

Termodinámica y Mecánica Estadística II

Notas 2010

Sergio Cannas

Estos capítulos fueron extraídos de la página:
<https://www.famaf.unc.edu.ar/~cannas/notas.html>

Capítulo 1

Introducción a la Teoría de Probabilidad

Para la mayoría de la gente, “probabilidad” es un término vago utilizado en el lenguaje cotidiano para indicar la posibilidad de ocurrencia de un evento futuro. Así, hablamos de la probabilidad de que llueva mañana, la probabilidad de que suba el dolar, etc. En ciencia el concepto de probabilidad se utiliza para hacer *inferencias*, esto es, predicciones acerca de ciertos procesos cuyo resultado es variable cada vez que se repiten. Así, por ejemplo, no se puede predecir con exactitud cual será la presión arterial de una persona en un momento dado, ni si la moneda arrojada caerá de un lado o del otro. Este tipo de procesos, dan un resultado diferente dentro de un cierto conjunto de valores cada vez que se repiten, pero la frecuencia relativa con que ocurre cada valor en una gran serie de observaciones es a menudo estable. Si arrojamus un número grande de veces la misma moneda (suponiendo que no este cargada) observaremos que aproximadamente la mitad de las veces sale cara y la otra mitad seca. Procesos que presentan estas propiedades se denominan **aleatorios** o **estocásticos**. Dicha frecuencia relativa estable nos dá una medida intuitiva (pero significativa) de la posibilidad de ocurrencia de un evento aleatorio en una observación futura y constituye la definición **estadística** de probabilidad. Supongamos que el resultado de un experimento esta descrito por una variable X , que puede tomar valores de un conjunto finito x_1, x_2, \dots, x_n . Supongamos que repetimos el experimento N veces y sea N_i el número de veces que la variable tomo el valor x_i , con $i = 1, \dots, n$. Podemos entonces definir la probabilidad P_i de observar el valor x_i como:

$$P_i = \lim_{N \rightarrow \infty} \frac{N_i}{N} \quad (1.1)$$

Dado que $\sum_{i=1}^n N_i = N$, tenemos que $0 \leq P_i \leq 1$ para $i = 1, \dots, n$ y $\sum_{i=1}^n P_i = 1$.

Notemos que la Ec.(1.1) es una definición *a posteriori*, esto es, para conocer la probabilidad de ocurrencia de un resultado en un experimento, primero debemos realizar el experimento un número grande de veces. Por otra parte, no tenemos manera de probar que este límite realmente existe y es independiente del conjunto de experimentos, lo cual debe ser asumido como un axioma. Vamos entonces a buscar una formulación que nos permita estimar probabilidades *a priori*, a partir de propiedades intrínsecas del sistema. Para ello vamos a plantear un conjunto diferente de axiomas. Sin embargo, el concepto anterior nos resultará útil como guía para una definición mas formal, ya que vamos a exigir que ambas definiciones coincidan en la circunstancias apropiadas. Supongamos entonces que un posible resultado de un experimento es A y que por algún método le asignamos una probabilidad de ocurrencia $P(A)$. Entonces si llevamos a cabo N experimentos idénticos, esperamos que el evento A ocurra aproximadamente un número $N P(A)$ de veces. De hecho, vamos a exigir que en el límite de N muy grande la fracción relativa de veces que ocurre A converja a $P(A)$.

1.1. Definiciones básicas y axiomas

Vamos a comenzar definiendo el **espacio muestral** de un experimento como un conjunto S , tal que cualquier resultado del experimento corresponda a uno o mas elementos de S . Un **evento** es cualquier subconjunto de S . Un **evento simple** es un elemento del conjunto. Dos eventos A y B se dicen **mutuamente excluyentes** si $A \cap B = \emptyset$; eventos simples son siempre mutuamente excluyentes.

La **probabilidad** de un evento $A \subset S$ se define como una función $P : S \rightarrow [0, 1]$ con las siguientes propiedades:

1. **Axioma 1:** $P(S) = 1$
2. **Axioma 2:** Para cualquier secuencia (finita o infinita) de eventos E_i mutuamente excluyentes entre sí, esto es, $E_i \cap E_j = \emptyset$ para todo par $i \neq j$, entonces

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Las siguientes propiedades se deducen de manera inmediata.

- Si tomamos como E_i todos los eventos simples (esto es, todos los elementos de S) sabemos que $S = \bigcup_i E_i$ y del axioma 2 se sigue que $P(S) = \sum_i P(E_i)$. Del axioma 1 tenemos entonces que

$$\sum_i P(E_i) = 1$$

Esta propiedad es válida en general para cualquier secuencia de eventos (subconjuntos) mutuamente excluyentes y exhaustiva, esto es, una secuencia tal que $S = \bigcup_i E_i$. Un caso particular de lo anterior es el siguiente. Sea E un evento cualquiera y sea \bar{E} su complemento, esto es, $S = E \cup \bar{E}$, con $E \cap \bar{E} = \emptyset$. Entonces se sigue que $P(\bar{E}) = 1 - P(E)$.

- Dado que el conjunto vacío es el complemento de S , del resultado anterior entonces $P(S) + P(\emptyset) = 1$. Por el axioma 1 se sigue que $P(\emptyset) = 0$.

Las propiedades anteriores tienen una interpretación bastante directa. $P = 1$ implica una certeza absoluta en el resultado del experimento. Un resultado del experimento con probabilidad uno es determinista. Así, la probabilidad de que *algo* ocurra $P(S) = 1$. Por otra parte, un resultado con probabilidad cero implica que el mismo no ocurriera nunca. Así, la probabilidad de que no ocurra *nada* es $P(\emptyset) = 0$, ya que el experimento siempre arroja algún resultado.

El axioma 2 nos da una regla de aditividad para las probabilidades. La probabilidad de eventos que no pueden ocurrir simultáneamente es la suma de las probabilidades de los eventos correspondientes.

Dados dos eventos A y B , $P(A \cup B)$ nos da la probabilidad de que ocurra A , B o ambos a la vez. Por otra parte, $P(A \cap B)$ es la probabilidad de ocurran ambos eventos simultáneamente y solo eso. Esta última se denomina **probabilidad conjunta** de A y B . Estas probabilidades están relacionadas a través de

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.2}$$

Esta propiedad puede demostrarse fácilmente mediante el diagrama de Venn de la Fig.1.1. El conjunto $P(A \cup B)$ puede dividirse en tres conjuntos disjuntos: *I*, *II* y *III* (eventos mutuamente

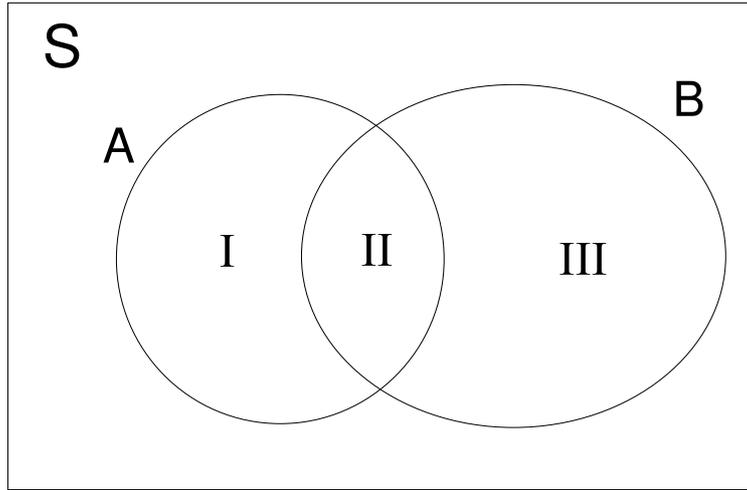


Figura 1.1: Diagrama de Venn para la unión de dos eventos.

excluyentes). En otras palabras, I es el conjunto de todos los elementos en A que no están en B , III lo inverso y II el conjunto de todos los elementos que pertenecen simultáneamente a A y B . Así $II = A \cap B$, $A = I \cup II$ y $B = III \cup II$. Además

$$A \cup B = I \cup II \cup III$$

Dado que I , II y III son mutuamente excluyentes

$$P(A \cup B) = P(I) + P(II) + P(III)$$

$$P(A) = P(I) + P(II)$$

$$P(B) = P(II) + P(III)$$

de donde se obtiene la Eq.(1.2).

En muchos experimentos resulta natural asumir que todos los resultados posibles (eventos simples) tienen la misma probabilidad de ocurrencia. Así, por ejemplo, si nuestro experimento es arrojar una moneda, el espacio muestral será $S = \{cara, seca\}$. Si la moneda no está cargada, es de esperar que ambos eventos simples tengan la misma probabilidad. Otro ejemplo puede ser el resultado de arrojar un dado de 6 caras. El espacio muestral consiste de 6 eventos simples $S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$, donde E_i corresponde al resultado i . Para un dado no cargado todos estos resultados deberían tener la misma probabilidad. Consideremos entonces un caso genérico de N posibles resultados E_i con $i = 1, \dots, N$. Dado que

$$\sum_{i=1}^N P(E_i) = 1$$

se sigue inmediatamente que

$$P(E_i) = \frac{1}{N}$$

Así, en el ejemplo de la moneda $P(E_i) = 1/2$ y para el ejemplo del dado $P(E_i) = 1/6$. De esto se sigue que, en un espacio muestral finito, para cualquier evento E

$$P(E) = \frac{\text{número de elementos en } E}{\text{número de elementos en } S} \quad (1.3)$$

En otras palabras, si asumimos que todos los resultados de un experimento son igualmente probables, la probabilidad de cualquier evento E es igual a la fracción de elementos en el espacio muestral contenidos en E . Por ejemplo, para el caso del dado, la probabilidad de que el resultado sea menor que 3 será $P(E) = 2/6 = 1/3$.

Un concepto de gran utilidad es la **probabilidad condicional** $P(A|B)$, la cual nos da la probabilidad de que ocurra el evento A *dado que tenemos la certeza de que ocurre* B . $P(A|B)$ se define por la ecuación

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.4)$$

Podemos entender esta ecuación si consideramos el ejemplo simple de un espacio muestral con N eventos equiprobables. Sea N_A , N_B y N_{AB} el número de elementos contenidos en A , B y $A \cap B$ respectivamente. Entonces, $P(A) = N_A/N$, $P(B) = N_B/N$ y $P(A \cap B) = N_{AB}/N$. Dado que tenemos la certeza de que ocurre B , para esta situación particular B es nuestro espacio muestral. Dado que todos los eventos en B a su vez son equiprobables, tenemos que $P(A|B) = N_{AB}/N_B$. Dividiendo numerador y denominador de la ecuación anterior por N se obtiene la Ec.(1.4). Dado que $P(A \cap B) = P(B \cap A)$, tenemos también que

$$P(A|B) P(B) = P(B|A) P(A) \quad (1.5)$$

El concepto de probabilidad condicional nos permite introducir el concepto de **eventos independientes**. Dos eventos A y B son independientes *si y solo si*:

$$P(A|B) = P(A) \quad (1.6)$$

De la Ec.(1.5) esto implica inmediatamente que $P(B|A) = P(B)$. Por otra parte, de la Ec.(1.4) tenemos que

$$P(A \cap B) = P(A) P(B) \quad (1.7)$$

esto es, si dos eventos son independientes, su probabilidad conjunta es el producto de las probabilidades individuales.

1.2. Cálculo de probabilidades en espacios finitos: elementos de análisis combinatorio

Consideremos el siguiente ejemplo. Se lanza una moneda perfecta tres veces. Cual es la probabilidad de obtener cara en dos de los tres lanzamientos? Un evento simple en este experimento consiste en un conjunto ordenado de tres elementos binarios C ó S (cara o seca), tal como CSS (cara en la primera tirada y seca en las dos siguientes). Dado que la moneda es perfecta, podemos asumir que los eventos simples son equiprobables. Así, el primer paso para resolver el problema es determinar el número total de puntos muestrales (eventos simples) y el segundo paso determinar el número de puntos muestrales que contiene dos C. La manera mas simple es enumerar todos los puntos muestrales:

CCC CCS CSC SCC

CSS SCS SSC SSS

Tenemos 8 puntos muestrales y por lo tanto la probabilidad de cada uno de ellos será $1/8$; de todos los puntos muestrales en tres hay dos caras y por lo tanto la probabilidad buscada es $3/8$.

Es evidente que esta técnica tan simple solo es util cuando el número posible de puntos muestrales es pequeño. En la mayoría de los casos de interés, los espacios muestrales son enormes y enumerar todos los puntos muestrales resulta imposible. Sin embargo, para calcular probabilidades de eventos solo necesitamos *contar* cuantos puntos muestrales tenemos y para ello no es necesario enumerarlos a todos. Esto se puede ver en el siguiente ejemplo.

Supongamos una reunión de 20 personas, en principio sin ninguna relación directa entre ellos. Cual es la probabilidad de que al menos 2 de ellas cumplan años el mismo día?

Para empezar a resolver este problema, podemos en primer lugar invertir la pregunta. Si consideramos el evento “todas las 20 personas cumplen años en días diferentes” es evidente que es el complemento del evento “al menos 2 de ellas cumplen años el mismo día”; si llamamos P_1 y P_2 a las probabilidades de ambos eventos tenemos que $P_1 + P_2 = 1$, o $P_2 = 1 - P_1$. Calculemos entonces P_1 (es mas facil que calcular directamente P_2).

Supongamos que numeramos los días del año de 1 a 365 (ignoramos por simplicidad los años bisiestos). Un punto muestral de este experimento consiste en un conjunto de 20 números (cada uno entre 1 y 365) donde el primero representa el cumpleaños de la primera persona, el segundo número el de la segunda persona, etc.. El número de puntos muestrales N_s será entonces el número de 20-adas diferentes que podemos formar con 365 números. Para cada persona tenemos 365 posibilidades diferentes. Dado que los números (fechas) pueden repetirse para diferentes personas $N_s = 365^{20}$. Calculemos ahora el número de puntos muestrales N_1 en los cuales no se repite ningún número entre los 20. Para la primera persona tenemos 365 posibilidades. Por cada una de ellas, tenemos solo 364 para la segunda persona (todas los números posibles, menos el que se selecciono antes). Para la tercera persona tenemos 363, etc. Así

$$N_1 = 365 \times 364 \times 363 \times \dots \times 346$$

Así

$$P_1 = \frac{N_1}{N_s} = \frac{365 \times 364 \times 363 \times \dots \times 346}{365^{20}} \approx 0,59$$

y por lo tanto $P_2 \approx 0,41$. Existe un 41 % de probabilidad de que en cualquier reunión de 20 personas al menos 2 cumplan años el mismo día. En un grupo de 50 personas esta probabilidad asciende a un 97 %.

Como ilustra el ejemplo anterior, al tratar con grandes cantidades de objetos a menudo tenemos que calcular el numero de diferentes maneras de combinar los mismos. Este tipo de cálculo entra dentro del análisis combinatorio, del cual veremos algunas nociones básicas.

Una **permutación** es cualquier arreglo de un conjunto de N objetos diferentes entre sí en un orden definido. Si queremos contar el número total de posibles permutaciones de N objetos, podemos pensar que tenemos N casilleros ordenados y distribuimos los objetos uno en cada casillero. Para el primer casillero tenemos N posibilidades. Por cada elección para el primer casillero procedemos a llenar los siguientes con los que quedan. Así, para el segundo tenemos $N - 1$ posibilidades por cada una de las N del primero. A su vez, por cada una de las $N - 1$ posibilidades del segundo tenemos $N - 2$ para el tercero y así sucesivamente. El número total de posibles permutaciones de N objetos es por lo tanto:

$$N \times (N - 1) \times (N - 2) \times \dots \times 2 \times 1 = N!$$

Supongamos ahora que queremos calcular el número de permutaciones diferentes de R objetos tomados entre un conjunto de $N > R$ (como en el ejemplo de los cumpleaños). Este número se denomina P_R^N : **permutaciones de N objetos tomados de a R** . El procedimiento para obtener

P_R^N es semejante al anterior. En este caso, podemos pensar que tenemos R casilleros ordenados y N objetos diferentes para llenarlos. Al igual que antes, tenemos N posibilidades para llenar el primero, $N - 1$ para el segundo y así hasta completar los R . Por lo tanto

$$P_R^N = N \times (N - 1) \times (N - 2) \times \cdots \times (N - R + 1) = \frac{N!}{(N - R)!}$$

Así, en el ejemplo de los cumpleaños $N_1 = P_{20}^{365} = 365!/(365 - 20)!$.

Una **combinación** es una selección de R objetos tomados entre un conjunto de N , pero sin importar *el orden* de los objetos. El número de posibles combinaciones de N objetos tomados de a R se denota C_R^N . Podemos obtener C_R^N a partir de P_R^N . En P_R^N cada selección de un conjunto particular de R objetos aparece repetida a través de todas las permutaciones posibles de los mismos, esto es, $R!$. Tenemos entonces que $P_R^N = R! C_R^N$ y por lo tanto

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

C_R^N también se suele llamar número combinatorio o coeficiente binomial, ya que es el coeficiente que aparece en el desarrollo del binomio de Newton:

$$(x + y)^N = \sum_{k=0}^N x^k y^{N-k} \binom{N}{k} \quad (1.8)$$

Algunas propiedades inmediatas de los coeficientes binomiales son

$$\binom{N}{0} = \binom{N}{N} = 1$$

$$\binom{N}{1} = N$$

En el ejemplo de las tres monedas, el número de eventos con dos caras es $C_2^3 = 3$.

El número de permutaciones de N objetos, en los cuales hay n_1 idénticos (por lo tanto no importa el orden) de un tipo, n_2 de otro tipo, ..., n_p elementos idénticos del tipo p , viene dado por el **coeficiente multinomial**:

$$\binom{N}{n_1, \dots, n_p} = \frac{N!}{n_1! \dots n_p!}$$

dado que los mismo aparecen en la fórmula multinomial

$$(x_1 + \cdots + x_p)^N = \sum x_1^{n_1} x_2^{n_2} \cdots x_p^{n_p} \binom{N}{n_1, \dots, n_p}$$

donde la suma comprende todos los valores de n_1, \dots, n_p entre cero y N , tal que $\sum_{i=1}^p n_i = N$.

1.2.1. Ejercicios

1. Una baraja francesa consta de 52 cartas, dividida en cuatro grupos (palos) con 13 figuras (números) cada grupo. En el juego de poker se reparten 5 cartas elegidas al azar.
 - a) Calcule la probabilidad de obtener un poker: cuatro cartas del mismo número.

