la ecuación de vínculo 4, se obtiene:

$$U = - \left. \frac{\partial(\ln Z_{IB}(\beta, \delta))}{\partial \beta} \right|_{\delta} \quad (46)$$

y también se obtiene

$$\bar{V} = - \left. \frac{\partial(\ln Z_{IB}(\beta, \delta))}{\partial \delta} \right|_{\beta} \quad (47)$$

Sí planteamos la ecuación 5, se obtiene:

$$S_{\text{termo}}^{IB} = k_B S_{\text{est}}^{\text{max}} = k_B (\ln Z_{IB} + \beta U + \delta \bar{V}) \quad (48)$$

Si planteamos la ecuación 6, obtenemos:

$$\beta = \left. \frac{\partial S_{\text{est}}^{\text{max}}}{\partial U} \right|_{\bar{V}} = \frac{1}{k_B} \left. \frac{\partial S_{\text{termo}}^{IB}}{\partial U} \right|_{\bar{V}} = \frac{1}{k_B T} \quad (49)$$

y además

$$\delta = \left. \frac{\partial S_{\text{est}}^{\text{max}}}{\partial \bar{V}} \right|_U = \frac{1}{k_B} \left. \frac{\partial S_{\text{termo}}^{IB}}{\partial \bar{V}} \right|_U = \frac{p}{k_B T} = \beta p \quad (50)$$

donde en las dos anteriores ecuaciones se utilizó el primer principio  $dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} d\bar{N}$ , y luego  $\left. \frac{\partial S}{\partial U} \right|_{V, \bar{N}} = \frac{1}{T}$  y  $\left. \frac{\partial S}{\partial \bar{V}} \right|_{U, N} = \frac{p}{T}$ . Y se utilizó que en ambas ocurre que  $N = \text{cte}$  ya que en el espacio muestral  $M$  ocurre que  $N = \text{cte}$ .

Además, si multiplicamos a  $S_{\text{termo}}^{IB}$  por  $T$  en la ecuación 48 y despejamos  $-k_B T \ln Z_{IB}$ , obtenemos (considerando  $k_B T \beta = 1$  y  $k_B T \delta = p$ ) la energía libre de Gibbs:

$$-k_B T \ln Z_{IB} = U - T S_{\text{termo}} + p \bar{V} = G = \mu N \quad (51)$$

Es decir, **la energía libre de Gibbs  $G$  es el potencial termodinámico** determinado por la elección **del ensemble gran canónico**.

Respecto a la última igualdad, si consideramos que  $U(S, V, N)$  es una función homogénea de primer orden  $U(bS, bV, bN) = bU(S, V, N) \forall b \in \mathbb{R}$  (ver ejercicio 5 de la guía 1), entonces es válido que  $dU = TdS - pdV + \mu dN \implies U = TS - pV + \mu N$ , y luego:  $G = U - TS + p\bar{V} = \mu N$ .

Por último, si empezamos a calcular derivadas segundas, considerando la ecuación 7, podemos obtener nuevamente  $\sigma_E^2/U^2 = k_B T^2 C_V/U^2 \propto 1/\bar{N} \rightarrow 0$  si  $\bar{N} \rightarrow \infty$ . Y también podemos obtener:

$$\sigma_V^2 = \frac{\partial^2(\ln Z_{IB})}{\partial \delta^2} \Big|_{\beta} = - \frac{\partial \bar{V}}{\partial \delta} \Big|_T = - \frac{\partial p}{\partial \delta} \Big|_T \frac{\partial \bar{V}}{\partial p} \Big|_T = -k_B T N \frac{\partial v}{\partial p} \Big|_T = k_B T \bar{V} \kappa_T$$

donde se utilizó que  $\bar{V} = - \frac{\partial(\ln Z_{IB})}{\partial \delta} \Big|_{\beta}$ , que  $\beta = \text{cte} \iff T = \text{cte}$ , que  $\delta = p/(k_B T)$ , que  $V = v\bar{N}$  y que en el ensamble gran canónico,  $N$  ya es de por sí constante en el espacio muestral  $M$ . También se utilizó la definición de la compresibilidad isotérmica  $\kappa_T \equiv - \frac{1}{\bar{V}} \frac{\partial \bar{V}}{\partial p} \Big|_T$ , que es una cantidad intensiva.

