

Termodinámica y Mecánica Estadística II

Apuntes de clase 2021 - Gustavo Castellano

Introducción

Para formular la teoría termodinámica aceptamos las variables macroscópicas como representativas de valores medios de diferentes magnitudes físicas (energía, magnetización, número de partículas, etc.). La termodinámica estudia las relaciones entre esos valores medios, pero no se inmiscuye con la física microscópica.

En este curso intentaremos deducir las propiedades macroscópicas a partir del conocimiento (escaso a veces) que tengamos de la física microscópica. Para ello tendremos que manipular grandes números de partículas, por lo que necesitaremos dominar el cálculo de valores promedios de manera eficiente, contando adecuadamente, para los sistemas analizados, las posibles configuraciones asociadas por ejemplo con una dada energía. Esto nos obliga entonces a revisar (o enterarnos de) algunas nociones sobre probabilidad.

1. Probabilidad (basado en el texto de Reichl)

1.1. Permutaciones y combinaciones

Cuando trabajamos con muchos objetos, estos conceptos aparecen frecuentemente. Una *permutación* es un arreglo de un conjunto de N objetos diferentes en un orden definido. El número de permutaciones diferentes de estos N objetos es $N!$; esto se ve fácilmente si pensamos que para la primera alternativa disponemos de los N elementos del conjunto, cada uno de los cuales puede complementarse con los $(N - 1)$ restantes como segunda opción, y así hasta llegar a la última elección, conformando el producto $N \cdot (N - 1) \cdot \dots \cdot 1 = N!$.

El número de permutaciones posibles al tomar R objetos del conjunto de N elementos será, siguiendo el mismo razonamiento,

$$N \cdot (N - 1) \cdot \dots \cdot (N - R + 1) = \frac{N!}{(N - R)!}.$$

Conviene enfatizar que también en este caso distinguimos subconjuntos que hayan sido escogidos en diferente orden. Una *combinación* C_R^N es una selección de R objetos sin importar el orden en que se escojan:

$$C_R^N = \frac{N!}{(N - R)! R!} \equiv \binom{N}{R}.$$

El factor $R!$ del denominador descuenta aquellas configuraciones que tienen los mismos elementos y sólo difieren en su ordenamiento.

Si un conjunto de N elementos contiene n_1 elementos idénticos de tipo 1, n_2 de tipo 2, \dots , n_k de tipo k , puede verse que el número de permutaciones posibles será

$$\frac{N!}{n_1! n_2! \dots n_k!} \quad \left(\sum_i n_i = N \right).$$

1.2. Definición de probabilidad

Aceptamos como *probabilidad* nuestra expectativa respecto del resultado de un experimento o “evento”. Si un posible resultado de cierto experimento es A , su probabilidad es $P(A)$ si al realizar N experimentos idénticos esperamos que el resultado A se obtenga en $N P(A)$ de ellos. Dicho de otra manera, a medida que $N \rightarrow \infty$, esperamos que la fracción de eventos con resultado A tienda a $P(A)$.

Un ejemplo sencillo pero que encontraremos muy frecuentemente es el de n resultados igualmente probables de los cuales m corresponden al caso A que nos interesa: en ese caso, $P(A) = m/n$.

Definimos *espacio muestral* asociado a un determinado experimento como el conjunto S constituido por todos los resultados posibles de ese experimento. Un evento plausible entonces corresponderá a un subconjunto de S .

Utilizando el lenguaje de teoría de conjuntos, denotamos como $A \cup B$ a la *unión* de A y B , que es el conjunto de todos los puntos que pertenecen a A o a B (o a ambos). La *intersección* $A \cap B$ de estos conjuntos se define como el conjunto de todos los puntos que pertenecen a A y a B simultáneamente. Si $A \cap B = \emptyset$ (no contiene puntos), A y B son *mutuamente excluyentes*.

Algunas probabilidades que resultan evidentes, en esta notación se expresan como

$$P(\emptyset) = 0, \quad P(S) = 1.$$

$P(A \cap B)$ representará entonces la probabilidad de tener ambos resultados, y $P(A \cup B)$, la probabilidad de tener como resultado un evento A o B (o ambos). En particular,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

En el caso de que A y B sean mutuamente excluyentes tendremos $P(A \cup B) = P(A) + P(B)$.

Si A_1, A_2, \dots, A_m son mutuamente excluyentes y exhaustivos, es decir, cubren todo el espacio muestral y por lo tanto $A_1 \cup A_2 \cup \dots \cup A_m = S$, decimos que $\{A_j\}$ es una *partición* de S en m subconjuntos. Cuando $\{A_j\}$ es una partición se cumple que $P(A_1) + P(A_2) + \dots + P(A_m) = 1$.

Dos observaciones diferentes A y B (pueden ser de eventos similares o completamente distintas) son *eventos independientes* si y sólo si $P(A \cap B) = P(A)P(B)$. Aquí es importante destacar que en esta notación, $P(A)$ corresponde al resultado A permitiendo *cualquier evento* B , y análogamente para el resultado B .

Se define *probabilidad condicional* $P(B|A)$ como la probabilidad de obtener el resultado A dado que también se obtiene B :

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

En esta expresión el denominador de alguna manera sugiere una nueva normalización, como si redujéramos el espacio muestral al evento B . Como $P(A \cap B) = P(B \cap A)$, puede inferirse la relación

$$P(A)P(A|B) = P(B)P(B|A).$$

Evidentemente, en el caso de que A y B sean independientes, se cumple que $P(B|A) = P(B)$.

1.3. Variables aleatorias o estocásticas

Una cantidad cuyo valor se determina como resultado de un experimento se denomina *variable estocástica*. En un determinado experimento, una variable estocástica X puede tomar diferentes valores x_i ; debe tenerse cuidado de distinguir la variable X de los distintos resultados $\{x_i\}$ posibles.

Para visualizar mejor este concepto pensemos en el ejemplo de sortear tres monedas: el número de caras que se obtiene es una variable estocástica X ; los resultados posibles son $x_i = 0, 1, 2, 3$, cada uno con diferente probabilidad. Vemos claramente que *el número de caras* no tiene un valor prefijado, y para determinarlo es necesario realizar el “experimento”. A lo sumo podemos decir cuál será el resultado menos probable, o cuántas caras obtenemos en promedio al cabo de muchos experimentos.

Variables estocásticas discretas

Si X es una variable estocástica que puede tomar un conjunto numerable de valores $\{x_1, x_2, \dots\}$, podemos asignar a cada x_i una probabilidad $f(x_i)$. Esta *distribución de probabilidades* no puede tomar valores negativos, y debe cumplir la condición de normalización $\sum_i f(x_i) = 1$ (sumando todas las x_i posibles).

La distribución de probabilidades $f(x_i)$ constituye toda la información posible sobre la variable estocástica X . A menudo no puede determinarse $f(x_i)$ sino algunos de sus *momentos n -ésimos*, definidos como

$$\langle X^n \rangle = \sum_i x_i^n f(x_i).$$

Los momentos de X dan información acerca de la forma de $f(x_i)$. En general, los más importantes son los primeros de ellos. El primer momento $\langle X \rangle$ suele señalarse como *valor esperado* porque nos provee el *valor medio* de la distribución. A partir del segundo momento se define la *varianza* de X como

$$\text{var}(X) = \langle X^2 \rangle - \langle X \rangle^2 \quad (= \langle (X - \langle X \rangle)^2 \rangle)$$

y de aquí, la *desviación estándar*

$$\sigma_X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \quad \left(= \sqrt{\langle (X - \langle X \rangle)^2 \rangle} \right),$$

que da una idea del “ancho” de la distribución.

Variables estocásticas continuas

Cuando la variable X puede tomar valores continuos, la distribución debe ser una función continua a trozos, de modo que la probabilidad de tener un valor $a \leq X \leq b$ sea el área bajo la curva $f_X(x)$ en ese intervalo:

$$P(a \leq X \leq b) = \int_a^b dx f_X(x).$$

En este caso, $f_X(x)$ se denomina *densidad de probabilidad* de X , y como en el caso discreto debe cumplirse $f_X(x) \geq 0$ además de la condición de normalización

$$\int dx f_X(x) = 1,$$

donde la integración debe extenderse sobre todo el rango de valores posibles para X .

Análogamente al caso de variables discretas, la definición del n -ésimo momento en este caso es

$$\langle X^n \rangle = \int dx x^n f_X(x).$$

Obviamente, el primer momento $\langle X \rangle$ vuelve a ser el valor medio (o valor esperado), y la desviación estándar σ_X se define como antes a partir del primer y el segundo momentos.

El conocimiento de todos los momentos $\langle X^n \rangle$ implica el conocimiento *completo* de $f_X(x)$. Esto se ve fácilmente introduciendo la *función característica*

$$\phi_X(k) \equiv \langle e^{ikX} \rangle = \int dx e^{ikx} f_X(x) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle X^n \rangle}{n!}, \quad (1)$$

donde el último término tiene sentido como herramienta siempre que $\langle X^n \rangle$ sea pequeño para n grande. Vemos que la densidad de probabilidad $f_X(x)$ es la transformada de Fourier de $\phi_X(k)$:

$$f_X(x) = \frac{1}{2\pi} \int dk e^{-ikx} \phi_X(k).$$

Así hemos visto que, efectivamente, si conocemos todos los momentos $\langle X^n \rangle$ queda determinada la función característica, y a través de esta última identidad podemos encontrar la densidad de probabilidad $f_X(x)$. Del mismo modo, cuando conocemos $\phi_X(k)$, los momentos se obtienen directamente a partir de la relación (1):

$$\langle X^n \rangle = \frac{1}{i^n} \left. \frac{d^n \phi_X(k)}{dk^n} \right|_{k=0}.$$

Esta relación es muy útil cuando resulta más fácil encontrar los momentos realizando la transformada de Fourier de $f_X(x)$ y luego derivando en lugar de utilizar la definición.

Otra alternativa que puede ser interesante, dependiendo del problema encarado, involucra una expansión de la función característica $\phi_X(k)$ en término de los *cumulantes de orden n* $C_n(X)$, definidos mediante la relación

$$\phi_X(k) = \exp \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right].$$

A menudo necesitamos conocer la densidad de probabilidad en términos de la variable estocástica $Y = g(X)$. Como los valores medios de cualquier función deben coincidir al calcularlos en términos de X o Y , el entrenado lector puede mostrar que debe satisfacerse la relación

$$f_Y(y) = \int dx \delta(y - g(x)) f_X(x),$$

donde la delta de Dirac es la delta de Dirac.

Probabilidades conjuntas

Si X e Y son variables estocásticas sobre un espacio muestral S , con $X(S) = \{x_1, x_2, \dots\}$ e $Y(S) = \{y_1, y_2, \dots\}$, el conjunto $X(S) \times Y(S)$ estará conformado por los pares ordenados $\{(x_1, y_1), (x_1, y_2), \dots, (x_2, y_1), (x_2, y_2), \dots, (x_i, y_j), \dots\}$ y será un “espacio de probabilidades”, de modo que a cada par ordenado (x_i, y_j) asignamos una *probabilidad conjunta* $f(x_i, y_j) = P(X = x_i, Y = y_j)$.

Por supuesto, esa definición vale también para variables continuas. Teniendo la precaución de sustituir integrales por sumatorias cuando corresponda, escribiremos el desarrollo que sigue para el caso de variables continuas. Cuando tenemos la probabilidad conjunta $f(x, y)$, podemos obtener la distribución para una de las variables integrando todos los valores posibles para la otra:

$$f_X(x) = \int dy f(x, y) .$$

Como en los casos anteriores, debe cumplirse siempre que $f(x, y) \geq 0$, y la condición de normalización se escribe

$$\int \int dx dy f(x, y) = 1 ,$$

La interdependencia entre estas dos variables estocásticas se evalúa mediante la *covarianza* de X e Y :

$$\begin{aligned} \text{cov}(X, Y) &= \int \int dx dy (x - \langle X \rangle) (y - \langle Y \rangle) f(x, y) \\ &= \int \int dx dy xy f(x, y) - \langle X \rangle \langle Y \rangle \\ &= \langle XY \rangle - \langle X \rangle \langle Y \rangle . \end{aligned}$$

También suele usarse la *correlación* de X e Y , definida como

$$\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} ,$$

pues tiene la ventaja de ser adimensional y presenta además las siguientes propiedades:

- a) $\text{cor}(X, Y) = \text{cor}(Y, X)$;
- b) $-1 \leq \text{cor}(X, Y) \leq 1$;
- c) $\text{cor}(X, X) = 1$; $\text{cor}(X, -X) = -1$;
- d) $\text{cor}(aX + b, cY + d) = \text{cor}(X, Y)$ (siempre que $a, c \neq 0$).

Es interesante destacar que pares de variables con idéntica distribución de probabilidad pueden tener diferentes correlaciones.

En el caso en que X e Y son independientes, se satisface que

- a) $f(x, y) = f_X(x) f_Y(y)$;
- b) $\langle XY \rangle = \langle X \rangle \langle Y \rangle$;
- c) $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$;
- d) $\text{cov}(X, Y) = 0$.

Aquí también vale la pena notar que la última propiedad, $\text{cov}(X, Y) = 0$, no necesariamente implica que X e Y sean independientes.

A veces partimos de varias variables estocásticas y necesitamos información sobre una nueva variable estocástica Z , función de aquéllas. Por ejemplo, supongamos que contamos con la densidad de probabilidad conjunta $f(x, y)$ para las variables X e Y , que están conectadas con Z mediante la relación $Z = g(X, Y)$. Como señalamos para el caso de una variable, puede mostrarse que

$$f_Z(z) = \int \int dx dy \delta(z - g(x, y)) f(x, y) .$$

Cuando X e Y son independientes,

$$f_Z(z) = \int \int dx dy \delta(z - g(x, y)) f_X(x) f_Y(y) .$$

En este caso, la función característica será

$$\phi_Z(k) = \int \int dx dy e^{ikg(x, y)} f_X(x) f_Y(y) .$$

En particular, cuando $g(x, y) = x + y$ resulta

$$\phi_Z(k) = \phi_X(k) \phi_Y(k) .$$

1.4. Distribución binomial

Frecuentemente nos enfrentamos con el caso de muchos experimentos que dan como resultado sólo dos valores posibles. Denotamos entonces p y q a las probabilidades para cada resultado, que satisfacen la condición de normalización $p + q = 1$.

Al cabo de un número N de experimentos, tendremos respectivamente n_1 y n_2 resultados, que cumplen con la relación $n_1 + n_2 = N$. Si estamos interesados en tener n_1 experimentos con el primer resultado *en un orden determinado*, la correspondiente probabilidad será $p^{n_1} q^{n_2}$. Si en cambio no nos interesa el orden, la probabilidad de tener cualquier combinación n_1, n_2 está dada por la llamada *distribución binomial*

$$P_N(n_1) = \binom{N}{n_1} p^{n_1} q^{n_2} .$$

Es fácil verificar que esta distribución está normalizada, pues

$$\sum_{n_1=0}^N P_N(n_1) = \sum_{n_1=0}^N \binom{N}{n_1} p^{n_1} q^{N-n_1} = (p + q)^N = 1 .$$

Podemos evaluar el primer momento, teniendo en cuenta que

$$\langle n_1 \rangle = \sum_{n_1=0}^N n_1 \binom{N}{n_1} p^{n_1} q^{N-n_1} = p \frac{\partial}{\partial p} \left[\sum_{n_1=0}^N P_N(n_1) \right] ;$$

como el término entre corchetes es $(p + q)^N$, resulta

$$\langle n_1 \rangle = p N .$$

Se deja como divertido ejercicio para el lector mostrar que, análogamente, el segundo momento resulta

$$\langle n_1^2 \rangle = (N p)^2 + N p q .$$

Con estos dos momentos evaluamos la varianza o la desviación estándar

$$\sigma_N = \sqrt{N p q} .$$

Mucho más interesante resulta analizar la desviación estándar relativa

$$\frac{\sigma_N}{\langle n_1 \rangle} = \sqrt{\frac{q}{p}} \sqrt{\frac{1}{N}} .$$

Aquí vemos que este valor se reduce a medida que crece N ; un valor pequeño de $\sigma_N / \langle n_1 \rangle$ indica que cualquier resultado n_1 tomará valores muy próximos al valor medio $p N$. Dicho en otras palabras, a medida que N crece, $n_1 / N \simeq \langle n_1 \rangle / N = p$.

1.4.1. Distribución gaussiana

Cuando N es suficientemente grande y p no es demasiado pequeña, de manera que $p N$ también sea grande, podemos utilizar la fórmula de Stirling para aproximar

$$n! \cong \sqrt{2\pi n} \left(\frac{n}{e}\right)^n ,$$

que es válida para $n > 10$. Si tenemos en cuenta que N y n_1 son muy grandes, reemplazando en la distribución binomial, podemos escribir

$$P_N(n_1) = \frac{1}{\sqrt{2\pi N}} \exp \left[-n_1 \ln \frac{n_1}{N} - (N - n_1) \ln \frac{N - n_1}{N} + n_1 \ln p + (N - n_1) \ln(1 - p) \right] . \quad (2)$$

Es sencillo demostrar que P_N es máximo cuando n_1 vale $p N = \langle n_1 \rangle$. Podemos desarrollar el exponente de P_N alrededor de $\langle n_1 \rangle$, obteniendo

$$P_N(n_1) = P_N(\langle n_1 \rangle) e^{-\frac{\epsilon^2}{2N p q} + \mathcal{O}(\epsilon^3)} ,$$

donde $\epsilon \equiv n_1 - \langle n_1 \rangle$. Como para N grande ϵ es chico, puede mostrarse que es una buena aproximación descartar los términos de orden superior al segundo. Recordando que $N p q = \sigma_N^2$ y evaluando correctamente $P_N(\langle n_1 \rangle)$ en la expresión (2), se obtiene

$$P_N(n_1) = \frac{1}{\sqrt{2\pi} \sigma_N} e^{-\frac{(n_1 - \langle n_1 \rangle)^2}{2\sigma_N^2}} \quad (3)$$

Esta es la conocida *distribución gaussiana*, y tiene la particularidad de que queda absolutamente determinada mediante $\langle n_1 \rangle$ y $\langle n_1^2 \rangle$.

1.4.2. Distribución de Poisson

En algunos casos, al crecer N , $p \rightarrow 0$, de modo que $pN = a$ (cte) $\ll N$. Entonces no es válido tomar n_1 grande, ya que sus valores serán próximos a $\langle n_1 \rangle = pN \ll N$. De todos modos, puede usarse la aproximación de Stirling en

$$\frac{N!}{(N - n_1)!} \approx \frac{\sqrt{N}}{\sqrt{N - n_1}} \frac{\left(\frac{N}{e}\right)^N}{\left(\frac{N - n_1}{e}\right)^{N - n_1}} = \frac{N^{n_1} e^{-n_1}}{\left(1 - \frac{n_1}{N}\right)^{N - n_1}} \approx \frac{N^{n_1} e^{-n_1}}{e^{-n_1}} = N^{n_1}$$

y

$$(1 - p)^{N - n_1} (= q^{n_2}) \approx (1 - p)^N = \left(1 - \frac{a}{N}\right)^N \rightarrow e^{-a},$$

de manera que

$$P_N(n_1) = \frac{a^{n_1} e^{-a}}{n_1!}$$

que es la llamada *distribución de Poisson*. Vemos que cumple la condición de normalización

$$\sum_{n_1=0}^{\infty} P_N(n_1) = e^a e^{-a} = 1$$

y queda completamente determinada por el primer momento $\langle n_1 \rangle = a$.

1.4.3. Caminata al azar

Veamos ahora una aplicación típica de la distribución binomial, consistente en un movimiento unidimensional con pasos de longitud fija, hacia adelante (con probabilidad p) o hacia atrás (con probabilidad q). Analicemos el caso particular en que ambos eventos son igualmente probables, es decir, $p = q = 1/2$.

Al cabo de un número grande N de pasos, la distribución de probabilidades para n_1 pasos hacia adelante será gaussiana, ya que pN también es grande. El desplazamiento neto será $m \equiv n_1 - n_2$, y como $n_1 + n_2 = N$, podemos reescribir $m = 2n_1 - N$. La distribución de probabilidades para m se obtiene de (3), aunque con esta notación, m sólo puede tomar los valores $-N, -N+2, \dots, N-2, N$. Si deseamos incluir valores intermedios de m (pues también son posibles), es necesario agregar un factor $1/2$ “repartiendo” la probabilidad, y evitando que se pierda la condición de normalización. De este modo obtenemos

$$P_N(m) = \sqrt{\frac{1}{2\pi N}} e^{-\frac{m^2}{2N}}$$

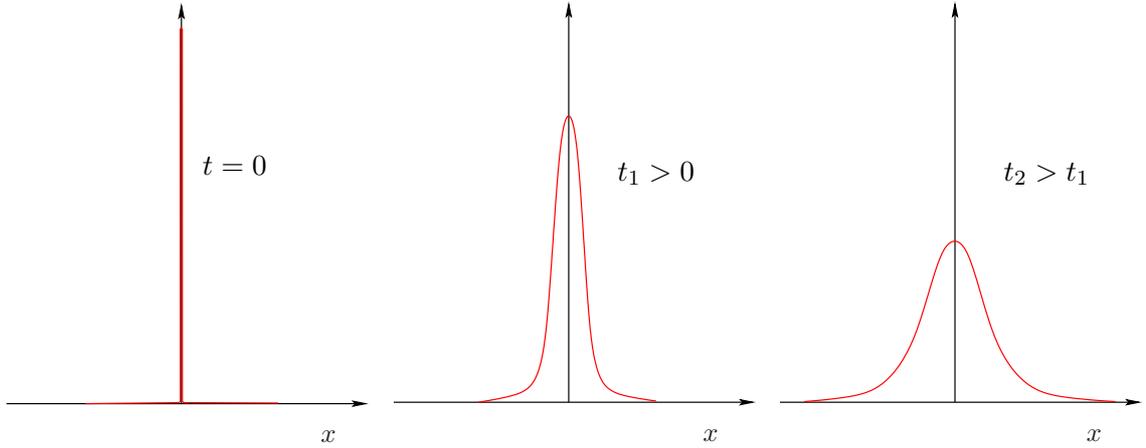
Si la longitud de cada paso es ℓ , el desplazamiento neto real será $x = m\ell$. Al cabo de N pasos, $\Delta x = \ell \Delta m$, de manera que, como $P_N(x) \Delta x = P_N(m) \Delta m = P_N(m) \Delta x / \ell$,

$$P_N(x) = \frac{1}{\sqrt{2\pi N \ell^2}} e^{-\frac{x^2}{2N \ell^2}}.$$

Cuando la partícula realiza n pasos por unidad de tiempo, $N = nt$; definiendo el *coeficiente de difusión* como $D = n \ell^2 / 2$ se obtiene

$$P(x, t) = \frac{1}{\sqrt{2\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

Este resultado indica que si en $t=0$ la partícula está localizada en $x=0$, la distribución inicial es $P(x,0)=\delta(x)$; a medida que transcurre el tiempo, la distribución se ensancha y se reduce su altura, de modo que el área bajo la curva se mantenga siempre igual a 1.



1.5. Teorema del límite central

Si tenemos una variable aleatoria X regida por $f_X(x)$, puede definirse una nueva variable como el promedio de N mediciones independientes de dicha magnitud, es decir,

$$y_N = \frac{x_1 + x_2 + \dots + x_N}{N}.$$

La pregunta que nos planteamos es ¿cómo será la distribución de probabilidades $f_Y(y_N)$ de esta nueva variable estocástica?

Para responder este ácido interrogante, estudiemos primero la variable $z_N = (x_1 + x_2 + \dots + x_N)/\sqrt{N}$, que es proporcional a y_N ; analicemos la función característica correspondiente a $f_Z(z_N)$

$$\begin{aligned} \Phi(k) &= \int dz_N e^{ikz_N} f_Z(z_N) = \int dx_1 \dots dx_N e^{ik(x_1 + x_2 + \dots + x_N)/\sqrt{N}} f_X(x_1) f_X(x_2) \dots f_X(x_N) \\ &= \left[\phi_X \left(\frac{k}{\sqrt{N}} \right) \right]^N. \end{aligned}$$

Tomando $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$, puede desarrollarse

$$\phi_X \left(\frac{k}{\sqrt{N}} \right) = e^{ik\langle X \rangle/\sqrt{N}} \int dx e^{ik(x - \langle X \rangle)/\sqrt{N}} f_X(x) = e^{ik\langle X \rangle/\sqrt{N}} \left[1 - \frac{k^2}{2N} \sigma^2 + \dots \right]$$

La exponencial oscilatoria en la integral nos permite ver que $\phi_X(k/\sqrt{N})$ decrece con k , de manera que $[\phi_X(k/\sqrt{N})]^N$ decrecerá aún más rápidamente. Si cuando $x \rightarrow \infty$ $f_X(x)$ decae de tal forma que todos los $\langle X^n \rangle$ son finitos,

$$\Phi(k) = e^{ik\sqrt{N}\langle X \rangle} \left[1 - \frac{k^2}{2N} \sigma^2 + \mathcal{O} \left(\frac{k}{\sqrt{N}} \right)^3 \right]^N \approx e^{ik\sqrt{N}\langle X \rangle} e^{-\frac{k^2 \sigma^2}{2}} \quad (N \rightarrow \infty).$$

De este modo, podemos reconstruir

$$f_Z(z_N) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikz_N} \Phi(k) \simeq \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ik(z_N - \sqrt{N}\langle X \rangle) - \frac{k^2 \sigma^2}{2}}.$$

Completando cuadrados puede resolverse la integral para obtener

$$f_Z(z_N) \simeq \frac{1}{2\pi} e^{-\frac{(z_N - \sqrt{N}\langle X \rangle)^2}{2\sigma^2}} \frac{\sqrt{2\pi}}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z_N - \sqrt{N}\langle X \rangle)^2}{2\sigma^2}},$$

de manera que la densidad de probabilidad para la variable Z es una gaussiana centrada en $\langle Z \rangle = \sqrt{N}\langle X \rangle$, con una desviación estándar σ , idéntica a la que tiene la variable X . Para la variable $Y = Z/\sqrt{N}$ es directo mostrar

que

$$f_Y(y_N) \simeq \frac{1}{\sqrt{2\pi}} \frac{\sqrt{N}}{\sigma} e^{-\frac{N(y_N - \langle X \rangle)^2}{2\sigma^2}}$$

Estos resultados se conocen como *teorema del límite central*, y nos dicen que, sin importar cuál es la forma de $f_X(x)$, el valor medio de la variable estocástica y_N coincide con $\langle X \rangle$ y la distribución $f_Y(y_N)$ es gaussiana, con $\sigma_Y = \sigma/\sqrt{N}$. Vale la pena resaltar que los requisitos que hemos pedido para nuestros N experimentos son: que sean independientes, que los momentos $\langle X^N \rangle$ sean finitos y que N sea grande.