- b)* Calcule la probabilidad de obtener un full: tres cartas del mismo número y las dos restantes iguales a otro número.
 - c)* Calcule la probabilidad de obtener un par simple: dos cartas del mismo número y las tres restantes con números diferentes entre sí y diferentes a los anteriores.
 - d)* Calcule la probabilidad de obtener un par doble: dos pares de cartas con números iguales en cada par pero diferentes entre pares y una quinta diferente de los dos grupos anteriores.
2. Se arroja una moneda 3 veces. Sea E_i el evento “sale cara en la tirada i -ésima” ($i = 1, 2, 3$). Muestre que los tres eventos son independientes entre sí.

Capítulo 2

VARIABLES ALEATORIAS Y PROBABILIDAD

Una variable cuyo valor está determinado por la ocurrencia de un evento aleatorio se denomina **variable aleatoria o estocástica**. En otras palabras, una variable aleatoria X es una función del espacio muestral S en los números reales. En un dado experimento, una variable aleatoria puede tomar diferentes valores. Debemos entonces tener cuidado en distinguir entre la variable (que denotaremos con letras mayúsculas) y sus posibles valores $\{x_i\}$ que puede tomar en cada realización del experimento. Por ejemplo, el número de caras que aparece en una tirada de tres monedas es una variable aleatoria X , cuyos posibles valores son $x = 0, 1, 2, 3$.

2.1. Variables aleatorias discretas

Una variable aleatoria X que puede tomar un conjunto numerable (finito o infinito) de valores $X(S) = x_1, x_2, \dots$ se dice **discreta**. Se define la **distribución de probabilidad** $P(x)$ de una variable aleatoria X como la probabilidad de que X tome el valor x , y viene dada por la suma de las probabilidades de todos los puntos muestrales en S para los cuales X toma el valor x . La misma satisface las propiedades

$$0 \leq P(x) \leq 1 \quad \forall x$$

$$\sum_x P(x) = 1$$

Se define el **valor esperado o valor medio** de la variable x como

$$\langle x \rangle \equiv \sum_x x P(x) \tag{2.1}$$

y el mismo representa un promedio pesado de la variable x . De la misma manera, se define el valor medio de una función arbitraria de $g(x)$ como

$$\langle g(x) \rangle \equiv \sum_x g(x) P(x) \tag{2.2}$$

La **variancia** $V(X)$ se define como

$$V(X) \equiv \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \tag{2.3}$$

y el **desvío estándar** como $\sigma_X = \sqrt{V(X)}$. El valor medio nos da una medida del promedio esperable de los valores de X si el experimento se repite muchas veces. El desvío estándar nos da una medida de cuán dispersos estarán estos resultados respecto del valor medio. Desarrollaremos estos conceptos con mayor detalle más adelante.

El momento n -ésimo de una variable X se define como

$$\langle x^n \rangle \equiv \sum_x x^n P(x) \quad (2.4)$$

Veremos a continuación algunos ejemplos de distribuciones de probabilidad que aparecen con frecuencia en la práctica.

2.1.1. La distribución de probabilidad binomial

Una de las aplicaciones más comunes de la Teoría de Probabilidades es el caso de un número n muy grande de experimentos, cada uno de los cuales tiene solo dos posibles resultados. Un ejemplo típico es una encuesta de opinión acerca de una votación por un plebiscito (votación por SI o por NO, donde el voto en blanco no está permitido). La empresa encuestadora selecciona una “muestra”, esto es un subconjunto, de n personas dentro de un espacio muestral enorme con $N \gg n$ elementos. Si bien cada persona tiene perfectamente definido su voto, los encuestados son elegidos completamente al azar. Supongamos entonces que una fracción p de los votantes votará por SI. Dado que solo hay dos posibilidades, una fracción $1 - p$ votará por NO. Si se elige entonces una persona al azar, la probabilidad de que vote por SI será justamente p (Ec.(1.3)). La pregunta entonces es: ¿Cuál es la probabilidad de que exactamente x entre los n voten por SI?

Un experimento binomial tiene entonces las siguientes características:

1. El experimento consta de n pruebas idénticas.
2. Cada prueba tiene dos resultados posibles. Llamaremos genéricamente *éxito* E y *fracaso* F .
3. La probabilidad de tener éxito en una sola prueba es p y permanece constante de prueba en prueba (la probabilidad de fracaso por lo tanto es $(1 - p)$).
4. Las pruebas son independientes entre sí.
5. La variable de estudio es X , el número de éxitos observados en las n pruebas.

Otro ejemplo de un experimento binomial sería arrojar n monedas perfectas y contar el número de veces que aparece cara. En este caso $p = 1/2$.

Los puntos muestrales de este experimento consisten en cadenas binarias del tipo

EEEEEEFF...FEFF

Supongamos una cadena particular conteniendo x valores E y $n - x$ valores F y calculemos su probabilidad. Este evento es la intersección de n eventos *independientes*, de los cuales x tienen probabilidad p y $n - x$ tienen probabilidad $1 - p$. Por lo tanto, la probabilidad de la intersección es $p^x(1 - p)^{n-x}$. Ahora bien, esa es la probabilidad de un conjunto *particular* conteniendo x valores E y $n - x$ valores F en un determinado orden. Si cambiamos el orden de las letras, sin alterar la cantidad de letras E y F , obtenemos otro conjunto diferente con la misma probabilidad. Dado que dichos eventos son mutuamente excluyentes, la probabilidad de la unión es la suma de las probabilidades. La probabilidad de obtener x valores E y $n - x$ valores F , sin importar el orden (es lo que buscamos) será entonces $p^x(1 - p)^{n-x}$ multiplicado por el número de *combinaciones* posibles de dichas letras, esto es, C_x^n . Así

$$P(x) = p^x(1 - p)^{n-x} \binom{n}{x} \quad \text{para } x = 0, 1, \dots, n \quad (2.5)$$

Esta es la **distribución binomial**. Podemos verificar que la misma está normalizada ($\sum_x P(x) = 1$):

$$\sum_{x=0}^n P(x) = \sum_{x=0}^n p^x (1-p)^{n-x} \binom{n}{x} = [p + (1-p)]^n = 1$$

donde hemos usado el binomio de Newton Ec.(1.8). Podemos calcular también el valor medio de X :

$$\begin{aligned} \langle x \rangle &= \sum_{x=0}^n x p^x (1-p)^{n-x} \binom{n}{x} \\ &= \sum_{x=1}^n p^x (1-p)^{n-x} \frac{n!}{(n-x)! (x-1)!} \\ &= np \sum_{x=1}^n p^{x-1} (1-p)^{n-x} \frac{(n-1)!}{(n-x)! (x-1)!} \\ &= np \sum_{y=0}^{n-1} p^y (1-p)^{n-1-y} \frac{(n-1)!}{(n-1-y)! y!} \\ &= np \sum_{y=0}^{n-1} p^y (1-p)^{n-1-y} \binom{n-1}{y} \\ &= np \end{aligned} \tag{2.6}$$

En forma semejante se puede demostrar que (Ej. 2.1.4-1)

$$V(X) = np(1-p). \tag{2.7}$$

2.1.2. La distribución de probabilidad geométrica

Supongamos ahora que realizamos una experimento semejante al binomial, pero en lugar de realizar n pruebas, terminamos el experimento cuando aparece por primer vez una E . La variable aleatoria que nos interesa aquí es el número X de la prueba para la cual se obtuvo el primer éxito. La aparición del primer éxito puede tener lugar en la primera prueba, en la segunda o nunca. Así, la variable X en este caso no está acotada. Los elementos del espacio muestral en este caso son: $E_1 = E$ (éxito en la primera prueba), $E_2 = FE$ (éxito en la segunda), $\dots, E_k = FFF \dots E$ (éxito en la k -ésima), etc. Dado que las pruebas son independientes, la probabilidad es

$$P(x) = (1-p)^{x-1} p \quad \text{para } x = 1, 2, \dots, \tag{2.8}$$

Esta se conoce como **distribución geométrica**. Dado que $(1-p) \leq 1$, vemos que la distribución geométrica decae exponencialmente con x , a menos que $p = 1$, en cuyo caso $P(1) = 1$ y $P(x) = 0$ en cualquier otro caso. Verifiquemos la normalización:

$$\sum_{x=1}^{\infty} P(x) = \sum_{x=1}^{\infty} p (1-p)^{x-1} = \frac{p}{1-p} \sum_{x=1}^{\infty} (1-p)^x = \frac{p}{1-p} \left[\frac{1}{[1 - (1-p)]} - 1 \right] = 1$$

donde hemos usado la suma de la serie geométrica $\sum_{x=0}^{\infty} a^x = 1/(1-a)$. Veamos el valor medio ($q \equiv 1-p$):

$$\langle x \rangle = p \sum_{x=1}^{\infty} x q^{x-1} = p \frac{d}{dq} \left(\sum_{x=1}^{\infty} q^x \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = \frac{1}{p}. \tag{2.9}$$

Este resultado es fácil de interpretar. Cuanto menor sea el valor de p , más lentamente decae $P(x)$ y por lo tanto mayor es el valor medio. En forma semejante, puede demostrarse que (Ej. 2.1.4-1)

$$V(X) = \frac{2}{p^2} - \frac{1}{p}. \quad (2.10)$$

2.1.3. La distribución de Poisson

Supongamos que tenemos una sustancia radioactiva y un contador Geiger. La sustancia emite una partícula aleatoriamente en el tiempo cada vez que un átomo decae y marca un conteo en el Geiger. Supongamos que el tiempo de vida media de la sustancia es muy grande comparado con el tiempo de observación, de manera que el número de conteos es relativamente pequeño (no estamos pensando en una reacción en cadena). Cada evento de decaimiento entonces puede considerarse independiente de los otros. Nos interesa entonces determinar la probabilidad de observar X conteos en un intervalo de tiempo dado τ . Podemos llevar a cabo este cálculo subdividiendo el intervalo de tiempo en n subintervalos, de tal manera que τ/n sea suficientemente pequeño para que la probabilidad de que ocurra más de un decaimiento en un subintervalo sea despreciable. Sea p la probabilidad de que ocurra un conteo en un subintervalo. Claramente, p y los resultados subsecuentes van a depender de n . Ahora bien, la manera de independizarnos de la longitud del subintervalo es tomar el límite en que este va a cero. Si bien p depende de la longitud del subintervalo, podemos asumir que es la misma para cualquiera de ellos (para una longitud τ dada). En ese caso, la distribución de probabilidad para X es binomial. Si bien no sabemos en principio como depende p de la longitud del subintervalo, parece razonable que p disminuya a medida que la misma disminuye, es decir, que n aumenta. La dependencia más simple que satisface estos criterios es $p = \lambda/n$, donde λ es una constante. En otras palabras, vamos a asumir que el valor medio de conteos $p n$ se mantiene constante a medida que aumentamos n . La distribución de probabilidad para X se obtiene entonces tomando el límite $n \rightarrow \infty$ de la distribución binomial con $p = \lambda/n$:

$$\begin{aligned} \lim_{n \rightarrow \infty} p^x (1-p)^{n-x} \binom{n}{x} &= \lim_{n \rightarrow \infty} \frac{n(n-1) \cdots (n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \frac{n(n-1) \cdots (n-x+1)}{n^x} \\ &= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \end{aligned}$$

Todos los factores a la derecha tienden a uno, mientras que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Así

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad (2.11)$$

Esta se conoce como **distribución de Poisson**. Esta distribución se aplica en general cada vez que tenemos un proceso que ocurre aleatoriamente en un intervalo de tiempo o espacio, cuando la probabilidad de ocurrencia de un evento es constante e independiente de los demás eventos. Algunos otros ejemplos de aplicación son el número de autos que pasan a través de un cierto punto en una ruta durante un periodo definido de tiempo, el número de llamadas telefónicas en una central telefónica por minuto, el número de animales muertos encontrados por unidad de longitud de ruta, etc. La distribución fue descubierta por Siméon-Denis Poisson (1781-1840) quien la publicó, junto

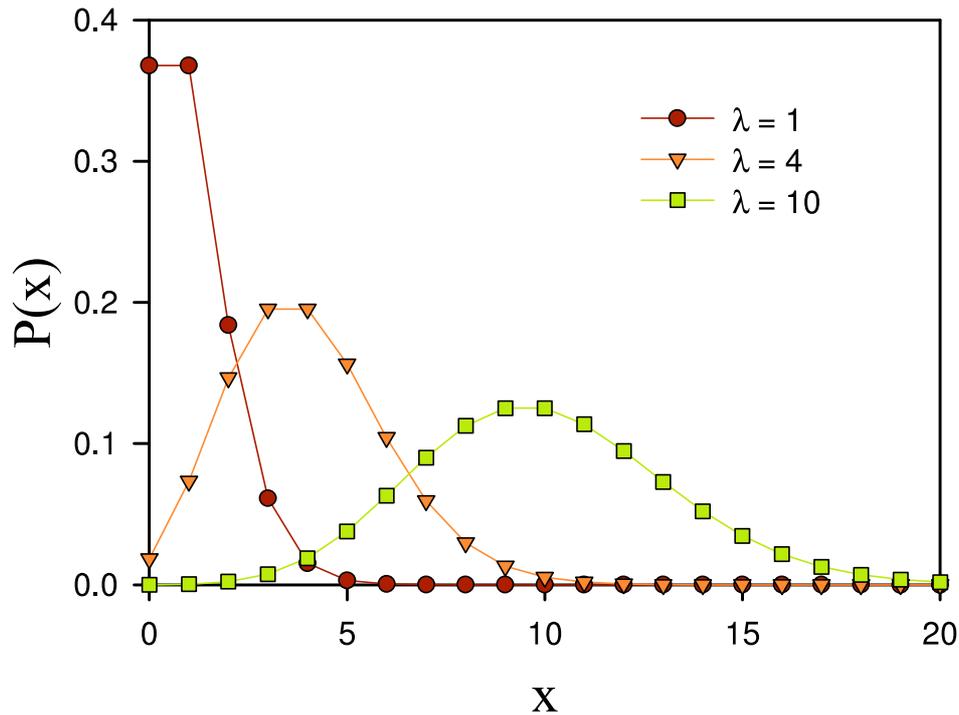


Figura 2.1: Distribución de Poisson para diferentes valores del valor medio λ .

con su teoría de probabilidad, en 1838 en su trabajo *Recherches sur la probabilité des jugements en matières criminelles et matière civile* ("Investigación sobre la probabilidad de los juicios en materias criminales y civiles").

Verifiquemos la normalización:

$$\sum_{x=0}^{\infty} P(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda} e^{-\lambda} = 1$$

y el valor medio

$$\langle x \rangle = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda \quad (2.12)$$

De la misma manera se puede demostrar que (Ej. 2.1.4-1)

$$V(X) = \lambda. \quad (2.13)$$

En la Fig.2.11 se muestran algunos ejemplos de la distribución de Poisson para diferentes valores de λ .

2.1.4. Ejercicios

1. Demuestre las ecuaciones (2.7), (2.10) y (2.13)

2.2. Variables aleatorias continuas

Una variable aleatoria X que puede tomar un conjunto no-numerable de valores en un dado intervalo del eje real se dice continua. Un intervalo (a, b) contenido en el dominio de definición de la

variable corresponde entonces a un evento. Vamos a introducir entonces una **densidad de probabilidad** $f_X(x)$, tal que la probabilidad de que X tome valores dentro del intervalo comprendido entre x y $x + dx$ esta dada por $f_X(x) dx$. La probabilidad de que X tome valores en un intervalo finito (a, b) viene entonces dada por

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

La densidad de probabilidad debe ser continua a tramos, satisfacer $f_X(x) \geq 0$ y

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

Notemos que, de acuerdo con esta definición, la probabilidad de que una variable continua tome un valor bien definido $P(X = x) = 0$.

Podemos englobar también dentro de esta definición a las variables aleatorias discretas. Si X es una variable aleatoria discreta que toma valores x_1, x_2, \dots , con probabilidad $p_i = P(x_i)$, entonces su densidad de probabilidad será

$$f_X(x) = \sum_i p_i \delta(x - x_i)$$

donde $\delta(x)$ es la función delta de Dirac.

Se define la **Función de distribución** $F_X(x)$ (no hay que confundirla con la distribución de probabilidad de una variable discreta) como la probabilidad $P(X \leq x)$:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(x') dx'$$

Se sigue entonces que $f_X(x) = dF_X(x)/dx$. Dado que f_X es no-negativa, la función de distribución es siempre no-decreciente. Por la normalización de f_X , la función de distribución toma los valores límite $F_X(-\infty) = 0$ y $F_X(\infty) = 1$. Para el caso de una variable aleatoria discreta tendremos que

$$F_X(x) = \sum_i p_i \Theta(x - x_i)$$

donde $\Theta(x)$ es la función escalón de Heaviside, esto es, $\Theta(x) = 0$ para $x \leq 0$ y $\Theta(x) = 1$ para $x > 0$.

El momento n-ésimo de la variable X se define entonces como

$$\langle x^n \rangle \equiv \int_{-\infty}^{\infty} x^n f_X(x) dx$$

Veamos un poco la interpretación de estas cantidades. Todas las propiedades de la variable X están contenidas en la función densidad (algo así como una “relación fundamental” para la variable). La probabilidad de que X tome valores en un dado intervalo está dada por el área bajo f_X para ese intervalo.

El primer momento $\langle x \rangle$ (**media o valor medio**) nos dá el “centro de masa” de la densidad f_X . Esta cantidad a menudo se confunde con otras dos cantidades: el **valor mas probable** x_p y la **mediana** x_m . El valor mas probable de X se define como el máximo de f_X . La mediana se define como el valor de x que divide el área bajo la curva $f_X(X)$ en partes iguales. En otras palabras, $F_X(x_m) = 1/2$. En algunos casos estas cantidades coinciden (como veremos mas adelante) pero en general son diferentes.

El segundo momento $\langle x^2 \rangle$ nos dá el “momento de inercia” de la densidad f_X respecto del origen. El desvío estándar $\sigma_X = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$ nos dá una medida de cuán lejos se dispersa la probabilidad respecto de la media $\langle x \rangle$, esto es, el desvío medio cuadrático.

Supongamos una variable tal que $\langle x \rangle = 0$ (siempre podemos hacer que esto ocurra desplazando el origen, esto es, restando la media de la variable). El tercer momento $\langle x^3 \rangle$ nos dá una medida de cuán asimétrica es la distribución respecto del origen. Esto es, si la densidad es simétrica respecto del origen, es una función par y por lo tanto el tercer momento se anula. Cuanto mayor sea el tercer momento, menos “simétrica” será f_X .