Además, si calculamos el error relativo al cuadrado de  $V$ , obtenemos:

$$\frac{\sigma_V^2}{\bar{V}^2} = k_B T \frac{1}{\bar{V}} \kappa_T \propto \frac{1}{N} \rightarrow 0 \text{ si } N \rightarrow +\infty \quad (52)$$

donde se recordó que  $T, \kappa_T$  son cantidades intensivas, y que  $\bar{V} = N\bar{v}$  es una variable extensiva. Y se aplicó al final el límite termodinámico  $N \rightarrow +\infty$ .

Por lo tanto, **en el límite termodinámico  $N \rightarrow +\infty$ , da igual considerar el ensamble canónico (con  $V = \text{cte}$ ) que el ensamble isobárico (con  $\langle V \rangle = \bar{V} = \text{cte}$ ):** Ambos son equivalentes entre sí, y **a su vez son equivalentes al ensamble microcanónico y al gran canónico.**

Un último comentario: A veces se tienen sistemas donde el trabajo externo es de la forma  $Yd\bar{X} = F_T dL$  (donde  $F_T$  es la tensión y  $L$  es la longitud), en vez de ser de la forma  $Yd\bar{X} = -pdV$ . Entonces tendríamos el análogo de un ensamble isobárico pero correspondiente al trabajo  $F_T dL$  en vez de al trabajo  $-pdV$  (con espacio muestral  $M = \{N = \text{cte}\}$  y vínculos no triviales  $\langle E \rangle = U$  y  $\langle L \rangle = \bar{L}$ ), si en todas las expresiones de ésta subsección reemplazamos  $V \rightarrow L$  y  $p \rightarrow -F_T$ .

## 4. Trucos importantes y aplicaciones

### 4.1. Indistinguibilidad y conteo correcto de Boltzmann

Notar que, en caso de que se tenga un sistema formado por  $N$  partículas distinguibles, entonces en el espacio muestral  $M$  dos estados que sólo difieren en una permutación de partículas son distinguibles entre sí: No son equivalentes.

En cambio, si se tiene un sistema formado por  $N$  partículas indistinguibles, entonces en el espacio muestral  $M$  dos estados que sólo difieren en una permutación de partículas son indistinguibles entre sí: Son equivalentes. Entonces es necesario considerar un nuevo espacio muestral en el cual los estados que sólo difieran en una permutación de partículas son el mismo

estado:  $M_I$ , donde los estados que difieren entre sí sólo en una permutación de partículas sólo se cuentan una vez en  $M_I$  (sólo se corresponden con un elemento  $r_I \in M_I$ ). No es difícil ver que la cantidad de estados equivalentes a cualquier estado  $r \in M$  es igual a la cantidad de permutaciones de  $N$  partículas: Es igual a  $N!$ . Para visualizar ésto, notemos que en caso de que  $M$  sea finito:  $\Omega_I = \#(M_I) = \#(M)/N! = \Omega/N!$ .

Entonces, en caso de que  $M$  sea discreto:

$$\sum_{r_I \in M_I} = \frac{1}{N!} \sum_{r \in M} \quad (53)$$

y en caso de que  $M$  sea continuo también ocurre lo mismo:

$$\int_{M_I} d^L R = \frac{1}{N!} \int_M d^L R \quad (54)$$

En definitiva, vemos que **plantear que las  $N$  partículas son indistinguibles** y plantear un espacio muestral  $M_I$ , es equivalente a **considerar el espacio muestral original  $M$  y añadir un factor  $\frac{1}{N!}$  siempre que se sume (o integre) sobre los posibles estados**. Añadir dicho factor es lo que se denomina como **conteo correcto de Boltzmann**.