2. Distribución de probabilidad en sistemas dinámicos

2.1. La evolución de sistemas clásicos (basado en el texto de Reichl)

Hemos dicho que la Mecánica Estadística estudia sistemas con muchos grados de libertad, como es el caso de N partículas en un recipiente, o N sitios de red interactuando entre sí. En un enfoque clásico, las ecuaciones de Hamilton pueden utilizarse para predecir la evolución del sistema, que estará caracterizado por $3N$ coordenadas y $3N$ impulsos generalizados (ignoramos grados de libertad internos).

Para realizar este análisis es útil introducir el concepto de *espacio de las fases* Γ , en este caso de $6N$ dimensiones, en el que cada punto \mathbf{X}^N contiene las $3N$ coordenadas y los $3N$ momentos ($\mathbf{q}^N, \mathbf{p}^N$). Suele llamarse a cada uno de estos posibles estados “microestados”, para hacer hincapié en el hecho de que dan información microscópica del sistema, a diferencia de los “macroestados” que estudia la termodinámica en general.

Si el valor de \mathbf{X}^N es conocido en algún instante, las ecuaciones de Hamilton nos proveen \mathbf{X}^N para cualquier instante posterior:

$$\dot{\mathbf{p}}_j = -\frac{\partial H^N}{\partial \mathbf{q}_j}; \quad \dot{\mathbf{q}}_j = \frac{\partial H^N}{\partial \mathbf{p}_j},$$

donde H^N es el hamiltoniano del sistema y el subíndice señala a la j -ésima partícula. Sabemos además que cuando este no depende explícitamente del tiempo, $H^N(\mathbf{X}^N) = E$ es una constante de movimiento, la energía, y se dice que el sistema es conservativo.

El microestado del sistema está representado entonces por un punto en el espacio de las fases que se mueve según las ecuaciones anteriores. Como el estado \mathbf{X}^N está perfectamente determinado por las condiciones iniciales para el sistema, sabemos que las trayectorias que describe nunca se cruzan en el espacio de las fases.

Cuando trabajamos con sistemas “termodinámicos” no se pueden especificar completamente sus microestados, de manera que siempre existe cierto grado de incertidumbre respecto de las condiciones iniciales. Por este motivo es razonable considerar que \mathbf{X}^N es una variable estocástica $6N$ -dimensional. Puede conocerse algo acerca de su valor medio, e incluso sobre su dispersión, pero hasta que no se realizara una “medición” (ideal) no se podría determinar el valor concreto que toma esta variable. Como toda variable estocástica, \mathbf{X}^N estará regida por una densidad de probabilidad $\rho(\mathbf{X}^N, t)$, de manera que $\rho(\mathbf{X}^N, t) d\mathbf{X}^N$ es la probabilidad de que el microestado del sistema esté en un entorno $d\mathbf{X}^N$ de \mathbf{X}^N en el instante t .

La condición de normalización

$$\int_{\Gamma} d\mathbf{X}^N \rho(\mathbf{X}^N, t) = 1$$

debe cumplirse para cualquier instante t , ya que el estado del sistema debe estar en alguna región de Γ . La probabilidad de tener en el instante t un estado en la región $R \subset \Gamma$ es

$$P(R) = \int_R d\mathbf{X}^N \rho(\mathbf{X}^N, t).$$

Si en algún momento la incertidumbre sobre \mathbf{X}^N es pequeña, ρ estará muy localizada alrededor del estado más probable, anulándose rápidamente al abandonar un entorno pequeño del mismo. A medida que transcurre el tiempo, es posible que la distribución permanezca muy ajustada, aunque el máximo quizás se desplace en el espacio de las fases, de manera que no se pierda información sobre el sistema; por el contrario, puede ocurrir que se esparza gradualmente, transformándose incluso en una distribución bastante uniforme, con lo cual se perdería la información por completo.

Aunque no es demasiado lo que conocemos sobre dinámica de sistemas continuos, veremos cómo puede establecerse una analogía entre la densidad de probabilidad en Γ y la masa en el caso de un fluido incompresible. Para comenzar, notemos que la condición de normalización equivale a la conservación de la materia. De la expresión anterior podemos analizar cómo cambia $P(R)$ con el tiempo, teniendo en cuenta que como un todo debe conservarse la probabilidad. Cualquier aumento en la probabilidad contenida en la región R debe igualarse a la cantidad de probabilidad que fluye hacia *dentro* de ese volumen en Γ :

$$\frac{dP(R)}{dt} = - \oint_S \rho(\mathbf{X}^N, t) \dot{\mathbf{X}}^N \cdot d\mathbf{S}^N,$$

donde $\dot{\mathbf{X}}^N \equiv (\dot{\mathbf{q}}^N, \dot{\mathbf{p}}^N)$, de manera que el último miembro representa el flujo saliente a través de la superficie S que encierra a R . Para evaluar este flujo, podemos recurrir al teorema de Gauss, reescribiendo la expresión anterior como

$$\int_R \frac{\partial \rho(\mathbf{X}^N, t)}{\partial t} d\mathbf{X}^N = - \int_R \nabla_{\mathbf{X}^N} \cdot (\rho(\mathbf{X}^N, t) \dot{\mathbf{X}}^N) d\mathbf{X}^N,$$

donde

$$\nabla_{\mathbf{X}^N} \equiv \left(\frac{\partial}{\partial \mathbf{q}_1}, \frac{\partial}{\partial \mathbf{q}_2}, \dots, \frac{\partial}{\partial \mathbf{q}_N}, \frac{\partial}{\partial \mathbf{p}_1}, \dots, \frac{\partial}{\partial \mathbf{p}_N} \right),$$

y $\nabla_{\mathbf{X}^N} \cdot (\)$ es la divergencia en Γ . Como la igualdad anterior debe cumplirse para cualquier $R \subset \Gamma$, entonces deben igualarse los integrandos, obteniéndose la celeberrima *ecuación de continuidad*

$$\frac{\partial}{\partial t} \rho(\mathbf{X}^N, t) + \nabla_{\mathbf{X}^N} \cdot (\rho(\mathbf{X}^N, t) \dot{\mathbf{X}}^N) = 0$$

El significado de esta ecuación es similar al que deducíamos en Electromagnetismo, donde habíamos partido considerando la *conservación de la carga*.

Para nuestro caso, en que ρ es la densidad de probabilidad en el espacio de las fases, $\nabla_{\mathbf{X}^N} \cdot \dot{\mathbf{X}}^N = 0$; es directo inferir este resultado explicitando

$$\nabla_{\mathbf{X}^N} \cdot \dot{\mathbf{X}}^N = \sum_{j=1}^N \left(\frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right),$$

que se anula en virtud de las ecuaciones de Hamilton que recordamos más arriba. Por lo tanto la ecuación de continuidad puede escribirse como

$$\frac{\partial \rho}{\partial t} = - \dot{\mathbf{X}}^N \cdot \nabla_{\mathbf{X}^N} \rho. \quad (4)$$

Esta derivada expresa los cambios temporales de ρ en un entorno fijo de Γ . Si deseamos analizar la evolución de ρ a medida que se desplaza el fluido, debemos recurrir a la derivada total o convectiva

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \dot{\mathbf{X}}^N \cdot \nabla_{\mathbf{X}^N}$$

En nuestro caso, esto se traduce como

$$\frac{d\rho}{dt} = 0.$$

Lo que significa esta identidad es que ρ se mantiene constante alrededor de un punto del “fluido” en movimiento. Es otra manera de identificar la densidad de probabilidad como fluido incompresible.

Utilizando las ecuaciones de Hamilton y definiendo

$$\hat{\mathcal{H}} \equiv \sum_{j=1}^N \left(\frac{\partial H^N}{\partial \mathbf{p}_j} \cdot \frac{\partial}{\partial \mathbf{q}_j} - \frac{\partial H^N}{\partial \mathbf{q}_j} \cdot \frac{\partial}{\partial \mathbf{p}_j} \right),$$

la ecuación de continuidad (4) puede reescribirse como

$$\frac{\partial \rho}{\partial t} = - \hat{\mathcal{H}} \rho = - [H^N, \rho],$$

donde $[\cdot, \cdot]$ representa el corchete de Poisson. La identidad anterior para la evolución temporal de ρ suele escribirse en términos del *operador de Liouville* $\hat{L} \equiv -i\hat{H}$:

$$i \frac{\partial \rho}{\partial t} = \hat{L} \rho$$

Esta expresión es conocida como *ecuación de Liouville*, y la solución formal puede escribirse como

$$\rho(\mathbf{X}^N, t) = e^{-it\hat{L}} \rho(\mathbf{X}^N, 0), \quad (5)$$

donde $\rho(\mathbf{X}^N, 0)$ es la condición inicial para ρ . Puede mostrarse que el operador \hat{L} es hermitiano, de manera que para la ecuación de autovectores

$$\hat{L} f_j = \lambda_j f_j,$$

sus autovalores λ_j son reales, y los autovectores f_j forman una base ortonormal en el espacio de Hilbert donde definimos ρ . Podemos expandir en esa base la condición inicial

$$\rho(\mathbf{X}^N, 0) = \sum_j C_j f_j(\mathbf{X}^N),$$

de manera que

$$\hat{L} \rho(\mathbf{X}^N, 0) = \sum_j C_j \lambda_j f_j(\mathbf{X}^N).$$

Entonces, la ecuación (5) toma la forma

$$\rho(\mathbf{X}^N, t) = \sum_j e^{-it\lambda_j} C_j f_j(\mathbf{X}^N).$$

Como los autovalores λ_j son reales, la densidad de probabilidad tendrá siempre un comportamiento oscilatorio, sin lograrse reproducir el decaimiento al equilibrio que se observa en los sistemas reales. Además, como las ecuaciones de Hamilton son reversibles en el tiempo, si en algún momento el sistema arribara al equilibrio, esta ecuación nos dice que podría recorrer el camino inverso, si se cambian todos los \mathbf{p} por $-\mathbf{p}$ hasta retornar a la condición inicial.

El problema de obtener el equilibrio termodinámico a partir de la dinámica que describen las ecuaciones de la mecánica clásica evidentemente no es posible. La ecuación de Liouville no nos permite conocer el camino mediante el cual se llega al equilibrio, aunque nos proporciona una condición que debe satisfacerse en un sistema que ya alcanzó un estado estacionario, pues en ese caso ρ no depende explícitamente del tiempo y por lo tanto $\hat{L}\rho=0$, o lo que es equivalente, $[H_N, \rho]=0$.

2.2. El operador densidad (basado en el texto de Nazareno)

Veremos a continuación cuánto se puede avanzar mediante una descripción cuántica, para lo cual introduciremos la idea de *ensamble*¹, definiéndolo como una colección de M sistemas físicos idénticos macroscópicamente, cada uno ocupando los diferentes microestados accesibles.

La idea es que el valor medio temporal de cualquier variable dinámica del sistema es igual al promedio de esa variable sobre el ensamble. Esta propiedad, denominada *ergodicidad* puede no ser válida en algunos sistemas, aunque sí lo será en todos los casos que estudiemos a lo largo de este curso. Si bien existen herramientas formales para demostrar cuándo un sistema es ergódico, nosotros descansaremos en el hecho de que en todos los casos que analizaremos, esta propiedad será siempre aceptable para el sentido común.

Cada sistema de nuestro ensamble tendrá un estado cuántico descrito por la función de onda $|k, t\rangle$, donde k representa todos los números cuánticos que caracterizan al sistema. El valor de expectación de una variable dinámica \hat{A} cuando el sistema se halla en ese estado estará dado por

$$A_k(t) = \langle k, t | \hat{A} | k, t \rangle,$$

siempre que tomemos estados normalizados ($\langle k, t | k, t \rangle = 1$). El promedio de \hat{A} sobre todo el ensamble será

$$\langle \hat{A} \rangle_M(t) = \frac{1}{M} \sum_{\text{ens}} \langle k, t | \hat{A} | k, t \rangle = \sum_k \omega_k A_k(t),$$

¹Pronto veremos que también se la utiliza en descripciones clásicas.

donde hemos introducido el peso estadístico $\omega_k = M_k/M$ (evidentemente, $\sum_k \omega_k = 1$); aquí M_k representa el número de sistemas que en el estado $|k, t\rangle$, para los cuales el valor de expectación de \hat{A} es A_k .

Si definimos el *operador densidad*

$$\hat{\rho} = \sum_k \omega_k |k, t\rangle \langle k, t|,$$

puede verse que para cualquier variable \hat{A} se puede escribir

$$\langle \hat{A} \rangle_M(t) = \text{Tr}(\hat{\rho}\hat{A}).$$

Es fácil mostrar que $\hat{\rho}^\dagger = \hat{\rho}$, es decir, es un operador hermitiano. También resulta inmediato verificar que $\text{Tr}(\hat{\rho}) = 1$. Otra propiedad de este operador es que, independientemente de la base elegida, son no-negativos sus elementos diagonales $\rho_i \equiv \rho_{ii}$ (la demostración es directa a partir de la definición, y también se deja como ejercicio). De esto resulta que como

$$\text{Tr}(\hat{\rho}) = 1 = \sum_i \rho_i \Rightarrow 0 \leq \rho_i \leq 1.$$

Por último destaquemos la propiedad de que como existe una base en la que $\hat{\rho}$ es diagonal, esta última desigualdad implica que $\text{Tr}(\hat{\rho}^2) \leq 1$, de manera que en una base arbitraria

$$\sum_{i,j} |\rho_{ij}|^2 \leq 1,$$

es decir, todos los elementos de $\hat{\rho}$ son acotados.

2.2.1. Estados puros y mezcla

Si todos los sistemas de nuestro ensamble están en el mismo estado $|k_o, t\rangle$, el valor de expectación de \hat{A} es $\langle k_o, t | \hat{A} | k_o, t \rangle = A_{k_o}$ en todos los sistemas. Por lo tanto,

$$\langle \hat{A} \rangle_M = A_{k_o} \quad \text{y} \quad \hat{\rho} = |k_o, t\rangle \langle k_o, t|.$$

Entonces se cumple que $\hat{\rho}^2 = \hat{\rho}$. Si lo expresamos en una base donde $\hat{\rho}$ es diagonal, esto significa que debe cumplirse $\rho_i^2 = \rho_i$, pero como $\text{Tr}(\hat{\rho}) = 1$ ($= \text{Tr}(\hat{\rho}^2)$), la única forma en que se pueden cumplir simultáneamente las condiciones $\sum_i \rho_i = 1$ y $\sum_i \rho_i^2 = 1$ con $0 \leq \rho_i \leq 1$ es que $\rho_\ell = 1$ y $\rho_{i \neq \ell} = 0$ para algún valor ℓ .

Como $\hat{\rho}$ siempre es diagonalizable, vemos que la diferencia entre un estado puro y un estado mezcla es justamente esta: en un estado mezcla, que es el caso más general, la representación diagonal de $\hat{\rho}$ contiene más de un elemento distinto de cero, y esta situación corresponde al caso en que no poseemos tanta información acerca del sistema como para predecir exactamente en qué estado se encuentra.

Como vemos, siempre es conveniente pensar algunas expresiones que involucran trazas en una base que diagonaliza a $\hat{\rho}$ (ya que la traza no depende de la representación elegida). Así evitaremos además la desazón de encontrar inesperadas la notación y la forma del operador densidad. Si $\hat{\rho}$ es diagonal, cada elemento ρ_i representa la probabilidad de que el sistema ocupe cierto estado $|i\rangle$, que anteriormente habíamos señalado como P_i . Entonces, el hecho de que $\text{Tr}(\hat{\rho}) = 1$, significa que $\sum_i P_i = 1$, que se convierte en una expresión mucho más familiar para nosotros. Por otro lado, en el caso en que $\hat{\rho}$ y \hat{A} conmuten, puede elegirse una base que diagonalice a ambos operadores: en este caso P_i representa también la probabilidad de que el sistema ocupe un estado para el cual \hat{A} posee su autovalor A_i . Esto implica que el valor medio $\langle \hat{A} \rangle_M$ presenta un aspecto bastante menos sorprendente, ya que $\text{Tr}(\hat{\rho}\hat{A}) = \sum_i \rho_i A_i = \sum_i P_i A_i$, reencontrándonos con las expresiones de nuestra infancia.

2.2.2. Ecuación de Liouville

Sabemos que, si \hat{H} es el hamiltoniano de nuestro sistema y \hat{A} es independiente de t , para cada estado $|k, t\rangle$ de nuestro ensamble

$$i\hbar \frac{d}{dt} \langle k, t | \hat{A} | k, t \rangle = \langle k, t | [\hat{A}, \hat{H}] | k, t \rangle,$$

donde no debe confundirse el conmutador con el corchete de Poisson. Por otro lado, como

$$\langle \hat{A} \rangle_M = \text{Tr}(\hat{\rho}\hat{A}) = \sum_k \omega_k \langle k, t | \hat{A} | k, t \rangle,$$

podemos escribir

$$i\hbar \frac{d}{dt} \langle \hat{A} \rangle_M = i\hbar \text{Tr}(\hat{\rho}\dot{\hat{A}}) = i\hbar \sum_k \omega_k \frac{d}{dt} \langle k, t | \hat{A} | k, t \rangle = \sum_k \omega_k \langle k, t | [\hat{A}, \hat{H}] | k, t \rangle = \langle [\hat{A}, \hat{H}] \rangle_M,$$

de modo que

$$i\hbar \operatorname{Tr} \left(\frac{\partial \hat{\rho}}{\partial t} \hat{A} \right) = \operatorname{Tr}(\hat{\rho} [\hat{A}, \hat{H}]) = \operatorname{Tr}([\hat{H}, \hat{\rho}] \hat{A})$$

(para verificar la última identidad sólo es necesario recordar propiedades simples de una traza, como $\operatorname{Tr}(\hat{A}\hat{B}) = \operatorname{Tr}(\hat{B}\hat{A})$). Como esta relación es válida para cualquier \hat{A} , entonces debe cumplirse la *ecuación de Liouville* cuántica (o *ecuación de Liouville - von Neumann*)

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}]$$

También en este caso podemos definir el operador de Liouville

$$\hat{L} = \frac{1}{\hbar} [\hat{H}, \cdot] \quad \Rightarrow \quad i \frac{\partial \hat{\rho}}{\partial t} = \hat{L} \hat{\rho} .$$

En este caso no diremos nada acerca de la hermiticidad de \hat{L} , aunque nuevamente es posible escribir la solución formal como

$$\hat{\rho}(t) = e^{-it\hat{L}} \hat{\rho}(0) = e^{-it\hat{H}/\hbar} \hat{\rho}(0) e^{it\hat{H}/\hbar} .$$

El último término puede obtenerse desarrollando $e^{-it\hat{L}}$ como una serie de potencias y luego explicitando el binomio de Newton correspondiente a las potencias del conmutador $[\hat{H}, \cdot]$. Si representamos a $\hat{\rho}$ en una base ortonormal de autofunciones de \hat{H} , $\{|E_n\rangle\}$, recordando que $\sum_n |E_n\rangle\langle E_n|$ es el operador identidad, resulta

$$\hat{\rho}(t) = \sum_{n,m} \langle E_n | \hat{\rho}(0) | E_m \rangle e^{-(i/\hbar)(E_n - E_m)t} | E_n \rangle \langle E_m | .$$

Esta expresión nos permite ver claramente que la descripción cuántica también provee soluciones oscilatorias para la evolución de nuestro sistema, ya que $\hat{\rho}$ consiste de una suma de términos que oscilan con distinta frecuencia.

En cambio podemos observar que cuando se ha alcanzado el estado estacionario debe cumplirse $\rho_{nm} = \rho_n \delta_{n,m}$, es decir $\hat{\rho}$ debe ser diagonal en esta representación, para que se anulen los términos oscilatorios. Esto es posible si en el equilibrio $\hat{\rho} = \hat{\rho}(\hat{H})$, lo cual era evidente de la ecuación de Liouville independiente de t , pues entonces $[\hat{H}, \hat{\rho}] = 0$.

3. Nociones sobre teoría de información (basado en el texto de Balian)

Hemos visto que las descripciones dinámicas que se pueden dar desde las teorías clásica o cuántica no predicen adecuadamente las relajaciones de los sistemas termodinámicos. La evolución de un sistema hacia un estado estacionario es un problema aún latente en la actualidad, y es por ello que la mecánica estadística aborda fundamentalmente los sistemas que ya han alcanzado el equilibrio termodinámico. Plantearemos entonces las bases para poder dar esa descripción de los sistemas en equilibrio, a partir de lo que se conoce como teoría de información. Si bien aquí la orientaremos hacia la mecánica estadística, la teoría de información hoy es de interés en inteligencia artificial, neurociencias, sistemas no lineales, cuántica, etc.

3.1. Entropía de información

Para avanzar con nuestra descripción, introduzcamos el operador $-k \ln \hat{\rho}$, donde k es una constante. El valor medio de este operador en un ensamble es

$$S \equiv \langle -k \ln \hat{\rho} \rangle_M = -k \operatorname{Tr}(\hat{\rho} \ln \hat{\rho}) = -k \sum_j \rho_j \ln \rho_j ;$$

el último miembro corresponde a una base que diagonaliza a $\hat{\rho}$. Vemos que $S \geq 0$, ya que los elementos diagonales de $\hat{\rho}$ cumplen $0 \leq \rho_j \leq 1$ (cualquiera sea la base escogida).

En el caso de tener un ensamble puro, en esta representación todos los elementos de $\hat{\rho}$ se anulan excepto uno que vale 1, de modo que $S = 0$. En otras palabras, cuando poseemos la máxima información sobre el sistema, S alcanza su valor mínimo. Por otro lado, en la próxima sección veremos que S alcanza el valor máximo cuando

todas las probabilidades ρ_j son iguales, es decir, cuando poseemos el mínimo de información sobre el estado de un sistema.

El valor de S cuantifica de alguna manera la falta de información o la incertidumbre respecto de un sistema. Esta magnitud se denomina *entropía de información* o *entropía estadística*², y en general para un sistema regido por una distribución de probabilidades $\{P_j\}$ se define directamente como

$$S \equiv -k \sum_{j=1}^M P_j \ln P_j, \quad (6)$$

Si los M estados accesibles para el sistema son equiprobables,

$$P_j = \frac{1}{M} \quad \Rightarrow \quad S = k \ln M,$$

de manera que S es creciente con M . Otra propiedad importante es la *aditividad*, para lo cual imaginamos dos sistemas independientes: el A , con ℓ resultados posibles cuyas probabilidades son p_i ($i = 1, \dots, \ell$) y el B , con m resultados posibles cuyas probabilidades son q_j ($j = 1, \dots, m$). Como los sistemas son independientes, la probabilidad de tener simultáneamente el resultado i para el sistema A y j para el B es $\pi_{ij} = p_i q_j$ ($\{p_i\}$ y $\{q_j\}$ están normalizadas, con lo cual $\{\pi_{ij}\}$ queda automáticamente normalizada). La entropía para el sistema conjunto es

$$S_{A+B} = -k \sum_{i,j=1}^{\ell,m} \pi_{ij} \ln \pi_{ij} = -k \sum_{i,j=1}^{\ell,m} p_i q_j \ln (p_i q_j) = -k \underbrace{\left(\sum_{i=1}^{\ell} p_i \ln p_i \right)}_{S_A} \underbrace{\sum_{j=1}^m q_j}_1 - k \underbrace{\left(\sum_{j=1}^m q_j \ln q_j \right)}_{S_B} \underbrace{\sum_{i=1}^{\ell} p_i}_1 = S_A + S_B.$$

Cuando consideremos sistemas termodinámicos, veremos que puede haber cierta interdependencia entre subsistemas, sin que esto afecte la aditividad de la entropía (ni de otras variables extensivas como la energía interna, el momento magnético total, etc.).

En el caso de trabajar con variables continuas el desarrollo anterior se mantiene, excepto por el correspondiente reemplazo de sumatorias por integrales.

3.2. Principio de máxima entropía (estadística)

Este principio queda prácticamente enunciado con el título que hemos elegido:

La entropía estadística de un sistema alcanza el máximo compatible con los vínculos impuestos.

El sentido de este principio ha quedado explicitado en los ejemplos anteriores: de algún modo señala que no hay razón para privilegiar un estado particular de antemano.