Veamos algunos ejemplos. Un variable tiene distribución **uniforme** si

$$f_X(x) = \begin{cases} A & \text{si } a \leq x \leq b \\ 0 & \text{en otro caso} \end{cases} \quad (2.14)$$

donde A es una constante. La normalización exige que $A = 1/(b - a)$. La probabilidad de que X tome valores en un intervalo (c, d) , con $a < c < d < b$ resulta $(d - c)/(b - a)$. Un cálculo directo muestra que $\langle x \rangle = (b + a)/2$, esto es, el centro del intervalo. De la misma forma es fácil mostrar que $\sigma_X = (b - a)/\sqrt{12}$. Un cálculo directo nos muestra que la probabilidad de que

$$P(\langle x \rangle - \sigma_X \leq x \leq \langle x \rangle + \sigma_X) \approx 0,58$$

Esto es, el intervalo $\pm \sigma_X$ alrededor de la media concentra aproximadamente el 60% de la probabilidad. La función de distribución se obtiene también fácilmente:

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ A(x - a) & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases} \quad (2.15)$$

De aquí puede verificarse inmediatamente que $x_m = \langle x \rangle$.

2.2.1. Distribución de Gauss

Otro ejemplo de gran importancia es la **distribución de Gauss** o **normal**, definida por:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (2.16)$$

definida para todo x real, con $\sigma > 0$. El lector puede verificar por integración directa que la misma está normalizada y que $\langle x \rangle = \mu$, $\sigma_X = \sigma$. Esta curva tiene su máximo en $x = \mu$ y es simétrica respecto del valor medio. Por lo tanto en este caso la media, la mediana y el valor más probable coinciden. Mediante integración numérica (o mediante valores de tablas), puede verificarse que la probabilidad de que la variable tome valores en un intervalo $\pm \sigma$ alrededor de la media es aproximadamente 68%, mientras que para un intervalo $\pm 2\sigma$ la probabilidad es de aproximadamente 95%.

2.2.2. Distribución exponencial

Un ejemplo de una distribución no simétrica es la **distribución exponencial**

$$f_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{\mu} e^{-x/\mu} & \text{si } x \geq 0 \end{cases} \quad (2.17)$$

El lector puede verificar por integración directa que la misma está normalizada y que $\langle x \rangle = \mu$, $\sigma_X = \mu$. Esta curva tiene su máximo en $x = 0$. Por otra parte, la función de distribución es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x/\mu} & \text{si } x \geq 0 \end{cases} \quad (2.18)$$

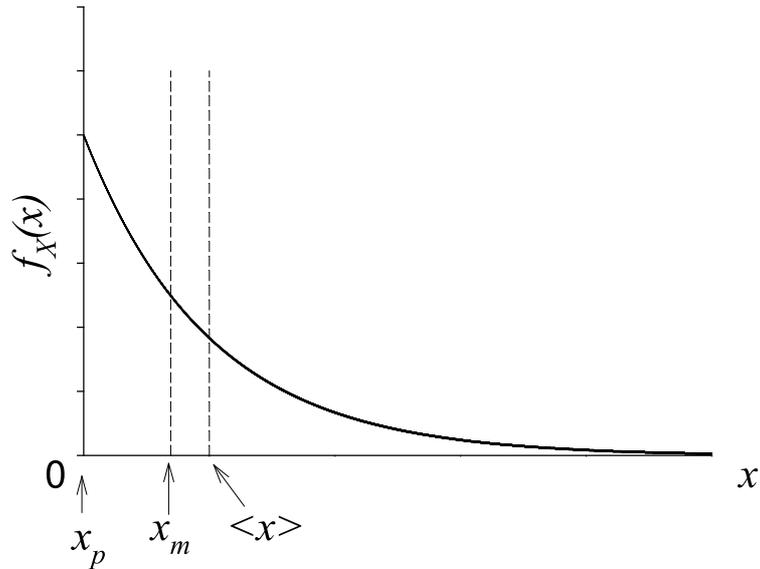


Figura 2.2: Distribución exponencial, media, mediana y valor mas probable.

Si resolvemos la ecuación $F(x_m) = 1/2$, obtenemos $x_m = \mu \ln 2$. Vemos que en este caso la media, la mediana y el valor mas probable no coinciden. Los mismos se muestran en la Fig.2.2 junto con la densidad.

2.2.3. Transformación de variables aleatorias

Sea $Y = g(X)$, donde $g(x)$ es una función arbitraria y X una variable aleatoria. Siendo X aleatoria, Y también lo es. Supongamos que conocemos la densidad de probabilidad $f_X(x)$. Nos preguntamos entonces cual es la densidad de probabilidad $f_Y(y)$. Para relacionar ambas cantidades notemos que

$$\langle y \rangle = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_{-\infty}^{\infty} y f_Y(y) dy \quad (2.19)$$

Resulta inmediato verificar que

$$f_Y(y) = \int_{-\infty}^{\infty} \delta(y - g(x)) f_X(x) dx \quad (2.20)$$

satisface la condición anterior. Usando la propiedad de la delta de Dirac:

$$\delta(h(z)) = \sum_i \frac{1}{|h'(z_i)|} \delta(z - z_i)$$

si $h'(z_i) \neq 0$, donde z_i son los ceros de $h(z)$, resulta

$$f_Y(y) = \sum_i \left| \frac{dg_i^{-1}(y)}{dy} \right| f_X[g_i^{-1}(y)] \quad (2.21)$$

donde $g_i^{-1}(y)$ son las diferentes ramas de la función inversa de $g(x)$ (estamos suponiendo el caso general en que la función $g(x)$ no tiene una inversa única). La Ec.(2.21) puede interpretarse fácilmente. Supongamos primero que la función $g(x)$ tiene inversa única. Entonces

$$f_Y(y) = \left| \frac{dx(y)}{dy} \right| f_X[x(y)]$$

Si $g(x)$ es creciente ($g'(x) > 0$) y por lo tanto $dx(y)/dy > 0$ entonces

$$f_Y(y) dy = f_X(x) dx$$

Esta ecuación nos dice que la probabilidad de que Y tome valores entre y e $y + dy$ es igual a la probabilidad de que X tome valores entre x y $x + dx$, donde $dy = g'(x) dx$. Si $g(x)$ es decreciente, su derivada es negativa y por lo tanto

$$f_Y(y) dy = -f_X(x) dx$$

ya que $f_Y(y)$ debe ser positiva. Si la función $g^{-1}(y)$ es multivaluada, entonces podemos repetir las consideraciones anteriores, pero para evaluar la probabilidad de que Y tome valores entre y e $y + dy$, debemos sumar las probabilidades de todas las ramas de $g^{-1}(y)$ correspondientes al intervalo de X entre x y $x + dx$.

2.2.4. Función característica y desarrollo en cumulantes

La función característica $\tilde{f}_X(k)$ correspondiente a una variable aleatoria X se define como

$$\tilde{f}_X(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} e^{ikx} f_X(x) dx \quad (2.22)$$

esto es, la transformada de Fourier de $f_X(x)$. Desarrollando en serie de Taylor la exponencial e integrando término a término tenemos:

$$\tilde{f}_X(k) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle x^n \rangle}{n!} \quad (2.23)$$

Es importante notar que el desarrollo anterior solo es válido si los momentos decrecen suficientemente rápido con n como para que la serie converja; en otras palabras, el intercambio entre la serie y la integración no siempre es válido. En el caso en que este desarrollo sea válido, vemos que si tenemos todos los momentos podemos reconstruir la densidad de probabilidad $f_X(x)$ antitransformando

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} \tilde{f}_X(k) dk \quad (2.24)$$

$\tilde{f}_X(k)$ es una función continua de k , compleja con las propiedades $\tilde{f}_X(k=0) = 1$, $|\tilde{f}_X(k)| \leq 1$ y $\tilde{f}_X^*(k) = \tilde{f}_X(-k)$ (* denota complejo conjugado). Si conocemos la función característica, podemos obtener los momentos por diferenciación:

$$\langle x^n \rangle = \lim_{k \rightarrow 0} (-i)^n \frac{d^n \tilde{f}_X(k)}{dk^n}$$

Tomemos por ejemplo la distribución de Gauss Ec.(2.16). De la definición (2.22) completando cuadrados en el exponente es fácil ver que

$$\tilde{f}_X(k) = e^{ik\mu - k^2\sigma^2/2} \quad (2.25)$$

Derivando una vez obtenemos:

$$\langle x \rangle = \lim_{k \rightarrow 0} (-i) (i\mu - k\sigma^2) \tilde{f}_X(k) = \mu$$

Derivando dos veces obtenemos:

$$\langle x^2 \rangle = \lim_{k \rightarrow 0} (-i)^2 [-\sigma^2 + (i\mu - k\sigma^2)^2] \tilde{f}_X(k) = \sigma^2 + \mu^2$$

Alternativamente al desarrollo en potencias de la función característica, podemos desarrollar en potencias el *logaritmo* de la misma. esto es

$$\ln \tilde{f}_X(k) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \quad (2.26)$$

Esto se conoce como **desarrollo en cumulantes**, donde el coeficiente de la potencia n -ésima $C_n(X)$ se conoce como *cumulante* de orden n . De la definición, tenemos que

$$\tilde{f}_X(k) = \exp \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right] \quad (2.27)$$

Usando el desarrollo en serie de Taylor de la exponencial en la Ec.(2.27) y el desarrollo (2.23) e igualando potencias de k , podemos expresar los cumulantes en términos de los momentos. Así, por ejemplo:

$$\begin{aligned} C_1(X) &= \langle x \rangle \\ C_2(X) &= \langle x^2 \rangle - \langle x \rangle^2 = V(X) \\ C_3(X) &= \langle x^3 \rangle - 3 \langle x \rangle \langle x^2 \rangle + 2 \langle x \rangle^3 \end{aligned}$$

y en general puede verse que el *cumulante* de orden n es función de todos los momentos de orden $l \leq n$. Si comparamos la función característica (2.25) para la distribución de Gauss con la expresión (2.27) vemos que para la misma se anulan todos los cumulantes de orden $n > 2$. Esto significa que todos los momentos de orden superior a 2 pueden ser expresados en función de los dos primeros momentos. Esto es una particularidad de la distribución de Gauss.

Capítulo 3

Distribuciones de probabilidad multivariadas

Sobre un dado espacio muestral podemos definir diferentes variables aleatorias. Por ejemplo, en un experimento binomial, X_1 podría ser la variable binomial (número de total de éxitos) y X_2 el número de éxitos en las k primeras pruebas. Si las variables aleatorias X_1, X_2, \dots, X_n están definidas sobre el mismo espacio muestral S se dice que están conjuntamente distribuidas.

3.1. Distribución de probabilidad conjunta

La función de distribución **conjunta** para las variables aleatorias X_1, X_2, \dots, X_n se define como

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) \equiv \text{Prob} \{X_1 < x_1, \dots, X_n < x_n\} \quad (3.1)$$

donde $\{X_1 < x_1, \dots, X_n < x_n\} = \{X_1 < x_1\} \cap \{X_1 < x_2\} \cap \dots \cap \{X_n < x_n\}$. En otras palabras, $F_{X_1, \dots, X_n}(x_1, \dots, x_n)$ es la probabilidad de que las variables aleatorias X_i tomen *simultaneamente* valores en los intervalos $\{-\infty < X_i < x_i\}$ con $i = 1, \dots, n$. La **densidad de probabilidad conjunta** $f_{X_1, \dots, X_n}(x_1, \dots, x_n)$ se define entonces como

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \frac{\partial^n F_{X_1, \dots, X_n}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n} \quad (3.2)$$

de tal manera que

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} dx'_1 \dots dx'_n f_{X_1, \dots, X_n}(x'_1, \dots, x'_n) \quad (3.3)$$

Consideremos por simplicidad de aquí en más solo dos variables aleatorias X e Y , con funciones de distribución y densidad $F_{X,Y}(x, y)$ y $f_{X,Y}(x, y)$ respectivamente. La función de distribución satisface $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = F_{X,Y}(-\infty, -\infty) = 0$ y $F_{X,Y}(\infty, \infty) = 1$. Además:

$$F_{X,Y}(x_2, y) - F_{X,Y}(x_1, y) = \text{Prob} \{x_1 < X_1 < x_2; Y < y\}$$

La densidad de probabilidad $f_{X,Y}(x, y)$ satisface las condiciones $f_{X,Y}(x, y) \geq 0$ y

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_{X,Y}(x, y) = 1 \quad (3.4)$$

Dadas dos variables X e Y distribuidas conjuntamente, la función de distribución **reducida** o **marginal** $F_X(x)$ para la variable X viene dada por

$$F_X(x) = F_{X,Y}(x, \infty) = \int_{-\infty}^x dx' \int_{-\infty}^{\infty} dy f_{X,Y}(x', y) \quad (3.5)$$

Y de la misma manera, la densidad de probabilidad marginal $f_X(x)$ para la variable X se define como

$$f_X(x) = \int_{-\infty}^{\infty} dy f_{X,Y}(x, y) \quad (3.6)$$

Las funciones de distribución y marginal para la variable Y se obtienen de manera semejante.

El momento n -ésimo de la variable X se define como

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy x^n f_{X,Y}(x, y) = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

Los momentos conjuntos para las variables X e Y se definen como

$$\langle x^n y^m \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy x^n y^m f_{X,Y}(x, y) \quad (3.7)$$

El momento conjunto más comúnmente utilizado en física es la **covariancia**

$$\text{Cov}(X, Y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \quad (3.8)$$

que nos da una medida de la correlación entre fluctuaciones de ambas variables. Alternativamente, suele usarse la **función de correlación** (también llamada coeficiente de Pearson):

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (3.9)$$

la cual es adimensional. La función de correlación satisface las siguientes propiedades (fácilmente demostrables):

- $\text{Cor}(X, Y) = \text{Cor}(Y, X)$
- $-1 \leq \text{Cor}(X, Y) \leq 1$
- $\text{Cor}(X, X) = 1$, $\text{Cor}(X, -X) = -1$
- $\text{Cor}(aX + b, cY + d) = \text{Cor}(X, Y)$ si $a, c \neq 0$

Dos variables aleatorias X e Y son **independientes** si

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \quad \forall x, y$$

Si X e Y son independientes se satisfacen las siguientes propiedades

- $\langle x^m y^n \rangle = \langle x^m \rangle \langle y^n \rangle$
- $\text{Cor}(X, Y) = 0$
- $V(X + Y) = V(X) + V(Y)$

Esta última propiedad se generaliza directamente al caso de n variables X_1, \dots, X_n todas independientes entre sí:

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) \quad (3.10)$$

Al tratar con muchas variables aleatorias, a menudo necesitamos encontrar la densidad de probabilidad de una nueva variable que es función de las anteriores. Por ejemplo, supongamos que conocemos la densidad conjunta $f_{X,Y}(x, y)$ y queremos obtener la densidad de probabilidad de la

variable $Z = G(X, Y)$, donde G es una función conocida de dos variables. En una generalización de la Ec.(2.20), tenemos que

$$f_Z(Z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \delta(z - G(x, y)) f_{X,Y}(x, y) \quad (3.11)$$

esta expresión se generaliza fácilmente al caso de n variables.

Tomemos un ejemplo. Sean X e Y dos variables aleatorias independientes, con distribuciones gaussianas de media nula y variancia unitaria, esto es

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \\ f_Y(y) &= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \end{aligned}$$

Supongamos que queremos conocer la densidad conjunta de las variables

$$\begin{aligned} V &= X + Y \\ W &= X - Y \end{aligned} \quad (3.12)$$

Tenemos que

$$f_{V,W}(v, w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \delta(v - v'(x, y)) \delta(w - w'(x, y)) e^{-(x^2+y^2)/2}$$

donde $v'(x, y) = x + y$ y $w'(x, y) = x - y$. Cambiando las variables de integración a v', w' :

$$f_{V,W}(v, w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dv' \int_{-\infty}^{\infty} dw' J \begin{pmatrix} x & y \\ v' & w' \end{pmatrix} \delta(v - v') \delta(w - w') e^{-(w'^2+v'^2)/4}$$

donde hemos usado que $x^2 + y^2 = (w'^2 + v'^2)/2$ y

$$J \begin{pmatrix} x & y \\ v' & w' \end{pmatrix} = \frac{1}{2}$$

es el Jacobiano de la transformación (3.12). Así,

$$f_{V,W}(v, w) = \frac{1}{4\pi} e^{-(w^2+v^2)/4}$$

Dado que $f_{V,W}(v, w)$ es factorizable, las variables V y W son independientes.

De la Ec.(3.11) la función característica para una variable $Z = G(X, Y)$ resulta

$$\tilde{f}_Z(k) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{ikG(x,y)} f_{X,Y}(x, y) \quad (3.13)$$

3.2. Distribuciones binomiales

3.2.1. Derivación mediante la función característica

Resulta instructivo derivar nuevamente la distribución binomial a partir de los conceptos que hemos visto en esta sección. En un experimento binomial podemos describir el resultado de cada prueba individual mediante una variable aleatoria X_i , donde el índice i corresponde al número de la

prueba. La variable X_i puede tomar los valores $x = 1$ con probabilidad p y $x = 0$ con probabilidad $1 - p$. La densidad de probabilidad para esta variable es

$$f_{X_i}(x) = p \delta(x - 1) + (1 - p) \delta(x) \quad (3.14)$$

y su función característica resulta

$$f_{X_i}(k) = (1 - p) + p e^{ik} \quad (3.15)$$

Dado que todas las pruebas son independientes, la densidad de probabilidad conjunta para las variables correspondientes a las n pruebas es

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \times \dots \times f_{X_n}(x_n)$$

La variable binomial (número de éxitos en las n pruebas) puede expresarse como $Y_n = \sum_{i=1}^n X_i$. De la Ec.(3.13), la función característica para la variable Y_n resulta

$$\begin{aligned} \tilde{f}_{Y_n}(k) &= \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n e^{ik(x_1 + \dots + x_n)} f_{X_1}(x_1) \times \dots \times f_{X_n}(x_n) \\ &= \tilde{f}_{X_1}(k) \times \dots \times \tilde{f}_{X_n}(k) = (q + p e^{ik})^n \end{aligned} \quad (3.16)$$

donde $q = 1 - p$. Expandiendo el binomio tenemos

$$\tilde{f}_{Y_n}(k) = \sum_{l=0}^n \binom{n}{l} p^l q^{n-l} e^{ikl} \quad (3.17)$$

de donde la densidad de probabilidad para la variable Y_n resulta

$$f_{Y_n}(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-iky} \tilde{f}_{Y_n}(k) = \sum_{l=0}^n \binom{n}{l} p^l q^{n-l} \delta(y - l) \quad (3.18)$$

que es equivalente a la expresión (2.5).