En particular, vemos:

$$Z_I = \frac{1}{N!} Z_D \quad (55)$$

## Aplicaciones particulares: Partículas clásicas en el espacio de fases

Consideremos que se tiene un **sistema de  $N$  partículas clásicas cuyo estado microscópico a un cierto tiempo  $t$  está dado por la posición  $(q, p) \equiv (q^{(3N)}, p^{(3N)}) = (\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N)$  en el espacio de fases de dimensión  $6N$  (donde  $(\vec{q}_i, \vec{p}_i)$  es la posición en el espacio de fases de dimensión 6 de la partícula clásica  $i$ )**. Equivalentemente podemos describir el estado microscópico a un cierto tiempo  $t$  como las  $N$  posiciones  $(\vec{q}_i, \vec{p}_i)$  de cada una de las  $N$  partículas en cada espacio de fases de dimensión 6 de la partícula clásica  $i$ . Por supuesto, los estados microscópicos en cuestión deben ser compatibles con los vínculos macroscópicos, por lo tanto en verdad sólo se consideran los puntos  $(q^{(3N)}, p^{(3N)})$  de una superficie  $M$  del espacio de fases de dimensión  $6N$  formada por aquellos estados que son compatibles con los vínculos.

Puesto que se debe adimensionalizar el volumen  $d^{3N} q d^{3N} p$ , se puede utilizar para ello el volumen mínimo que debe tener un microestado según mecánica cuántica: Como  $\delta q_i \delta p_i \sim h$  (tomamos  $h$  en vez de  $\hbar$  y en vez de  $\hbar/2$ ), luego tomamos el volumen  $\omega_0 \equiv h^{3N}$  para normalizar el elemento de volumen (ver ecuación 10). Es decir, **adimensionalizamos el volumen**

$$\int_M d^{3N} q d^{3N} p \rightarrow \int_M \frac{d^{3N} q d^{3N} p}{h^{3N}}.$$

Y como las partículas clásicas son indistinguibles, debemos **añadir un factor  $\frac{1}{N!}$  siempre que se integre sobre los posibles estados**. Luego, reemplazamos en todos lados (desde la ecuación

1 hasta la ecuación 9)

$$\sum_{r \in M} \rightarrow \frac{1}{N! h^{3N}} \int_M d^{3N} q d^{3N} p \quad (56)$$

y considerar que si se parte de una densidad  $\rho(q, p)$  (no necesariamente normalizada) entonces debemos considerar en verdad la distribución de probabilidad  $f(q, p) = \rho_N(q, p) h^{3N} = \frac{\rho(q, p)}{\int_M \rho(q, p) d^{3N} q d^{3N} p} N! h^{3N} = \frac{\rho(q, p)}{\int_M \rho(q, p) d^{3N} q d^{3N} p / (N! h^{3N})}$  normalizada y adimensionalizada.

## 4.2. Ensamblen en términos de otros ensamblen más chicos

Supongamos que queremos calcular la función de partición  $Z_n$  de un cierto ensamble con  $n$  vínculos  $\langle A_k \rangle = a_k$  para  $1 \leq k \leq n$ , siendo  $n \in \mathbb{N}$  (notar que consideramos  $n > 0$ ). Y supongamos que conocemos la función de partición  $Z_{n-1}(C)$  de un ensamble con  $n-1$  vínculos  $\langle A_k \rangle = a_k$  para  $1 \leq k \leq n-1$  y un espacio muestral  $M_C$  en el cual la cantidad  $A_n = C = \text{cte}$  es constante y vale  $C$ , es decir un ensamble que posee un vínculo menos que el anterior porque hay una magnitud que ahora está fija y antes sólo su valor medio estaba fijo.