El procedimiento habitual para maximizar la entropía consiste en recurrir al método variacional, imponiendo los vínculos mediante multiplicadores de Lagrange. Veamos primero el ejemplo de hallar la distribución de probabilidad en nuestro conjunto de M eventos bajo la única imposición de la normalización, $\sum_j P_j = 1$. Debemos exigir que alrededor de la distribución $\{P_j\}$ que maximiza (6) las variaciones virtuales se anulen:

$$\delta \left[-k \sum_{j=1}^M P_j \ln P_j + \alpha \sum_{j=1}^M P_j \right] = 0,$$

donde α es el multiplicador de Lagrange asociado con la normalización. Permitimos entonces variaciones infinitesimales δP_j en las probabilidades (como si fueran independientes), obteniendo

$$\sum_{j=1}^M \delta P_j [-k \ln P_j + (\alpha - k)] = 0.$$

²Este concepto fue utilizado en problemas de comunicación por Shannon en 1948, y formalmente aplicado a la mecánica estadística por Jaynes en 1957; sin embargo, Gibbs ya había propuesto esta expresión como entropía estadística en 1901.

Como esto debe anularse para cualquier conjunto de variaciones $\{\delta P_j\}$ arbitrarias, entonces debe anularse el término entre corchetes para cada j , con lo cual resulta

$$P_j = e^{\frac{\alpha}{k} - 1} ,$$

es decir, todos los estados tienen la misma probabilidad, que por la condición de normalización no pueden sino valer $1/M$. En efecto, escribiendo explícitamente $\sum_j P_j = 1$ podemos además encontrar el valor para el multiplicador α :

$$e^{\frac{\alpha}{k} - 1} = \frac{1}{M} \quad \Rightarrow \quad \alpha = k(1 - \ln M) .$$

En el caso de contar con algún vínculo adicional

$$\bar{f} = \sum_{j=1}^M P_j f_j , \quad (7)$$

el proceso de maximización es similar, sólo que debemos acompañar esta nueva restricción con otro multiplicador de Lagrange β , resultando

$$\sum_{j=1}^M \delta P_j [-k \ln P_j + (\alpha - k) + \beta f_j] = 0 .$$

Obtenemos entonces

$$P_j = e^{(\frac{\alpha}{k} - 1) + \frac{\beta}{k} f_j} ,$$

y cuando imponemos la condición de normalización vemos que debe cumplirse

$$e^{\frac{\alpha}{k} - 1} \sum_{j=1}^M e^{\frac{\beta}{k} f_j} = 1 .$$

Definimos entonces la *función partición*

$$Z \equiv e^{1 - \frac{\alpha}{k}} = \sum_{j=1}^M e^{\frac{\beta}{k} f_j} . \quad (8)$$

Sustituyendo en la expresión anterior, obtenemos

$$P_j = \frac{e^{\frac{\beta}{k} f_j}}{\sum_{j=1}^M e^{\frac{\beta}{k} f_j}} ,$$

donde se evidencia que la distribución $\{P_j\}$ está normalizada. Falta aún determinar el valor del multiplicador β , que puede obtenerse utilizando el vínculo (7)

$$\bar{f} = \frac{\sum_{j=1}^M f_j e^{\frac{\beta}{k} f_j}}{\sum_{j=1}^M e^{\frac{\beta}{k} f_j}} .$$

Despejando de aquí β , podemos encontrar también el valor de α a través de (8). La generalización para el caso de mayor número de vínculos es inmediata, y sólo debe tenerse la precaución de no sobredeterminar el sistema.

4. Ensamble microcanónico (basado en los textos de Reichl y Huang)

Abordamos a partir de este capítulo el estudio de sistemas termodinámicos en equilibrio. Como dijimos anteriormente, en estos casos la mecánica estadística plantea el principio de máxima entropía utilizando el operador densidad en la expresión de la entropía de información:

$$S = -k \operatorname{Tr}(\hat{\rho} \ln \hat{\rho}) .$$

Veremos que la asociación de esta magnitud con la entropía termodinámica nos provee una adecuada descripción, tanto para sistemas aislados como para aquellos acoplados de alguna manera con su entorno.

4.1. Sistemas cerrados y aislados

Comenzamos analizando un sistema que no intercambia ningún tipo de energía con su entorno, ni puede incorporar o liberar materia. Si bien su energía E es fija, dispone de muchos microestados compatibles con la condición $E=U$ (constante). El ensamble asociado con esta situación se denomina *microcanónico*.

La maximización de S con el único vínculo $\operatorname{Tr}(\hat{\rho}) = 1$ implica un procedimiento similar al de la sección anterior, aunque en este caso representaremos como $\delta\hat{\rho}$ a las variaciones que se provoquen en el operador $\hat{\rho}$ que maximiza la entropía:

$$\delta \{-k \operatorname{Tr}(\hat{\rho} \ln \hat{\rho}) + \alpha \operatorname{Tr}(\hat{\rho})\} = 0 \quad \Rightarrow \quad \operatorname{Tr} \left\{ \delta\hat{\rho} \left[-k \ln \hat{\rho} + (\alpha - k) \hat{I} \right] \right\} = 0 ,$$

donde \hat{I} es el operador identidad. Nuevamente, esto debe cumplirse para variaciones $\delta\hat{\rho}$ arbitrarias, de manera que debe anularse el factor entre corchetes, lo que implica

$$\hat{\rho} = e^{\frac{\alpha}{k} - 1} \hat{I} . \quad (9)$$

Esto significa que $\hat{\rho}$ es un operador diagonal, con todos sus elementos ρ_i idénticos. Si el número de estados accesibles al sistema con energía E es $W(E)$, entonces

$$\rho_i = \frac{1}{W(E)} .$$

Este resultado es análogo al de la sección anterior, y es equivalente al *postulado de igual probabilidad a priori*, que indica que no hay motivo para privilegiar ningún estado si la única restricción que se impone es que el sistema tenga energía E . Muchas formulaciones de la mecánica estadística comienzan aceptando este postulado en lugar del principio de máxima entropía, que es el camino que nosotros hemos escogido.

Reemplazando los valores para ρ_i en la expresión para la entropía obtenemos nuevamente

$$S = k \ln W(E) \quad (10)$$

que puede tomarse como *relación fundamental* para nuestro sistema si aceptamos que se corresponde con la entropía termodinámica. En particular, la conexión directa con la temperatura se escribe como

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

y a partir de aquí, por supuesto, pueden encontrarse magnitudes como el calor específico u otras funciones respuesta.³

En el caso de que el sistema analizado pueda describirse adecuadamente mediante un enfoque clásico, tendremos una densidad de probabilidad $\rho(\mathbf{X}^N)$, y la entropía se escribe

$$S = -k \int_{E < H(\mathbf{X}^N) < E + \Delta E} d\mathbf{X}^N \rho(\mathbf{X}^N) \ln \left[C^N \rho(\mathbf{X}^N) \right] ,$$

donde la constante C^N se introduce para tener las unidades correctas y corresponde al volumen en el espacio de las fases Γ ocupado por un estado del sistema. Este volumen sólo puede hallarse mediante argumentos

³Aquí la variable X representa los parámetros extensivos involucrados en trabajos que pueden realizarse sobre nuestro sistema termodinámico (volumen, momento magnético total, etc.)

cuánticos, y pronto mostraremos que $C^N = h^{3N}$, donde h es la constante de Planck⁴. Aceptando este volumen finito ocupado en Γ por cada estado, la integración de la expresión anterior abarca estados con energías levemente superiores a E en una cantidad ΔE pequeña, aunque el volumen abarcado en el espacio de las fases es mucho mayor que el de un único estado.

Al maximizar la entropía en esta descripción clásica se obtiene

$$\rho(\mathbf{X}^N) = \frac{1}{\omega(E, V, N)},$$

donde $\omega(E, V, N)$ es el volumen en Γ comprendido entre las superficies de energía E y $E + \Delta E$

$$\omega(E, V, N) = \int_{E < H(\mathbf{X}^N) < E + \Delta E} d\mathbf{X}^N,$$

que también puede representarse como el área $\Sigma(E, V, N)$ de la hipersuperficie con energía E multiplicada por el ancho ΔE , $\omega(E, V, N) = \Sigma(E, V, N)\Delta E$. La densidad de probabilidad se anula si el estado analizado no se encuentra comprendido en $\omega(E, V, N)$. Reemplazando la expresión hallada para $\rho(\mathbf{X}^N)$ en la definición de entropía, se obtiene

$$S(E, V, N) = k \ln \frac{\omega(E, V, N)}{C^N} \quad (11)$$

Esta expresión es equivalente a la desarrollada para el caso cuántico: si $\omega(E, V, N)$ es el volumen en Γ ocupado por todos los estados accesibles y C^N es el volumen asociado a un solo estado, $\omega(E, V, N)/C^N$ es el número de estados compatibles con la energía E prefijada. Por otro lado, la expresión encontrada para S puede tomarse como relación fundamental a partir de la cual pueden derivarse todas las propiedades termodinámicas del sistema considerado.

Es interesante verificar que esta expresión para la entropía resulta *extensiva*. Para ello consideremos un sistema con N partículas distinguibles dividido en dos partes no interactuantes, conteniendo cada subsistema α y $\beta = N - \alpha$ partículas respectivamente. El hamiltoniano del sistema conjunto puede escribirse como

$$H(\mathbf{X}^N) = H(\mathbf{X}^\alpha) + H(\mathbf{X}^\beta)$$

y el espacio de las fases es el producto cartesiano de los respectivos espacios para cada subsistema, $\Gamma^{(N)} = \Gamma^{(\alpha)} \times \Gamma^{(\beta)}$. Un estado con energía total E se logra combinando un estado del subsistema α cuya energía parcial es E_i con un estado de β cuya energía es $E - E_i$, lo que en el espacio $\Gamma^{(N)}$ implica un volumen

$$\omega^{(\alpha)}(E_i) \cdot \omega^{(\beta)}(E - E_i).$$

Como esta no es la única alternativa, ya que existen muchos valores de E_i que se corresponden con la misma energía total E , el volumen en el espacio $\Gamma^{(N)}$ que debemos evaluar resulta de sumar todas las combinaciones posibles, es decir

$$\omega(E) = \sum_{i=1}^{E/\Delta E} \left[\omega^{(\alpha)}(E_i) \cdot \omega^{(\beta)}(E - E_i) \right].$$

Si el mayor sumando en esta expresión se da cuando $E_i = \bar{E}_\alpha$ ($\bar{E}_\beta = E - \bar{E}_\alpha$), podemos acotar ω por arriba reemplazando en la sumatoria todos los términos por ese valor máximo, y por abajo eliminando todos los otros términos (que son no negativos):

$$\omega^{(\alpha)}(\bar{E}_\alpha) \cdot \omega^{(\beta)}(\bar{E}_\beta) \leq \omega(E) \leq \frac{E}{\Delta E} \omega^{(\alpha)}(\bar{E}_\alpha) \cdot \omega^{(\beta)}(\bar{E}_\beta).$$

Tomando logaritmos y multiplicando por k obtenemos

$$k \ln \left[\frac{\omega^{(\alpha)}(\bar{E}_\alpha)}{C^\alpha} \cdot \frac{\omega^{(\beta)}(\bar{E}_\beta)}{C^\beta} \right] \leq S(E, V, N) \leq k \ln \left[\frac{\omega^{(\alpha)}(\bar{E}_\alpha)}{C^\alpha} \cdot \frac{\omega^{(\beta)}(\bar{E}_\beta)}{C^\beta} \right] + k \ln \frac{E}{\Delta E},$$

donde hemos identificado $C^N = C^\alpha C^\beta$ a partir de su definición. Cualquier volumen en el espacio α (o β) es proporcional a cierta cantidad (dependiente del volumen en cuestión) elevada a una potencia 6α (o 6β), de

⁴Más adelante veremos que para el caso de partículas indistinguibles debe corregirse esta constante para conseguir resultados adecuados.

manera que el primer término del miembro de la derecha es proporcional a $\alpha + \beta = N$, mientras que el segundo es proporcional a $\ln N$ (pues $E \propto N$). En el *límite termodinámico*, es decir cuando N es muy grande (y también α y β), la contribución del último término es despreciable frente al primero, coincidiendo entonces la cota inferior con la superior. Esto implica que, en el límite termodinámico,

$$S^{(N)} = S^{(\alpha)} + S^{(\beta)} ,$$

de modo que la entropía es aditiva. La extensividad que deseábamos verificar es consecuencia directa de la aditividad.

Con un razonamiento similar al anterior puede mostrarse que si $\Omega(E, V, N) = \sum_{i=0}^{E/\Delta E} \omega(E_i)$ representa el volumen contenido dentro de la hipersuperficie $\Sigma(E, V, N)$, la expresión para la entropía puede escribirse también como

$$S(E, V, N) = k \ln \frac{\Omega(E, V, N)}{C^N} \quad (12)$$

Matemáticamente esta expresión suele resultar más sencilla en la práctica, de modo que la utilizaremos en lugar de la fórmula (11) cuando sea conveniente.

Paralelamente a la demostración anterior, hemos resaltado el hecho de que la energía E_α del subsistema α tiene en realidad un valor bien definido, ya que —siempre en el límite termodinámico— la sumatoria puede reemplazarse por un único sumando: el término que maximiza $\omega^{(\alpha)} \cdot \omega^{(\beta)}$ bajo la restricción $E_\alpha + E_\beta = E$ (cte). Como el logaritmo es una función monótonamente creciente, la condición de máximo puede escribirse como

$$\delta \left[\ln \left(\omega^{(\alpha)} \cdot \omega^{(\beta)} \right) \right] = 0 \quad \text{con} \quad \delta E_\alpha + \delta E_\beta = 0 .$$

Dado que los subsistemas son independientes, $\omega^{(j)}$ sólo depende de E_j , de modo que la condición anterior implica

$$\left[\frac{\partial \ln \omega^{(\alpha)}}{\partial E_\alpha} - \frac{\partial \ln \omega^{(\beta)}}{\partial E_\beta} \right] \delta E_\alpha = 0$$

para δE_α arbitrario, es decir

$$\left. \frac{\partial S_\alpha}{\partial E_\alpha} \right|_{E_\alpha} = \left. \frac{\partial S_\beta}{\partial E_\beta} \right|_{E_\beta} .$$

Esta identidad es un nuevo indicio de la “razonabilidad” de asociar la entropía de información con la entropía termodinámica, pues si $\partial S / \partial U = 1/T$, la identidad anterior significa que

$$T_\alpha = T_\beta ,$$

es decir, en el equilibrio las temperaturas de ambos subsistemas deben igualarse. Por conveniencia, en la definición de la entropía se elige la constante de Boltzmann $k = 1,380649 \cdot 10^{-23} \text{J/K}$ ⁵, con el objeto de que esta temperatura coincida con la temperatura absoluta que se introduce en la termodinámica. De todos modos, en la expresión (9) se ha explicitado que la elección de esta magnitud es de algún modo irrelevante, ya que está ligada a los valores de los multiplicadores de Lagrange, y modificarla solo representaría un cambio de unidades.

Finalmente, es fácil ver que nuestra expresión para la entropía crece con la energía interna: al aumentar U , el volumen Ω que debe evaluarse en la expresión (12) crece, pues pasamos a abarcar, además de los estados anteriores, nuevos estados con energía interna algo mayor. Como el logaritmo es una función creciente, esto significa que S crece con U . La expresión que aceptamos para la entropía estadística satisface entonces los requisitos que exigimos al introducir esta función para desarrollar la teoría termodinámica: por un lado, toma el máximo valor compatible con U , V (o la variable extensiva X que corresponda) y N ; por otro, es una función monótonamente creciente con U , y además cumple con la propiedad de aditividad.

4.1.1. Gas ideal clásico

Consideremos como ejemplo el caso del gas ideal, que consiste en un sistema de N moléculas libres no interactuantes contenidas en un recipiente de volumen V . En particular, supondremos que estas moléculas no

⁵Mediante el número de Avogadro N_A , esta constante se relaciona con la constante universal de los gases $R = N_A k$.

poseen grados de libertad internos, y para un enfoque clásico las consideraremos como partículas distinguibles de masa m . En cartesianas, el hamiltoniano para este sistema es

$$H(\mathbf{X}^N) = \frac{1}{2m} \sum_{i=1}^N \mathbf{p}_i^2 .$$

El volumen en el espacio de las fases correspondiente a estados con energía menor que E es

$$\Omega(E, V, N) = \int_{H(\mathbf{X}^N) < E} d^3\mathbf{p}_1 \dots d^3\mathbf{p}_N d^3\mathbf{r}_1 \dots d^3\mathbf{r}_N .$$

Teniendo en cuenta que el hamiltoniano no depende de \mathbf{r} , la integración sobre cada coordenada espacial es simplemente el volumen V , por lo que resulta

$$\Omega(E, V, N) = V^N \Omega_{3N}(R) ,$$

donde denotamos $\Omega_n(R)$ al volumen de una esfera n -dimensional de radio $R = \sqrt{2mE}$. Este volumen debe ser proporcional a R^n ,

$$\Omega_n(R) = b_n R^n ,$$

de manera que nuestro problema ahora se reduce a encontrar una expresión para b_n . Con ese objeto, pensemos que si $S_n(R) = d\Omega_n(R)/dR$ es el área de la superficie de la esfera n -dimensional de radio R , podemos recurrir a la identidad

$$\int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n e^{-(x_1^2 + \dots + x_n^2)} = \int_0^{\infty} dR S_n(R) e^{-R^2} ,$$

para escribir

$$\left[\int_{-\infty}^{\infty} dx e^{-x^2} \right]^n = n b_n \int_0^{\infty} dR R^{n-1} e^{-R^2} .$$

Sustituyendo $x = R^2$ en la última integral, obtenemos la definición de la función Γ , con lo cual

$$\pi^{n/2} = \frac{n b_n}{2} \Gamma\left(\frac{n}{2}\right) \quad \Rightarrow \quad b_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} .$$

Utilizando la aproximación de Stirling para la función Γ , se obtiene en nuestro caso

$$\ln(b_{3N}) \simeq \frac{3N}{2} \ln \pi - \frac{3N}{2} \ln \frac{3N}{2} + \frac{3N}{2} \quad (N \rightarrow \infty) ,$$

de manera que

$$S(E, V, N) = \frac{3}{2} Nk + Nk \ln \left[V \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} \right] .$$

De esta relación fundamental pueden derivarse las conocidas ecuaciones de estado

$$T = \frac{2}{3} \frac{U}{Nk} ; \quad P = \frac{NkT}{V} .$$

También puede hallarse para el calor específico (por partícula) la relación $c_v = \frac{3}{2} k$.

Paradoja de Gibbs

Podemos escribir la entropía en notación más compacta como

$$S = N s_o + Nk \ln \left(V u^{3/2} \right) ,$$

donde

$$s_o = \frac{3}{2} k \left(1 + \ln \frac{4\pi m}{3h^2} \right) , \quad u = \frac{3}{2} kT .$$

Esta expresión es muy similar a la que habíamos obtenido al desarrollar la teoría termodinámica. Analicemos el caso en que un recipiente contiene dos gases diferentes a temperatura T separados por una pared diatérmica, de modo que en un compartimiento hay N_1 moléculas en un volumen V_1 , mientras que en el otro hay N_2 moléculas

en un volumen V_2 ($V_1 + V_2 = V$, $N_1 + N_2 = N$). Al remover la pared divisoria, como U no cambia, T tampoco varía, de modo que al mezclar habrá un aumento de entropía dado por

$$\frac{\Delta S}{k} = N_1 \ln \frac{V}{V_1} + N_2 \ln \frac{V}{V_2} > 0.$$

Cuando los gases son diferentes, esta expresión aparentemente provee una adecuada predicción. Sin embargo, si los gases son idénticos, aun cuando el sistema sea homogéneo antes de remover la pared divisoria ($V_1/N_1 = V_2/N_2 = V/N$), esta expresión señala que habrá un aumento en la entropía del sistema. Por supuesto, este resultado es absurdo, y para resolver este problema Gibbs introdujo una modificación en la constante C^N para el caso de partículas indistinguibles, agregándole un factor $N!$, es decir $C^N = h^{3N} N!$. De este modo, utilizando la aproximación de Stirling la expresión para la entropía resulta

$$S = \frac{3}{2} Nk \left(\frac{5}{3} + \ln \frac{4\pi m}{3h^2} \right) + Nk \ln \left(\frac{V}{N} u^{3/2} \right)$$

Esta expresión se conoce como *ecuación de Sackur-Tetrode*, y ahora es completamente coincidente con la que habíamos obtenido en *Termino I* (en particular, a diferencia de la expresión anterior, esta entropía sí es extensiva). La predicción del aumento de entropía se mantiene, excepto para el caso en que el sistema conjunto esté originalmente en equilibrio termodinámico, con igual temperatura y presión a ambos lados de la pared divisoria.

No hay una respuesta convincente acerca de por qué corresponde dividir ω/h^{3N} u Ω/h^{3N} por $N!$ dentro del formalismo clásico. Quizás la motivación original surgió de pensar que si las partículas son indistinguibles, existen $N!$ permutaciones que proveen el mismo estado, aunque está claro que dentro de la teoría clásica todas las partículas son distinguibles. La verdadera respuesta es en realidad provista por la cuántica, en el marco de la llamada “segunda cuantización”: como veremos más adelante, para describir N partículas idénticas, las funciones de onda conjuntas deben ser *simétricas* o *antisimétricas* ante permutaciones de partículas, y cuando las condiciones indican que la descripción debe coincidir con el enfoque clásico, se ve que entonces ambas formulaciones concuerdan gracias a ese denominador $N!$.

4.1.2. Dos ejemplos de sistemas cuánticos

El problema central en el caso cuántico también es contar los estados de nuestro sistema compatibles con el valor de energía E preestablecido.

Veremos en primer lugar el caso de N espines $1/2$ no interactuantes en un sólido bajo la acción de un campo externo B . Las energías posibles para cada espín son $\pm\mu B/2$, de modo que las energías totales para todo el sistema serán $-N\mu B/2, (-N/2+1)\mu B, \dots, +N\mu B/2$. Denotando el valor de la energía como

$$E_M = \left(-\frac{N}{2} + M \right) \mu B \quad (0 \leq M \leq N),$$

debemos entonces evaluar el número de formas $W(M)$ en que pueden combinarse M espines alineados con el campo con $N - M$ desalineados. Se deja como ejercicio resolver este problema combinatorio, válido para cualquier sistema en el que sus componentes pueden acceder solo a 2 estados cuyas energías suelen denotarse como 0 y ϵ . Este problema es conocido como *sistema de dos niveles* (por razones desconocidas), y puede verse que para temperaturas extremas se anulan los valores asintóticos para el calor específico por distintos motivos. En el límite de bajas temperaturas, la misma definición de calor específico $\Delta U = Nc\Delta T$ nos hace ver que la cuántica prohíbe a nuestro sistema absorber cantidades infinitesimales de energía (calor) puesto que existe un umbral discreto entre el nivel fundamental y el primer nivel excitado para nuestro sistema⁶. Para temperaturas altas, como los niveles de energía son acotados, una vez que el sistema absorbió el máximo de energía posible, no importa si el baño térmico que lo circunda aumenta su temperatura, pues ya no puede seguir incrementando su energía U .

El segundo ejemplo que abordaremos aquí consiste en el modelo de Einstein para el sólido cristalino. La descripción termodinámica que se busca para el sistema sólo tiene en cuenta las oscilaciones colectivas del sólido (ignorando la contribución de los electrones, efectos de superficie, etc.). Esto significa considerar únicamente las interacciones armónicas entre los N núcleos del sólido descrito, que se representan mediante un sistema de $3N$ osciladores acoplados. Sabemos que esto significa que habrá $3N$ modos normales (osciladores desacoplados), y la simplificación de este modelo es que todos los osciladores poseen la misma frecuencia ω_o ⁷.

⁶Para energías mayores también el cambio en U es discreto, pero es válido tomar como diferenciales los saltos $\Delta U \ll U$.

⁷Esta hipótesis se basa en el hecho de que son muchos más los modos que poseen frecuencia más alta frente a los de frecuencias bajas, y ω_o representa una especie de promedio pesado de todas las frecuencias.

El j -ésimo oscilador puede tomar valores de energía $E_j = (n_j + 1/2) \hbar\omega_o$ con $n_j = 0, 1, \dots$. La energía total del sistema será entonces

$$E_M = \left(M + \frac{3N}{2} \right) \hbar\omega_o, \quad \text{con} \quad M = \sum_{j=1}^{3N} n_j.$$

La pregunta entonces, dentro del esquema microcanónico es ¿de cuántas maneras $W(M)$ se puede lograr el valor de energía E_M ? En otras palabras, ¿de cuántas maneras pueden sumarse $3N$ números enteros mayores o iguales que 0 para obtener el valor M (fijo)? Otra vez la respuesta se logra mediante un análisis combinatorio que se deja como ejercicio. También en este caso se verifica que el calor específico se anula para $T = 0$, tal como lo predice la cuántica. Para el límite de altas temperaturas puede verse que el calor específico se vuelve constante, tal como debemos esperar de acuerdo con el teorema de equipartición de la energía que veremos en el próximo capítulo.

5. Ensamble canónico (basado en el texto de Reichl)

5.1. Sistemas cerrados

Los sistemas que pueden intercambiar calor con un baño térmico sin permitir el ingreso o egreso de partículas se describen mediante el *ensamble canónico*. Para construir este ensamble consideramos microestados con cualquier valor para la energía, aun cuando sabemos que macroscópicamente el valor medio $\langle E \rangle$ es una magnitud muy precisa y debe coincidir con la energía interna U .