3.2.2. Límite $n \rightarrow \infty$

En muchas aplicaciones vamos a trabajar con distribuciones binomiales con n muy grande. Queremos entonces obtener el comportamiento asintótico de la distribución binomial cuando $n \rightarrow \infty$. Tanto el valor medio como la variancia divergen en este caso. Así, resulta práctico trabajar con la variable aleatoria normalizada

$$Z_n = \frac{Y_n - \langle y_n \rangle}{\sigma_{Y_n}} = \frac{Y_n - pn}{\sqrt{pqn}} \quad (3.19)$$

La densidad de probabilidad para la variable Z_n es

$$f_{Z_n}(z) = \int_{-\infty}^{\infty} dy \delta\left(z - \frac{y - pn}{\sqrt{pqn}}\right) f_{Y_n}(y)$$

y la función característica

$$\begin{aligned} \tilde{f}_{Z_n}(k) &= \int_{-\infty}^{\infty} dz e^{ikz} f_{Z_n}(z) = \int_{-\infty}^{\infty} dy \exp\left(ik \frac{y - pn}{\sqrt{pqn}}\right) f_{Y_n}(y) \\ &= e^{-ik\sqrt{pn/q}} \tilde{f}_{Y_n}\left(\frac{k}{\sqrt{pqn}}\right) = \left(q e^{-ik\sqrt{p/qn}} + p e^{ik\sqrt{q/pn}}\right)^n \end{aligned} \quad (3.20)$$

donde hemos usado la Ec.(3.16) y que $q = 1 - p$. A continuación desarrollamos la cantidad entre paréntesis en potencias de k . Es facil ver que los dos primeros términos no nulos son $1 - k^2/2n$. Así, podemos expresar

$$\tilde{f}_{Z_n}(k) = \left(1 - \frac{k^2}{2n}(1 + R_n)\right)^n \quad (3.21)$$

donde

$$R_n = 2 \sum_{m=3}^{\infty} \frac{1}{m!} \left(\frac{ik}{\sqrt{n}}\right)^{m-2} \frac{pq^m + q(-p)^m}{(pq)^{m/2}}$$

Para $n \rightarrow \infty$ tenemos que $R_n \rightarrow 0$. Sea $Z = \lim_{n \rightarrow \infty} Z_n$. Entonces

$$\tilde{f}_Z(k) \equiv \lim_{n \rightarrow \infty} \tilde{f}_{Z_n}(k) = \lim_{n \rightarrow \infty} \left(1 - \frac{k^2}{2n}\right)^n = e^{-k^2/2} \quad (3.22)$$

Así, en el límite $n \rightarrow \infty$ obtenemos

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikz} e^{-k^2/2} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad (3.23)$$

La variable Z tiene una distribución gaussiana con media nula y variancia $\sigma_Z = 1$. Para el caso $n \gg 1$, pero finito, aún podemos aproximar $f_{Z_n}(z) \approx (1/\sqrt{2\pi}) \exp(-z^2/2)$. Transformando nuevamente a la variable $y = \sigma_Y z + \langle y \rangle$ tenemos finalmente

$$f_{Y_n}(y) \approx \int_{-\infty}^{\infty} dz \delta(y - \sigma_Y z + \langle y \rangle) f_Z(z) = \frac{1}{\sigma_Y \sqrt{2\pi}} \exp\left(-\frac{(y - \langle y \rangle)^2}{2\sigma_Y^2}\right) \quad (3.24)$$

3.2.3. Caminata aleatoria

El problema de la caminata aleatoria es un ejemplo de un problema binomial en física. Este es un modelo simplificado para el movimiento Browniano, esto es, el movimiento de una partícula bajo la acción de fuerzas aleatorias sin correlación entre sí y constituye la explicación mas básica del fenómeno de difusión. Veremos el caso mas simple de una caminata unidimensional.

Consideremos el caso de una partícula restringida a moverse a lo largo del eje x . A cada intervalo de tiempo τ la partícula tiene una probabilidad $p = 1/2$ de dar un paso de longitud Δ hacia la derecha y una probabilidad $q = 1/2$ de dar un paso equivalente hacia la izquierda. Supongamos que la partícula ha dado N pasos y que los pasos son estadísticamente independientes (la probabilidad de un paso a la derecha o a la izquierda a cada instante es independiente de todos los anteriores). El paso i -ésimo tiene entonces asociada una variable X_i que toma los valores $x = \pm\Delta$ con probabilidades $p = q = 1/2$ y por lo tanto su densidad de probabilidad es

$$f_{X_i}(x) = \frac{1}{2} (\delta(x - \Delta) + \delta(x + \Delta))$$

y la función característica resulta $\tilde{f}_{X_i}(k) = \cos(k\Delta)$.

El desplazamiento neto Y_N de la partícula al cabo de N pasos esta dado por $Y_N = \sum_{i=1}^N X_i$, suponiendo que la partícula parte del origen. De la Ec.(3.16) tenemos que la función característica para la variable Y_N es

$$\tilde{f}_{Y_N}(k) = (\cos(k\Delta))^N \quad (3.25)$$

Vamos a analizar ahora el límite cuando tanto el desplazamiento neto en un paso Δ como el tiempo entre pasos τ se vuelven infinitesimalmente pequeños. En dicho límite tanto Y_N como el

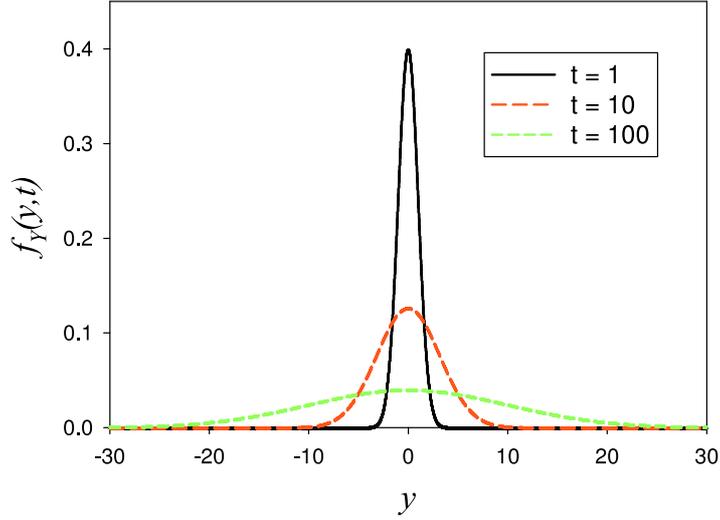


Figura 3.1: Densidad de probabilidad para la caminata aleatoria unidimensional con $D = 1/2$.

tiempo transcurrido $t = N\tau$ pueden aproximarse por variables continuas. Podemos escribir entonces $\tilde{f}_{Y_N}(k) = \tilde{f}_Y(k, t)$. Dado que la partícula parte del origen ($y = 0$), tenemos que $f_Y(y, t = 0) = \delta(y)$ y por lo tanto $\tilde{f}_Y(k, 0) = 1$. De la Ec.(3.25) podemos escribir

$$\tilde{f}_Y(k, t + \tau) - \tilde{f}_Y(k, t) = (\cos(k\Delta) - 1) \tilde{f}_Y(k, t) = \left(-\frac{k^2\Delta^2}{2} + \dots \right) \tilde{f}_Y(k, t) \quad (3.26)$$

Si tomamos el límite para $N \rightarrow \infty$ y $\tau \rightarrow 0$ manteniendo $t = N\tau$ constante, tenemos

$$\lim_{N \rightarrow \infty} \lim_{\tau \rightarrow 0} \frac{\tilde{f}_Y(k, t + \tau) - \tilde{f}_Y(k, t)}{\tau} = \frac{\partial \tilde{f}_Y(k, t)}{\partial t} \quad (3.27)$$

De la Ec.(3.26) tenemos entonces

$$\frac{\partial \tilde{f}_Y(k, t)}{\partial t} = \lim_{N \rightarrow \infty} \lim_{\tau \rightarrow 0} \lim_{\Delta \rightarrow 0} = \frac{1}{\tau} \left(-\frac{k^2\Delta^2}{2} + \dots \right) \tilde{f}_Y(k, t) \quad (3.28)$$

La única manera de que este límite sea finito (y distinto de cero) es que

$$\lim_{\tau \rightarrow 0} \lim_{\Delta \rightarrow 0} \frac{\Delta^2}{\tau} = D < \infty$$

en cuyo caso obtenemos la ecuación diferencial

$$\frac{\partial \tilde{f}_Y(k, t)}{\partial t} = -D k^2 \tilde{f}_Y(k, t), \quad (3.29)$$

cuya solución con la condición inicial $\tilde{f}_Y(k, 0) = 1$ es

$$\tilde{f}_Y(k, t) = e^{-Dk^2t} \quad (3.30)$$

La densidad de probabilidad es

$$f_Y(y, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} e^{-Dk^2t} = \sqrt{\frac{1}{4\pi Dt}} \exp\left(-\frac{y^2}{4Dt}\right) \quad (3.31)$$

Así, la probabilidad de encontrar al caminante en una posición entre y e $y + dy$ es gaussiana con media nula y desviación estándar $\sigma = \sqrt{2Dt}$. La densidad de probabilidad se “desparrama” en el tiempo manteniendo el área unitaria, como se muestra en la Fig.3.1. Notemos que $\sqrt{\langle y^2 \rangle} = \sigma \propto \sqrt{t}$. Este es el comportamiento característico de un movimiento difusivo: el desplazamiento de una partícula browniana de su posición inicial crece con \sqrt{t} . La constante D se conoce como **coeficiente de difusión**. De hecho, notemos que de la definición

$$f_Y(y, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} \tilde{f}_Y(k, t)$$

tenemos que

$$\frac{\partial f_Y(y, t)}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iky} \frac{\partial f_Y(k, t)}{\partial t}$$

y

$$\frac{\partial^2 f_Y(y, t)}{\partial y^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-k^2) e^{-iky} \tilde{f}_Y(k, t)$$

Así, antitransformando la Ec.(3.29) obtenemos

$$\frac{\partial f_Y(k, t)}{\partial t} = D \frac{\partial^2 f_Y(y, t)}{\partial y^2} \quad (3.32)$$

Esta última ecuación diferencial se conoce como **ecuación de difusión** (tiene la misma estructura que la ecuación de difusión de calor) y su solución viene dada por la Ec.(3.31).

3.3. El Teorema del Límite Central

Consideremos un conjunto de N variables aleatorias X_i , todas independientes entre sí e idénticamente distribuidas, esto es, la densidad de probabilidad $f_X(x)$ es la misma para todas. Un ejemplo de esto podría ser la repetición independiente de un mismo experimento N veces. Sean $\mu = \langle x \rangle$ y $\sigma^2 = V(X)$ el valor medio y la variancia correspondientes a la densidad $f_X(x)$.

Consideremos ahora la variable Y_N , definida como la desviación del **promedio aritmético** de las N variables respecto de μ , esto es

$$Y_N \equiv \frac{1}{N} \sum_{i=1}^N X_i - \mu = \sum_{i=1}^N Z_i \quad (3.33)$$

donde

$$Z_i \equiv \frac{1}{N}(x_i - \mu) \quad (3.34)$$

donde $\langle Z_i \rangle = 0$ y $V(Z_i) = \sigma^2/N^2$. Resulta inmediato que

$$\left\langle \sum_{i=1}^N X_i \right\rangle = N\mu$$

con lo cual

$$\langle Y_N \rangle = 0 \quad (3.35)$$

Pensemos en que queremos medir una cierta cantidad, la cual fluctua de experimento en experimento (X_i). Habitualmente queremos estimar el valor medio de esta variable (μ), para lo cual repetimos

el experimento N veces y aproximamos μ por el promedio aritmético de las N mediciones. Sin embargo, ese promedio es a su vez una variable aleatoria (si repetimos la serie de N experimentos, el nuevo promedio también fluctúa). La Ec.(3.35) nos dice que el valor medio de este promedio coincide con μ , con lo que es en principio un estimador aceptable. Sin embargo, querríamos evaluar de alguna manera el error cometido en esta estimación. Para eso tenemos que analizar las fluctuaciones del promedio aritmético respecto de su valor medio, las cuales están descritas por la variable Y_N . Si recordamos la ecuación (3.10), vemos que, bajo el supuesto de que las mediciones son completamente independientes entre sí, tenemos que

$$V(Y_N) = \sum_{i=1}^N V(Z_i) = \frac{\sigma^2}{N} \quad (3.36)$$

con lo cual $\sigma_{Y_N} = \sigma/\sqrt{N}$. Podemos concluir a partir de este resultado que la dispersión del promedio disminuye con N ? Estrictamente no, ya que podría ocurrir que momentos de orden superior al segundo se hagan muy grandes. Para dar una respuesta a la pregunta anterior debemos obtener la distribución de probabilidad para la variable Y_N . Para ello calculamos primero la función característica para las variables Z_i . Tenemos que

$$\tilde{f}_Z(k) = \int_{-\infty}^{\infty} dx e^{i(k/N)(x-\mu)} f_X(x) = 1 - \frac{1}{2} \frac{k^2}{N^2} \sigma^2 + \dots \quad (3.37)$$

donde hemos desarrollado en serie de Taylor la exponencial e integrado término a término. Si este desarrollo es válido, es decir, si los momentos de $f_X(x)$ son **finitos**, los términos superiores pueden despreciarse para N suficientemente grande (notemos que el carácter oscilatorio del integrando asegura que la integral se va a cero para valores grandes de k). La función característica para Y_N es por lo tanto

$$\tilde{f}_{Y_N}(k) = \left(1 - \frac{1}{2} \frac{k^2}{N^2} \sigma^2 + \dots\right)^N \rightarrow \exp\left(-\frac{k^2 \sigma^2}{2N}\right) \quad \text{para } N \rightarrow \infty \quad (3.38)$$

y la densidad de la variable Y_N :

$$f_{Y_N}(x) \rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{iky} \exp\left(-\frac{k^2 \sigma^2}{2N}\right) = \sqrt{\frac{N}{2\pi\sigma^2}} \exp\left(-\frac{Ny^2}{2\sigma^2}\right) \quad (3.39)$$

Independientemente de la forma particular de $f_X(x)$, si sus momentos son finitos, el promedio de un número grande de mediciones de X tiene una distribución gaussiana centrada en $\langle x \rangle$ y con desvío estándar igual al desvío de X dividido por \sqrt{N} . Este es el Teorema del Límite Central.

Capítulo 4

Fundamentos de la Mecánica Estadística

4.1. Introducción: Relación entre la descripciones microscópica y macroscópica de los fenómenos físicos.

La termodinámica constituye un formalismo poderoso de gran generalidad, construido sobre la base de unas pocas y simples hipótesis. El concepto fundamental introducido en dichas hipótesis es la *entropía*, la cual entra en la formulación de manera abstracta, a través de un principio variacional que determina los estados de equilibrio. En el formalismo resultante, no obstante, la entropía es uno más entre un conjunto de parámetros extensivos, junto con la energía, los números molares, el volumen, la magnetización, etc.. Mientras que las últimas cantidades poseen una interpretación física clara y fundamental, la entropía, desde la termodinámica, adolece de la misma.

Por otra parte, la termodinámica nos brinda una descripción macroscópica (esto es, en términos de unos pocos parámetros globales que adoptan valores a escalas humanas) del comportamiento de sistemas compuestos por un número gigantesco N de partículas interactuantes¹ (del orden de un número de Avogadro $\sim 10^{23}$). Dichas partículas (átomos, moléculas, etc.) obedecen las leyes de la mecánica (leyes de Newton, en una descripción clásica o la ecuación de Schrödinger, en una descripción cuántica). Una descripción microscópica, por lo tanto, debe basarse en la dinámica de un sistema de muchos cuerpos, los cuales pueden ser idealizados como partículas puntuales o como pequeños cuerpos con unos pocos grados internos de libertad.

La Mecánica Estadística establece la conexión entre estos dos niveles de descripción. La cantidad fundamental para establecer dicha conexión es nuevamente la entropía. Veremos como, a partir de unos pocos postulados fundamentales sencillos, esta vez a nivel microscópico, es posible derivar el principio variacional sobre el cual se fundamenta la termodinámica. Esto nos permitirá una interpretación física clara para la entropía, a la vez de una receta para calcularla y con ello derivar relaciones fundamentales a partir de primeros principios.

Al tratar un un número muy grande de partículas, lo primero a destacar es la necesidad de una descripción probabilística del estado de los sistemas.

Pensemos en un sistema aislado. Por simplicidad, podemos considerar un fluido encerrado en un recipiente adiabático con paredes perfectamente rígidas y reflectantes. Los vínculos del sistema son entonces el volumen V , la energía U (dado que el sistema está aislado, podemos identificar la energía interna con la energía mecánica del sistema) y el número de partículas N . El estado microscópico del sistema deberá ser compatible con estos vínculos. Evidentemente, existen muchísimos estados mecánicos microscópicos compatibles con los valores especificados de U , V y N . El conjun-

¹Durante el curso de termodinámica llamamos N al número de moles; de aquí en más N designará el número de partículas.

to de los estados microscópicos compatible con dichos valores constituye un estado *macroscópico* ó *macroestado*. Para un dado macroestado el sistema puede estar en cualquiera de dichos estados microscópicos permitidos.

Resulta evidente que es imposible determinar con precisión el estado microscópico para cualquier instante de tiempo. Para ello deberíamos ser capaces de resolver las ecuaciones de movimiento para las N partículas. Aún en el caso en que fuéramos capaces (hecho evidentemente falso) para ello necesitamos conocer con exactitud las condiciones iniciales del sistema, ya sea de las posiciones y velocidades iniciales de todas las partículas, en una descripción clásica, o la función de onda de las N partículas a un tiempo dado, en una descripción cuántica.