Suponiendo que todas las magnitudes  $A_k$  toman valores discretos, y luego que  $M$  es discreto:

$$\begin{aligned} Z_n &= \sum_{r \in M} \exp \left( - \sum_{k=1}^n \lambda_k A_{k,r} \right) = \sum_{A_n} \sum_{r' \in M_{A_n}} \exp \left( - \sum_{k=1}^n \lambda_k A_k(\vec{R}) \right) \\ &= \sum_{A_n} e^{-\lambda_n A_n} \sum_{r' \in M_{A_n}} \exp \left( - \sum_{k=1}^{n-1} \lambda_k A_k(\vec{R}) \right) = \sum_{A_n} e^{-\lambda_n A_n} Z_{n-1}(A_n) \end{aligned}$$

En definitiva, obtenemos:

$$Z_n = \sum_{A_n} e^{-\lambda_n A_n} Z_{n-1}(A_n) \quad (57)$$

No es difícil ver que ésto último es válido si la magnitud  $A_n$  toma valores discretos, sin importar si el resto de las magnitudes  $A_k$  con  $1 \leq k \leq n-1$  toman valores discretos o continuos.

### Aplicaciones particulares

- En particular, podemos aplicar ésto para obtener la función de partición del ensamble gran canónico en términos del ensamble canónico en el que  $N = \text{cte}$ :

$$Z_{GC} = \sum_{N=0}^{\infty} e^{-\gamma N} Q(N) = \sum_{N=0}^{\infty} z^N Q(N) \quad (58)$$

- También podemos aplicarlo para obtener la función de partición del ensamble canónico en términos del ensamble microcanónico en el que  $E = \text{cte}$ , considerando que hay niveles de energía discretos:

$$Q = \sum_E e^{-\beta E} \Omega(E) \quad (59)$$

### 4.3. Trucos y suposiciones para simplificar el cálculo de $Z$

Supongamos un ensamble en el cual el número de partículas (o celdas)  $N$  sea constante en  $M$ . Por ahora supondremos que las partículas/celdas son distinguibles pero después veremos que pasa si son indistinguibles. Supongamos que un microestado  $r \in M$  de todo el sistema (de  $N$  partículas/celdas) consiste en especificar el microestado de cada una de las  $N$  partículas/-celdas:  $r = (r_1, \dots, r_N)$  donde  $r_i \in M_i$  es el microestado de la partícula/celda  $i$ . Supongamos luego que las magnitudes  $A_k$  de las partículas/celdas son no interactuantes: Es decir, se puede expresar  $A_k(r) = A_k(r_1, \dots, r_N) = \sum_{i=1}^N A_k^{(i)}(r_i)$  donde  $A_k^{(i)}(r_i)$  es el valor de la magnitud para la partícula/celda  $i$  en el estado  $r_i$ . Entonces podemos expresar la función de partición  $Z_D$  (donde la  $D$  es para enfatizar que las  $N$  partículas/celdas son distinguibles) como:

$$\begin{aligned} Z_D &= \sum_{r \in M} \exp \left( - \sum_{k=1}^n \lambda_k A_k(r) \right) = \sum_{r_1, (\dots), r_N} \exp \left( - \sum_{k=1}^n \lambda_k A_k(r_1, \dots, r_N) \right) \\ &= \sum_{r_1, (\dots), r_N} \exp \left( - \sum_{k=1}^n \lambda_k \left( \sum_{i=1}^N A_k^{(i)}(r_i) \right) \right) = \sum_{r_1, (\dots), r_N} \prod_{i=1}^N \exp \left( - \sum_{k=1}^n \lambda_k A_k^{(i)}(r_i) \right) \end{aligned} \quad (60)$$

En el primer renglón se usó que  $r = (r_1, \dots, r_N)$ . Y en el segundo renglón se usó que las magnitudes  $A_k$  de las partículas/celdas son no interactuantes.