5.1.1. Descripción clásica

Cuando el sistema analizado se describe adecuadamente mediante un enfoque clásico, debemos maximizar la expresión

$$S = -k \int_{\Gamma} d\mathbf{X}^N \rho(\mathbf{X}^N) \ln[C^N \rho(\mathbf{X}^N)],$$

agregando a la condición de normalización

$$\int_{\Gamma} d\mathbf{X}^N \rho(\mathbf{X}^N) = 1$$

la restricción para el valor medio de la energía

$$\int_{\Gamma} d\mathbf{X}^N H(\mathbf{X}^N) \rho(\mathbf{X}^N) = U.$$

Nuevamente utilizamos el método variacional acompañando estos dos vínculos con multiplicadores de Lagrange α_o y α_E respectivamente. Exigimos entonces que las variaciones virtuales alrededor del máximo correspondan a un extremo

$$\delta \left\{ \int_{\Gamma} d\mathbf{X}^N [\alpha_o \rho(\mathbf{X}^N) + \alpha_E H(\mathbf{X}^N) \rho(\mathbf{X}^N) - k \rho(\mathbf{X}^N) \ln[C^N \rho(\mathbf{X}^N)] \right\} = 0.$$

Como las variaciones $\delta\rho(\mathbf{X}^N)$ deben ser arbitrarias, esta condición se traduce como

$$\alpha_o + \alpha_E H(\mathbf{X}^N) - k \ln[C^N \rho(\mathbf{X}^N)] - k = 0, \quad (13)$$

de modo que

$$\rho(\mathbf{X}^N) = \frac{1}{C^N} e^{\left(\frac{\alpha_o}{k} - 1\right) + \frac{\alpha_E}{k} H(\mathbf{X}^N)}.$$

En esta expresión aún desconocemos cuánto deben valer los multiplicadores α_o y α_E . Para determinarlos es necesario recurrir a las condiciones de vínculo. En primer lugar, la condición de normalización nos conduce a definir la *función partición*

$$Z_N(\alpha_E, V) \equiv e^{1 - \frac{\alpha_o}{k}} = \frac{1}{C^N} \int_{\Gamma} d\mathbf{X}^N e^{\frac{\alpha_E}{k} H(\mathbf{X}^N)}.$$

Vemos que esta expresión relaciona los dos multiplicadores desconocidos, y para determinar el valor de α_E , multiplicamos la expresión (13) por $\rho(\mathbf{X}^N)$ y luego integramos en Γ , obteniendo

$$(\alpha_o - k) + \alpha_E U + S = -k \ln Z_N + \alpha_E U + S = 0 .$$

Esta relación entre U y S evoca la definición que aceptamos para el potencial de Helmholtz $F - U + TS = 0$, lo que sugiere identificar

$$\alpha_E = -\frac{1}{T} \quad \text{y} \quad F(T, V, N) = -kT \ln Z_N(T, V) .$$

De esta manera, introduciendo el parámetro $\beta \equiv 1/(kT)$, la función partición toma la forma

$$Z_N(T, X) = \frac{1}{C^N} \int_{\Gamma} d\mathbf{X}^N e^{-\beta H(\mathbf{X}^N)}$$

donde hemos reemplazado V por X para incluir el caso general de variable extensiva asociada con el trabajo mecánico. El ensamble canónico queda completamente determinado por la expresión anterior y la distribución de probabilidades, que finalmente toma la forma

$$\rho(\mathbf{X}^N) = \frac{1}{C^N Z_N(T, X)} e^{-\beta H(\mathbf{X}^N)}$$

La ecuación $F(T, X, N) = -kT \ln Z_N(T, X)$ provee toda la información termodinámica que posee nuestro sistema, pues se trata de la *relación fundamental* en la representación de Helmholtz. Esto significa que el objetivo de nuestra descripción de un sistema en el ensamble canónico se reduce a encontrar la función partición, ya que de allí se establece toda la conexión con la teoría termodinámica.

Es interesante destacar el comportamiento del sistema para temperaturas extremas, según la descripción dada por la expresión anterior. Por un lado, cuando $T \rightarrow 0$, es decir $\beta \rightarrow \infty$, $\rho \rightarrow 0$ para cualquier estado excepto el fundamental. Por otro lado, cuando $T \rightarrow \infty$ ($\beta \rightarrow 0$), ρ se hace constante, de modo que todos los estados microscópicos del sistema son igualmente probables: habrá estados con idéntica energía, pero ningún estado se puebla con preferencia cuando la temperatura se vuelve lo suficientemente alta.

Respecto de la expresión que hemos dado para la partición, también vale la pena enfatizar que deben agotarse todos los estados accesibles para el sistema, y no solo recorrer los niveles de energía. En otras palabras, también aquí puede ocurrir que muchos estados compartan la misma energía, por lo que en la integral involucrada en Z_N se repetirán muchas contribuciones idénticas. Suele escribirse la función partición como

$$Z_N(T, X) = \frac{1}{C^N} \int_{-\infty}^{\infty} dE g(E) e^{-\beta E} ,$$

donde $g(E)$ es la densidad de estados en el espacio de las fases con energía E , a menudo llamada *degeneración*: $g(E) dE$ da la probabilidad de que un estado elegido al azar en Γ tenga energía entre E y $E + dE$. Con esta notación, en términos de la energía E del sistema, la densidad de probabilidad se escribe análogamente como

$$\rho(E) = \frac{1}{C^N Z_N(T, X)} g(E) e^{-\beta E} .$$

Equipartición de la energía

En esta sección analizaremos una propiedad importante de sistemas *clásicos* para los cuales el hamiltoniano es *cadrático* en cada variable del espacio de las fases. Denotando como x_1, \dots, x_n a las componentes de $(\mathbf{q}^N, \mathbf{p}^N)$ que intervienen en el hamiltoniano (pueden ser solo algunas de ellas), podemos expresar

$$H(\mathbf{X}^N) = \sum_{j=1}^n \frac{1}{2} \alpha_j x_j^2 .$$

Los valores de las constantes α_j ($\neq 0$ todas) dependen de la situación particular de que se trate: si x_j representa un momento lineal, entonces α_j corresponde a la recíproca de la masa involucrada, $1/m$; si x_j designa una coordenada generalizada, entonces α_j es la constante elástica asociada a la fuerza generalizada de restitución.

El *teorema de equipartición de la energía* puede enunciarse de la siguiente manera:

Cada grado de libertad x_j contribuye a la energía interna con $\frac{1}{2} kT$, sin importar cuánto vale α_j .

En otras palabras, $U = nkT/2$, es decir la energía está igualmente distribuida en cada grado de libertad, y es proporcional a la temperatura. La demostración del teorema es muy sencilla, ya que la función partición se escribe como

$$Z_N = \frac{1}{C^N} \int_{\Gamma} d\mathbf{X}^N e^{-\frac{\beta}{2} \sum \alpha_j x_j^2} = B(N, X) \prod_{j=1}^n \sqrt{\frac{2\pi}{\beta \alpha_j}}.$$

La función $B(N, X)$ resulta de integrar aquellas q y p que no intervienen en el hamiltoniano, y esencialmente aquí nos interesa que no depende de la temperatura. Dejando como ejercicio probar que $U = -\partial(\ln Z_N)/\partial\beta$, puede utilizarse esta identidad para ver que

$$U = \frac{\partial}{\partial\beta} \left(\sum_{j=1}^n \ln \sqrt{\beta} \right) = \frac{n}{2} kT,$$

que es lo que se deseaba demostrar. A menudo el teorema de equipartición se verifica al analizar los calores específicos, que satisfacen la relación

$$N c_X = \left(\frac{\partial U}{\partial T} \right)_{X, N} = \frac{nk}{2}.$$

Para considerar algunos ejemplos sencillos, recurrimos primero al caso del gas ideal, cuyo hamiltoniano involucra los $3N$ impulsos lineales. Se deja como ejercicio ver que la función partición resulta

$$Z_N(T, V) = \frac{V^N}{C^N} (2\pi mkT)^{3N/2},$$

con lo cual

$$U = \frac{3}{2} NkT \quad ; \quad c_V = \frac{3}{2} k.$$

Este resultado coincide con el previsto por el teorema de equipartición de la energía, ya que en este caso los grados de libertad x_j del enunciado anterior son los $3N$ impulsos lineales.

En el caso de moléculas diatómicas clásicas, a los grados de libertad traslacionales deben agregarse las rotaciones, que a temperatura ambiente están presentes y pueden ser descriptas clásicamente. El hamiltoniano clásico suma entonces dos términos cuadráticos a cada molécula, que involucran los impulsos generalizados correspondientes a las dos coordenadas angulares que determinan la orientación de la molécula. La energía interna agrega ahora la contribución de estos $2N$ términos, de modo que $U = 5NkT/2$. Cuando además se activan los modos vibracionales, en el hamiltoniano se incorporan dos términos cuadráticos por molécula, pues las oscilaciones pueden separarse del movimiento del centro de masa mediante la coordenada relativa y el impulso generalizado correspondiente. De este modo, la energía interna resulta en este caso igual a $7NkT/2$.

5.1.2. Descripción cuántica

Para desarrollar la mecánica estadística desde un enfoque cuántico, el procedimiento es similar al del caso clásico. Ahora debe maximizarse la entropía

$$S = -k \text{Tr} (\hat{\rho} \ln \hat{\rho})$$

recurriendo a las expresiones adecuadas para los vínculos

$$\text{Tr} \hat{\rho} = 1 \quad \text{y} \quad \text{Tr} (\hat{H}_N \hat{\rho}) = U.$$

Se deja como ejercicio verificar que si acompañamos los vínculos con los mismos multiplicadores de Lagrange que en la sección anterior, la condición de máximo implica

$$\text{Tr} \left\{ \delta \hat{\rho} \left[(\alpha_o - k) \hat{I} + \alpha_E \hat{H}_N - k \ln \hat{\rho} \right] \right\} = 0.$$

También en este caso esta igualdad debe cumplirse para cualquier variación $\delta\hat{\rho}$, de modo que arribamos a las siguientes expresiones análogas a las del caso clásico:

$$Z_N(T, X) = \text{Tr} \left(e^{-\beta\hat{H}_N} \right) \quad ; \quad F(T, X, N) = -kT \ln Z_N(T, X) .$$

La solución para el operador densidad resulta

$$\hat{\rho} = \frac{e^{-\beta\hat{H}_N}}{\text{Tr} \left(e^{-\beta\hat{H}_N} \right)} .$$

En caso de tener una representación diagonal⁸, el elemento r -ésimo de $\hat{\rho}$ se escribe

$$\rho_r = \frac{e^{-\beta E_r}}{Z_N}$$

y representa la probabilidad de que el sistema se encuentre en el estado r .⁹

5.2. Fluctuaciones de la energía

Ahora que hemos desarrollado el formalismo correspondiente al ensamble canónico vale la pena preguntarnos en qué cambiará la descripción termodinámica provista desde este formalismo en relación a la dada en el ensamble microcanónico. Para ello es útil analizar cuánto puede fluctuar la energía de nuestro sistema al considerar en el ensamble canónico.

A partir de la expresión

$$\text{Tr} \left\{ e^{\beta[F(T, X, N) - \hat{H}_N]} \right\} = 1$$

derivamos con respecto a β y obtenemos

$$\text{Tr} \left\{ \left[\left(\frac{\partial \beta F}{\partial \beta} \right)_{X, N} - \hat{H}_N \right] e^{\beta(F - \hat{H}_N)} \right\} = 0 .$$

Derivando nuevamente, llegamos a la relación

$$\text{Tr} \left\{ \left[\left(\frac{\partial^2 \beta F}{\partial \beta^2} \right)_{X, N} + \left(\left(\frac{\partial \beta F}{\partial \beta} \right)_{X, N} - \hat{H}_N \right)^2 \right] e^{\beta(F - \hat{H}_N)} \right\} = 0 .$$

Recordando que $\partial(\beta F)/\partial\beta = \langle E \rangle$, vemos que las fluctuaciones de la energía se rigen según

$$\langle E^2 \rangle - \langle E \rangle^2 = - \left(\frac{\partial^2 \beta F}{\partial \beta^2} \right)_{X, N} = - \left(\frac{\partial U}{\partial \beta} \right)_{X, N} = kT^2 N c_x .$$

Como $\langle E \rangle = U$ es extensiva y por ende proporcional a N , en el límite termodinámico

$$\frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} \sim \frac{1}{\sqrt{N}} \rightarrow 0 .$$

Esto significa que las fluctuaciones relativas de la energía son muy pequeñas cuando $N \rightarrow \infty$, lo que implica que los únicos estados del sistema con probabilidad no nula son aquellos con $E \cong U$. Esto nos permite concluir que el ensamble canónico es equivalente al microcanónico, y la elección de uno u otro ensamble quedará supeditada a la conveniencia de la descripción en el formalismo escogido.

Físicamente podíamos prever esta equivalencia, ya que es idéntico pensar que prefijamos los valores de U , X y N (ensamble microcanónico) y procuramos el valor de T a través de una ecuación de estado (por ej., en la representación entropía), o imaginar que conocemos T , X y N (ensamble canónico) y de la relación fundamental obtenemos el valor de U en el equilibrio. La situación termodinámica no cambia en absoluto, aunque los microestados permitidos para el sistema ahora pueden abarcar diferentes energías.

5.3. Revisión de los postulados de la termodinámica

Antes de abordar otros temas, repasemos los postulados que habíamos introducido en el desarrollo de la teoría termodinámica.

⁸Recordemos que en el equilibrio $[\hat{H}_N, \hat{\rho}] = 0$, como vimos en §2.2.2

⁹Insistimos con el hecho de que puede haber muchos estados diferentes con la misma energía.

Postulado 1. *Existen estados de equilibrio caracterizados por la energía U , el volumen V (o la variable extensiva X asociada con el trabajo externo), y los números de moles $\{n_j\}$ de los componentes químicos del sistema.*

Hemos basado nuestro desarrollo aceptando este postulado, con la única salvedad de caracterizar la cantidad de materia por los números de partículas $\{N_j\}$ en lugar de los números de moles de cada componente, sabiendo que ambas representaciones son equivalentes. En realidad las dudas pueden haber surgido a partir de la aceptación de la entropía de información como la entropía termodinámica. Los postulados restantes son los que en realidad nos interesan en este sentido.

Postulado 2. *Existe la función entropía S , función de U, X y $\{N_j\}$, que toma el máximo valor posible en el equilibrio.*

En nuestro caso utilizamos la entropía de información, que toma el máximo valor admitido según los vínculos macroscópicos que se imponen, de manera que nuestro desarrollo es perfectamente compatible con este postulado.

Postulado 3. *La entropía de un sistema compuesto es aditiva, continua, derivable y creciente con U .*

En lo referente a la aditividad, hemos hecho el análisis correspondiente al presentar el ensamble microcanónico, y al relacionar la temperatura absoluta con la derivada de S respecto de U , convinimos que la entropía es creciente con U .

Postulado 4. *La entropía se anula a medida que T se aproxima al cero absoluto.*

La verificación de este postulado en nuestra formulación puede realizarse tanto en el ensamble microcanónico como en el canónico, pues la expresión para la entropía en este último caso involucra únicamente al estado fundamental, ya que a $T=0$ solo se puebla este estado. La entropía

$$S = k \ln W$$

queda determinada por la degeneración del estado fundamental. Esta puede valer 1 en el caso en que no esté degenerado hasta algún valor que a lo sumo será $W \sim N$, de manera que la entropía por partícula se anula en el límite termodinámico:

$$\frac{S}{N} \rightarrow 0 \quad \text{para} \quad N \rightarrow \infty .$$

El postulado de Nernst entonces también se satisface a partir de la asociación de la entropía estadística con la entropía termodinámica.

5.4. Sólido cristalino

5.4.1. Descripción clásica

Hemos visto que la expresión clásica para la función partición

$$Z_N(T, X) = \frac{1}{C^N} \int_{\Gamma} d\mathbf{X}^N e^{-\beta H(\mathbf{X}^N)}$$

provee una correcta conexión con la termodinámica en el caso de un gas ideal, verificándose además el teorema de equipartición de la energía. Para analizar el caso de las oscilaciones colectivas de una red cristalina, debemos recordar que los $3N$ osciladores acoplados pueden describirse mediante coordenadas normales como un sistema de $3N$ osciladores armónicos desacoplados (modos normales)

$$H(\mathbf{X}^N) = \sum_{i=1}^{3N} \left(\frac{1}{2m} p_i^2 + \frac{m\omega_i^2}{2} x_i^2 \right) .$$

La integración de la función partición es directa, teniendo en cuenta que todas las coordenadas e impulsos x_i y p_i son independientes:

$$Z_N(T, V) = \frac{1}{h^{3N}} \int d\mathbf{x}^{3N} d\mathbf{p}^{3N} e^{-\beta \sum \left(\frac{1}{2m} p_i^2 + \frac{m\omega_i^2}{2} x_i^2 \right)} = \prod_{i=1}^{3N} \left[\frac{1}{h} \int dp_i dx_i e^{-\beta \left(\frac{1}{2m} p_i^2 + \frac{m\omega_i^2}{2} x_i^2 \right)} \right] .$$

El hecho de que los osciladores sean independientes se traduce en la factorización de la función partición como producto de las “particiones individuales” z_i correspondientes a cada oscilador. Esta factorización podrá

realizarse en todo sistema que pueda separarse en componentes no interactuantes. En nuestro caso particular las “particiones individuales” se calculan mediante la expresión

$$z_i = \frac{1}{h} \int_{-\infty}^{+\infty} dp e^{-\beta \frac{p^2}{2m}} \int_{-L/2}^{+L/2} dx e^{-\beta \frac{m\omega_i^2}{2} x^2} .$$

Como la exponencial decae rápidamente para posiciones alejadas del equilibrio, podemos realizar la segunda integración entre $-\infty$ y $+\infty$, de manera que

$$z_i = \frac{1}{h} \sqrt{2\pi mkT} \sqrt{\frac{2\pi kT}{m\omega_i^2}} = \frac{2\pi}{h\omega_i} kT .$$

Entonces, independientemente de los valores que puedan tomar los ω_i , la descripción clásica del sólido cristalino predice que

$$U = 3NkT \quad \text{y} \quad c_v = 3k .$$

Estos resultados podrían haberse predicho mediante el teorema de equipartición de la energía.

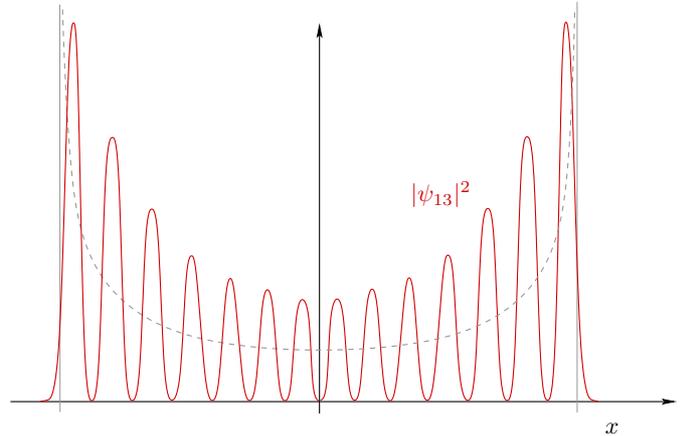
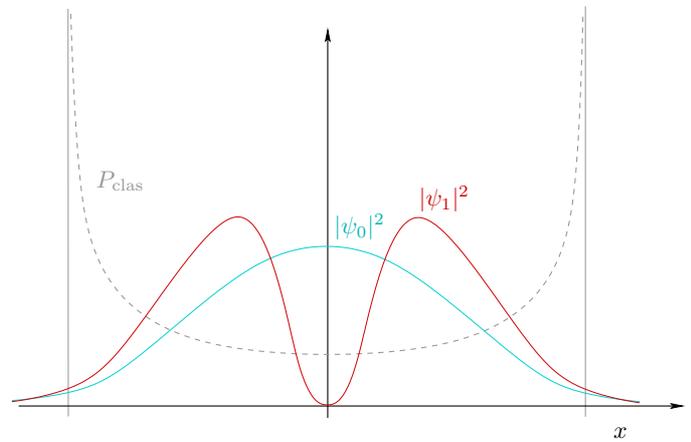
Vemos que esta descripción no reproduce satisfactoriamente el comportamiento para bajas temperaturas, pues el calor específico de un sólido debe anularse a medida que se reduce su temperatura, tal como se deduce a partir de consideraciones cuánticas. A temperaturas altas los estados más energéticos del oscilador cuántico se pueblan cada vez más, de manera que la correspondiente “distribución de presencia” se hace cada vez más parecida a la predicha por la clásica.

Para aclarar un poco esta idea recordemos que mientras la probabilidad de encontrar a la partícula clásica que oscila es proporcional a la inversa de la velocidad $v(x)$

$$P_{\text{clás}}(x) \propto \frac{1}{|v(x)|} = \frac{1}{+\omega \sqrt{x_o^2 - x^2}} ,$$

donde x_o es el punto de retorno clásico, la correspondiente probabilidad cuántica es el módulo de la función de onda al cuadrado $|\psi_n(x)|^2$. Para los primeros estados cuánticos, esta probabilidad se parece muy poco a la del oscilador clásico, tal como se muestra en la figura (aquí se ha normalizado el rango de coordenadas espaciales para que coincidan los puntos de retorno clásico.)

Para estados con n mayor, la distribución $|\psi_n(x)|^2$ se parece cada vez más a $P_{\text{clás}}(x)$, por lo cual ambas descripciones pueden resultar adecuadas a temperaturas altas, pero no a bajas temperaturas donde se privilegian los estados de menor energía.



5.4.2. Modelo de Einstein

La descripción cuántica retoma la idea de $3N$ osciladores desacoplados (distinguidos), para los cuales los autovalores de los hamiltonianos \hat{H}_i resultan $E_i = (n_i + 1/2) \hbar\omega_i$, o bien, en términos del operador número,

$$\hat{H}_i = \hbar\omega_i \left(\hat{n}_i + \frac{1}{2} \right) .$$

(Podemos pensar que el operador número se define a partir de esta relación, y se satisface que $[\hat{H}_i, \hat{n}_i] = 0$ y $\hat{n}_i |n_i\rangle = n_i |n_i\rangle$).

Dijimos que el modelo de Einstein consiste en hacer $\omega_i = \omega$ ($\forall i$), de manera que obtenemos

$$Z_N = \text{Tr} \left[e^{-\beta \hbar\omega \sum (\hat{n}_i + 1/2)} \right] .$$

Como las funciones de onda conjuntas que describen a los $3N$ osciladores independientes son simplemente productos tensoriales

$$|n_1, n_2, \dots, n_{3N}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \dots \otimes |n_{3N}\rangle,$$

la expresión para la partición se evalúa fácilmente en esta base, utilizando el hecho de que las funciones de onda individuales son ortonormales ($\langle n|m\rangle = \delta_{n,m}$):

$$Z_N = \sum_{n_1=0}^{\infty} \dots \sum_{n_{3N}=0}^{\infty} e^{-\beta\hbar\omega \sum (n_i+1/2)} = \prod_{i=1}^{3N} \left[e^{-\beta\hbar\omega/2} \sum_{n_i=0}^{\infty} (e^{-\beta\hbar\omega})^{n_i} \right],$$

es decir

$$Z_N(T, V) = \left(\frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \right)^{3N}. \quad (14)$$

Podemos entonces expresar la energía libre de Helmholtz como

$$F(T, V, N) = \frac{3}{2} N\hbar\omega + 3NkT \ln(1 - e^{-\beta\hbar\omega}).$$

El primer término representa la energía libre a $T = 0$, y no es cero, tal como lo predice el principio de incertidumbre de Heisenberg. Se sugiere como ejercicio verificar que esta expresión satisface la tercera ley de la termodinámica ($S \rightarrow 0$ para $T \rightarrow 0$).

Podemos evaluar la energía interna a partir de la expresión anterior:

$$U = \left(\frac{\partial \beta F}{\partial \beta} \right)_{V, N} = \frac{3}{2} N\hbar\omega + \frac{3N\hbar\omega}{e^{\beta\hbar\omega} - 1}.$$

El ingenuo lector no dudará en verificar fácilmente que esta expresión reproduce los resultados clásicos para temperaturas elevadas; para $T=0$, la energía interna nuevamente está dada por el primer término.

También puede obtenerse el calor específico por núcleo

$$c_v = \frac{1}{N} \left(\frac{\partial U}{\partial T} \right)_{V, N} = \frac{3(\hbar\omega)^2}{kT^2} \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2}.$$

En el régimen de bajas temperaturas, es decir, pensando $\beta \rightarrow \infty$, el comportamiento asintótico está dado por

$$c_v \sim 3k(\hbar\omega)^2 \frac{\beta^2}{e^{\beta\hbar\omega}} \quad (\rightarrow 0);$$

para temperaturas altas ($\beta \rightarrow 0$), al tomar el límite surgen un par de indeterminaciones que se resuelven mediante la regla de L'Hôpital, y puede verificarse (¡tú puedes hacerlo!) que se obtiene el comportamiento clásico $c_v = 3k$. Por supuesto, los comportamientos para temperaturas altas podían predecirse recurriendo al teorema de equipartición de la energía.