Desde un punto de vista macroscópico, sabemos que la especificación de una pocas variables globales es suficiente para determinar todas las propiedades de un sistema en equilibrio. En otras palabras, fijar las condiciones iniciales macroscópicas permite especificar completamente el problema. Sin embargo, si pensamos los sistemas como una colección de partículas, es evidente que la especificación de las condiciones iniciales macroscópicas no determina de manera unívoca las condiciones mecánicas iniciales. Existe un número enorme de microestados compatibles con casi cualquier condición macroscópica que podamos fijar. Si repetimos un mismo experimento muchas veces, la probabilidad de preparar el sistema en las mismas condiciones iniciales microscópicas es infinitesimal. No obstante, la experiencia demuestra que los fenómenos observados a escala macroscópica son insensibles a esas diferencias. Todas las condiciones iniciales compatibles con los vínculos macroscópicos son, en cierto sentido, equivalentes y deben ser tratadas en pie de igualdad. Una manera de expresar esto matemáticamente es asignar un *peso* o *probabilidad* a todos los posibles estados del sistema al tiempo cero. Por ejemplo, podemos asignar peso cero a todos los estados incompatibles con los vínculos macroscópicos y un peso diferente de cero, pero igual para todos ellos, a los estados que son compatibles. Podemos entonces definir las magnitudes macroscópicas (por ejemplo, la temperatura) como un promedio pesado sobre todos dichos estados de alguna cantidad dependiente de las variables microscópicas (en el ejemplo de la temperatura, sería el promedio de la energía cinética por partícula). Mediante este procedimiento es claro que, para cualquier condición inicial, el resultado para cualquier magnitud macroscópica será siempre el mismo: tenemos una explicación para la reproducibilidad de experimentos macroscópicos.

Podemos cuestionarnos acerca del poder predictivo de una teoría de este tipo. Esta teoría nos da una predicción acerca del resultado promedio de un número grande de experimentos realizados bajo condiciones idénticas, pero no nos brinda una predicción precisa del resultado del mismo. Es claro que habrá fluctuaciones. No obstante, la teoría puede predecir las propiedades estadísticas de dichas fluctuaciones, tal como por ejemplo el desvío medio cuadrático del promedio.

Llegamos así a la idea de la descripción estadística de un sistema de muchos cuerpos. El objeto representativo del sistema no es mas un punto en el espacio de las fases (o un vector de onda en un espacio de Hilbert), sino una colección de puntos en el espacio de las fases, cada uno con un peso estadístico o probabilidad. En otras palabras, el estado microscópico del sistema está descrito por una variable aleatoria (vectorial). Los diferentes puntos en el espacio de las fases en los cuales el sistema puede encontrarse constituyen los valores posibles de dicha variable aleatoria.

El programa de la Mecánica Estadística se basa en determinar cual es la distribución de probabilidad correcta de esta variable. Dicha distribución de probabilidad debe depender de las propiedades mecánicas microscópicas y tiene que ser independiente del tiempo para un sistema en equilibrio.

4.2. La densidad de probabilidad clásica y el concepto de ensemble.

Comenzaremos considerando un sistema clásico, esto es, un sistema de N partículas encerradas en una caja de volumen V , cuya dinámica esta gobernada por las ecuaciones de Hamilton. Un estado

microscópico del sistema en este caso está determinado por las posiciones y los momentos de todas las partículas en un instante dado, las cuales determinan de manera unívoca el estado del sistema en cualquier otro instante. Así, para un sistema tridimensional, tendremos $6N$ variables reales independientes (3 coordenadas y 3 componentes del momento lineal por cada partícula). Vamos a denotar genéricamente dicho estado por (p, q) , donde p y q son dos vectores de $3N$ dimensiones, cuyas componentes son las de los momentos y las coordenadas respectivamente. Dado que estamos considerando $N \sim 10^{23}$ partículas y $V \sim 10^{23}$ volúmenes moleculares, resultará conveniente considerar el sistema en el límite

$$N \rightarrow \infty \quad V \rightarrow \infty \quad \text{con} \quad V/N = v = \text{cte}$$

Esto se conoce como el **límite termodinámico**.

Sea $H = H(p, q)$ el Hamiltoniano del sistema. La evolución de las variables (p, q) está dada por las ecuaciones de Hamilton

$$\begin{aligned} \frac{dq_i}{dt} &= \frac{\partial H(p, q)}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H(p, q)}{\partial q_i} \end{aligned} \quad (4.1)$$

Si el sistema está aislado, H no depende explícitamente del tiempo y el sistema es conservativo, esto es, H es una constante de movimiento

$$H(p, q) = E \quad (4.2)$$

donde E es la energía total del sistema. En este contexto, las paredes del recipiente deben ser idealizadas como paredes reflectantes perfectas.

La Ec.(4.2) nos define una hipersuperficie en el espacio de las fases, la cual llamaremos genéricamente la *superficie energía*. Todos los puntos de esta superficie constituyen los microestados accesibles del sistema compatibles con el macroestado definido por la energía E (la superficie también depende del volumen V y obviamente de N). A medida que el sistema evoluciona, el punto (p, q) describirá una trayectoria sobre la superficie energía. Dado que las trayectorias en mecánica clásica están determinadas unívocamente por las condiciones iniciales, se sigue que las trayectorias en el espacio de las fases, o bien son cerradas (cíclicas) o nunca se cruzan.

Tenemos entonces un número infinito de posibles microestados, los cuales vamos a pensar como los posibles resultados de una variable aleatoria (p, q) . Sea entonces $\rho(p, q, t)$ la densidad de probabilidad de dicha variable aleatoria, esto es, $\rho(p, q, t) dp dq$ nos da la probabilidad de encontrar el sistema en un elemento de volumen $dp dq$ centrado en (p, q) al tiempo t . Podemos entonces pensar en un conjunto denso de puntos sobre la superficie energía, cuya distribución está dada por $\rho(p, q, t) dp dq$ (fracción de puntos en un volumen elemental centrado en (p, q)). Cada uno de dichos puntos representa una copia del mismo sistema en un microestado diferente. A esta imagen mental de un conjunto de copias del mismo sistema en todos los posibles diferentes estados se la denomina **ensemble estadístico**. Este concepto fue introducido por W. Gibbs.

La densidad de probabilidad debe cumplir con ciertas restricciones. En primer lugar, la misma tiene que estar normalizada para todo tiempo

$$\int \rho(p, q, t) dp dq = 1 \quad (4.3)$$

donde la integral se extiende a todo el espacio de las fases. La probabilidad $P(R, t)$ de encontrar el sistema en un volumen finito R del espacio de las fases al tiempo t está entonces dada por

$$P(R, t) = \int_R \rho(p, q, t) dp dq \quad (4.4)$$

Consideremos un volúmen arbitrario V_0 centrado en un punto fijo sobre la superficie energía. A medida que transcurre el tiempo, la probabilidad de encontrar el sistema en dicho volúmen puede cambiar. Pero dado que la probabilidad total es constante, el cambio en dicha probabilidad por unidad de tiempo debe ser igual al flujo de probabilidad a través de la superficie S que encierra V_0 . En términos de ensemble, podemos pensar que el mismo se comporta como un fluido en el espacio de las fases. A medida que el sistema evoluciona, cada punto se mueve a lo largo de una trayectoria. Dado que esas trayectorias no se cruzan, podemos pensarlas como líneas de flujo. Y dado que el número total de puntos representativos en el ensemble es constante, el número de trayectorias también lo es.

Sea $\vec{v} \equiv (\dot{p}, \dot{q})$ la velocidad de un punto en el espacio de las fases y sea $\vec{J} \equiv \vec{v} \rho$ una corriente de probabilidad. El flujo a través de un elemento de superficie viene entonces dado por $\vec{n} \cdot \vec{J} dS$, donde \vec{n} es un vector normal a la superficie y dS es el elemento de área. Tenemos entonces la ley de conservación:

$$\frac{dP(V_0)}{dt} = \frac{\partial}{\partial t} \int_{V_0} \rho(p, q, t) dp dq = - \oint_S \vec{n} \cdot \vec{J} dS \quad (4.5)$$

Usando el Teorema de Gauss, podemos cambiar la integral de superficie a una integral de volúmen obteniendo

$$\int_{V_0} \frac{\partial}{\partial t} \rho(p, q, t) dp dq = - \int_{V_0} \nabla \cdot (\vec{v} \rho(p, q, t)) dp dq \quad (4.6)$$

Dado que V_0 es arbitrario, los integrandos de ambos lados de la ecuación anterior deben ser iguales y así obtenemos la ecuación de balance para la probabilidad:

$$\frac{\partial \rho(p, q, t)}{\partial t} + \nabla \cdot (\vec{v} \rho(p, q, t)) = 0 \quad (4.7)$$

Ahora bien, de las ecuaciones de movimiento (4.1) tenemos que

$$\nabla \cdot \vec{v} = \sum_{i=1}^{3N} \left(\frac{\partial \dot{p}_i}{\partial p_i} + \frac{\partial \dot{q}_i}{\partial q_i} \right) = 0$$

Reemplazando en la Ec.(4.7) tenemos

$$\frac{\partial \rho(p, q, t)}{\partial t} = -\vec{v} \cdot \nabla \rho(p, q, t) \quad (4.8)$$

Esta ecuación se conoce como **Teorema de Liouville**. En particular, nos dice que la *derivada total ó convectiva* de $d\rho/dt = 0$ para cualquier punto en cualquier instante. La interpretación geométrica es la siguiente. Si seguimos el movimiento de un punto cualquiera, veremos que la densidad en su entorno permanece constante en el tiempo. En otras palabras, los puntos del ensemble se mueven como un *fluido incompresible*. Por otra parte, usando nuevamente las ecuaciones de movimiento, la Ec.(4.8) puede ser re-escrita como

$$\frac{\partial \rho(p, q, t)}{\partial t} = -\{\rho, H\} \quad (4.9)$$

donde

$$\{\rho, H\} = \frac{\partial \rho}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial \rho}{\partial p} \frac{\partial H}{\partial q}$$

es el paréntesis de Poisson de ρ con H . En el caso de una distribución estacionaria, esto es, una distribución donde la densidad no depende explícitamente del tiempo, tenemos entonces la condición

$$\{\rho, H\} = 0 \quad (4.10)$$

La densidad de probabilidad que estamos buscando para describir los estados termodinámicos tiene que satisfacer esta propiedad. Una forma bastante general de que esta condición se satisfaga es que ρ dependa de p y q a través de una dependencia explícita en H , esto es, $\rho(p, q) = \rho[H(p, q)]$. La dependencia más sencilla de este tipo (luego veremos otras) es $\rho = cte$ en la superficie energía y cero fuera de ella.

Estamos ahora en condiciones de enunciar el postulado fundamental de la Mecánica Estadística, pero antes debemos hacer una consideración. La energía no puede ser estrictamente constante, ya que en rigor un sistema totalmente aislado es una idealización (existen siempre campos externos aleatorios débiles de largo alcance, tales como gravitacional, electromagnético, etc, que se acoplan con el sistema). Vamos a permitir entonces cierta incerteza en la energía, esto es, considerar que los estados del ensemble tienen una energía entre E y $E + \Delta$ (pensemos en una cáscara), donde $\Delta \ll E$. Esto es más bien una conveniencia conceptual que simplifica los cálculos y, como veremos en el próximo capítulo, no afecta el resultado final (es independiente de Δ).

Postulado de igual probabilidad a priori: *En un sistema aislado en equilibrio todos los microestados accesibles son igualmente probables.* Esto es

$$\rho(p, q) = \begin{cases} \frac{1}{\Gamma(E)} & \text{para } E \leq H(p, q) \leq E + \Delta \\ 0 & \text{en caso contrario} \end{cases} \quad (4.11)$$

donde

$$\Gamma(E) = \int_{E \leq H(p, q) \leq E + \Delta} dp dq \quad (4.12)$$

es el volumen ocupado por el ensemble en el espacio de las fases. Notemos que ρ satisface la normalización (4.3). El ensemble definido por la densidad (4.11) se conoce como **ensemble microcanónico** y lo estudiaremos en detalle en el siguiente capítulo.

4.3. La Hipótesis Ergódica

El postulado de igual probabilidad a priori puede justificarse a través de lo que se conoce como **hipótesis ergódica**, la cual se refiere a las propiedades dinámicas de los sistemas bajo consideración. Sin embargo, es importante notar que el postulado debe ser considerado como un axioma de la teoría y como tal su validez última se basa en la contrastación de las predicciones que del mismo se derivan con los resultados experimentales. En última instancia, al asumir la hipótesis ergódica estamos reemplazando un axioma por otro de nivel más fundamental.

Un sistema Hamiltoniano define un flujo en el espacio de las fases $(p(t), q(t))$. Sea una función arbitraria $f(p, q)$. Se define el promedio *temporal* de la función f como

$$\langle f \rangle_T = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(p(t), q(t)) dt \quad (4.13)$$

esto es, el promedio a lo largo de una trayectoria. Por otra parte, el promedio en ensemble de la función f viene dado por

$$\langle f \rangle = \int f(p, q) \rho(p, q) dp dq = \frac{1}{\Gamma(E)} \int_{E \leq H(p, q) \leq E + \Delta} f(p, q) dp dq \quad (4.14)$$

Un sistema es **ergódico** si para casi toda trayectoria $(p(t), q(t))$ (excepto un conjunto de medida nula) se cumple que

$$\langle f \rangle = \langle f \rangle_T \quad (4.15)$$

La **hipótesis ergódica**, introducida por Ludwig Boltzmann, consiste en asumir que para $N \gg 1$ los sistemas son ergódicos. La interpretación es la siguiente. Debido a la inercia de los aparatos de medición, lo que realmente se mide en el laboratorio son promedios temporales de magnitudes físicas, tomados sobre períodos de tiempo grandes comparados con los tiempos característicos de evolución de los sistemas. Si el sistema es ergódico, la inmensa mayoría de las trayectorias compatibles con los vínculos macroscópicos barren de manera casi uniforme toda la superficie energía. Esto último puede verse de la siguiente manera. Tomemos un volúmen arbitrario pequeño R sobre la superficie energía y sea τ_R el tiempo que el sistema pasa en dicho volúmen. Podemos estimar τ_R introduciendo la función

$$\phi_R(p, q) = \begin{cases} 1 & \text{si } (p, q) \in R \\ 0 & \text{caso contrario} \end{cases}.$$

De esta manera

$$\tau_R = \int_{t_0}^{t_0+T} \phi(p(t), q(t)) dt$$

y si el sistema es ergódico

$$\lim_{T \rightarrow \infty} \frac{\tau_R}{T} = \langle \phi \rangle_T = \frac{1}{\Gamma(E)} \int_{E \leq H(p, q) \leq E+\Delta} \phi(p, q) dp dq = \frac{\Gamma(R)}{\Gamma(E)}$$

donde $\Gamma(R)$ es el volúmen ocupado por R en la superficie energía. Esto es, si el sistema es ergódico, la fracción de tiempo que el sistema pasa en una región cualquiera depende solo de su volúmen y no de su posición en la superficie energía. Así, el sistema recorre de manera uniforme toda la superficie y tomar un promedio temporal es equivalente a tomar un promedio en el ensemble microcanónico. Establecer si un sistema es ergódico o nó es un problema altamente complejo (la teoría ergódica es una rama de la matemática que excede el presente curso). Hasta el presente, la hipótesis ergódica solo ha podido ser verificada de manera rigurosa para unos pocos sistemas simples. Sin embargo, la computadoras actuales permiten resolver numéricamente las ecuaciones de movimiento para un número bastante grande de partículas con potenciales de interacción arbitrarios. Las mismas en general verifican la hipótesis ergódica.

4.4. El Operador Densidad en Mecánica Cuántica.

Vamos a extender ahora las ideas desarrolladas en la sección anterior al tratamiento de sistemas cuánticos. A diferencia del caso clásico, la Mecánica Cuántica tiene un carácter estadístico intrínseco debido al principio de incertidumbre: aún cuando poseemos la información máxima posible del sistema, solo podemos hacer predicciones de carácter estadístico.

En Mecánica Cuántica un microestado de un sistema es caracterizado por una función de onda $\Psi(q)$ (donde q denota al igual que antes las posiciones de las N partículas), la cual es solución de la ecuación de Schrödinger para el sistema de N cuerpos. Esta función de onda puede ser desarrollada en términos de un conjunto ortonormal completo $\phi_n(q)$ de autofunciones de algún operador (por ejemplo, la energía):

$$\Psi(q) = \sum_n c_n \phi_n(q) \quad (4.16)$$

donde $|c_n|^2$ nos da la probabilidad de encontrar el sistema en el autoestado ϕ_n . De la normalización de la función de onda tenemos que $\sum_n |c_n|^2 = 1$. El valor promedio (o valor esperado) cuántico de un operador cualquiera \hat{O} (observable) en el microestado $\Psi(q)$ viene dado por

$$\langle \Psi | \hat{O} | \Psi \rangle = \int \Psi^*(q) \hat{O} \Psi(q) dq = \sum_{m,n} c_m^* c_n O_{mn} \quad (4.17)$$

donde $O_{mn} = \langle \phi_m | \hat{O} | \phi_n \rangle$ son los elementos de matriz del operador \hat{O} en la base considerada.

Todos los resultados anteriores son válidos si conocemos exactamente el microestado $\Psi(q)$, esto es, si tenemos un estado puro. No obstante, las mismas consideraciones del caso clásico se aplican aquí. Dado un conjunto de vínculos macroscópicos, tendremos una cantidad enorme de microestados $\Psi^{(i)}(q)$ compatibles con los mismos y resulta imposible predecir con exactitud en cual de ellos se encuentra el sistema. Lo máximo que podemos aspirar es a asignarle al sistema una probabilidad p_i de estar en el estado $\Psi^{(i)}(q)$. Esta es la generalización del concepto de ensemble al caso cuántico.