Si suponemos ahora que los microestados  $r_1, \dots, r_N$  son independientes entre sí y no poseen ningún vínculo entre ellos, entonces podemos separar la sumatoria en la anterior ecuación:

$$\begin{aligned} Z_D &= \sum_{r_1, (\dots), r_N} \prod_{i=1}^N \exp \left( - \sum_{k=1}^n \lambda_k A_k^{(i)}(r_i) \right) \\ &= \sum_{r_1 \in M_1} (\dots) \sum_{r_N \in M_N} \prod_{i=1}^N \exp \left( - \sum_{k=1}^n \lambda_k A_k^{(i)}(r_i) \right) \\ &= \prod_{i=1}^N \sum_{r_i \in M_i} \exp \left( - \sum_{k=1}^n \lambda_k A_k^{(i)}(r_i) \right) \\ &= \prod_{i=1}^N Z_i \end{aligned} \quad (61)$$

donde  $Z_i = \sum_{r_i \in M_i} \exp \left( - \sum_{k=1}^n \lambda_k A_k^{(i)}(r_i) \right)$  es la función de partición de la partícula  $i$ .

Y si las partículas/celdas son idénticas, tienen entonces el mismo espacio muestral, por lo tanto:  $Z_i = Z_1 \forall 1 \leq i \leq N$ . Entonces podemos expresar:

$$Z_D = Z_1^N \quad (62)$$

Y en caso de que las partículas/celdas sean indistinguibles, sólo hay que añadir un factor  $1/N!$  de acuerdo al conteo correcto de Boltzmann:

$$Z_I = \frac{1}{N!} Z_D = \frac{1}{N!} Z_1^N \quad (63)$$

En resumen, si realizamos las siguientes suposiciones en orden, en cuyo caso diremos que el ensamble es separable en las  $N$  partículas (o celdas):

1. El **número de partículas (o celdas)  $N$  es constante** en el espacio muestral  $M$ .
2. Un **microestado  $r \in M$  de todo el sistema** (de  $N$  partículas/celdas) **consiste en especificar el microestado de cada una de las  $N$  partículas/celdas:**  $r = (r_1, \dots, r_N)$  donde  $r_i \in M_i$  es el microestado de la partícula/celda  $i$ .
3. **Las magnitudes  $A_k$  de las partículas/celdas son no interactuantes:** Es decir, se puede expresar  $A_k(r) = A_k(r_1, \dots, r_N) = \sum_{i=1}^N A_k^{(i)}(r_i)$  donde  $A_k^{(i)}(r_i)$  es el valor de la magnitud para la partícula/celda  $i$  en el estado  $r_i$ .
4. **Los microestados  $r_1, \dots, r_N$  son independientes entre sí y no poseen ningún vínculo entre ellos,** lo que nos permite separar la sumatoria  $\sum_{r_1, (\dots), r_N} = \sum_{r_1 \in M_1} (\dots) \sum_{r_N \in M_N}$
5. **Las partículas/celdas son idénticas,** tienen entonces el mismo espacio muestral, por lo tanto:  $Z_i = Z_1 \forall 1 \leq i \leq N$ .

Entonces, **después de realizar las anteriores suposiciones en orden, podemos expresar la función de partición del sistema de  $N$  partículas/celdas como:**

$$Z = \begin{cases} Z_D = Z_1^N & \text{si las } N \text{ partículas/celdas son distinguibles} \\ Z_I = \frac{1}{N!} Z_1^N & \text{si las } N \text{ partículas/celdas son indistinguibles} \end{cases} \quad (64)$$

siendo  $Z_1 = \sum_{r_1 \in M_1} \exp\left(-\sum_{k=1}^n \lambda_k A_k^{(1)}(r_1)\right)$  la función de partición de una de las  $N$  partículas/celdas.