5.4.3. Modelo de Debye

Aunque cualitativamente la descripción dada por el modelo de Einstein es correcta, de acuerdo con los resultados experimentales c_v no debería decaer exponencialmente cuando $\beta \rightarrow \infty$ ($T \rightarrow 0$). Este inconveniente se resuelve evitando la simplificación de frecuencias idénticas para todos los modos normales:

$$\hat{H}_N = \sum_{i=1}^{3N} \hbar\omega_i \left(\hat{n}_i + \frac{1}{2} \right).$$

La idea de Debye fue suponer que los osciladores “conforman” una red de dimensiones características L_x, L_y, L_z , de modo que las longitudes de onda λ permitidas en estas oscilaciones serán $2L_x/m_x, 2L_y/m_y, 2L_z/m_z$, con $m_\alpha = 1, 2, \dots$. Admitiendo que entre las frecuencias ω y los vectores de onda \mathbf{k} , con $k \equiv |\mathbf{k}| = 2\pi/\lambda$, la relación de dispersión unidimensional es $\omega = ck$, donde c es la velocidad del sonido en el medio, en tres dimensiones tendremos

$$\omega_i^2 = c^2 \left[\left(\frac{\pi m_{xi}}{L_x} \right)^2 + \left(\frac{\pi m_{yi}}{L_y} \right)^2 + \left(\frac{\pi m_{zi}}{L_z} \right)^2 \right]$$

(estas son las ω_i que teníamos en la expresión anterior: por supuesto, algunas ternas (m_{xi}, m_{yi}, m_{zi}) definen frecuencias repetidas). Dado que el límite termodinámico implica $L_\alpha \rightarrow \infty$, estas ω_i pueden tomarse como

variables continuas. En el espacio ω , la distancia según cada coordenada α es $\pi c/L_\alpha$, de modo que el volumen correspondiente a cada modo será $(\pi c)^3/V$. Si ω_D es la frecuencia máxima posible, el volumen del octavo de esfera de radio ω_D ($\omega > 0$) dividido el volumen correspondiente a un estado es igual al número de modos normales posibles

$$\frac{\frac{1}{8} \left(\frac{4}{3} \pi \omega_D^3 \right)}{\frac{(\pi c)^3}{V}} = 3N \quad \Rightarrow \quad \omega_D^3 = \frac{18N\pi^2}{V} c^3 .$$

En realidad, hay en general 3 modos normales posibles con cada ω : uno *longitudinal* y dos *transversales*, con diferentes velocidades de propagación, c_ℓ y c_t . La igualdad anterior debe escribirse entonces como la suma de los modos transversales más el longitudinal para totalizar los $3N$ modos normales

$$\frac{1}{8} \left(\frac{4}{3} \pi \omega_D^3 \right) \left[2 \frac{V}{(\pi c_t)^3} + 1 \frac{V}{(\pi c_\ell)^3} \right] = 3N \quad \Rightarrow \quad \omega_D^3 = \frac{18N\pi^2}{V} \left(\frac{2}{c_t^3} + \frac{1}{c_\ell^3} \right)^{-1} .$$

La frecuencia ω_D se denomina *frecuencia de Debye*. La expresión anterior es también válida para obtener el número de modos $N(< \omega)$ con frecuencias menores que cualquier ω

$$N(< \omega) = \frac{1}{8} \left(\frac{4}{3} \pi \omega^3 \right) \left[2 \frac{V}{(\pi c_t)^3} + 1 \frac{V}{(\pi c_\ell)^3} \right] = \frac{3N}{\omega_D^3} \omega^3 ,$$

de manera que el número de modos con frecuencias entre ω y $\omega + d\omega$ es

$$dN(\omega) = \frac{9N}{\omega_D^3} \omega^2 d\omega . \quad (15)$$

Volviendo a la función partición, la expresión que reemplaza a la ecuación (14) es

$$Z_N(T, V) = \prod_{i=1}^{3N} \left[e^{-\beta \hbar \omega_i / 2} \sum_{n_i=0}^{\infty} (e^{-\beta \hbar \omega_i})^{n_i} \right] = \prod_{i=1}^{3N} \left(\frac{e^{-\beta \hbar \omega_i / 2}}{1 - e^{-\beta \hbar \omega_i}} \right) ,$$

de modo que tendremos

$$F(T, V, N) = \sum_{i=1}^{3N} \frac{\hbar \omega_i}{2} + kT \sum_{i=1}^{3N} \ln (1 - e^{-\beta \hbar \omega_i}) .$$

Teniendo presente que los valores de ω se tornan continuos, y utilizando la identidad (15), podemos hacer el reemplazo de sumatorias por integrales obteniendo la expresión

$$F(T, V, N) = \frac{9N\hbar}{2\omega_D^3} \int_0^{\omega_D} d\omega \omega^3 + \frac{9NkT}{\omega_D^3} \int_0^{\omega_D} d\omega \omega^2 \ln (1 - e^{-\beta \hbar \omega}) .$$

El primer término se integra directamente, y para la energía interna puede escribirse

$$U = \left(\frac{\partial \beta F}{\partial \beta} \right)_{V, N} = \frac{9}{8} N \hbar \omega_D + \frac{9N}{\omega_D^3} \int_0^{\omega_D} d\omega \frac{\hbar \omega^3}{e^{\beta \hbar \omega} - 1} .$$

Sustituyendo $x = \hbar \omega / kT$ podemos escribir la expresión anterior en términos de la *temperatura de Debye*¹⁰ $\theta \equiv \hbar \omega_D / k$

$$U = \frac{9}{8} N \hbar \omega_D + \frac{9N}{\omega_D^3} \frac{(kT)^4}{\hbar^3} \int_0^{\theta/T} dx \frac{x^3}{e^x - 1} .$$

La expresión para el calor específico resulta entonces

$$c_v = \frac{9}{\omega_D^3} \frac{k^4 T^3}{\hbar^3} \int_0^{\theta/T} dx \frac{x^4 e^x}{(e^x - 1)^2} .$$

Nuevamente, la predicción para altas temperaturas coincide con la clásica, pero la diferencia fundamental con el modelo de Einstein se da para bajas temperaturas, pues en ese caso

$$c_v \sim \frac{12}{5} \frac{k\pi^4}{\theta^3} T^3 ,$$

describiendo adecuadamente el comportamiento de las determinaciones experimentales.

¹⁰Algunos valores representativos de θ : para diamante, 3000 K; para aluminio, 400 K; para plomo, 90 K.

5.5. Moléculas diatómicas: el caso del hidrógeno (basado en el texto de Balian)

El análisis de las moléculas diatómicas resulta de especial interés porque permite avanzar en la comprensión de diferentes situaciones. Para ello, consideramos las moléculas de un gas diatómico como rotadores rígidos, conformados por dos átomos puntuales mantenidos a distancia fija. A temperatura ambiente, los grados de libertad electrónicos y de vibración no están excitados, de manera que esta aproximación resulta muy adecuada.

Los grados de libertad traslacionales se corresponden con los de una partícula cuántica en una caja (tridimensional) de lado a , de modo que los niveles de energía estarán dados por $[h^2/(8ma^2)](k^2+\ell^2+m^2)$ ($k, \ell, m \in \mathbb{Z} \geq 0$). Para una masa atómica del orden de 10^{-23} g y una caja de lado $a=1$ cm, la energía térmica relacionada con el salto entre dos niveles consecutivos de energía corresponde a una temperatura de 10^{-14} K, de modo que a temperatura ambiente, la contribución de los modos traslacionales a la energía interna puede calcularse mediante un enfoque clásico.

Nos concentramos entonces en la energía cinética de rotación, cuya expresión clásica, en términos del momento de inercia I es

$$\frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I \sin^2 \theta}.$$

Para masas m_1 y m_2 , si r_o es la distancia entre ambas, resulta

$$I = \frac{m_1 m_2}{m_1 + m_2} r_o^2.$$

Sabemos que los niveles de energía para un rotador cuántico se traducen directamente como

$$\epsilon_j = \frac{\hbar^2}{2I} j(j+1),$$

con degeneración $g_j = 2j+1$. Definiendo la *temperatura rotacional* $\Theta \equiv \hbar^2/(2Ik)$, la función partición por molécula estará dada por

$$Z_1 = \sum_{j=0}^{\infty} (2j+1) e^{-j(j+1)\Theta/T}.$$

Para el caso $T \gg \Theta$, podemos hacer la sustitución $x = j(j+1)\Theta/T$, resultando una variable continua en este rango de temperaturas. De este modo, obtenemos

$$Z_1 \approx \frac{T}{\Theta} \int_0^{\infty} dx e^{-x} = \frac{T}{\Theta}.$$

Si en cambio $T \ll \Theta$, tenemos

$$Z_1 \approx 1 + 3e^{-2\Theta/T} + 5e^{-6\Theta/T} + \dots$$

Así, para la energía media por molécula $\bar{\epsilon} = -\partial(\ln Z_1)/\partial\beta$ encontramos:

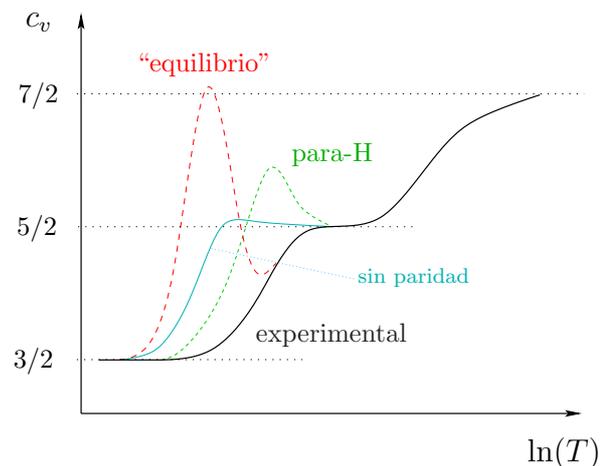
$$\begin{aligned} \cdot T \gg \Theta : \quad \bar{\epsilon} &= kT \\ \cdot T \ll \Theta : \quad \bar{\epsilon} &= k\Theta [6e^{-2\Theta/T} - 18e^{-4\Theta/T} + \dots] \end{aligned}$$

De este modo, los calores específicos resultan

$$\begin{aligned} \cdot T \gg \Theta : \quad c_v &= k \\ \cdot T \ll \Theta : \quad c_v &= k \left(\frac{\Theta}{T}\right)^2 [12e^{-2\Theta/T} - 72e^{-4\Theta/T} + \dots] \end{aligned}$$

Para el caso de la molécula de hidrógeno, los valores predichos para c_v sobreestiman demasiado los valores experimentales. Esta discrepancia *no* se debe al hecho de haber despreciado las energías de vibración y estados electrónicos excitados, no solo porque la diferencia se notaría únicamente a T muy altas, sino porque además en ese caso la predicción debería ir por debajo de las mediciones (pues faltaría agregar alguna contribución a c_v).

En cambio, lo que no se ha tenido en cuenta es el espín $1/2$ de los núcleos de hidrógeno, ni tampoco el hecho de que el espín total σ de las moléculas puede ser 0 (núcleos antiparalelos) o 1 (paralelos). El *para*-estado ($\sigma=0$) tiene degeneración $2\sigma+1=1$, mientras que el *orto*-estado ($\sigma=1$) tiene degeneración 3.



Cuando los átomos son distintos, en muy buena aproximación, la degeneración de cada estado no influye en el valor de la energía, porque esta será casi independiente de σ . En ese caso, la función partición debe multiplicarse por un factor constante $(2s_1+1)(2s_2+1)$ ($=4$ en nuestro caso).

Sin embargo, cuando los átomos son idénticos, debe tenerse en cuenta la *indistinguibilidad* de las partículas. El postulado de simetrización exige que la función de onda conjunta de la molécula sea antisimétrica al intercambiar los átomos. Las partes correspondientes a vibración y a estados electrónicos son simétricas en el estado base (no están excitadas). En el para-estado, la función de onda de espín es antisimétrica, mientras que en el orto-estado es simétrica. La función de onda rotacional es simétrica o antisimétrica según el número cuántico j sea par o impar ¹¹. O sea que tendremos:

$$\begin{aligned} \text{para-hidrógeno } (\sigma=0): & \quad j = 0, 2, 4, \dots & \text{ degeneración } 1; \\ \text{orto-hidrógeno } (\sigma=1): & \quad j = 1, 3, 5, \dots & \text{ degeneración } 3. \end{aligned}$$

Así, la función partición debe calcularse teniendo en cuenta que debe sumarse sobre todos los estados posibles:

$$Z_1 = Z_{\text{para}} + 3 Z_{\text{orto}},$$

donde Z_{para} abarca los j pares y Z_{orto} , los j impares. Teniendo en cuenta todos estos aspectos, la predicción para c_v sigue estando para el *para*. La solución definitiva se basa en analizar la razón de moléculas en ambos estados

$$r = \frac{N_{\text{orto}}}{N_{\text{para}}} = \frac{3 \sum_{j \text{ impar}} (2j+1) e^{-j(j+1)\Theta/T}}{1 \sum_{j \text{ par}} (2j+1) e^{-j(j+1)\Theta/T}}.$$

Según esta expresión, cuando $T \rightarrow 0$, $r \rightarrow 0$, y cuando $T \gg \Theta$, $r = 3$ (para calcularla, es válido el reemplazo de la sumatoria por una integral tanto en el numerador como en el denominador). A temperatura ambiente, $r = 3$ pues $T \simeq 4\Theta$ ($\approx 4 \times 85$ K). Al reducir la temperatura, r debería bajar, pero no es esto lo que sucede: la transición entre los estados orto y para involucra la inversión de espines, que tiene una probabilidad sumamente baja de ocurrir, y puede llevar años alcanzar el equilibrio en este sentido. Esto significa que típicamente las determinaciones experimentales se realizan *fuera del equilibrio*, con $r = 3$. Entonces puede considerarse al sistema como una mezcla de gases independientes, de manera que la contribución al calor específico se computa como la suma de los dos gases

$$c_v = c_v^{\text{para}} + c_v^{\text{orto}}.$$

Ahora sí, las predicciones concuerdan con el experimento.

Hay otras alternativas para verificar que el cálculo mediante este método es correcto. Una de ellas es medir el calor específico de mezclas con diferentes valores para r . Para hacerlo, el gas se prepara pasándolo por un catalizador (carbono activado) a diferentes temperaturas, con el objeto de acelerar la transición orto-para: en esos experimentos las predicciones concuerdan con los datos medidos.

Las mismas consideraciones que hemos hecho valen para el deuterio diatómico D_2 . En el caso del HD, por supuesto, no surgen inconvenientes, ya que los dos átomos son diferentes y las cuentas con que comenzamos eran válidas.

5.6. Acoplamiento mecánico con el exterior: sistemas magnéticos

Consideremos ahora el caso en que el sistema analizado, además de estar en contacto con un baño térmico, se encuentra bajo un campo Y constante, la variable intensiva relacionada con el trabajo externo. Tendremos que agregar ahora un nuevo vínculo al maximizar la entropía de información, correspondiente a mantener el valor medio $\langle X \rangle$ de la variable extensiva constante.

Los siguientes ejemplos se encuadran dentro de esta situación:

- gas en un recipiente a P constante \rightarrow se impone el vínculo sobre el valor medio $\langle V \rangle$ del volumen en el recipiente;
- sustancia magnética (o dieléctrica) a inducción externa B (campo eléctrico E) constante \rightarrow debe considerarse la restricción de momento magnético total M (polarización eléctrica \mathcal{P}) constante;
- cadena (molecular) a tensión constante \rightarrow la magnitud macroscópica que debe agregarse como vínculo es la longitud media $\langle L \rangle$.

¹¹Recordemos que ante inversiones, los armónicos esféricos se transforman según $Y_{jm}(\pi - \theta, \phi + \pi) = (-1)^j Y_{jm}(\theta, \phi)$.

Para afianzar ideas vamos a abordar el ejemplo concreto de sistemas magnéticos bajo inducción externa B . En el caso de que los núcleos interactúen entre ellos, una descripción sencilla del hamiltoniano del sistema acoplado con el campo externo corresponde al denominado *modelo de Ising*, cuyos autovalores son

$$H' = - \sum_{ij} J_{ij} \sigma_i \sigma_j - \mu_B B \sum_j \sigma_j ,$$

donde las magnitudes σ_i corresponden a los autovalores de los operadores de espín S_z (individuales) según la dirección del campo B , y μ_B es el momento magnético correspondiente a cada núcleo —en la discusión que sigue veremos la importancia de señalar este hamiltoniano como H' en lugar de H . El cálculo de la función partición nos lleva, como siempre, a recorrer todos los estados accesibles para el sistema

$$Z = \sum e^{-\beta H'} ,$$

y la energía libre de Gibbs queda determinada por

$$G = -kT \ln Z .$$

¿Por qué el potencial termodinámico que debe asociarse es el de Gibbs y no el de Helmholtz? Al maximizar la entropía, además de la condición de normalización y de fijar $U = \langle H \rangle$, agregamos el vínculo $M = \mu_B \sum \langle \sigma_i \rangle = N \mu_B \langle \sigma_i \rangle$. Dejando los detalles como ejercicio, la identificación que debemos hacer en el proceso variacional, es

$$-kT \ln Z = U - TS - BM ,$$

que es precisamente la energía libre de Gibbs. Para arribar a este resultado, es necesario distinguir el valor medio del hamiltoniano del sistema $U \equiv \langle H \rangle$ del valor medio $\langle H' \rangle$: en H' está incluido el campo B como constante, de manera que cualquier configuración $\{\sigma_j\}$ no modifica el valor de B , aunque drene o entregue energía magnética al enderezar espines. El hecho de que H' incluya las fuentes de B significa que

$$\langle H' \rangle = U - BM ,$$

es decir, está asociado con la *entalpía magnética*. Entonces, para el hamiltoniano “del sistema”, lo correcto es escribir

$$H = - \sum_{ij} J_{ij} \sigma_i \sigma_j ,$$

teniendo la adecuada asociación $\langle H \rangle = U$.

Conviene hacer hincapié en el hecho de que algunos textos presentan esta asociación de manera equívoca, refiriendo el potencial termodinámico como energía libre de Helmholtz; no obstante, se utilizan siempre las relaciones

$$M = - \left(\frac{\partial G}{\partial B} \right)_{T,N} \quad \text{o} \quad S = - \left(\frac{\partial G}{\partial T} \right)_{B,N} .$$

En el caso particular de espines no interactuantes, $J_{ij} = 0$, de manera que $U \equiv 0$. A diferencia de un gas ideal, al reducir el campo externo el sistema no guarda para sí ninguna energía, mientras que en el caso del gas, las moléculas conservan la energía mecánica que absorben como energía cinética.

6. Ensamble gran canónico (basado en el texto de Reichl)

6.1. Sistemas abiertos

En este capítulo consideramos sistemas que además de intercambiar calor con un baño térmico pueden intercambiar partículas con un reservorio. Estos sistemas están descriptos mediante el *ensamble gran canónico*, en el cual debemos considerar microestados que además de acceder a cualquier valor para la energía, pueden estar conformados por un número de partículas variable.

El procedimiento para maximizar la entropía es similar al de los casos anteriores, y sólo desarrollaremos aquí el formalismo cuántico, dejándose como ejercicio la obtención de la expresión adecuada para sistemas que pueden describirse clásicamente. Debemos considerar los vínculos impuestos por la condición de normalización

y el valor medio del hamiltoniano asociado con la energía interna del sistema, a los cuales se suma el número medio de partículas a través de la restricción

$$\text{Tr}(\hat{\rho} \hat{N}) = \langle N \rangle ,$$

donde \hat{N} es el operador número total de partículas. La condición de extremo para la entropía de información implica

$$\delta \left\{ \text{Tr} \left[\alpha_o \hat{\rho} + \alpha_E \hat{H} \hat{\rho} + \alpha_N \hat{N} \hat{\rho} - k \hat{\rho} \ln \hat{\rho} \right] \right\} = 0 ,$$

es decir

$$\text{Tr} \left\{ \delta \hat{\rho} \left[(\alpha_o - k) \hat{I} + \alpha_E \hat{H} + \alpha_N \hat{N} - k \ln \hat{\rho} \right] \right\} = 0 .$$

Como siempre, esta condición debe cumplirse para cualquier variación $\delta \hat{\rho}$, de manera que debe anularse todo el corchete de la expresión anterior:

$$(\alpha_o - k) \hat{I} + \alpha_E \hat{H} + \alpha_N \hat{N} - k \ln \hat{\rho} = 0 . \quad (16)$$

Al igual que en el ensamble canónico, la condición de normalización puede utilizarse para definir la función *gran partición*

$$Z(\alpha_E, \alpha_N) \equiv e^{1 - \frac{\alpha_o}{k}} = \text{Tr} \left[e^{\frac{\alpha_E}{k} \hat{H} + \frac{\alpha_N}{k} \hat{N}} \right] .$$

Esta expresión relaciona los tres multiplicadores de Lagrange que hemos introducido: las otras dos condiciones de vínculo deben bastar para encontrar los valores de α_E y α_N . Para hacerlo multiplicamos la ecuación (16) por $\hat{\rho}$ y tomamos la traza, obteniendo

$$-k \ln Z(\alpha_E, \alpha_N) + \alpha_E U + \alpha_N \langle N \rangle + S = 0 ,$$

o, lo que es equivalente,

$$-kT \ln Z(\alpha_E, \alpha_N) = U - TS - \mu \langle N \rangle \equiv \Omega(T, V, \mu) ,$$

si realizamos las identificaciones

$$\alpha_E = -\frac{1}{T} \quad \text{y} \quad \alpha_N = \frac{\mu}{T} .$$

De este modo, la función gran partición puede escribirse como

$$Z_\mu(T, V) \equiv \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} \right]$$

y el operador densidad toma la forma

$$\hat{\rho} = \frac{e^{-\beta(\hat{H} - \mu \hat{N})}}{\text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} \right]} .$$

Como dijimos más arriba, la conexión con la termodinámica se da a través del gran potencial

$$\Omega(T, V, \mu) = -kT \ln Z_\mu(T, V) .$$

6.2. Fluctuaciones en el ensamble gran canónico

Analicemos ahora las fluctuaciones en el número de partículas de nuestro sistema en equilibrio termodinámico. Teniendo presente que

$$\langle N \rangle = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V} ,$$

a partir de la igualdad

$$\text{Tr} \hat{\rho} = \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N} - \Omega)} \right] = 1 ,$$

derivamos dos veces con respecto a μ . Obtenemos entonces

$$\left\langle (N - \langle N \rangle)^2 \right\rangle = -kT \left(\frac{\partial^2 \Omega}{\partial \mu^2} \right)_{T, V} = kT \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T, V} .$$

Como la última derivada conserva el carácter extensivo, será proporcional a $\langle N \rangle$, de manera que las fluctuaciones *relativas* satisfacen

$$\frac{\sqrt{\left\langle (N - \langle N \rangle)^2 \right\rangle}}{\langle N \rangle} \sim \frac{1}{\sqrt{\langle N \rangle}} .$$

Esto significa que en el límite termodinámico, las fluctuaciones del número de partículas son excesivamente pequeñas. Este resultado suele invocarse como *equivalencia de ensambles* y nos permite elegir con libertad la descripción de un sistema en equilibrio a partir del ensamble que resulte más conveniente.

La expresión anterior puede explicitarse aun más considerando la densidad de partículas por unidad de volumen

$$n \equiv \frac{\langle N \rangle}{V} = \left(\frac{\partial \langle N \rangle}{\partial V} \right)_{T, \mu} = \left(\frac{\partial \langle N \rangle}{\partial V} \right)_{T, P}.$$

Valiéndonos del potencial de Helmholtz, cuya expresión diferencial en el contexto del ensamble gran canónico es $dF = -S dT - P dV + \mu d\langle N \rangle$, hallamos la relación de Maxwell

$$\left(\frac{\partial \mu}{\partial V} \right)_{T, \langle N \rangle} = - \left(\frac{\partial P}{\partial \langle N \rangle} \right)_{T, V} \quad \Leftrightarrow \quad \left(\frac{\partial V}{\partial \mu} \right)_{T, \langle N \rangle} = - \left(\frac{\partial \langle N \rangle}{\partial P} \right)_{T, V}.$$

Podemos escribir entonces

$$\left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T, V} = - \left(\frac{\partial \langle N \rangle}{\partial V} \right)_{T, \mu} \left(\frac{\partial V}{\partial \mu} \right)_{T, \langle N \rangle} = +n \left(\frac{\partial \langle N \rangle}{\partial P} \right)_{T, V} = -n \left(\frac{\partial \langle N \rangle}{\partial V} \right)_{T, P} \left(\frac{\partial V}{\partial P} \right)_{T, \langle N \rangle} = \frac{\langle N \rangle^2}{V} \kappa_T,$$

de manera que

$$\langle (N - \langle N \rangle)^2 \rangle = \frac{kT \langle N \rangle^2}{V} \kappa_T.$$

El hecho de que la compresibilidad κ_T no diverge, excepto en las proximidades de un punto crítico, nos permite así confirmar que las fluctuaciones (relativas) alrededor de $\langle N \rangle$ son despreciables.

6.3. Partículas idénticas (basado en el texto de Balescu)

En un sistema de N partículas indistinguibles, debe exigirse que la función de onda conjunta provea la misma densidad de probabilidad al intercambiar dos (o más) partículas cualesquiera k y ℓ :

$$|\Psi(\dots, x_k, \dots, x_\ell, \dots; t)|^2 = |\Psi(\dots, x_\ell, \dots, x_k, \dots; t)|^2.$$

Esto implica que la función de onda debe permanecer inalterada o a lo sumo cambiar de signo, condición que suele resumirse mediante la notación

$$\Psi(\dots, x_k, \dots, x_\ell, \dots; t) = \theta \Psi(\dots, x_\ell, \dots, x_k, \dots; t),$$

con $\theta = +1$ para el caso de bosones (partículas con espín entero) y $\theta = -1$ para fermiones (espín semientero). Esta condición es conocida como *postulado de simetrización*: en el caso de bosones la función de onda conjunta debe ser *simétrica* ante permutaciones de las partículas, y para fermiones, *antisimétrica*.