El paso siguiente es establecer como se calcula el promedio de un observable cuando el sistema es especificado en esta forma estadística. Para ello desarrollamos cada uno de los estados accesibles en la base ortonormal $\phi_n(q)$:

$$\Psi^{(i)}(q) = \sum_n c_n^{(i)} \phi_n(q). \quad (4.18)$$

El promedio *cuántico* del operador \hat{O} en uno de los estados $\Psi^{(i)}(q)$ está dado por

$$\langle \Psi^{(i)} | \hat{O} | \Psi^{(i)} \rangle = \sum_{m,n} c_m^{(i)*} c_n^{(i)} O_{mn} \quad (4.19)$$

Esta cantidad tiene una probabilidad p_i de ocurrencia. Para obtener el **promedio en ensemble** debemos promediar a su vez sobre los microestados del mismo:

$$\langle \hat{O} \rangle = \sum_i p_i \langle \Psi^{(i)} | \hat{O} | \Psi^{(i)} \rangle = \sum_i p_i \sum_{m,n} c_m^{(i)*} c_n^{(i)} O_{mn} \quad (4.20)$$

donde se asume que $p_i \geq 0$ y $\sum_i p_i = 1$. Introduciendo ahora una matriz ρ cuyos elementos de matriz son

$$\rho_{nm} = \sum_i p_i c_n^{(i)} c_m^{(i)*} \quad (4.21)$$

podemos expresar

$$\langle \hat{O} \rangle = \sum_{m,n} \rho_{nm} O_{mn} \quad (4.22)$$

Denotemos ahora $\hat{\rho}$ el operador cuyos elementos de matriz en la base $\phi_n(q)$ son ρ_{mn} :

$$\rho_{mn} = \int dq \phi_m(q)^* \hat{\rho} \phi_n(q) \equiv \langle \phi_m | \hat{\rho} | \phi_n \rangle; \quad (4.23)$$

$\hat{\rho}$ se conoce como el **operador densidad de probabilidad**. Es facil probar que el mismo es hermitiano. De la Ec.(4.22) tenemos que

$$\langle \hat{O} \rangle = \sum_n (\hat{\rho} \hat{O})_{nn} = \sum_n \langle \phi_n | \hat{\rho} \hat{O} | \phi_n \rangle \quad (4.24)$$

Así, el promedio en ensemble de un observable resulta igual a la suma de los elementos diagonales ó *traza* del operador $\hat{\rho} \hat{O}$. Pero la traza de una matriz u operador es una propiedad independiente de la base escogida, de manera que podemos expresar

$$\langle \hat{O} \rangle = \text{Tr} \hat{\rho} \hat{O} = \text{Tr} \hat{O} \hat{\rho} \quad (4.25)$$

donde hemos usado la propiedad de invariancia de la traza de un producto de matrices (operadores) ante permutaciones cíclicas de los factores. Notemos que

$$\text{Tr } \hat{\rho} = \sum_i p_i \sum_n |c_n|^2 = \sum_i p_i = 1$$

De la Ec.(4.21) podemos escribir el operador $\hat{\rho}$ como un producto matricial

$$\hat{\rho} = \sum_i p_i \Psi^{(i)} \Psi^{(i)\dagger} \quad (4.26)$$

donde

$$\Psi^{(i)} = \begin{pmatrix} c_1 \\ \vdots \\ c_l \\ \vdots \end{pmatrix} \quad \Psi^{(i)\dagger} = (c_1^*, \dots, c_l^*, \dots)$$

Los microestados $\Psi^{(i)}$ son soluciones de la ecuación de Schrödinger

$$i\hbar \frac{\partial \Psi^{(i)}}{\partial t} = \hat{H} \Psi^{(i)} \quad (4.27)$$

donde \hat{H} es el operador Hamiltoniano. Trasponiendo la ecuación anterior y tomando el complejo conjugado

$$-i\hbar \frac{\partial \Psi^{(i)\dagger}}{\partial t} = \Psi^{(i)\dagger} \hat{H} \quad (4.28)$$

donde hemos usado que el Hamiltoniano es hermitiano $\hat{H} = \hat{H}^\dagger$. Derivando la Ec.(4.26) obtenemos

$$\begin{aligned} i\hbar \frac{\partial \hat{\rho}}{\partial t} &= i\hbar \sum_i p_i \left(\Psi^{(i)} \frac{\partial \Psi^{(i)\dagger}}{\partial t} + \frac{\partial \Psi^{(i)}}{\partial t} \Psi^{(i)\dagger} \right) \\ &= - \sum_i p_i \left(\Psi^{(i)} \Psi^{(i)\dagger} \hat{H} - \hat{H} \Psi^{(i)} \Psi^{(i)\dagger} \right) \\ &= -(\hat{\rho} \hat{H} - \hat{H} \hat{\rho}) \end{aligned}$$

esto es

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = -[\hat{\rho}, \hat{H}] \quad (4.29)$$

donde $[\cdot, \cdot]$ denota el conmutador. Esta ecuación es el equivalente del Teorema de Liouville en Mecánica Cuántica. Si el sistema se encuentra en equilibrio tenemos que

$$[\hat{\rho}, \hat{H}] = 0$$

lo cual implica que siempre podemos encontrar una base de autoestados de \hat{H} en la cual $\hat{\rho}$ es también diagonal. Sea entonces ϕ_n una base de autoestados del operador Hamiltoniano \hat{H} :

$$\hat{H} \phi_n = E_n \phi_n \quad (4.30)$$

donde E_n es el autovalor del Hamiltoniano correspondiente al autoestado ϕ_n . Si $\hat{\rho}$ es diagonal en esta base, de la Ec.(4.22) tenemos

$$\langle \hat{O} \rangle = \sum_n \rho_{nn} O_{nn} \quad (4.31)$$

donde $\rho_{nn} = \sum_i p_i |c_n^{(i)}|^2 \geq 0$ y $\sum_n \rho_{nn} = 1$. ρ_{nn} puede interpretarse como la probabilidad de encontrar el sistema en el autoestado ϕ_n .

Podemos enunciar ahora el postulado de igual probabilidad a priori en su versión cuántica. Consideremos un sistema aislado con energía E (o energía entre E y $E + \Delta$, según convenga). Sea ϕ_l con $l = 1, 2, \dots, M(E)$ el subconjunto de autoestados correspondientes al autovalor E . Entonces

$$\rho_{ll} = \frac{1}{M(E)} \text{ para } l = 1, 2, \dots, M(E) \quad (4.32)$$

Capítulo 5

El ensemble microcanónico

El postulado de igual probabilidad a priori nos dice que, en un sistema aislado con energía E , la probabilidad de encontrar el sistema en un microestado es la misma para todos los microestados compatibles con los vínculos macroscópicos e igual a la inversa del número total de microestados (en el caso clásico, el volumen de la hipersuperficie de energía constante es una medida del número de microestados). Así, dado el Hamiltoniano del sistema (clásico o cuántico) podemos calcular la distribución de probabilidad (ya veremos varios ejemplos en breve). Una vez que tenemos la distribución de probabilidad podemos calcular promedios de magnitudes tales como la energía, magnetización, etc, ya que las mismas están directamente asociadas a propiedades mecánicas. Sin embargo, hay una cantidad que hasta aquí el formalismo no nos dice como calcularla: la entropía. Para establecer las bases de la termodinámica debemos comenzar entonces por postular una forma microscópica para la entropía.

5.1. La entropía de Boltzmann-Gibbs

Si asumimos que toda la información está contenida en la distribución de probabilidad, la entropía debería ser calculable a partir de la misma. Dado que la distribución de probabilidad es una función del número de microestados la entropía también lo será.

Llamaremos genéricamente $W(E)$ al número de microestados con energía E . En el caso cuántico $W(E)$ es directamente el número de autoestados $M(E)$ con energía E . En el caso clásico estamos tentados a identificar $W(E)$ con $\Gamma(E)$, pero debemos ser cuidadosos. En primer lugar, W tiene que ser adimensional y $\Gamma(E)$ tiene dimensiones de $(\text{distancia} \times \text{momento})^{3N}$ (o $(\text{momento angular})^{3N}$). Por otra parte, sabemos que una descripción clásica debe resultar consistente con la descripción cuántica en algún límite apropiado (como veremos, esto corresponde a los límites de baja densidad y/o altas temperaturas). Para ello debemos incorporar la restricción impuesta por el principio de incertidumbre, por el cual la máxima precisión que podemos tener en una determinación simultánea de cada componente de posición y momento es $\Delta p \Delta q \approx h$. Así, resulta razonable suponer que el espacio de las fases está dividido en celdas de volumen h^{3N} , dentro de las cuales el punto no puede ser definido con precisión (h tiene dimensiones de momento angular). Estas celdas pueden ponerse en correspondencia uno a uno con los estados cuánticos del sistema y por lo tanto el número de microestados correspondería al número de dichas celdas $W(E) = \Gamma(E)/h^{3N}$. Volvamos entonces a la relación entre la entropía y $W(E)$. La definición microscópica de entropía que adoptemos tiene que satisfacer dos condiciones: a) S tiene que ser extensiva y b) Tiene que cumplir el principio variacional enunciado en el segundo postulado de la termodinámica.

5.1.1. La formulación de Boltzmann

Consideremos un sistema en el cual un vínculo es removido, por ejemplo, abriendo una válvula que permite al sistema expandirse a un volumen mayor, hasta alcanzar el equilibrio. Desde el punto de vista macroscópico esto implica un aumento en la entropía. Desde el punto de vista microscópico, aumentar el volumen accesible al sistema implica un aumento del número de microestados accesibles. Esto sugiere la identificación de la entropía con W . Sin embargo, esto presenta la dificultad de que W es una cantidad *multiplicativa*, mientras que la entropía es una cantidad *aditiva* (el número de estados accesibles de dos sistemas es el producto del número de estados de cada subsistema, al igual que el número de estados de dos dados $6 \times 6 = 36$). Necesitamos entonces una función $S = f(W)$, con la propiedad $f(W_1 \times W_2) = f(W_1) + f(W_2)$ para todo par de valores W_1, W_2 . Es posible demostrar que la *única* función continua que satisface esta propiedad es $f(x) = k \ln x$.

Llegamos así al segundo postulado de la Mecánica Estadística mediante la definición

$$S(E) = k_B \ln W(E) \quad (5.1)$$

Esta forma de la entropía fué introducida por Ludwig Boltzmann y se la suele llamar **entropía de Boltzmann**; la misma se encuentra inscrita en el túmulo de L. Boltzmann en Viena. El prefactor en la Ec.(5.1) es adimensional (recordemos que la entropía lo es) y se elige de manera de obtener un acuerdo con la escala Kelvin de temperatura $T^{-1} = \partial S / \partial E$. Veremos que tal acuerdo se obtiene eligiendo que esta sea la constante de Boltzmann $k_B = R/N_A = 1,2807 \times 10^{-23}$ J/K; esta identificación fue demostrada por primera vez por Plank, quien la bautizó con su nombre actual. Vamos a ver que esta definición satisface las propiedades requeridas.

Supongamos por simplicidad un sistema clásico. De acuerdo a nuestra definición, la entropía es

$$S(E) = k_B \ln \left(\frac{\Gamma(E)}{h^{3N}} \right) \quad (5.2)$$

donde $\Gamma(E)$ es el volumen de una cáscara encerrada entre las superficies de energía constante $H = E$ y $H = E + \Delta$.

Consideremos el caso típico de equilibrio térmico de dos sistemas (1) y (2) (pensemos dos fluidos). Inicialmente tenemos cada sistema en un recipiente perfectamente rígido, separados por una pared adiabática y rígida, de tal manera que cada subsistema tiene un volumen fijo V_1 y V_2 , así como un número de partículas fijo N_1 y N_2 respectivamente. Si la pared es perfectamente adiabática los subsistemas no interactúan y el Hamiltoniano total del sistema será

$$H(p, q) = H_1(p_1, q_1) + H_2(p_2, q_2) \quad (5.3)$$

donde (p_1, q_1) y (p_2, q_2) denotan respectivamente las coordenadas y momentos de las partículas contenidas en cada subsistema.

Supongamos que inicialmente el subsistema (1) tiene energía entre E_1^0 y $E_1^0 + \Delta$ y el subsistema (2) entre E_2^0 y $E_2^0 + \Delta$. El número de microestados de cada sistema en su ensemble respectivo es $\Gamma_1(E_1^0)/h^{3N_1}$ y $\Gamma_2(E_2^0)/h^{3N_2}$, de manera que las entropías de cada subsistema, de acuerdo con la definición (5.2), son

$$S_1(E_1^0, V_1, N_1) = k_B \ln \left(\frac{\Gamma_1(E_1^0)}{h^{3N_1}} \right)$$

$$S_2(E_2^0, V_2, N_2) = k_B \ln \left(\frac{\Gamma_2(E_2^0)}{h^{3N_2}} \right)$$

Si los subsistemas estan aislados, el ensemble del sistema compuesto esta dado por todos los puntos del espacio de las fases tal que (a) las coordenadas de las N_1 partículas están contenidas

en el volúmen V_1 ; (b) las coordenadas de las N_2 partículas están contenidas en el volúmen V_2 ; (c) la energía del subsistema i esta entre E_i^0 y $E_i^0 + \Delta$. La energía del sistema total está entre¹ E y $E + 2\Delta$, con $E = E_1^0 + E_2^0$. El volúmen de este ensemble es claramente

$$\Gamma(E) = \Gamma_1(E_1^0) \times \Gamma_2(E_2^0)$$

de donde se sigue que

$$S(E, V, N) = S_1(E_1^0, V_1, N_1) + S_1(E_2^0, V_2, N_2)$$

donde

$$V = V_1 + V_2$$

$$N = N_1 + N_2$$

Supongamos que ahora se remueve el vínculo adiabático, de manera que los sistemas pueden intercambiar energía a través de la pared, manteniendo los volúmenes macroscópicos V_1 y V_2 fijos. Podemos pensar en colisiones inelásticas de las partículas que transmiten energía a la pared, produciendo vibraciones que no alteran los volúmenes promedio, pero pueden transmitir energía a su vez al otro subsistema. La presencia de la pared diatérmica introduce una interacción efectiva entre las partículas de ambos subsistemas, con lo cual deberíamos adicionar un término $H_{int}(p_1, q_1, p_2, q_2)$ al Hamiltoniano (5.3). Sin embargo, si la interacción es de corto alcance, dicho término tendrá peso únicamente para partículas en las cercanías de la interface entre ambos subsistemas, siendo por lo tanto proporcional a su área. En el límite termodinámico $V_1 \rightarrow \infty$ y $V_2 \rightarrow \infty$ la relación superficie/volúmen tiende a cero y dicho término puede despreciarse en comparación con las energías internas de cada subsistema. Es importante notar que, desde un punto de vista microscópico, esto sólo es válido si no existen potenciales de interacción fuertes de largo alcance²

Así, el efecto principal de remover el vínculo es el de permitir el intercambio de energía entre los subsistemas. El ensemble para este sistema resulta entonces **ampliado** respecto del sistema no interactuante, ya que ahora cada subsistema podrá adoptar **cualquier** par de valores E_1 y E_2 que satisfagan $E_1 + E_2 = E$. Para cada par de valores posibles (E_1, E_2) que satisfacen esta condición el número de estados accesibles del sistema completo con energía entre E y $E + 2\Delta$ es

$$W(E) = \frac{\Gamma(E)}{h^{3N}} = \frac{1}{h^{3N}} \Gamma_1(E_1) \Gamma_2(E - E_1)$$

Para obtener el número total de estados tenemos entonces que sumar esta cantidad sobre todos los pares posibles de valores de E_1 y E_2 compatibles con la energía total E . Dado que E_1 y E_2 son valores posibles de los Hamiltonianos $H_1(p_1, q_1)$ y $H_2(p_2, q_2)$, su espectro de valores esta acotado por debajo ya que de otra manera los subsistemas no serían estables. Por simplicidad, asumiremos que la cota inferior para ambos espectros es cero. Si dividimos cada espectro de energía en intervalos de tamaño Δ , entonces entre cero y E tenemos E/Δ intervalos en cada espectro. Para $\Delta \ll E$ podemos aproximar:

¹Como veremos, en el límite termodinámico el valor de la constante Δ resulta irrelevante para las predicciones de la teoría. Así, no tiene importancia que la cáscara del sistema total sea mas “gruesa” que las de los subsistemas

²Este sería el caso de partículas cargadas electricamente que interactúan via un potencial de Coulomb. En principio, los sistemas de partículas con interacciones de largo alcance no son extensivos y la cuestión del equilibrio termodinámico es un problema delicado y en general no bien resuelto. En el caso de las interacciones gravitatorias, las mismas pueden despreciarse (al menos para sistemas moleculares) ya que, si bien de largo alcance, son extremadamente débiles comparadas con otros potenciales de interacción moleculares, tales como las fuerzas de Van der Waals.

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i) \Gamma_2(E - E_i) \quad (5.4)$$

donde E_i es la energía en el centro del intervalo i -ésimo. Esta es una suma sobre E/Δ términos. Vamos a mostrar que en el límite termodinámico esta suma esta dominada por un único término. $\Gamma(E)$ crece monótonamente con la energía, ya que a medida que aumenta la misma aparecen nuevos estados accesibles. Así, mientras $\Gamma_1(E_i)$ crece, $\Gamma_2(E - E_i)$ decrece y viceversa. La función $\Gamma_1(E_i) \Gamma_2(E - E_i)$ debe por lo tanto presentar un máximo para algún valor de $0 \leq E_i \leq E$. Sean \bar{E}_1 y $\bar{E}_2 = E - \bar{E}_1$ los valores de energías para los cuales el producto $\Gamma_1(E_1) \Gamma_2(E_2)$ es máximo. Tenemos entonces que

$$\Gamma_1(\bar{E}_1) \Gamma_2(\bar{E}_2) \leq \Gamma(E) \leq \frac{E}{\Delta} \Gamma_1(\bar{E}_1) \Gamma_2(\bar{E}_2)$$

ó

$$\ln [\Gamma_1(\bar{E}_1) \Gamma_2(\bar{E}_2)] \leq \ln \Gamma(E) \leq \ln [\Gamma_1(\bar{E}_1) \Gamma_2(\bar{E}_2)] + \ln \frac{E}{\Delta} \quad (5.5)$$

En general uno observa que el número de microestados de un sistema molecular crece exponencialmente con el número de partículas, esto es, para $N_1 \rightarrow \infty$, $N_2 \rightarrow \infty$ se tiene que

$$\ln \Gamma_1 \propto N_1$$

$$\ln \Gamma_2 \propto N_2$$

lo cual resulta consistente con la extensividad de la entropía, mientras que

$$E \propto N = N_1 + N_2.$$

Por lo tanto, el término $\ln(E/\Delta)$ en la Ec.(5.5) puede ser despreciado, ya que Δ es independiente de N . Por lo tanto

$$S(E, V, N) = S_1(\bar{E}_1, V_1, N_1) + S_2(\bar{E}_2, V_2, N_2) + \mathcal{O}(\ln N) \quad (5.6)$$

Hemos obtenido entonces que la entropía de Boltzmann es extensiva (en el límite termodinámico) y que los subsistemas en equilibrio adoptan energías bien definidas \bar{E}_1 y \bar{E}_2 . Entendamos mejor esto. La hipótesis fundamental es que

$$\Gamma_i \sim e^{N_i F_i(E_i)} \quad i = 1, 2$$

donde las funciones F_i son intensivas (esto es, se vuelven independientes de N para $N \rightarrow \infty$); esto, como veremos en numerosos ejemplos, es absolutamente genérico. Asumamos por simplicidad que $N_1 = N_2 = N/2$. Tenemos entonces que

$$\Gamma_1(E_1) \Gamma_2(E - E_1) \sim e^{N g(E_1)}$$

donde la función g es también intensiva y tiene su máximo en $E_1 = \bar{E}_1$. Así, podemos expresar la suma (5.4) como

$$\begin{aligned} \Gamma(E) &\sim \sum_{E_1} e^{N g(E_1)} \\ &\sim e^{N g(\bar{E}_1)} \left[1 + \sum_{E_1 \neq \bar{E}_1} e^{-N(g(\bar{E}_1) - g(E_1))} \right] \end{aligned} \quad (5.7)$$

Dado que $g(\bar{E}_1) - g(E_1) > 0 \forall E_1 \neq \bar{E}_1$, todos los términos de la suma decaen rápidamente a cero. Cada valor de E_1 define un macroestado particular del sistema total, en el cual el subsistema 1 tiene energía macroscópica E_1 y el subsistema 2 energía E_2 . Dicho macroestado particular define un subensamble. Dado que todos los microestados del sistema total son equiprobables, la probabilidad de que el sistema se encuentre en un macroestado particular con energías E_1 y E_2 viene dada por el cociente entre el número de estados del subensamble asociado $\Gamma_1(E_1)\Gamma_2(E_2)$ y $\Gamma(E)$. Llamemos entonces R al subensamble correspondiente a las energías \bar{E}_1 y \bar{E}_2 . La probabilidad de encontrar el sistema en el mismo es

$$P(R) = \frac{\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)}{\Gamma(E)} \sim \frac{1}{1 + \sum_{E_1 \neq \bar{E}_1} e^{-N(g(\bar{E}_1) - g(E_1))}}$$

y por lo tanto $P(R) \rightarrow 1$ cuando $N \rightarrow \infty$. El crecimiento exponencial en N del número de microestados hace que, para N suficientemente grande, la inmensa mayoría de los estados accesibles correspondan a R . Todos estos argumentos se aplican exactamente igual al caso cuántico.