### Aplicaciones particulares

- **Si las suposiciones 1 a 5 aplican para el ensamble canónico de un sistema con  $N$  partículas,** para lo cual debe ocurrir (entre otras cosas) que  $E(r) = E(r_1, \dots, r_N) = \sum_{i=1}^N E^{(i)}(r_i)$  donde  $E^{(i)}(r_i)$  es la energía que le corresponde a la partícula  $i$  en el estado  $i$  (es decir que no hay términos de interacción entre las distintas partículas), entonces podemos aplicar lo anterior y obtener **la función**

de partición en el ensamble canónico del sistema de  $N$  partículas como:

$$Q = \begin{cases} Q_D = Q_1^N & \text{si las } N \text{ partículas son distinguibles} \\ Q_I = \frac{1}{N!} Q_1^N & \text{si las } N \text{ partículas son indistinguibles} \end{cases} \quad (65)$$

siendo  $Q_1 = \sum_{r_1 \in M_1} \exp(-\beta E^{(1)}(r_1))$  la función de partición de una de las  $N$  partículas.

- También, si las suposiciones 1 a 5 aplican para el ensamble isobárico de un sistema con  $N$  partículas, para lo cual debe ocurrir (entre otras cosas) que el volumen (además de la energía) también se puede expresar como  $V(r) = V(r_1, \dots, r_N) = \sum_{i=1}^N V^{(i)}(r_i)$ , entonces podemos aplicar lo anterior y obtener la función de partición en el ensamble isobárico del sistema de  $N$  partículas como:

$$Z_{IB} = \begin{cases} Z_{IB;D} = (Z_{IB;1})^N & \text{si las } N \text{ partículas son distinguibles} \\ Z_{IB;I} = \frac{1}{N!} (Z_{IB;1})^N & \text{si las } N \text{ partículas son indistinguibles} \end{cases} \quad (66)$$

siendo  $Z_{IB;1} = \sum_{r_1 \in M_1} \exp(-\beta E^{(1)}(r_1) - \delta V^{(1)}(r_1))$  la función de partición de una de las  $N$  partículas.

- Por último, consideremos que las suposiciones 1 a 5 aplican para el ensamble gran canónico de un sistema con  $M$  celdas, donde cada celda  $i$  posee un estado  $r_i$  descrito por un cierto número de partículas  $N^{(i)}(r_i)$  y una cierta energía  $E^{(i)}(r_i)$ . Para ello, debe ocurrir (entre otras cosas) que el número de partículas y la energía de todas las  $M$  celdas se pueden expresar como  $E(r) = E(r_1, \dots, r_M) = \sum_{i=1}^M E^{(i)}(r_i)$  y  $N(r) = N(r_1, \dots, r_M) = \sum_{i=1}^M N_i(r_i)$ . Entonces podemos aplicar lo anterior y obtener la función de partición en el ensamble gran canónico del sistema de  $M$  celdas como:

$$Z_{GC} = \begin{cases} Z_{GC;D} = (Z_{GC;1})^M & \text{si las } M \text{ celdas son distinguibles} \\ Z_{GC;I} = \frac{1}{M!} (Z_{GC;1})^M & \text{si las } M \text{ celdas son indistinguibles} \end{cases} \quad (67)$$

siendo  $Z_{GC;1} = \sum_{r_1 \in M_1} \exp(-\beta E^{(1)}(r_1) - \gamma N^{(1)}(r_1))$  la función de partición de una de las  $M$  celdas.