Usualmente comenzamos nuestra descripción escribiendo Ψ en términos de productos tensoriales de las funciones de onda individuales

$$\Psi(x_1, \dots, x_N; t) = \sum_{m_1} \cdots \sum_{m_N} c(m_1, \dots, m_N; t) \varphi_{m_1}(x_1) \cdots \varphi_{m_N}(x_N). \quad (17)$$

Aquí los m_j pueden ser vectores de números cuánticos que representan los posibles estados individuales, y las $\varphi_m(x)$ son las autofunciones estacionarias individuales. Como el producto tensorial $\varphi_{m_1}(x_1) \cdots \varphi_{m_N}(x_N)$ (o $|m_1\rangle \otimes |m_2\rangle \otimes \cdots \otimes |m_N\rangle$ en la notación de Dirac) no satisface de antemano ninguna simetría, sobre los coeficientes $c(m_1, \dots, m_N; t)$ debe imponerse la condición

$$c(\dots, m_k, \dots, m_\ell, \dots; t) = \theta c(\dots, m_\ell, \dots, m_k, \dots; t),$$

para satisfacer el postulado de simetrización. De aquí se deduce directamente el *principio de exclusión de Pauli*, pues en el caso de fermiones la única posibilidad para que se cumpla la identidad anterior cuando dos estados son iguales a m es $c(\dots, m_k = m, \dots, m_\ell = m, \dots) = 0$, con lo cual es imposible que dos fermiones ocupen el mismo estado.

En este contexto, vemos que lo que importa es cuántas partículas n_ℓ se encuentran en el estado m_ℓ , pues no podemos distinguir cuál es cuál. En otros términos, las variables “naturales” para caracterizar un estado son los *números de ocupación* n_ℓ . La densidad de probabilidad en términos de los $\{n_\ell\}$ debe ser equivalente a la descripción basada en el conjunto de números cuánticos $\{m_j\}$, es decir

$$|c(n_1, \dots, n_i, \dots; t)|^2 = \sum |c(m_1, \dots, m_N; t)|^2,$$

donde la sumatoria de la derecha debe abarcar todos los estados que generan el mismo conjunto de números de ocupación de la izquierda. Como las partículas son indistinguibles, los $N!/\prod_i(n_i!)$ sumandos son iguales, de modo que la condición anterior puede escribirse

$$|C(n_1, \dots, n_i, \dots; t)|^2 = \frac{N!}{\prod_i(n_i!)} |c(m_1, \dots, m_N; t)|^2.$$

Esta relación nos permite pasar del formalismo habitual a la representación de los números de ocupación. Esta última es comúnmente referida como “segunda cuantización”.

En casi todos los casos de interés los n_ℓ son infinitos, ya que el número de niveles accesibles a una partícula también lo es. Por supuesto, en cada estado, sólo algunos n_ℓ serán distintos de cero, ya que debe satisfacerse la condición

$$\sum_\ell n_\ell = N.$$

El formalismo de la segunda cuantización

La función de onda del sistema debe satisfacer la ecuación de Schrödinger

$$i\hbar \frac{\partial}{\partial t} \Psi(x_1, \dots, x_N; t) = \hat{H} \Psi(x_1, \dots, x_N; t).$$

En el caso en que el hamiltoniano sea separable

$$\hat{H} = \sum_{j=1}^N \hat{I}_1 \otimes \hat{I}_2 \cdots \otimes \hat{h}^{(j)} \otimes \cdots \otimes \hat{I}_N = \sum_{j=1}^N \hat{H}^{(j)} \quad (\text{simplificando notación}),$$

podemos escribir la ecuación de Schrödinger en la notación de Dirac

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = \sum_{j=1}^N \left[\hat{I}_1 \otimes \hat{I}_2 \cdots \otimes \hat{h}^{(j)} \otimes \cdots \otimes \hat{I}_N \right] |\Psi\rangle.$$

Utilizando la propiedad (ejercicio)

$$\langle m_1, \dots, m_N | \hat{H} | m'_1, \dots, m'_N \rangle = \sum_{j=1}^N \langle m_j | \hat{h}^{(j)} | m'_j \rangle \prod_{r=1, r \neq j}^N \delta_{m_r, m'_r},$$

se obtiene, a partir de (17), la ecuación equivalente

$$i\hbar \frac{\partial}{\partial t} c(m_1, \dots, m_N; t) = \sum_{j=1}^N \sum_{m'_j} \langle m_j | \hat{h}^{(j)} | m'_j \rangle c(m_1, \dots, m'_j, \dots, m_N; t).$$

Esta identidad nos dice que la evolución temporal de los coeficientes $c(m_1, \dots, m_N; t)$ está determinada por las probabilidades de transición de cada partícula j para pasar del estado m'_j al estado m_j (siempre sumando sobre todas las transiciones posibles).

Esta misma idea puede darse en términos de los números de ocupación. Pensemos en el caso en que antes de la transición hay $n_{m'_j}$ partículas en el estado m'_j y n_{m_j} en el estado m_j , y después de la transición hay $(n_{m'_j} - 1)$ partículas en el estado m'_j y $(n_{m_j} + 1)$ en el estado m_j . Resulta natural pensar este proceso asociándolo con la *aniquilación* de una partícula en el estado m'_j juntamente con la *creación* de otra en el estado m_j . Los operadores básicos del formalismo de la segunda cuantización actúan precisamente mediante esos procesos: creación y aniquilación.

Si adoptamos una base ortonormal $\{\phi(\{n_m\})\}$ con un producto interno asociado de manera que

$$(\phi(\dots, n_m, \dots), \phi(\dots, n'_m, \dots)) = \prod_m \delta_{n_m, n'_m},$$

el *operador aniquilación* se define mediante la relación

$$\hat{a}_m \phi(\dots, n_m, \dots) = \sqrt{n_m} \phi(\dots, n_m - 1, \dots)$$

y el *operador creación* a través de

$$\hat{a}_m^\dagger \phi(\dots, n_m, \dots) = \sqrt{n_m + 1} \phi(\dots, n_m + 1, \dots).$$

(De este modo, para cada estado m hay un par de operadores $\hat{a}_m, \hat{a}_m^\dagger$.)

Si bien la notación puede parecer enrevesada, *adan arap se ol on*. Se deja como ejercicio verificar que cada uno de estos operadores es el conjugado hermitiano del otro. Evidentemente, esto significa que \hat{a}_m y \hat{a}_m^\dagger no son hermitianos.

Con estos nuevos elementos, resulta natural incorporar el operador

$$\hat{\mathcal{H}} = \sum_{m,m'} \langle m | \hat{h}^{(j)} | m' \rangle \hat{a}_m^\dagger \hat{a}_{m'}$$

para producir las transiciones que procurábamos. En esta última expresión hemos eliminado los índices j en virtud de que las partículas son indistinguibles. Centraremos nuestra atención en el *operador número* para el estado m , cuya definición natural es

$$\hat{a}_m^\dagger \hat{a}_m \phi(\dots, n_m, \dots) = n_m \phi(\dots, n_m, \dots).$$

Las autofunciones de este operador son estados con números de ocupación $\{n_j\}$ bien definidos.

La conservación del número de partículas permite definir el operador número total

$$\sum_m \hat{a}_m^\dagger \hat{a}_m = \hat{N},$$

que, al actuar sobre funciones con $\{n_j\}$ bien definidos nos devuelve el número total de partículas del sistema.

Como

$$\hat{a}_m \hat{a}_m^\dagger \phi(\dots, n_m, \dots) = (n_m + 1) \phi(\dots, n_m, \dots),$$

es obvio que \hat{a}_m y \hat{a}_m^\dagger no conmutan. Para agilizar la pluma, se deja como ejercicio sencillo verificar que se cumplen las siguientes relaciones:

$$\begin{aligned} [\hat{a}_m, \hat{a}_{m'}] &= 0 \\ [\hat{a}_m^\dagger, \hat{a}_{m'}^\dagger] &= 0 \\ [\hat{a}_m, \hat{a}_{m'}^\dagger] &= \delta_{m,m'} \end{aligned}$$

Un hecho importante de este formalismo es que existe un “estado vacío” $\phi_o \equiv \phi(0, 0, \dots)$. Cuando consideramos este estado, se cumple que $\hat{a}_m \phi_o = 0$ para cualquier valor de m . La base $\phi(\dots, n_m, \dots)$ puede construirse aplicando sucesivamente \hat{a}^\dagger a ϕ_o . Por ejemplo, $\hat{a}_m^\dagger \phi_o$ es un estado normalizado con 1 partícula; $(1/\sqrt{1 + \delta_{m,m'}}) \hat{a}_m^\dagger \hat{a}_{m'}^\dagger \phi_o$ es un estado normalizado correspondiente a dos partículas, etc.

Estos estados son automáticamente simétricos ante permutaciones de partículas en estados m y m' arbitrarios, y vemos que no hay ninguna restricción sobre el número de partículas n_m que pueden habitar un estado cualquiera. Esta característica permite construir directamente el caso general para un sistema de bosones:

$$\Psi = c_o \phi_o + \sum_m c_m \hat{a}_m^\dagger \phi_o + \sum_{m,m'} c(m, m') \hat{a}_m^\dagger \hat{a}_{m'}^\dagger \phi_o + \dots$$

Este formalismo se denomina *representación de Fock*, y la expresión anterior es válida para cualquier caso en que el número total de partículas fluctúe: resulta muy adecuada para la descripción de fotones y fonones, y también para los estados que nos permiten construir el ensamble gran canónico. En el caso en que N es constante, la expresión anterior se reduce a

$$\Psi = \sum_{m_1} \dots \sum_{m_N} \frac{1}{\sqrt{\prod_i n_i!}} c(m_1, \dots, m_N) \hat{a}_{m_1}^\dagger \dots \hat{a}_{m_N}^\dagger \phi_o.$$

En esta suma puede haber términos con m_i iguales; los denominadores $\sqrt{n_i!}$ provienen del hecho de que al aplicar n_i veces el operador $\hat{a}_{m_i}^\dagger$ no se tiene un estado normalizado, como habíamos planteado en un comienzo.

Para el caso de fermiones deben redefinirse los operadores creación y aniquilación de manera que sólo permitan que los números de ocupación adopten los valores 0 ó 1. Si ordenamos los estados individuales de manera que puedan tomarse $\varepsilon_{m_1} \leq \varepsilon_{m_2} \leq \dots \leq \varepsilon_{m_j} \dots$, el número de estados ocupados por debajo del nivel m puede escribirse como

$$s_m = \sum_{j=m_1}^m n_j,$$

con lo cual los operadores aniquilación y creación pueden definirse respectivamente como

$$\begin{aligned}\hat{\alpha}_m \phi(\dots, n_m, \dots) &= (-1)^{s_m} n_m \phi(\dots, n_m - 1, \dots) \quad \text{y} \\ \hat{\alpha}_m^\dagger \phi(\dots, n_m, \dots) &= (-1)^{s_m} (1 - n_m) \phi(\dots, n_m + 1, \dots).\end{aligned}$$

Puede verse fácilmente que de esta manera se respetan adecuadamente las reglas de exclusión y la antisimetría requerida para la función de onda Ψ al reproducir la construcción que hicimos para bosones.

6.4. Gases ideales cuánticos (basado en el texto de Reichl)

Hemos visto ya que, al estudiar sistemas abiertos en equilibrio termodinámico mediante el ensamble gran canónico, las fluctuaciones en el número de partículas son despreciables, por lo que es equivalente el planteo de cualquier problema en el ensamble canónico o en el gran canónico. Esto nos permite abordar el estudio de los gases ideales cuánticos en el formalismo gran canónico, evitando inconvenientes algebraicos muy complicados para resolver en el ensamble canónico (o microcanónico).

Partimos entonces de la expresión para la gran partición

$$Z_\mu(T, V) = \text{Tr} \left[e^{-\beta(\hat{H} - \mu\hat{N})} \right].$$

En la representación de Fock, esta traza se evalúa fácilmente, pues \hat{H} y \hat{N} son diagonales. Si $\epsilon_\ell = \hbar^2 k_\ell^2 / (2m)$ es la energía cinética de una partícula en el estado $\mathbf{p}_\ell = \hbar \mathbf{k}_\ell$, para el caso general obtenemos

$$Z_\mu(T, V) = \sum_{n_o=0}^{\infty} \cdots \sum_{n_j=0}^{\infty} \cdots \sum_{n_\infty=0}^{\infty} S(N; n_o, \dots, n_\infty) e^{-\beta \sum_{\ell=0}^{\infty} n_\ell (\epsilon_\ell - \mu)},$$

donde el factor $S(N; n_o, \dots, n_\infty)$ representa el número de configuraciones diferentes con esos números de ocupación o poblaciones y debe cumplirse $\sum_{\ell=0}^{\infty} n_\ell = N$. La diferencia entre los comportamientos que se describen reside en este factor estadístico S , que tiene en cuenta si deben satisfacerse las estadísticas de Maxwell-Boltzmann, Bose-Einstein o Fermi-Dirac.

Definimos las partículas de Maxwell-Boltzmann como aquellas cuya función de onda conjunta no satisface el postulado de simetrización, sino que pueden pensarse como partículas *distinguibles* que obedecen la cuántica, a las que aplicamos el conteo correcto de Boltzmann introducido en §4.1.1 (ese misterioso denominador $N!$). Si bien este enfoque no representa ningún sistema físico, reviste cierto interés académico que se evidenciará pronto.

En el caso de partículas distinguibles, la expresión para el factor estadístico se obtiene de manera directa:

$$S_{\text{dist}}(N; n_o, \dots, n_\infty) = \frac{N!}{n_o! \dots n_j! \dots n_\infty!}.$$

Con este factor la función gran partición adopta una expresión algo difícil de evaluar en virtud del numerador $N!$, por lo que en este caso generalmente se utiliza el ensamble canónico. Sin embargo, para el caso de Maxwell-Boltzmann, el conteo correcto simplifica las cuentas, resultando

$$Z_\mu^{(\text{MB})}(T, V) = \sum_{n_o=0}^{\infty} \cdots \sum_{n_\infty=0}^{\infty} \frac{1}{n_o! \dots n_\infty!} e^{-\beta \sum_{\ell=0}^{\infty} n_\ell (\epsilon_\ell - \mu)}. \quad (18)$$

En la próxima sección veremos que esta expresión puede evaluarse de manera muy directa.

Para el caso de bosones, cualquier conjunto $\{n_j\}$ representa un único estado posible, sin restricción alguna sobre los valores de los números de ocupación, lo que implica que siempre el factor estadístico S vale 1:

$$Z_\mu^{(\text{BE})}(T, V) = \sum_{n_o=0}^{\infty} \cdots \sum_{n_\infty=0}^{\infty} e^{-\beta \sum_{\ell=0}^{\infty} n_\ell (\epsilon_\ell - \mu)}. \quad (19)$$

Para temperaturas suficientemente altas, todos los números de ocupación valen a lo sumo 1, ya que todos los estados son igualmente probables, de modo que coinciden las descripciones de Maxwell-Boltzmann y la de Bose-Einstein.

Para fermiones, también cada conjunto $\{n_j\}$ representa un único estado, pero ahora deben restringirse los números de ocupación posibles a 0 ó 1:

$$Z_\mu^{(\text{FD})}(T, V) = \sum_{n_o=0}^1 \cdots \sum_{n_\infty=0}^1 e^{-\beta \sum_{\ell=0}^{\infty} n_\ell (\epsilon_\ell - \mu)}. \quad (20)$$

Claramente, todas las particiones coinciden a T suficientemente altas.

6.4.1. Partículas de Maxwell-Boltzmann

La expresión (18) puede factorizarse fácilmente, pues se trata de una suma sobre índices independientes sobre la exponencial de una sumatoria, obteniendo

$$Z_{\mu}^{(\text{MB})}(T, V) = \prod_{\ell=0}^{\infty} \exp \left[e^{-\beta(\epsilon_{\ell} - \mu)} \right].$$

A partir de aquí, el gran potencial resulta

$$\Omega_{\text{MB}}(T, V, \mu) = -kT \ln Z_{\mu}^{(\text{MB})}(T, V) = -kT \sum_{\ell=0}^{\infty} e^{-\beta(\epsilon_{\ell} - \mu)}.$$

Para V suficientemente grande (límite termodinámico), la distribución de estados se hace continua, pues el espaciamiento de los autovalores de \mathbf{p}_{ℓ} permitidos es inversamente proporcional a V , de modo que las sumatorias se convierten en integrales

$$\sum_{\ell} \rightarrow \frac{V}{(2\pi)^3} \int d^3\mathbf{k} = \frac{V}{h^3} \int d^3\mathbf{p}.$$

La expresión para el gran potencial puede entonces escribirse

$$\Omega_{\text{MB}}(T, V, \mu) = -\frac{kTV}{h^3} \int d^3\mathbf{p} e^{-\beta(\frac{p^2}{2m} - \mu)} = -\frac{kTV}{h^3} 4\pi \int_0^{\infty} dp p^2 e^{-\beta(\frac{p^2}{2m} - \mu)},$$

es decir

$$\Omega_{\text{MB}}(T, V, \mu) = -kTV e^{\beta\mu} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2}.$$

Suele introducirse aquí la *longitud de onda térmica*

$$\lambda \equiv \sqrt{\frac{2\pi\hbar^2}{mkT}},$$

que mide el “desparramo” de la función de onda de las partículas. Un ejercicio muy loable es comparar esta magnitud con la longitud de onda de De Broglie. Si λ es mucho menor que el espaciamiento medio entre partículas, todas las estadísticas predicen los mismos resultados: al no haber solapamientos de las funciones de onda individuales, no influye la indistinguibilidad de las partículas en la física del sistema.

Utilizando esta definición, la expresión para el gran potencial resulta

$$\Omega_{\text{MB}}(T, V, \mu) = -\frac{kTV e^{\beta\mu}}{\lambda^3}.$$

Esta identidad puede tomarse como relación fundamental, de la que pueden derivarse las diferentes ecuaciones de estado. Por ejemplo, para el número medio de partículas tenemos

$$\langle N \rangle = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V} = V e^{\beta\mu} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} = \frac{V e^{\beta\mu}}{\lambda^3}.$$

Despejando el potencial químico de esta igualdad, se obtiene

$$\mu = kT \ln \left(\frac{\langle N \rangle}{V} \lambda^3 \right).$$

Para la entropía tenemos

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V, \mu} = \left(\frac{5}{2}k - \frac{\mu}{T} \right) \frac{V e^{\beta\mu}}{\lambda^3} = \frac{5k\langle N \rangle}{2} - k\langle N \rangle \ln \left(\frac{\langle N \rangle}{V} \lambda^3 \right).$$

Reobtenemos de este modo la cautivante ecuación de Sackur-Tetrode.

Finalmente, para la presión del gas resulta

$$P = - \left(\frac{\partial \Omega}{\partial V} \right)_{T, \mu} = \frac{kT e^{\beta\mu}}{\lambda^3} = \frac{\langle N \rangle kT}{V}.$$

Afortunadamente, los resultados son los mismos que predijimos en otras ocasiones con la misma convicción.

7. Gas de Bose-Einstein (basado en los textos de Reichl y Huang)

Hemos visto que la gran partición para un sistema de partículas regido por la estadística de Bose-Einstein está dada por la ecuación (19), la que puede reescribirse como

$$Z_{\mu}^{(\text{BE})}(T, V) = \prod_{\ell'=0}^{\infty} \left[\sum_{n_{\ell'}=0}^{\infty} e^{-\beta(\epsilon_{\ell'}-\mu)n_{\ell'}} \right],$$

pues los índices $n_{\ell'}$ se suman independientemente unos de otros. Veremos en seguida que para garantizar la convergencia de estas sumas debe cumplirse que $\mu < \epsilon_{\ell'} \forall \ell'$, de manera que podemos tomar el corchete como una suma geométrica, resultando

$$Z_{\mu}^{(\text{BE})}(T, V) = \prod_{\ell'=0}^{\infty} \left[\frac{1}{1 - e^{-\beta(\epsilon_{\ell'}-\mu)}} \right].$$

Conviene aclarar aquí que ℓ' debe recorrer cada uno de los estados accesibles a las partículas. Como la energía depende solamente de los autovalores del impulso y no de la proyección del espín s de las partículas, podemos pensar que la sumatoria recorre los distintos estados con dos índices: ℓ para \mathbf{p}_{ℓ} y $m_s (= -s, -s+1, \dots, s)$ para las coordenadas de espín. Por otro lado, las sumatorias no involucran a m_s , de manera que tendremos un producto de $g_s = 2s+1$ sumatorias idénticas, lo que puede resumirse en un exponente g_s en el último corchete:

$$Z_{\mu}^{(\text{BE})}(T, V) = \prod_{\ell=0}^{\infty} \left[\frac{1}{1 - e^{-\beta(\epsilon_{\ell}-\mu)}} \right]^{g_s}.$$

De aquí en adelante trataremos el caso de partículas sin espín ($s=0$), aunque la generalización para $s \neq 0$ es directa. Obtenemos entonces para el gran potencial

$$\Omega_{\text{BE}}(T, V, \mu) = kT \sum_{\ell=0}^{\infty} \ln \left[1 - e^{-\beta(\epsilon_{\ell}-\mu)} \right],$$

y para el número medio de partículas

$$\langle N \rangle = - \left(\frac{\partial \Omega_{\text{BE}}}{\partial \mu} \right)_{T, V} = \sum_{\ell=0}^{\infty} \frac{1}{e^{\beta(\epsilon_{\ell}-\mu)} - 1} = \sum_{\ell=0}^{\infty} \langle n_{\ell} \rangle,$$

donde $\langle n_{\ell} \rangle$ es el número medio de partículas en el estado \mathbf{p}_{ℓ} . Definiendo la *fugacidad* $z \equiv e^{\beta\mu}$, identificamos

$$\langle n_{\ell} \rangle = \frac{z}{e^{\beta\epsilon_{\ell}} - z}.$$

Como las poblaciones medias $\langle n_{\ell} \rangle$ no pueden ser negativas, debe cumplirse $e^{\beta\epsilon_{\ell}} > z \forall \ell$. Dado que el mínimo valor para ϵ_{ℓ} es cero, podemos establecer cotas para la fugacidad y, por ende, para el potencial químico

$$0 < z < 1 \quad \Rightarrow \quad \mu < 0.$$

En particular, vemos que la población del estado fundamental

$$\langle n_o \rangle = \frac{z}{1 - z}$$

puede diverger si z se aproxima a 1. Si ocurriera, ese estado se poblaría *macroscópicamente*, lo que, como veremos, se traduce como una transición de fase.

Podemos reescribir las expresiones para el gran potencial y el número medio de partículas tomando el límite termodinámico, de manera que convertimos las sumatorias en integrales siguiendo el procedimiento habitual, en términos del impulso lineal p o de la energía individual $\epsilon = p^2/(2m)$

$$\sum_{\ell} \rightarrow \frac{V}{h^3} \int d^3\mathbf{p} = \frac{4\sqrt{2}\pi m^{3/2}V}{h^3} \int \sqrt{\epsilon} d\epsilon.$$

Obtenemos entonces

$$\frac{\Omega_{\text{BE}}(T, V, \mu)}{V} = \frac{kT}{V} \ln(1 - z) - \frac{kT}{\lambda^3} g_{5/2}(z)$$

y para la recíproca del volumen específico v por partícula

$$\frac{1}{v} \equiv \frac{\langle N \rangle}{V} = \frac{\langle n_o \rangle}{V} + \frac{1}{\lambda^3} g_{3/2}(z), \quad (21)$$

donde

$$g_{5/2}(z) = -\frac{4}{\sqrt{\pi}} \int_0^{\infty} dx x^2 \ln(1 - ze^{-x^2}) = \sum_{j=1}^{\infty} \frac{z^j}{j^{5/2}}$$

y

$$g_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^{\infty} dx \frac{x^2}{z^{-1}e^{x^2} - 1} = z \frac{\partial}{\partial z} g_{5/2}(z) = \sum_{j=1}^{\infty} \frac{z^j}{j^{3/2}}.$$

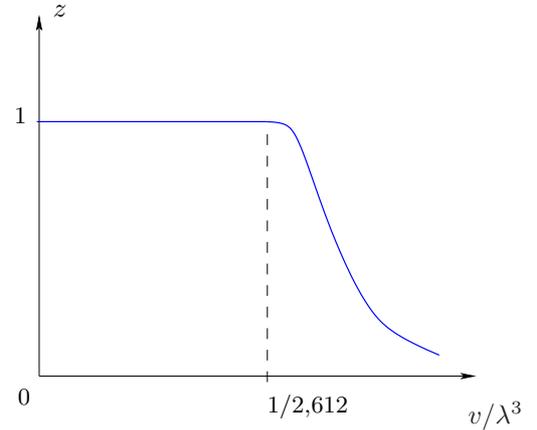
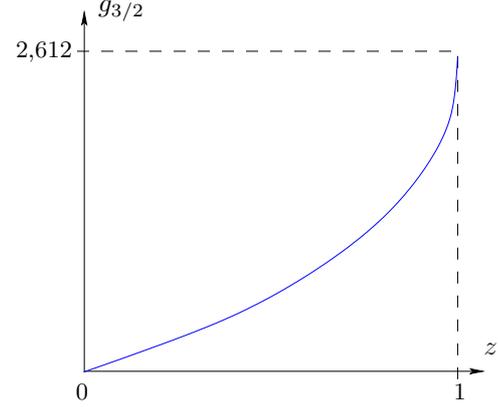
A estas expresiones integrales se arriba realizando la sustitución $x^2 = \beta p^2 / (2m)$. Para temperaturas suficientemente elevadas o valores de v bastante grandes (bajas densidades), los términos correspondientes a $\ell = 0$ son despreciables, por lo que la separación de esos sumandos es innecesaria y pueden omitirse en las expresiones anteriores. Sin embargo, la función $g_{3/2}(z)$, positiva y monótonamente creciente, es acotada: analizando la expresión (21), vemos que para alguna situación con T ó v suficientemente bajos, z alcanzará su máximo valor y, consiguientemente, lo mismo ocurrirá con $g_{3/2}$ ($= 2,612$). En ese caso, el término $g_{3/2}(z)/\lambda^3$ del segundo miembro no es suficiente para mantener la identidad, de manera que la población media del estado $\mathbf{p} = 0$ deja de ser “diferencial” para poblarse macroscópicamente: de esta manera, en el límite termodinámico ($V \rightarrow \infty$), el cociente $\langle n_o \rangle / V \neq 0$. Este es el verdadero motivo por el cual se separan las contribuciones correspondientes a $\mathbf{p} = 0$ en las expresiones anteriores. La población macroscópica del estado con impulso cero se conoce como *condensación de Bose-Einstein*, y en seguida veremos que se corresponde con una transición de fase.