Los valores que adoptan entonces las energías de cada subsistema en ausencia del vínculo son aquellos que maximizan el producto $\Gamma_1(E_1)\Gamma_2(E_2)$ bajo la restricción $E_1 + E_2 = E$, esto es

$$\delta [\Gamma_1(E_1)\Gamma_2(E_2)], \quad \delta E_1 + \delta E_2 = 0$$

lo cual lleva a la condición

$$\Gamma_2(E_2) \frac{\partial \Gamma_1(E_1)}{\partial E_1} \delta E_1 + \Gamma_1(E_1) \frac{\partial \Gamma_2(E_2)}{\partial E_2} \delta E_2 = 0 \quad (5.8)$$

ó

$$\begin{aligned} \left[\frac{1}{\Gamma_1(E_1)} \frac{\partial \Gamma_1(E_1)}{\partial E_1} \right]_{E_1=\bar{E}_1} &= \left[\frac{1}{\Gamma_2(E_2)} \frac{\partial \Gamma_2(E_2)}{\partial E_2} \right]_{E_2=\bar{E}_2} \\ \left[\frac{\partial \ln \Gamma_1(E_1)}{\partial E_1} \right]_{E_1=\bar{E}_1} &= \left[\frac{\partial \ln \Gamma_2(E_2)}{\partial E_2} \right]_{E_2=\bar{E}_2} \\ \left[\frac{\partial S_1(E_1)}{\partial E_1} \right]_{E_1=\bar{E}_1} &= \left[\frac{\partial S_2(E_2)}{\partial E_2} \right]_{E_2=\bar{E}_2} \end{aligned} \quad (5.9)$$

Es decir, los valores que adoptan las energías de cada subsistema son aquellos **maximizan** la entropía total $S(E) = S_1(E_1) + S_2(E_2)$ entre todos los conjuntos de valores que conservan la energía total. Esto no es ni más ni menos que el enunciado del II Postulado de la Termodinámica. Desde el punto de vista microscópico, el macroestado que maximiza la entropía no es el único posible, pero la probabilidad de encontrar el sistema en otro macroestado es extremadamente pequeña cuando N es suficientemente grande.

Tomado la definición termodinámica de temperatura

$$\frac{\partial S(E)}{\partial E} \equiv \frac{1}{T} \quad (5.10)$$

la Ec.(5.9) nos da la conocida condición de equilibrio térmico

$$\frac{1}{T_1} = \frac{1}{T_2} \quad (5.11)$$

Las condiciones de equilibrio mecánico (es decir, al liberar el vínculo que mantiene la pared fija, además de volverla diatérmica) se obtienen de manera semejante, tomando en cuenta que el número de microestados es claramente una función creciente del volumen (queda como ejercicio para el lector). Vemos así que la entropía de Boltzmann satisface todos los requisitos establecidos por los postulados de la termodinámica y por lo tanto

$$S(E, V, N) = k_B \ln W(E, V, N)$$

constituye una **relación fundamental**. Sin embargo, conviene recalcar que esto es cierto **solamente en el límite termodinámico**. Solo en dicho límite la entropía anterior es una función homogénea de primer orden de los parámetros extensivos. Por otra parte, tenemos una interpretación estadística clara para el segundo postulado: el macroestado que maximiza la entropía es aquel que contiene el mayor número de microestados, solo por el hecho de que este es el más probable. Otros macroestados no están excluidos, pero su probabilidad se vuelve infinitesimalmente pequeña cuando el número de partículas se vuelve suficientemente grande.

Antes de continuar con el desarrollo de la teoría, veremos algunos ejemplos.

5.2. Algunos ejemplos

5.2.1. El gas ideal monoatómico clásico

Al igual que en termodinámica, el gas ideal monoatómico clásico es el sistema más simple que podemos considerar: N partículas puntuales idénticas no interactuantes en un volumen V . El Hamiltoniano del sistema es

$$H(p, q) = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2 \quad (5.12)$$

donde m es la masa de las partículas. Tenemos que calcular el volumen en el espacio de las fases $\Gamma(E)$ dado por la Ec.(4.12). Para ello resulta conveniente introducir la función

$$\Sigma(E) = \int_{H(p,q) \leq E} dp dq \quad (5.13)$$

Dado que la energía dada por la Ec.(5.12) está acotada por abajo, esta integral es finita (para N finito). Tenemos entonces que

$$\Gamma(E) = \Sigma(E + \Delta) - \Sigma(E)$$

Si Δ se elige tal que $\Delta \ll E$, tenemos que

$$\Gamma(E) = \omega(E) \Delta \quad (5.14)$$

donde

$$\omega(E) = \frac{\partial \Sigma(E)}{\partial E} \quad (5.15)$$

Un cálculo semejante al utilizado para obtener la relación (5.6) muestra que las siguientes definiciones son equivalentes, a menos de un término aditivo de orden $\ln N$ o menor:

$$S = k_B \ln[\Gamma(E)/h^{3N}] \quad (5.16)$$

$$S = k_B \ln[\Sigma(E)/h^{3N}] \quad (5.17)$$

$$S = k_B \ln[\omega(E)/h^{3N}] \quad (5.18)$$

Calculemos entonces $\Sigma(E)$ para el gas ideal. Dado que el Hamiltoniano es independiente de las posiciones, la integración de las mismas en la Ec.(5.13) es directamente la integral de la posición de cada partícula en el volumen V y por lo tanto

$$\Sigma(E) = V^N \int_{H(p,q) \leq E} dp_1 \dots dp_{3N} \quad (5.19)$$

Definiendo $R = \sqrt{2mE}$, la condición $H = E$ resulta

$$\sum_{i=1}^{3N} p_i^2 = R^2$$

y así

$$\Sigma(E) = V^N \Omega_{3N}(R) \quad (5.20)$$

donde

$$\Omega_n(R) = \int_{x_1^2 + \dots + x_n^2 \leq R^2} dx_1 \dots dx_n \quad (5.21)$$

es el volúmen de una hiperesfera de radio R en n dimensiones. Para calcular Ω_n notemos primero que

$$\Omega_n = C_n R^n \quad (5.22)$$

donde C_n es una constante independiente de R . Para calcular C_n vamos a usar la identidad

$$\int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n e^{-(x_1^2 + \dots + x_n^2)} = \left(\int_{-\infty}^{\infty} dx e^{-x^2} \right)^n = \pi^{n/2} \quad (5.23)$$

Sea ahora $f(r)$ una función arbitraria del radio $r = \sqrt{x_1^2 + \dots + x_n^2}$. Tenemos que

$$\int_{x_1^2 + \dots + x_n^2 \leq R^2} f(r) dx_1 \dots dx_n = \int_0^R S_n(r) f(r) dr \quad (5.24)$$

donde $S_n(r)$ es la superficie de una hiperesfera de radio R en n dimensiones. Así

$$\Omega_n(R) = \int_0^R S_n(r) dr \quad (5.25)$$

y por lo tanto

$$S_n(R) = \frac{d\Omega_n(R)}{dR} = n C_n R^{n-1} \quad (5.26)$$

El lado izquierdo de la Ec.(5.23) puede ser entonces expresado como

$$\begin{aligned} \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_n e^{-(x_1^2 + \dots + x_n^2)} &= \int_0^{\infty} dr S_n(r) e^{-r^2} \\ &= n C_n \int_0^{\infty} dr r^{n-1} e^{-r^2} \\ &= \frac{1}{2} n C_n \int_0^{\infty} dt t^{(n/2)-1} e^{-t} = \frac{1}{2} n C_n \left(\frac{n}{2} - 1 \right)! \end{aligned} \quad (5.27)$$

donde hemos usado que la última integral es la función gamma de argumento $n/2$. Comparando las Ecs.(5.23) y (5.27) obtenemos finalmente

$$C_n = \frac{2 \pi^{n/2}}{n (n/2 - 1)!} \quad (5.28)$$

Por lo tanto

$$\Sigma(E) = C_{3N} \left[V (2mE)^{3/2} \right]^N \quad (5.29)$$

Notemos que $\Sigma(E)$ (y por lo tanto $\Gamma(E)$) crece exponencialmente con N , como habíamos supuesto en la sección anterior. De la Ec.(5.17) obtenemos entonces la entropía para el gas ideal

$$S(E, V, N) = k_B \left[\ln C_{3N} + N \ln \left(\frac{V}{h^3} (2mE)^{3/2} \right) \right] \quad (5.30)$$

De la Ec.(5.28), utilizando la aproximación de Stirling

$$\ln m! \sim m \ln m - m \quad \text{para } m \rightarrow \infty \quad (5.31)$$

obtenemos finalmente

$$S(E, V, N) = N k_B \ln \left[V \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} \right] + \frac{3}{2} N k_B \quad (5.32)$$

Invirtiéndola, podemos despejar E en función de los restantes parámetros. Llamando a la función resultante $U(S, V, N)$ tenemos

$$U(S, V, N) = \left(\frac{3h^2}{4\pi m} \right) \frac{N}{V^{2/3}} \exp \left(\frac{2S}{3Nk_B} - 1 \right) \quad (5.33)$$

La temperatura es

$$T = \left(\frac{\partial U}{\partial S} \right)_V = \frac{2}{3} \frac{U}{Nk_B} \quad (5.34)$$

de donde se sigue que la capacidad calorífica a volumen constante es (para obtener el calor específico hay que dividir por el número molar)

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = \frac{3}{2} Nk_B \quad (5.35)$$

Finalmente,

$$P = - \left(\frac{\partial U}{\partial V} \right)_S = \frac{2U}{3V} = \frac{Nk_B T}{V} \quad (5.36)$$

Expresando $N = N_A N_m$, donde N_m es el número molar y N_A es el número de Avogadro se sigue de las expresiones anteriores que $k_B = R/N_A$, donde R es la constante universal de los gases.

A pesar de que hemos obtenido las ecuaciones de estado del gas ideal correctamente, el lector debe notar un hecho importantísimo. Si multiplicamos E , V y N por un factor λ arbitrario, de la Ec.(5.32) obtenemos que $S(\lambda E, \lambda V, \lambda N) \neq \lambda S(E, V, N)$. Esta entropía no es extensiva! Si calculamos la ecuación de estado para el potencial químico μ , veremos que el mismo no es una función intensiva de los parámetros termodinámicos.

Gibbs resolvió este problema de manera empírica, postulando que el cálculo del número de microestados está sobre-estimado y proponiendo de manera *ad hoc* la corrección $\Sigma(E) \rightarrow \Sigma(E)/N!$. Al hacer esto, debemos sustraer de la Ec.(5.32) el término $\ln N! \sim N \ln N - N$, con lo cual obtenemos

$$S(E, V, N) = N k_B \ln \left[\frac{V}{N} \left(\frac{E}{N} \right)^{3/2} \right] + \frac{3}{2} N k_B \left(\frac{5}{3} + \ln \frac{4\pi m}{3h^2} \right) \quad (5.37)$$

Esta expresión se conoce como *Fórmula de Sackur y Tetrode* y ha sido verificada experimentalmente. La corrección de dividir el número de microestados por $N!$ se conoce como **contaje**

correcto de Boltzmann³. La misma no afecta las ecuaciones de estado para P y T , ya que el término sustraído en la entropía no depende de E ni de V .

No es posible entender clásicamente porque debemos dividir $\Sigma(E)$ por $N!$ para obtener el conteo correcto de estados. La razón es inherentemente cuántica y está en la **indistinguibilidad** de partículas idénticas en Mecánica Cuántica.

Si las funciones de onda de dos partículas idénticas se pueden superponer (como en el caso de un gas) no existe manera de seguirlas individualmente, ya que las mismas no son entidades localizadas. Lo que tenemos es un único estado de dos partículas. Esto significa que, dada la función de onda de las N partículas, si permutamos las posiciones de dos cualesquiera de ellas la función de onda cambia a lo sumo en un factor de fase (ya que lo que tiene sentido físico es el módulo de la función de onda). Esto no existe en mecánica clásica, en donde cada partícula es una entidad con una posición espacial perfectamente definida y por lo tanto en principio es posible seguir el comportamiento individual de cada una. Supongamos un estado microscópico en el cual la partícula A se encuentra en la posición q_1 con momento p_1 y la partícula B se encuentra en la posición q_2 con momento p_2 ; si ahora consideramos el mismo microestado (esto es, con las restantes partículas con las mismas posiciones y momentos) pero con la partícula A en la posición q_2 con momento p_2 y la B en la posición q_1 con momento p_1 , ambos estados son diferentes. De esta manera, al considerar que las partículas en el gas son distinguibles, por cada conjunto particular de valores de posiciones y momentos contamos todas las posibles permutaciones de partículas entre dichos estados individuales ($N!$). Dado que en caso cuántico todas esas permutaciones constituyen un único estado, debemos dividir $\Sigma(E)$ por $N!$.

La argumentación anterior claramente no es una derivación del conteo correcto de Boltzmann, solo es una justificación general de su origen. A medida que avancemos en la materia veremos como el mismo emerge naturalmente del tratamiento cuántico correcto de partículas indistinguibles.

Como ya mencionamos, a temperaturas altas y bajas densidades los sistemas se comportan de acuerdo con la Mecánica Clásica. Esto es cierto en todo, excepto en la cuestión de la indistinguibilidad. Así, arribamos a la siguiente regla general: a altas temperaturas y bajas densidades, los sistemas de partículas indistinguibles admiten una descripción clásica (considerando las partículas como distinguibles), con la corrección dada por el conteo correcto de Boltzmann. Este enfoque suele denominarse Estadística de Maxwell-Boltzmann.

Hay que recalcar que partículas cuánticas idénticas pueden ser también **distinguibles** en determinadas situaciones, por ejemplo, átomos en un sólido cristalino. Dado que cada átomo tiene una posición fija bien definida, el permutar los estados individuales de dos átomos diferentes conduce a un nuevo estado microscópico. Veremos un ejemplo de este tipo a continuación

5.2.2. El sólido de Einstein

Con frecuencia los materiales en estado sólido presentan una estructura regular, en la cual los átomos vibran alrededor de posiciones relativas fijas, las cuales se ordenan en forma de arreglos periódicos (podemos pensar en una grilla). Estos arreglos periódicos se conocen como cristales ó **redes de Bravais**.

En 1906 Einstein propuso un modelo microscópico simple para explicar el comportamiento característico del calor específico de los sólidos con la temperatura: el mismo tiende a cero para $T \rightarrow 0$ y se vuelve constante para altas temperaturas (esto último se conoce como ley de Dulong y Petit). Si bien el calor específico es el resultado de diversas contribuciones (por ejemplo, electrónica) es posible ver que la principal contribución proviene de las vibraciones atómicas. El modelo de Einstein consiste entonces en asociar un oscilador armónico tridimensional independiente a cada uno

³Algunos autores llaman con justicia **factor de Gibbs** a la corrección $1/N!$. Sin embargo, he preferido mantener la denominación de “conteo correcto de Boltzmann” para evitar confusión, ya que existe una larga tradición en la literatura de llamarlo así.

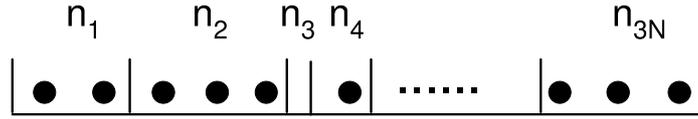


Figura 5.1: M bolas distribuidas en $3N$ cajas. n_i nos da el número de bolas en la caja i -ésima. En el ejemplo $n_1 = 2$, $n_2 = 3$, $n_3 = 0$, $n_4 = 1$, etc.

de los N sitios de un cristal. Los osciladores no interactúan entre ellos y se asume que todos tienen la misma frecuencia ω . Un autoestado de este sistema es entonces el producto de los autoestados de cada uno de los osciladores (los osciladores son independientes y distinguibles). En esta base el Hamiltoniano del sistema (y por lo tanto la matriz densidad) es diagonal y viene dado por

$$H = \hbar\omega \sum_{i=1}^{3N} (n_i + 1/2) = \hbar\omega \sum_{i=1}^{3N} n_i + \frac{3}{2} \hbar\omega N \quad (5.38)$$

donde $n_i = 0, 1, 2, \dots$ es el número cuántico que describe el autoestado individual del oscilador i -ésimo. Un microestado queda entonces especificado por un conjunto particular de valores para los $3N$ números cuánticos n_i . El número de microestados $W(E)$ para a un valor dado de la energía total E se obtiene entonces a partir de la condición

$$\hbar\omega \sum_{i=1}^{3N} n_i + \frac{3}{2} \hbar\omega N = E$$

o bien

$$\sum_{i=1}^{3N} n_i = M = \frac{E}{\hbar\omega} - \frac{3}{2} N$$

Esto claramente nos impone una restricción sobre los posibles valores de E , ya que M tiene que ser entero. En otras palabras, el sistema total tiene un espectro discreto. No obstante, $E \sim N$ y la variación relativa de E será $\Delta E/E \sim \hbar\omega/N \rightarrow 0$ para $N \rightarrow \infty$, con lo cual el espectro puede considerarse aproximadamente continuo.