#### 4.4. Truco para un ensamble con uno de sus vínculos: $\langle N \rangle = \bar{N}$

Ahora planteemos una situación distinta: Supongamos un ensamble en el cual el número de partículas (o celdas)  $N$  no es constante en  $M$  (en principio), pero uno de los vínculos es que su valor medio es constante:  $\langle N \rangle = \bar{N}$ , y los otros vínculos son  $\langle A_k \rangle = a_k$  para  $1 \leq k \leq n-1$ . Entonces, tomando  $A_n = N$ , se puede aplicar lo razonado en la anterior subsección puesto que la magnitud  $A_n = N$  toma valores discretos  $N \in \mathbb{N}_0$ :

$$Z = \sum_{N=0}^{\infty} e^{-\gamma N} Z_{n-1}(N) = \sum_{N=0}^{\infty} z^N Z_{n-1}(N) \quad (68)$$

donde  $Z_{n-1}(N)$  es la función de partición que corresponde al ensamble con  $n-1$  vínculos  $\langle A_k \rangle = a_k$  para  $1 \leq k \leq n-1$  y un espacio muestral  $M_N$  en el cual la cantidad  $A_n = N = \text{cte}$  es constante y vale  $N$ , es decir un ensamble que posee un vínculo menos que el anterior (que es precisamente el número de partículas/celdas) porque hay una magnitud que ahora está fija y antes sólo su valor medio estaba fijo.

Si asumimos ahora que las anteriores suposiciones (1 a 5) se verifican para el ensamble con  $n-1$  vínculos y con  $N$  fijo (luego la suposición 1 es válida siempre), entonces podemos utilizar lo razonado anteriormente para expresar:

$$Z = \sum_{N=0}^{\infty} z^N Z_{n-1}(N) = \begin{cases} \sum_{N=0}^{\infty} z^N Z_{n-1;1}^N & \text{si las } N \text{ partículas/celdas son distinguibles} \\ \sum_{N=0}^{\infty} z^N \frac{1}{N!} Z_{n-1;1}^N & \text{si las } N \text{ partículas/celdas son indistinguibles} \end{cases}$$

Recordando que  $\sum_{N=0}^{\infty} x^N = \frac{1}{1-x}$  si  $|x| < 1$  (sino, lo anterior diverge), y que  $\sum_{N=0}^{\infty} \frac{1}{N!} x^N = e^x$ , entonces podemos expresar finalmente:

$$Z = \sum_{N=0}^{\infty} z^N Z_{n-1}(N) = \begin{cases} \frac{1}{1 - z Z_{n-1;1}} & \text{si las } N \text{ partículas/celdas son distinguibles y si } z Z_{n-1;1} < 1 \\ e^{z Z_{n-1;1}} & \text{si las } N \text{ partículas/celdas son indistinguibles} \end{cases} \quad (69)$$

#### Aplicaciones particulares

Podemos **aplicar ésto** sin mucha dificultad **al ensamble gran canónico de un sistema termodinámico**, considerando la ecuación 58:

$$Z_{GC} = \sum_{N=0}^{\infty} e^{-\gamma N} Q(N) = \sum_{N=0}^{\infty} z^N Q(N)$$

Y considerando que **podemos aplicar las suposiciones 1 a 5 al ensamble canónico del mismo sistema termodinámico con  $N$  partículas** (es decir, que el ensamble canónico es

separable), **obtenemos**:

$$Z_{GC} = \sum_{N=0}^{\infty} z^N Q(N) = \begin{cases} \frac{1}{1 - zQ_1} & \text{si las } N \text{ partículas/celdas son distinguibles y si } zQ_1 < 1 \\ e^{zQ_1} & \text{si las } N \text{ partículas/celdas son indistinguibles} \end{cases} \quad (70)$$

Notar en particular que  $Q_1 = \sum_{r_1 \in M_1} e^{-\beta E_{r_1}^{(1)}} \geq e^{-\beta \epsilon_0}$ , donde  $\epsilon_0$  es la menor energía posible de una partícula (nivel fundamental de energía).

Si  $\epsilon_0 = 0$ , entonces  $Q_1 \geq 1 \iff 0 < (Q_1)^{-1} \leq 1$ , y por lo tanto pedir  $zQ_1 < 1$  es equivalente a pedir  $z < (Q_1)^{-1}$ . Ésto implica luego que  $z < (Q_1)^{-1} \leq 1 \implies z < 1$ .