La expresión (21) debe entonces desdoblarse para distinguir las situaciones mencionadas, de manera que tendremos

$$\frac{1}{v} = \begin{cases} \frac{1}{\lambda^3} g_{3/2}(z) & \text{si } \frac{\lambda^3}{v} \leq g_{3/2}(1) = 2,612 \\ \frac{1}{V} \frac{z}{1-z} + \frac{1}{\lambda^3} g_{3/2}(1) & \text{si } \frac{\lambda^3}{v} > g_{3/2}(1) \end{cases}$$

Dicho de otro modo, el valor de la fugacidad se obtiene según qué caso corresponda:

$$z = \begin{cases} \text{despejar } g_{3/2}(z) = \frac{\lambda^3}{v} & \text{si } \frac{\lambda^3}{v} \leq g_{3/2}(1) = 2,612 \\ 1 & \text{si } \frac{\lambda^3}{v} > g_{3/2}(1) \end{cases}$$



Para una dada densidad $1/v$, hay una *temperatura crítica* T_c a la cual ocurre el cambio, y está dada por

$$\lambda^3 = g_{3/2}(1) v \quad \Rightarrow \quad T_c = \frac{2\pi\hbar^2}{mk} \left(\frac{1}{2,612 v} \right)^{2/3}.$$

Del mismo modo, para una dada T , el *volumen específico crítico* resulta

$$v_c = \frac{1}{2,612} \left(\frac{2\pi\hbar^2}{mkT} \right)^{3/2}.$$

De las expresiones anteriores puede obtenerse la fracción de partículas en el estado $\mathbf{p} = 0$

$$\frac{\langle n_0 \rangle}{\langle N \rangle} = 1 - \frac{2,612}{\lambda^3} \frac{V}{\langle N \rangle} = 1 - \left(\frac{T}{T_c} \right)^{3/2}$$

(si $T > T_c$ esta fracción se anula).

En virtud de la ecuación de Euler $U = TS - PV + \mu \langle N \rangle$, $\Omega = -PV$, de manera que la ecuación de estado para la presión resulta

$$\begin{cases} P^> = \frac{kT}{\lambda^3} g_{5/2}(z) \\ P^< = \frac{kT}{\lambda^3} g_{5/2}(1) \end{cases}$$

donde hemos indicado con un supraíndice “>” la presión por encima de la zona de condensación ($T > T_c$ ó $v > v_c$), y “<” para la zona de condensación. Es interesante destacar que debajo del punto de condensación la presión resulta *independiente* de v . También es interesante notar que $P^< \rightarrow 0$ cuando $T \rightarrow 0$, es decir que los bosones en el estado $\mathbf{p} = 0$ (fase condensada) no aportan a la presión; esto podía preverse, ya que estas partículas no pueden transferir ningún impulso para contribuir a la presión en el sistema.

Las isothermas en un diagrama P - v tienen el aspecto de la figura. Como en el caso del cambio de fase líquido-gas, la región horizontal puede pensarse como una región de coexistencia: alguna inestabilidad ha sido subsanada en el potencial de Helmholtz uniendo dos estados mediante una recta, cuya pendiente (constante) corresponde al valor de la presión en la coexistencia.

La presión de vapor representada por la curva cortada de la figura, está dada por la condición

$$P_o(T) = \frac{kT}{\lambda^3} g_{5/2}(1) \quad \left(= \frac{2\pi\hbar^2 g_{5/2}(1)}{m [g_{3/2}(1)]^{5/3}} \frac{1}{v_c^{5/3}} \right).$$

Derivando respecto de T se obtiene

$$\frac{dP_o}{dT} = \frac{5}{2} \frac{k}{\lambda^3} \frac{g_{5/2}(1)}{T} = \frac{1}{T v_c} \left[\frac{5}{2} kT \frac{g_{5/2}(1)}{g_{3/2}(1)} \right].$$

Vemos que se satisface la ecuación de Clausius-Clapeyron si notamos que la discontinuidad Δv en el volumen específico es igual al volumen específico crítico en cada caso, ya que en la fase condensada $v = 0$ (¿por qué?). En ese caso, como podemos verificar a partir de la expresión para la entropía obtenida más abajo, el calor latente de la transformación (siempre por partícula) debe ser

$$\ell = \frac{5}{2} kT \frac{g_{5/2}(1)}{g_{3/2}(1)}.$$

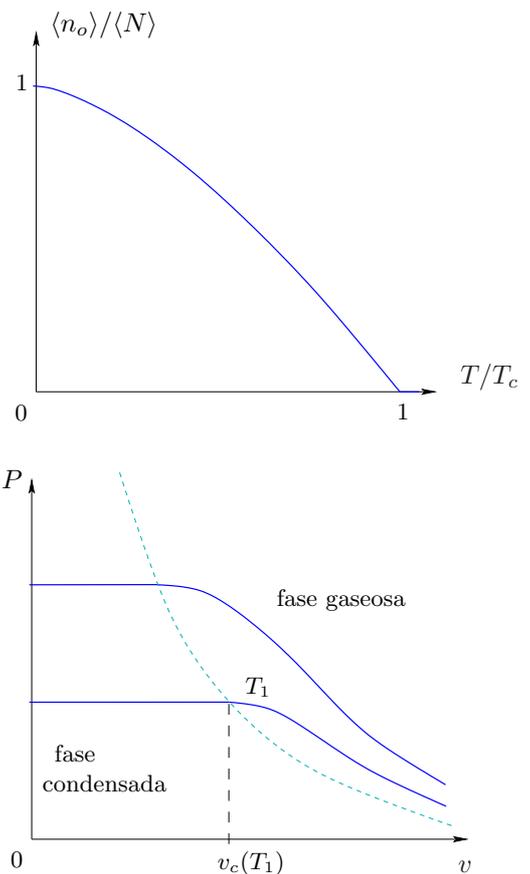
Estas dos discontinuidades en derivadas primeras de la energía libre de Gibbs hacen que este cambio de fase resulte una transición de primer orden.

Para el potencial de Gibbs se obtiene

$$\begin{cases} G^> = \langle N \rangle kT \ln z \\ G^< = 0 \end{cases}$$

y para la entropía

$$\begin{cases} S^> = \frac{5}{2} \frac{kV}{\lambda^3} g_{5/2}(z) - \langle N \rangle k \ln z \\ S^< = \frac{5}{2} \frac{kV}{\lambda^3} g_{5/2}(1) \end{cases}$$



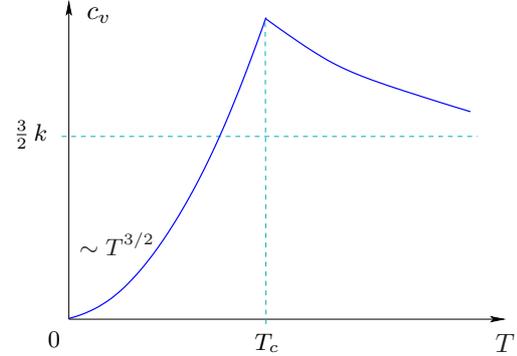
Estas expresiones para la entropía nos permiten calcular explícitamente el calor latente de la transformación; se deja como ejercicio comprobar que afortunadamente coincide con el valor identificado más arriba.

Combinando las ecuaciones anteriores encontramos la energía interna

$$U = \frac{3}{2}PV \Rightarrow \begin{cases} U^> = \frac{3}{2} \frac{kTV}{\lambda^3} g_{5/2}(z) \\ U^< = \frac{3}{2} \frac{kTV}{\lambda^3} g_{5/2}(1) \end{cases}$$

Podemos calcular de aquí el calor específico (por partícula) a volumen constante, tomando la precaución de derivar la entropía manteniendo V y $\langle N \rangle$ constantes

$$\begin{cases} c_v^> = \frac{15}{4} \frac{k}{\lambda^3} g_{5/2}(z) - \frac{9}{4} k \frac{g_{3/2}(z)}{g_{1/2}(z)} \\ c_v^< = \frac{15}{4} \frac{k}{\lambda^3} g_{5/2}(1) \end{cases}$$



El aspecto de la curva de c_v ha hecho que este cambio de fase se denomine “transición lambda”.

Como habrá notado el avezado lector, hemos recurrido a la ecuación (21) para $1/v$ o a la anterior para Ω_{BE}/V , incluyendo o despreciando la contribución de las partículas en el estado con $\mathbf{p} = 0$, según nuestra conveniencia. Sin embargo, puede demostrarse que siempre puede incluirse esta contribución sin que pierdan validez los resultados encontrados. Se sugiere como ejercicio realizar esta verificación.

Para el caso de un *gas de fotones*, como el número de partículas no se conserva, la maximización de la entropía estadística no conlleva el vínculo asociado con $\langle N \rangle$, de manera que no aparece el correspondiente multiplicador de Lagrange, y las poblaciones de los estados \mathbf{p}_ℓ toman la forma $1/[e^{\beta\epsilon_\ell} - 1]$ en lugar de $1/[e^{\beta(\epsilon_\ell - \mu)} - 1]$. Esto es equivalente a aceptar que para fotones $\mu = 0$, y está claro que en este caso el gran potencial Ω coincidirá con la energía libre de Helmholtz $F = U - TS$. El hecho de que el potencial químico se anule en este caso, podría esperarse de antemano aceptando que ocurre una “reacción química” en las paredes de la cavidad, en la que el intercambio de energía de la radiación con las mismas implica la creación o absorción de fotones. La gran partición que se obtiene de este modo coincide con la partición correspondiente al sistema de osciladores de frecuencias respectivas ω_ℓ : en ese caso, cada oscilador se encuentra en el estado cuántico n_ℓ , en lugar de pensar que el modo electromagnético de frecuencia ω_ℓ posee un número variable de fotones n_ℓ .

8. Gas de Fermi-Dirac (basado en los textos de Reichl y Huang) (y Schwabl)

En la sección 6.4 habíamos obtenido la función gran partición para un sistema de partículas regido por la estadística de Fermi-Dirac. En la correspondiente expresión (20), el índice ℓ debe señalar unívocamente el estado cuántico accesible a cada partícula: por ejemplo, para partículas de espín $1/2$, este índice debe abarcar los posibles estados de impulso \mathbf{p}_ℓ y también las proyecciones ‘up’ (\uparrow) o ‘down’ (\downarrow) para el espín. Para el caso de partículas libres, los autovalores individuales para el hamiltoniano no dependen de la orientación del espín s , por lo que deben repetirse $g_s = (2s + 1)$ veces las sumatorias sobre los diferentes estados de impulso. Como esta suma solo puede abarcar dos números de ocupación posibles en cada estado (0 y 1, pues se trata de fermiones), la gran partición puede expresarse como

$$Z_\mu^{\text{FD}}(T, V) = \prod_{\ell=0}^{\infty} (1 + z e^{-\beta\epsilon_\ell})^{g_s} .$$

Podemos entonces escribir para el gran potencial

$$\Omega_{\text{FD}}(T, V, \mu) = -kT g_s \sum_{\ell=0}^{\infty} \ln(1 + z e^{-\beta\epsilon_\ell}) ,$$

y de aquí derivamos el número medio de partículas

$$\langle N \rangle = - \left(\frac{\partial \Omega_{\text{FD}}}{\partial \mu} \right)_{T,V} = g_s \sum_{\ell=0}^{\infty} \frac{1}{z^{-1} e^{\beta \epsilon_{\ell}} + 1} .$$

Como el número total de partículas se construye sumando las contribuciones medias de todos los niveles \mathbf{p}_{ℓ} , obtenemos para el número medio de partículas en el estado ℓ

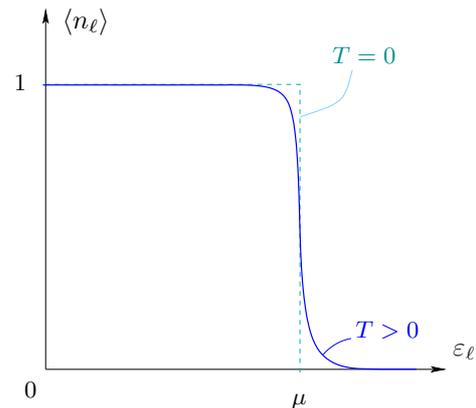
$$g_s \langle n_{\ell} \rangle = \frac{g_s}{z^{-1} e^{\beta \epsilon_{\ell}} + 1} ,$$

donde hemos separado la probabilidad $\langle n_{\ell} \rangle$ de ocupar el estado ℓ , debida solo al carácter estadístico de las partículas bajo estudio (en este caso fermiones). Vemos en primer lugar que siempre se cumple $g_s \langle n_{\ell} \rangle \leq g_s$, tal como exige el principio de exclusión, por lo que evidentemente no habrá condensación en el caso de fermiones. Por otro lado, es interesante analizar la población media del estado fundamental,

$$\langle n_0 \rangle = \frac{1}{z^{-1} + 1} ,$$

que nunca diverge, pues $z \geq 0$.

Vale la pena notar que la expresión para las poblaciones medias $\langle n_{\ell} \rangle$ nos indica que si la temperatura alcanza el cero absoluto ($\beta = \infty$), todos los niveles se despueblan excepto aquellos para los que $\epsilon_{\ell} < \mu$. Cuando la temperatura es apenas superior, la distribución se aparta suavemente de la correspondiente a $T = 0$, como se indica en la figura. La derivada de esta distribución se anula rápidamente cuando ϵ_{ℓ} se aleja apenas de μ , como notaremos más adelante.



Al tomar el límite termodinámico los autovalores ϵ para el hamiltoniano individual de las partículas toman valores continuos, y es frecuente escribir la población estadística como

$$n(\epsilon) = \frac{1}{e^{\beta(\epsilon - \mu)} + 1} .$$

En la expresión anterior para el gran potencial, cuando tomamos el límite termodinámico podemos reemplazar la sumatoria sobre ℓ por una integral en el espacio ahora continuo de \mathbf{p} (o de ϵ), obteniendo

$$\Omega_{\text{FD}}(T, V, \mu) = - \frac{g_s kTV}{\lambda^3} f_{5/2}(z) ,$$

donde

$$f_{5/2}(z) \equiv \frac{4}{\sqrt{\pi}} \int_0^{\infty} dx x^2 \ln(1 + z e^{-x^2}) \quad \left[= \sum_{j=1}^{\infty} \frac{(-1)^{j+1} z^j}{j^{5/2}} \quad \text{si } z < 1 \right] .$$

Para llegar a esta expresión integral es necesario realizar el cambio de variable $x^2 = p^2/(2mkT)$, mientras que la sumatoria se obtiene al hacer un desarrollo en serie del integrando, resolviendo luego término a término.

Es interesante señalar que podemos escribir la gran partición en términos de β , V y z como variables independientes, en lugar de T , V y μ . De este modo, con un mínimo de inspiración el lector podrá mostrar por ejemplo la siguiente relación

$$U = - \left(\frac{\partial \ln Z(\beta, V, z)}{\partial \beta} \right)_{V,z} ,$$

teniendo la precaución de que al derivar deben permanecer constantes V y z (y no μ). Se obtiene así

$$U = \frac{3}{2} g_s \frac{kTV}{\lambda^3} f_{5/2}(z) ,$$

mientras que para un gas de bosones con degeneración g' la energía interna resulta

$$U = \frac{3}{2} g'_s \frac{kTV}{\lambda^3} g_{5/2}(z) .$$

Comparando estas expresiones con $\Omega = -PV$, se ve que al igual que para la estadística de Bose-Einstein se cumple

$$U = \frac{3}{2} PV .$$

De manera semejante puede expresarse el número medio de partículas

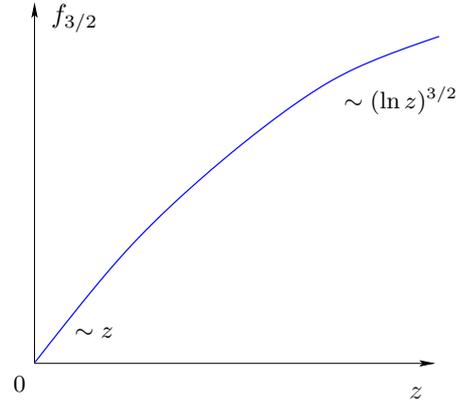
$$\langle N \rangle = z \left(\frac{\partial \ln Z(\beta, V, z)}{\partial z} \right)_{\beta, V} = \frac{g_s V}{\lambda^3} f_{3/2}(z), \quad (22)$$

donde

$$f_{3/2}(z) \equiv \frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \frac{1}{z^{-1} e^{x^2} + 1} = z \frac{\partial f_{5/2}(z)}{\partial z} \quad \left[= \sum_{j=1}^{\infty} \frac{(-1)^{j+1} z^j}{j^{3/2}} \quad \text{si } z < 1 \right].$$

Como la función $f_{3/2}(z)$ es creciente con z , de la expresión anterior para $\langle N \rangle$ puede verse que, para una dada densidad, cuando crece la temperatura debe decrecer z , de manera que para altas temperaturas $\mu \rightarrow -\infty$. Por otro lado, cuando la temperatura se hace muy pequeña, z debe tender a ∞ , es decir $\mu/T \rightarrow \infty$, de donde inferimos que a bajas temperaturas $\mu > 0$, aunque su valor puede ser acotado, pues el denominador T garantiza de todos modos la divergencia.

Para estudiar el comportamiento de este sistema a bajas temperaturas debemos analizar entonces el comportamiento de $f_{3/2}(z)$ para valores de z grandes. Seguiremos para ello el método de Sommerfeld, notando que estamos interesados en propiedades termodinámicas como la energía interna del sistema



$$U = \sum_{\ell} g_s \langle n_{\ell} \rangle \epsilon_{\ell} \rightarrow \frac{4\sqrt{2} \pi g_s m^{3/2} V}{h^3} \int_0^\infty d\epsilon \sqrt{\epsilon} \epsilon n(\epsilon) = \int_0^\infty d\epsilon g(\epsilon) \epsilon n(\epsilon), \quad (23)$$

o el número medio total de partículas

$$\langle N \rangle = \sum_{\ell} g_s \langle n_{\ell} \rangle \rightarrow \int_0^\infty d\epsilon g(\epsilon) n(\epsilon).$$

Al cambiar de variables, la densidad de estados $g(\epsilon)$ representa la fracción de estados individuales diferentes que comparten ese autovalor de energía ϵ . En general entonces, tenemos integrales de la forma

$$I = \int_0^\infty d\epsilon f(\epsilon) n(\epsilon).$$

El desarrollo de Sommerfeld consiste en separar esta integral en dos partes

$$I = \int_0^\mu d\epsilon f(\epsilon) + \int_0^\infty d\epsilon f(\epsilon) \left[n(\epsilon) - \Theta(\mu - \epsilon) \right],$$

donde $\Theta(x)$ es la función escalón de Heaviside. De este modo, retenemos en el primer término del miembro de la derecha el comportamiento que sobrevivirá para el límite $T \rightarrow 0$. Puede extenderse la última integral hasta $-\infty$ definiendo $f(-\epsilon) = f(\epsilon)$ para $\epsilon < 0$, teniendo en cuenta que $n(\epsilon)$ solo difiere apreciablemente de $\Theta(\mu - \epsilon)$ en un entorno cercano de $\epsilon = \mu$, pues para $\epsilon < 0 < \mu(T \rightarrow 0)$, $n(\epsilon) \approx 1 + \mathcal{O}(e^{-\beta(\mu - \epsilon)})$

$$I \approx \int_0^\mu d\epsilon f(\epsilon) + \int_{-\infty}^\infty d\epsilon f(\epsilon) \left[n(\epsilon) - \Theta(\mu - \epsilon) \right].$$

Como ejercicio sencillo, se recomienda demostrar que la función $n(\epsilon) - \Theta(\mu - \epsilon)$ es impar alrededor de $\epsilon = \mu$. Podemos tomar entonces la expansión en serie de Taylor de $f(\epsilon)$ alrededor de μ y realizar el cambio de variable $x = (\epsilon - \mu)/(kT)$, notando que las potencias pares de x no contribuirán a la integral

$$I = \int_0^\mu d\epsilon f(\epsilon) + \int_{-\infty}^\infty dx \left[\frac{1}{e^x + 1} - \Theta(x) \right] \times \left(f'(\mu) (kT)^2 x + \frac{f'''(\mu)}{3!} (kT)^4 x^3 + \dots \right).$$

El producto del factor entre corchetes y las potencias impares de x es una función par, y dado que $\Theta(-x) = 0$ para $x > 0$,

$$I = \int_0^\mu d\epsilon f(\epsilon) + \frac{\pi^2}{6} (kT)^2 f'(\mu) + \frac{7\pi^4}{360} (kT)^4 f'''(\mu) + \dots, \quad (24)$$

donde hemos utilizado la función ζ de Riemann, relacionada con la función Γ mediante la identidad

$$\int_0^\infty du \frac{u^{y-1}}{e^u + 1} = (1 - 2^{1-y}) \zeta(y) \Gamma(y) ;$$

en particular, $\zeta(2) = \pi^2/6$ y $\zeta(4) = \pi^4/90$. La expansión de Sommerfeld (24) permite reescribir el número medio de partículas

$$\langle N \rangle = \sum_\ell g_s \langle n_\ell \rangle \rightarrow \frac{g_s m^{3/2} V}{\sqrt{2} \pi^2 \hbar^3} \int_0^\infty d\epsilon \sqrt{\epsilon} n(\epsilon) = \frac{g_s V}{\lambda^3} f_{3/2}(z) .$$

sustituyendo la integral en ϵ , de modo que puede despejarse el comportamiento de $f_{3/2}(z)$ para valores grandes de z :

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}} (\ln z)^{3/2} \left[1 + \frac{\pi^2}{8} (\ln z)^{-2} + \dots + \mathcal{O}((\ln z)^{-4}) \right] . \quad (25)$$

Entonces, para el caso de bajas temperaturas o altas densidades, es decir $\lambda^3/v \gg 1$, la ecuación (22) puede aproximarse como

$$\frac{1}{v} = \frac{4}{3\sqrt{\pi}} (\ln z)^{3/2} \frac{g_s}{\lambda^3} \left[1 + \frac{\pi^2}{8} (\ln z)^{-2} + \dots \right]$$

En particular, se define como *energía de Fermi*

$$\mu(T=0) \equiv \epsilon_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{g_s v} \right)^{2/3} .$$

Completamos ahora el comentario de más arriba diciendo que a $T=0$ si $\epsilon_\ell > \epsilon_F$, $n_\ell = 0$, y si $\epsilon_\ell < \epsilon_F$, $n_\ell = 1$.

Se define también el impulso de Fermi p_F escribiendo la relación clásica $\epsilon_F = p_F^2/2m$. Más frecuentemente se utiliza la *temperatura de Fermi* $T_F \equiv \epsilon_F/k$, de manera que si $T \ll T_F$ se dice que el gas está “degenerado”, porque todas las partículas tienden a ocupar los estados de menor energía, según lo permitido por la degeneración g_s .

La expansión (25) puede “invertirse”, resultando en esta región

$$\mu (= kT \ln z) = \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] .$$

Podemos obtener una expresión para la energía interna de este sistema desarrollando análogamente la expansión (23), siguiendo también el método de Sommerfeld. Se deja como ejercicio verificar

$$U = \frac{3}{5} \langle N \rangle \epsilon_F \left[1 + \frac{5}{12} \pi^2 \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] .$$

El primer término corresponde a la energía del estado fundamental, y puede verificarse fácilmente que

$$\sum_{|\mathbf{p}| < p_F} \frac{p^2}{2m} = \frac{3}{5} \langle N \rangle \epsilon_F .$$

Con estos elementos puede calcularse el calor específico

$$c_v \approx \frac{\pi^2}{2} \frac{k^2}{\epsilon_F} T .$$

Aquí se pone en evidencia que se cumple la tercera ley de la termodinámica, ya que este calor específico tiende a cero en el límite de bajas temperaturas.

A partir de la igualdad $U = (3/2)PV$ puede escribirse

$$P \approx \frac{2}{5} \frac{\epsilon_F}{v} \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] .$$

Podemos ver que a $T=0$, $P \neq 0$, como consecuencia del principio de exclusión, pues no todas las partículas pueden tener el mínimo valor de impulso.

Analicemos a continuación el régimen de temperaturas altas o bajas densidades, es decir, $\lambda^3/v \ll 1$, donde esperamos que los efectos cuánticos se “disimulen”. Recordando que en este caso z toma valores muy próximos a 0, aproximamos

$$\frac{\lambda^3}{g_s v} = f_{3/2}(z) = z - \frac{z^2}{2^{3/2}} + \dots$$

Como antes, “invirtiendo” esta serie tenemos

$$z = \frac{\lambda^3}{g_s v} + \frac{1}{2^{3/2}} \left(\frac{\lambda^3}{g_s v} \right)^2 + \dots$$

En los casos en que $\lambda^3/(g_s v)$ es extremadamente pequeño, podemos quedarnos solo con el primer término, coincidiendo nuestro resultado con el correspondiente al gas de Maxwell-Boltzmann. Por otro lado, puede verificarse que las poblaciones de los distintos niveles también coinciden con el caso de Maxwell-Boltzmann. Finalmente, la ecuación de estado

$$\frac{Pv}{kT} = \frac{4\pi g_s v}{h^3} \int_0^\infty dp p^2 \ln \left(1 + z e^{-\frac{\beta p^2}{2m}} \right)$$

puede aproximarse como

$$\frac{Pv}{kT} = \frac{g_s v}{\lambda^3} \left(z - \frac{z^2}{2^{5/2}} + \dots \right) = 1 + \frac{1}{2^{5/2}} \frac{\lambda^3}{g_s v} + \dots$$

Nuevamente, el resultado coincide con el de Maxwell-Boltzmann en la situación límite.

8.1. Diamagnetismo de Landau (basado en el texto de Huang)

Ya hemos visto en uno de los ejercicios el teorema de Van Leewen, que establece que no existe el diamagnetismo (y el magnetismo en general) en estadística clásica. Sin embargo Landau en 1930 mostró que es posible explicar el fenómeno del diamagnetismo mediante la cuantización de las órbitas de partículas cargadas en un campo magnético.

Las propiedades magnéticas de un material se analizan a través de la susceptibilidad magnética *por unidad de volumen*

$$\chi = \left(\frac{\partial M'}{\partial B} \right)_T .$$

Aquí M' representa el momento magnético por unidad de volumen, y B , la inducción *externa* aplicada. Si el análisis es llevado adelante en el ensamble canónico, sabemos que cuando el hamiltoniano incorpora el acoplamiento con el campo externo, la función partición estará relacionada con la energía libre de Gibbs

$$G(T, B, N) = -kT \ln Z_N(T, B) \quad \Rightarrow \quad M' = -\frac{1}{V} \left(\frac{\partial G}{\partial B} \right)_{T, N} = \frac{kT}{V} \frac{\partial}{\partial B} \ln Z_N .$$

En el gran canónico el análisis es similar, resultando

$$M' = - \left[\frac{\partial}{\partial B} \left(\frac{\tilde{\Omega}}{V} \right) \right]_{T(V), z} = kT \left[\frac{\partial}{\partial B} \left(\frac{\ln Z_\mu}{V} \right) \right]_{T(V), z} .$$

Hemos denotado el gran potencial como $\tilde{\Omega}$ para recordar que a diferencia de los casos anteriores, éste involucra también la transformación de Legendre de M por B (otra vez, en virtud de que el hamiltoniano utilizado para el cálculo de la gran partición incluye el acoplamiento con la fuente de campo). Por otro lado, aunque en general las variables P y V no intervienen en las ecuaciones de estado, se ha indicado a V como otra posible variable natural de nuestro potencial.

Un sistema es *diamagnético* cuando $\chi < 0$, y *paramagnético* cuando $\chi > 0$. Los núcleos que constituyen una red cristalina en general no aportan a la respuesta magnética de los materiales, pues su momento magnético intrínseco es mucho menor que el de los electrones (típicamente 1000 veces más chico). El alineamiento del espín de los electrones con el campo externo da lugar al paramagnetismo, que compite con el movimiento orbital de los mismos. Para llevar adelante nuestro análisis, pensamos en un gas de electrones libres, ignorando entonces sus coordenadas de espín al estudiar su traslación.

Niveles de Landau

En un campo magnético externo asociado a un potencial vector \mathbf{A} , el hamiltoniano no relativista correspondiente a un electrón “libre” es

$$\hat{H}_1 = \frac{1}{2m} \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right)^2 .$$

La ecuación de Schrödinger independiente de t es invariante ante transformaciones de gauge

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &\longrightarrow \mathbf{A}(\mathbf{r}) - \nabla\omega(\mathbf{r}) , \\ \psi(\mathbf{r}) &\longrightarrow e^{-\frac{ie}{\hbar c}\omega(\mathbf{r})}\psi(\mathbf{r}) , \end{aligned}$$

donde $\omega(\mathbf{r})$ es cualquier función continua. Si $\mathbf{B} = B\hat{k}$, podemos elegir $A_x = -By$; $A_y = A_z = 0$, con lo cual obtenemos

$$\hat{H}_1 = \frac{1}{2m} \left[\left(p_x - \frac{eB}{c} y \right)^2 + p_y^2 + p_z^2 \right] .$$

Proponiendo $\psi(x, y, z) = e^{i(k_x x + k_z z)} f(y)$, donde k_α es el autovalor respectivo para p_α , la ecuación de Schrödinger estacionaria correspondiente al autovalor de energía ϵ implica

$$\left[\frac{1}{2m} p_y^2 + \frac{1}{2} m \omega_o^2 (y - y_o)^2 \right] f(y) = \epsilon' f(y) ,$$

donde $\omega_o \equiv eB/(mc)$ es la *frecuencia ciclotrónica* correspondiente a las órbitas circulares clásicas de una masa m con carga e bajo un campo B uniforme, $y_o \equiv \hbar c/(eB) k_x$ y $\epsilon' \equiv \epsilon - \hbar^2 k_z^2/(2m)$. La anterior ecuación para $f(y)$ no es otra que la entrañable ecuación de un oscilador armónico, de manera que los autovalores de \hat{H}_1 son los *niveles de Landau*

$$\epsilon = \frac{p_z^2}{2m} + \hbar\omega_o \left(j + \frac{1}{2} \right) ,$$

donde $j = 0, 1, 2, \dots$ y $p_z \equiv \hbar k_z$ son los autovalores de partícula libre según la dirección z .

Notemos que los niveles de Landau son independientes de k_x , lo que significa que habrá una degeneración para ellos igual al número de valores de k_x permitidos en cada caso, de manera que el valor de y_o no escape del volumen que contiene a los electrones. Suponiendo que se trata de un material cúbico de lado L , e imponiendo condiciones periódicas a ψ , los valores permitidos para k_x resultan $2\pi n_x/L$, donde n_x es un número natural. En ese caso,

$$y_o = \frac{\hbar c}{eB} \frac{2\pi}{L} n_x ,$$

de modo que la condición $y_o \leq L$ implica $n_x \leq eB/(hc) L^2$, o equivalentemente, la correspondiente degeneración será

$$g = \frac{eB}{hc} L^2 .$$

Vemos entonces que cuando aparece un campo externo $B \neq 0$, las energías asociadas con el movimiento en el plano x - y dejan de ser continuas para tomar los valores discretos que hemos encontrado aquí.

Para obtener la función gran partición debemos desarrollar adecuadamente la expresión

$$Z_\mu = \prod_{\ell} (1 + z e^{-\beta \epsilon_\ell})^g ,$$

donde aquí los estados posibles son señalados por los números cuánticos $\ell = (p_z, j)$ o $\ell' = (p_z, j, \alpha)$ en nuestra notación previa, representando los valores posibles para k_x con el índice $\alpha = 1, 2, \dots, g$. Tenemos entonces

$$\ln Z_\mu = \sum_{\alpha=1}^g \sum_{j=0}^{\infty} \sum_{p_z} \ln \left[1 + z e^{-\beta \epsilon(p_z, j)} \right] .$$

Como $\epsilon(p_z, j)$ no depende de α , sumar sobre este índice simplemente implica repetir g veces el resto de la expresión. En el límite termodinámico, los autovalores de p_z pasan a ser continuos, de modo que la suma se transforma en integral mediante el procedimiento habitual

$$\ln Z_\mu = \frac{2gL}{h} \sum_{j=0}^{\infty} \int_0^{\infty} dp \ln \left[1 + z e^{-\beta \epsilon(p, j)} \right] ,$$

donde hemos reducido el intervalo de integración sobre p a partir de la paridad del integrando. El número medio de partículas resulta entonces

$$\langle N \rangle = z \frac{\partial}{\partial z} \ln Z_\mu = \frac{2gL}{h} \sum_{j=0}^{\infty} \int_0^{\infty} dp \frac{1}{z^{-1} e^{\beta \epsilon(p,j)} + 1}. \quad (26)$$

Para temperaturas altas, z se aproxima a cero, de modo que el valor de $\langle N \rangle$ no diverja. Utilizando entonces el hecho de que z es pequeño, podemos aproximar

$$\ln Z_\mu \simeq \frac{2gzL}{h} \sum_{j=0}^{\infty} \int_0^{\infty} dp e^{-\beta \left[\frac{p^2}{2m} + \hbar\omega_o \left(j + \frac{1}{2} \right) \right]}.$$

Ahora la integración y la suma se resuelven independientemente, y definiendo $x \equiv \hbar\omega_o/(2kT)$ obtenemos

$$\ln Z_\mu = \frac{gzL}{\lambda} \frac{e^{-x}}{1 - e^{-2x}}.$$

Como x es pequeño para temperaturas altas, podemos desarrollar en serie de Taylor, teniendo el cuidado de agrupar pertinentemente los términos del mismo orden. Así obtenemos

$$\ln Z_\mu \simeq \frac{gzL}{\lambda} \frac{1}{2x} \left(1 - \frac{x^2}{6} \right) = \frac{zV}{\lambda^3} \left[1 - \frac{1}{24} \left(\frac{\hbar\omega_o}{kT} \right)^2 \right],$$

de donde puede obtenerse una estimación para la susceptibilidad

$$\chi = -\frac{z}{3kT\lambda^3} \left(\frac{e\hbar}{2mc} \right)^2.$$

La primera gran sorpresa es que este tratamiento cuántico nos ha permitido obtener una respuesta magnética no nula. Pero además, el hecho de que χ sea negativo implica que estamos dando cuenta de un comportamiento *diamagnético* del material en cuestión, lo que nos mantendrá por un tiempo en un éxtasis irrefrenable.

Para z pequeños la ecuación (26) provee una expresión que coincide con $\ln Z_\mu$, de manera que podemos aproximar $\langle N \rangle/V \simeq z/\lambda^3$. Reemplazando z en la expresión para χ obtenemos

$$\chi = -\frac{1}{3kTv} \left(\frac{e\hbar}{2mc} \right)^2,$$

que en particular satisface la ley de Curie ($\chi \propto 1/T$).

8.2. Efecto de Haas - Van Alphen

Analicemos ahora el comportamiento para temperaturas bajas. Como sabemos, los electrones ‘intentan’ ocupar los estados de energía más baja, aunque también sabemos que a medida que se reduce el valor de B , correspondientemente caben menos en el estado fundamental, porque la degeneración g disminuye linealmente con B . Esto trae como consecuencia oscilaciones en la magnetización, fenómeno que se conoce como efecto *de Haas - Van Alphen*, a partir de los resultados de un famoso experimento realizado por estos investigadores en 1930, en una muestra de bismuto a 14,2 K entre 0,5 y 2 Tesla.

Al interesarnos en el caso $kT \ll \hbar\omega_o$, tomamos directamente $T = 0$ en las expresiones anteriores, e ignoramos el movimiento en la dirección z , pues no aporta a la susceptibilidad magnética.

Definiendo $\mu_o \equiv e\hbar/(2mc)$, los niveles de Landau pueden escribirse como $\epsilon_j = 2\mu_o B(j + 1/2)$, y en términos de $B_o \equiv nhc/e$, donde $n = N/L^2$ es el número de partículas por unidad de área, su degeneración es $g = NB/B_o$. La magnitud B_o representa el valor mínimo para que todos los electrones quepan en cada uno de los niveles de Landau. A $T = 0$, cuando $B > B_o$ todos los electrones se acomodan en el nivel fundamental $j = 0$; en ese caso, la energía por partícula será $E_o/N = \mu_o B$. Si en cambio $B < B_o$, algunas partículas deben ocupar niveles más altos: la condición para que los primeros $j + 1$ niveles de Landau (del 0 al j) estén ocupados completamente, el $(j + 1)$ parcialmente lleno y los otros vacíos es

$$(j + 1)g < N < (j + 2)g,$$

o bien

$$(j + 1)N \frac{B}{B_o} < N < (j + 2)N \frac{B}{B_o} \quad \Rightarrow \quad \frac{1}{j + 2} < \frac{B}{B_o} < \frac{1}{j + 1}.$$

Cuando el valor de B yaga en ese intervalo la energía fundamental por partícula será

$$\frac{E_o}{N} = \frac{g}{N} \sum_{k=0}^j \epsilon_k + \left[1 - \frac{(j+1)g}{N} \right] \epsilon_{j+1} = \mu_o B \left[2j + 3 - (j+1)(j+2) \frac{B}{B_o} \right],$$

es decir,

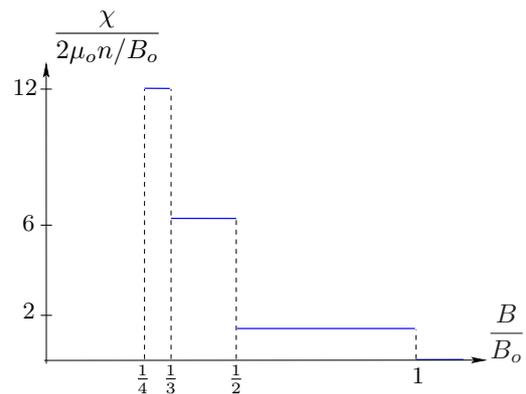
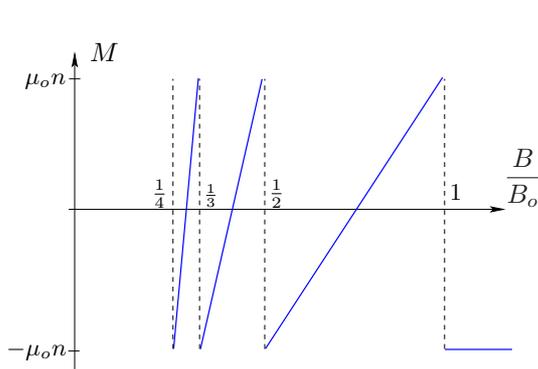
$$\frac{E_o}{N}(B) = \begin{cases} \mu_o B_o x & \text{si } x \equiv \frac{B}{B_o} > 1 \\ \mu_o B_o x [2j + 3 - (j+1)(j+2)x] & \text{si } \frac{1}{j+2} < x < \frac{1}{j+1} \quad (j = 0, 1, 2, \dots) \end{cases}$$

Puede obtenerse (¿cómo y por qué?) a partir de estas expresiones

$$M = \begin{cases} -\mu_o n & \text{si } x > 1 \\ \mu_o n [2(j+1)(j+2)x - (2j+3)] & \text{si } \frac{1}{j+2} < x < \frac{1}{j+1} \end{cases}$$

y

$$\chi = \begin{cases} 0 & \text{si } x > 1 \\ \frac{2\mu_o n}{B_o} (j+1)(j+2) & \text{si } \frac{1}{j+2} < x < \frac{1}{j+1} \end{cases}$$



8.3. Ferromagnetismo: Modelo de Ising

Uno de los fenómenos más interesantes en física del estado sólido es el ferromagnetismo: en algunos metales como el hierro o el níquel, una fracción finita de los espines asociados con cada núcleo se polariza espontáneamente en determinada dirección, generando un momento magnético *macroscópico* en ausencia de campo externo. Esto ocurre solo a bajas temperaturas; a temperaturas elevadas los espines se orientan al azar, produciendo una magnetización resultante nula.

Para intentar una descripción, supongamos que cada uno de los N puntos de la red cristalina aloja un átomo al que asociamos un espín $1/2$ y momento magnético μ_B (magnetón de Bohr en unidades de \hbar). Bajo la acción de una inducción magnética externa \mathbf{B} , el hamiltoniano del sistema (incluyendo la interacción) es

$$\hat{H} = - \sum_{\langle ij \rangle} J_{ij} \hat{\sigma}_i \cdot \hat{\sigma}_j - \mu_B \mathbf{B} \cdot \sum_{i=1}^N \hat{\sigma}_i,$$

donde $\hat{\sigma} = (\hat{\sigma}^x, \hat{\sigma}^y, \hat{\sigma}^z)$ son las matrices de Pauli y J_{ij} , denominada “integral de intercambio”, depende de la distancia entre los núcleos i - j : en nuestra descripción consideraremos que es distinta de cero para los núcleos más cercanos en la red, motivo por el cual la primera sumatoria barre solamente los pares $\langle ij \rangle$ de primeros vecinos. Cuando $J_{ij} > 0$ se dice que el sistema es *ferromagnético*, mientras que si $J_{ij} < 0$ el sistema es *antiferromagnético*.

El hamiltoniano de la expresión anterior se conoce como *modelo de Heisenberg* o *isotrópico*. Cuando la respuesta del sistema es relevante solo en una orientación ($\mathbf{B} = B\hat{k}$), la expresión anterior se transforma en el

modelo de Ising o anisotrópico al considerar $J_{ij} = J$ (pues solo se incluyen las interacciones entre primeros vecinos):

$$\hat{H} = -J \sum_{\langle ij \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z - \mu_B B \sum_{i=1}^N \hat{\sigma}_i^z .$$

En adelante obviaremos el supraíndice z , sobreentendiendo que $\hat{\sigma}$ representa el operador $\hat{\sigma}^z$. Si bien originalmente el modelo de Ising se desarrolló para analizar el ferromagnetismo, se lo puede adaptar fácilmente para describir otros cambios de fase. Así es que han surgido aplicaciones para estudiar transiciones orden-desorden (aleaciones), vidrios de espín, el llamado gas de red, y recientemente se lo ha utilizado también en redes neuronales.

8.3.1. Modelo de Ising unidimensional

La descripción de un sistema unidimensional se simplifica notablemente cuando se imponen condiciones de contorno periódicas, es decir, el estado del núcleo $N + 1$ coincide con el del núcleo 1. Entonces podemos escribir

$$\hat{H} = -J \sum_{i=1}^N \hat{\sigma}_i \hat{\sigma}_{i+1} - \frac{\mu_B B}{2} \sum_{i=1}^N (\hat{\sigma}_i + \hat{\sigma}_{i+1}) .$$

Para calcular la función partición, solo debe tenerse la precaución de barrer todos los autovalores de este hamiltoniano, a través de la acción de los operadores $\hat{\sigma}_i$, que actúan cada uno en el espacio de Hilbert correspondiente, con autovalores $\sigma_i = \pm 1$

$$Z_N(T, B) = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} e^{\beta \sum_{i=1}^N \left[J \sigma_i \sigma_{i+1} + \frac{\mu_B B}{2} (\sigma_i + \sigma_{i+1}) \right]} .$$

Definimos ahora el operador \hat{P} con elementos de matriz

$$\langle \sigma_i | \hat{P} | \sigma_{i+1} \rangle = e^{\beta \left[J \sigma_i \sigma_{i+1} + \frac{\mu_B B}{2} (\sigma_i + \sigma_{i+1}) \right]} .$$

Eligiendo la base de manera que el autovalor $\sigma = 1$ corresponda al elemento $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\sigma = -1$ al $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, la representación matricial de este operador es

$$\hat{P} = \begin{pmatrix} e^{\beta(J + \mu_B B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J - \mu_B B)} \end{pmatrix}$$

De este modo, la función partición resulta

$$Z_N(T, B) = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \langle \sigma_1 | \hat{P} | \sigma_2 \rangle \langle \sigma_2 | \hat{P} | \sigma_3 \rangle \cdots \langle \sigma_N | \hat{P} | \sigma_1 \rangle .$$

Como los estados $|+1\rangle$ y $|-1\rangle$ forman una base,

$$\sum_{\sigma = \pm 1} |\sigma\rangle \langle \sigma| = I$$

de manera que

$$Z_N(T, B) = \sum_{\sigma_1 = \pm 1} \langle \sigma_1 | \hat{P}^N | \sigma_1 \rangle = \text{Tr} (\hat{P}^N)$$

Al ser \hat{P} un operador simétrico, puede diagonalizarse exigiendo

$$\begin{vmatrix} e^{\beta(J + \mu_B B)} - \lambda & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J - \mu_B B)} - \lambda \end{vmatrix} = 0 ,$$

resultando la ecuación característica

$$\lambda^2 - 2\lambda e^{\beta J} \cosh(\beta \mu_B B) + 2 \sinh(2\beta J) = 0 ,$$

de donde

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta \mu_B B) \pm \sqrt{e^{-2\beta J} + e^{2\beta J} \sinh^2(\beta \mu_B B)} .$$

Como $\text{Tr}(\hat{P}^N)$ no depende de la base

$$\text{Tr}(\hat{P}^N) = \lambda_+^N + \lambda_-^N \quad \Rightarrow \quad G(T, B, N) = -kT \ln(\lambda_+^N + \lambda_-^N),$$

Es fácil ver que en el límite termodinámico solo sobrevive la contribución de $\lambda_+ (> \lambda_-)$ analizando la energía libre de Gibbs por núcleo

$$g(T, B) = \lim_{N \rightarrow \infty} \frac{1}{N} G(T, B, N) = -kT \ln \lambda_+.$$

El momento magnético total para el sistema resulta

$$\langle M \rangle = -N \left(\frac{\partial g}{\partial B} \right)_T = \frac{N \mu_B \sinh(\beta \mu_B B)}{\sqrt{e^{-4\beta J} + \sinh^2(\beta \mu_B B)}}.$$

Es evidente entonces que cuando $B = 0$, $\langle M \rangle$ también se anula, es decir, no hay magnetización espontánea en una dimensión.

8.3.2. Modelo de Weiss

En dos dimensiones, el modelo de Ising también puede resolverse analíticamente, lográndose encontrar respuesta ferromagnética bajo campos externos. Sin embargo, el desarrollo se torna excesivamente complicado para sistemas tridimensionales, recurriéndose a métodos aproximados. El que veremos aquí se conoce como *método de Weiss* o modelo de *campo molecular*.

En esta descripción se reemplaza la interacción con los γ primeros vecinos de la red por un promedio, de modo que el hamiltoniano resulta

$$\hat{H}_W = -\frac{J\gamma}{2} \sum_{i=1}^N \hat{\sigma}_i \langle \sigma \rangle - \mu_B B \sum_{i=1}^N \hat{\sigma}_i.$$

El denominador 2 del primer miembro se introduce para evitar contar dos veces la interacción correspondiente a cada par $\langle ij \rangle$. Vemos entonces que el efecto de los primeros vecinos se traduce como la producción de un campo efectivo $J\gamma \langle \sigma \rangle / (2\mu_B)$. En esta aproximación de *campo medio* puede entonces escribirse

$$\hat{H}_W = \sum_{i=1}^N \hat{H}_i, \quad \text{con} \quad \hat{H}_i = - \left(\frac{J\gamma \langle \sigma \rangle}{2} + \mu_B B \right) \hat{\sigma}_i,$$

es decir, como un sistema de espines no interactuantes. Al ser independientes los espines, sabemos que la función partición del sistema total Z_N se evalúa a partir de la partición correspondiente a cada espín Z_1 , pues $Z_N = Z_1^N$. En nuestro caso conocemos el resultado

$$Z_1 = 2 \cosh \left[\beta \left(\frac{J\gamma \langle \sigma \rangle}{2} + \mu_B B \right) \right]$$

y, a partir de los autovalores H_1 del hamiltoniano para cada espín,

$$\langle \sigma \rangle = \langle \sigma_i \rangle = \frac{1}{Z_1} \sum_{\sigma=\pm 1} \sigma e^{-\beta H_1} = \text{tgh} \left[\beta \left(\frac{J\gamma \langle \sigma \rangle}{2} + \mu_B B \right) \right].$$

Esta ecuación trascendente puede tener solución no nula para $B = 0$, como puede verificarse gráficamente. También se deja como ejercicio comprobar que existe una *temperatura crítica* por encima de la cual las soluciones no triviales no existen, cuyo valor es $T_c = J\gamma / (2k)$.

Por supuesto como esta descripción del fenómeno del ferromagnetismo no depende de la dimensión de nuestro problema, resulta inadecuada para un sistema unidimensional.

De la “ecuación de Curie-Weiss” correspondiente a $B = 0$

$$\langle \sigma \rangle = \text{tgh} \left(\frac{J\gamma}{2kT} \langle \sigma \rangle \right)$$

podemos hallar (numéricamente) el comportamiento del momento magnético por núcleo $m \equiv \mu_B \langle \sigma \rangle$ en función de la temperatura. Pueden encontrarse las aproximaciones asintóticas para T bajas

$$m / \mu_B \simeq 1 - 2e^{-\frac{J\gamma m}{\mu_B k T}}$$

y para $T \rightarrow T_c$ (por debajo)

$$m/\mu_B \simeq \sqrt{3 \left(1 - \frac{T}{T_c}\right)}.$$

De este modo pueden analizarse todos los *exponentes críticos*, relacionados con los comportamientos de los diferentes parámetros cerca del punto crítico.