El problema de la determinación de $W(E)$ consiste entonces en, dado un entero positivo M , calcular el número de posibles combinaciones de $3N$ enteros positivos n_1, \dots, n_{3N} tales que

$$\sum_{i=1}^{3N} n_i = M \quad (5.39)$$

Podemos resolver este problema combinatorio mediante el siguiente esquema mental. Supongamos que tenemos M bolas alineadas en una caja unidimensional. Supongamos además que la caja permite poner paredes divisorias entre cada par de bolas y que disponemos de $3N - 1$ de tales paredes. Una configuración particular de paredes va a dividir la caja completa en $3N$ subcajas, cada una conteniendo n_1, \dots, n_{3N} bolas, como se muestra esquemáticamente en la Fig.5.1. Vamos a suponer además que un número arbitrario de paredes pueden ubicarse en la misma posición, de manera que puede haber “cajas” que no contienen ninguna bola. Por ejemplo, podríamos poner todas las paredes en la primera posición, de manera que $n_1 = 1$, $n_{3N} = M - 1$ y $n_i = 0$ para los restantes casos. Por construcción, tenemos que $\sum_{i=1}^{3N} n_i = M$. Para cada configuración particular de posiciones de las paredes tenemos una configuración diferente de enteros n_1, \dots, n_{3N} que satisfacen la condición (5.39). Moviendo las paredes de lugar podemos recorrer todas las posibles combinaciones de enteros que satisfacen (5.39). Tenemos $3N - 1 + M$ objetos de dos clases (paredes

y bolas). W puede calcularse como el número de permutaciones de dichos objetos, en las cuales se descuentan las permutaciones de objetos del mismo tipo, esto es, el número combinatorio:

$$W(E) = \binom{3N + M - 1}{M} = \frac{(3N + M - 1)!}{M! (3N - 1)!} \quad (5.40)$$

y

$$S(E, N) = k_B \ln \left(\frac{(3N + M - 1)!}{M! (3N - 1)!} \right) \quad (5.41)$$

(para este modelo el volúmen no es una variable termodinámica independiente, ya que $V = aN$, donde a es una constante que depende de la estructura cristalina particular). Definiendo la entropía y la energía por átomo como $s = S/N$ y $u = E/N$ respectivamente, encontramos en el límite termodinámico $N \rightarrow \infty$ que

$$s(u) = 3k_B \left[\left(\frac{u}{\epsilon} + \frac{1}{2} \right) \ln \left(\frac{u}{\epsilon} + \frac{1}{2} \right) - \left(\frac{u}{\epsilon} - \frac{1}{2} \right) \ln \left(\frac{u}{\epsilon} - \frac{1}{2} \right) \right] \quad (5.42)$$

donde hemos usado la fórmula de Stirling y $\epsilon \equiv 3\hbar\omega$. Esta es la relación fundamental para el sólido de Einstein. La ecuación de estado para la temperatura viene dada por

$$\frac{1}{T} = \frac{\partial s}{\partial u} = \frac{3k_B}{\epsilon} \left[\ln \left(\frac{u}{\epsilon} + \frac{1}{2} \right) - \ln \left(\frac{u}{\epsilon} - \frac{1}{2} \right) \right] \quad (5.43)$$

Invirtiendola esta ecuación, obtenemos la energía por átomo en función de la temperatura:

$$u(T) = \frac{3}{2} \hbar\omega + \frac{3\hbar\omega}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1} \quad (5.44)$$

En el límite de bajas temperaturas $k_B T \ll \hbar\omega$ obtenemos la energía cuántica del punto cero $u \rightarrow (3/2)\hbar\omega$ (cada oscilador armónico se encuentra en su estado fundamental). En el límite de altas temperaturas $k_B T \gg \hbar\omega$ tenemos que $u \sim k_B T$. El calor específico de este modelo viene dado por

$$c(T) = \frac{\partial u}{\partial T} = 3k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \frac{1}{\left[\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1 \right]^2} \quad (5.45)$$

y se muestra en la Fig.5.2. En el límite de altas temperaturas $T \gg T_E = \hbar\omega/k_B$, donde la temperatura de Einstein es del orden de los 100 K para la mayoría de los sólidos cristalinos mas comunes, encontramos que $c \rightarrow 3k_B$. Esta es la ley de Dulong y Petit. A bajas temperaturas $T \ll T_E$ el calor específico se anula exponencialmente con la temperatura:

$$c \sim 3k_B \left(\frac{\hbar\omega}{k_B T} \right)^2 \exp\left(-\frac{2\hbar\omega}{k_B T}\right)$$

El modelo de Einstein describe de manera cualitativamente correcta los aspectos globales del comportamiento del calor específico. Sin embargo, cuando años después del trabajo pionero de Einstein se midió con alta precisión esta cantidad, se encontró que la misma tiende a cero como T^3 . Un modelo mas sofisticado para explicar este comportamiento fue introducido por Debye y lo veremos mas adelante.

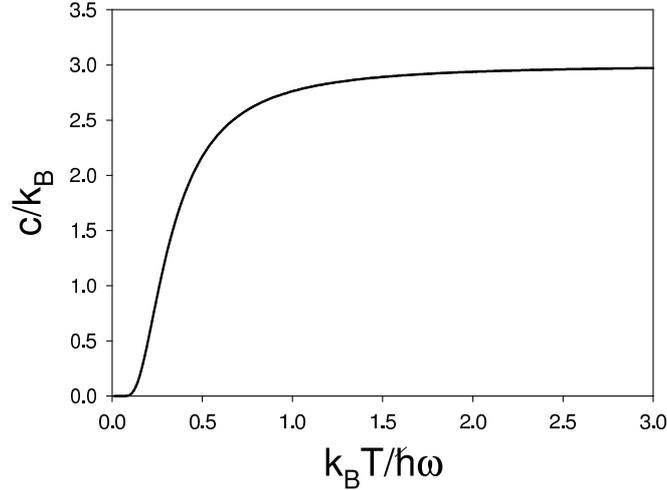


Figura 5.2: Calor específico del modelo de Einstein en función de la temperatura.

5.2.3. El calor específico de un sólido clásico

Resulta instructivo considerar la versión clásica del sólido de Einstein. El Hamiltoniano clásico de un conjunto de $3N$ osciladores armónicos independientes viene dado por

$$H(p, q) = \frac{1}{2} \sum_{i=1}^{3N} \left[\frac{1}{m} p_i^2 + m \omega^2 q_i^2 \right] \quad (5.46)$$

Calculemos

$$\Sigma(E) = \int_{H(p,q) \leq E} dp dq$$

Haciendo el cambio de variables $x_i = p_i / \sqrt{2m}$ para $i = 1, \dots, 3N$ y $x_i = \omega q_i \sqrt{m/2}$ para $i = 3N + 1, \dots, 6N$ tenemos que

$$\Sigma(E) = \left(\frac{2}{\omega} \right)^{3N} \int_{x_1^2 + \dots + x_{6N}^2 \leq R^2} dx_1 \cdots dx_{6N} = \left(\frac{2}{\omega} \right)^{3N} \Omega_{6N}(\sqrt{E}) = \left(\frac{2}{\omega} \right)^{3N} C_{6N} E^{3N}$$

donde $\Omega_n(R)$ es el volumen de la hipersfera en n dimensiones definida en la sección 5.2.1. Usando la Ec.(5.28) y la aproximación de Stirling tenemos en el límite termodinámico

$$s = S/N = \frac{k_B}{N} \ln \left(\frac{\Sigma(E)}{h^{3N}} \right) \rightarrow 3k_B \left[1 + \ln \left(\frac{u}{3\hbar\omega} \right) \right] \quad (5.47)$$

y por lo tanto

$$\frac{1}{T} = \frac{\partial s}{\partial u} = \frac{3k_B}{u} \quad \text{ó} \quad u = 3k_B T$$

de donde se deriva la ley de Dulong y Petit: $c = 3k_B$. Vemos así que el modelo clásico reproduce el límite de altas temperaturas de su correspondiente modelo cuántico. Como veremos a lo largo del curso, esta es una propiedad absolutamente general.

Notemos que la energía interna (y por lo tanto el calor específico) es totalmente independiente de las características de los osciladores (masa y frecuencia). Siguiendo los pasos anteriores, es fácil demostrar entonces que para un Hamiltoniano clásico

$$H = \sum_{i=1}^{fN} [A_i p_i^2 + B_i q_i^2]$$

la energía por partícula resulta

$$u = f k_B T$$

Cada grado de libertad armónico en el Hamiltoniano contribuye con $\frac{1}{2}k_B T$ a la energía. Esto se conoce como **Teorema de equipartición de la energía**.

5.3. La formulación de Gibbs

J. Gibbs propuso una expresión para la entropía alternativa a la de Boltzmann que, si bien enteramente equivalente a la anterior, permite formular la teoría en base a un principio variacional. Mas allá de la elegancia de este tipo de formulación, veremos que la misma permite extender de manera simple los resultados del ensemble microcanónico (sistema aislado) a situaciones mas generales (sistemas no aislados). Supongamos un sistema en equilibrio, con lo cual la probabilidad de encontrarlo en un dado microestado esta descrita por una probabilidad ρ independiente del tiempo, la cual podemos asumir (en base al Teorema de Liouville) una función del Hamiltoniano. La entropía de Gibbs tiene la forma

$$S = -k_B \langle \ln(C\rho) \rangle \quad (5.48)$$

donde ρ es la densidad de probabilidad en equilibrio (en el caso clásico) ó el operador densidad (en el caso cuántico) y C es una constante que vale $C = 1$ en el caso cuántico y $C = h^{3N}$ en el caso clásico. El promedio en la expresión anterior es un promedio en ensemble, es decir que se calcula con la misma ρ . Así, en el caso clásico tenemos que

$$S = -k_B \int \rho(p, q) \ln[h^{3N} \rho(p, q)] dp dq \quad (5.49)$$

donde la integral se extiende a todo el espacio de las fases; en el caso cuántico tenemos

$$S = -k_B \text{Tr} [\hat{\rho} \ln \hat{\rho}] \quad (5.50)$$

o en la base en que $\hat{\rho}$ es diagonal

$$S = -k_B \sum_n \rho_{nn} \ln \rho_{nn} \quad (5.51)$$

donde n indexa un conjunto completo de números cuánticos. Dado que $0 \leq \rho_{nn} \leq 1$, la entropía definida de esta manera es $S \geq 0$. Esta entropía puede pensarse como una funcional de ρ . Diferentes elecciones de ρ darán lugar a diferentes valores de la entropía. Por ejemplo, podríamos considerar un estado totalmente ordenado, esto es un macroestado en el cual el sistema puede encontrarse en un único microestado n_0 (esto se denomina un estado cuántico *puro*). En ese caso, la probabilidad debe ser igual a uno para dicho microestado y cero para los demás. Tomando entonces el límite $\rho_{nn} \rightarrow 0$ para todo estado $n \neq n_0$ (con lo cual $\rho_{n_0 n_0} \rightarrow 1$) tenemos que $S = 0$ para este macroestado. En el caso clásico es mas delicado hacer este cómputo, ya que involucra deltas de Dirac, pero podemos pensar que partimos el espacio de las fases en volúmenes mínimos de tamaño h^{3N} , con lo cual se reduce al cálculo anterior. Dado que la entropía de Gibbs es definida no negativa, un estado puro tiene entropía mínima .

Vamos a **postular** que la densidad de probabilidad (u operador densidad) que describe los estados de equilibrio del sistema es aquella que **maximiza** la entropía de Gibbs, compatible con los vínculos macroscópicos del sistema. A estos vínculos, debemos añadir siempre la condición de normalización

$$\int \rho(p, q) dp dq = 1 \quad (5.52)$$

ó

$$\text{Tr } \hat{\rho} = 1 \quad (5.53)$$

Para un sistema aislado, los vínculos son que E , V y N son constantes y por lo tanto ρ debe anularse para cualquier microestado que no cumpla estas condiciones. Calculemos entonces el máximo de S sujeto a la condición de normalización, la cual podemos imponer introduciendo un multiplicador de Lagrange que llamaremos α_0 . Veamos el caso clásico. La condición de extremo se obtiene pidiendo que la variación

$$\delta \left(S[\rho] - \alpha_0 \int \rho dp dq \right) = 0 \quad (5.54)$$

para variaciones pequeñas arbitrarias $\delta\rho$, donde las integrales se llevan a cabo en la cáscara $E \leq H(p, q) \leq E + \Delta$. Tenemos

$$\begin{aligned} \delta \left(S[\rho] - \alpha_0 \int \rho dp dq \right) &= \left(S[\rho + \delta\rho] - \alpha_0 \int (\rho + \delta\rho) dp dq \right) - \left(S[\rho] - \alpha_0 \int \rho dp dq \right) \\ &= -k_B \int \left((\rho + \delta\rho) \ln[h^{3N}(\rho + \delta\rho)] - \rho \ln[h^{3N}\rho] \right) dp dq - \alpha_0 \int \delta\rho dp dq \end{aligned}$$

Desarrollando el integrando de la primera integral a primer orden en $\delta\rho$ y despreciando las variaciones de orden superior:

$$\int \left(-k_B - k_B \ln[h^{3N}\rho] - \alpha_0 \right) \delta\rho dp dq = 0$$

Dado que las variaciones $\delta\rho$ son arbitrarias obtenemos

$$\rho = \frac{1}{h^{3N}} e^{-(\alpha_0/k_B + 1)} = cte \quad (5.55)$$

El multiplicador de Lagrange se obtiene de reemplazar la expresión anterior en la condición de normalización (5.52) de donde $\rho = 1/\Gamma(E)$ para $E \leq H(p, q) \leq E + \Delta$. Analizando las variaciones a segundo orden es posible mostrar que este extremo es efectivamente un máximo.

La distribución equiprobable es la que maximiza la entropía de Gibbs para un sistema aislado. Si reemplazamos ρ en la ecuación (5.49) obtenemos la entropía de Boltzmann Ec.(5.2).

5.4. Ejercicios

1. Demuestre que las definiciones (5.16)-(5.18) son equivalentes en el límite termodinámico.
2. Obtenga el operador densidad cuántico a partir del principio variacional de Gibbs para un sistema aislado. Calcule la expresión final de la entropía de Gibbs. Ayuda: trabaje en la base de autoestados de \hat{H} en que $\hat{\rho}$ es diagonal.

Capítulo 6

El ensemble canónico

En el formalismo termodinámico es posible trabajar en diferentes representaciones (mediante la transformada de Legendre) las cuales, dependiendo de las circunstancias, resultan mas convenientes que la representación entropía. De la misma manera, en la Mecánica Estadística es posible utilizar otros ensembles diferentes del microcanónico, caracterizados por diferentes parámetros macroscópicos.

6.1. Densidad de probabilidad: la función partición

Consideremos el caso de un sistema simple con volumen V y número de partículas N en equilibrio con un reservorio térmico a temperatura T , con el cual puede intercambiar calor pero no trabajo. Nos preguntamos entonces cual es la densidad de probabilidad para un dado microestado. Dado que el sistema no está aislado, su energía no es constante y fluctúa (puede intercambiar energía con el reservorio). Pero dado que el sistema está en equilibrio, podemos asumir que, para un dado valor de la temperatura del reservorio, el *valor medio* de su energía esta fijo en

$$U = \langle E \rangle = \text{Tr}(\hat{\rho}\hat{H}) = \sum_n \rho_{nn} E_n \quad (6.1)$$

donde n representa un conjunto de números cuánticos que indexa una base de autoestados del Hamiltoniano de N partículas y E_n son los correspondientes autovalores. Podemos encontrar la densidad de probabilidad aplicando el principio variacional de Gibbs que enunciamos el capítulo pasado. Para ello debemos maximizar la entropía de Gibbs

$$S = -k_B \text{Tr}(\hat{\rho} \ln \hat{\rho}) = -k_B \sum_n \rho_{nn} \ln \rho_{nn} \quad (6.2)$$

sujeta al vínculo (6.1) y a la normalización

$$\text{Tr} \hat{\rho} = \sum_n \rho_{nn} = 1. \quad (6.3)$$

Introduciendo dos multiplicadores de Lagrange α_0 y α_1 asociados a cada vínculo tenemos

$$\begin{aligned} \delta \left[\text{Tr}(\alpha_0 \hat{\rho} + \alpha_1 \hat{\rho}\hat{H} - k_B \hat{\rho} \ln \hat{\rho}) \right] &= \delta \sum_n (\alpha_0 \rho_{nn} + \alpha_1 \rho_{nn} E_n - k_B \rho_{nn} \ln \rho_{nn}) \\ &= \sum_n [(\alpha_0 - k_B) + \alpha_1 E_n - k_B \ln \rho_{nn}] \delta \rho_{nn} = 0 \end{aligned} \quad (6.4)$$

a primer orden en las variaciones $\delta \rho_{nn}$. Para variaciones arbitrarias $\delta \rho_{nn}$ esto nos lleva a la condición

$$(\alpha_0 - k_B) + \alpha_1 E_n - k_B \ln \rho_{nn} = 0 \quad (6.5)$$

o bien

$$\rho_{nn} = \exp \left[\left(\frac{\alpha_0}{k_B} - 1 \right) + \frac{\alpha_1}{k_B} E_n \right] \quad (6.6)$$

El multiplicador de Lagrange α_0 puede eliminarse aplicando la condición de normalización (6.3):

$$\exp \left(\frac{\alpha_0}{k_B} - 1 \right) \sum_n \exp \left(\frac{\alpha_1}{k_B} E_n \right) = 1.$$

Definiendo la **función partición**

$$Z_N(\alpha_1) \equiv \exp \left(1 - \frac{\alpha_0}{k_B} \right) = \sum_n \exp \left(\frac{\alpha_1}{k_B} E_n \right) = \text{Tr} \left(e^{\alpha_1 \hat{H}/k_B} \right) \quad (6.7)$$

tenemos

$$\rho_{nn} = \frac{e^{\alpha_1 E_n/k_B}}{Z_N(\alpha_1)} = \frac{e^{\alpha_1 E_n/k_B}}{\sum_n e^{\alpha_1 E_n/k_B}} \quad (6.8)$$

A continuación determinamos el multiplicador α_1 . Para ello multiplicamos la Ec.(6.5) por ρ_{nn} y sumamos sobre n :

$$(\alpha_0 - k_B) + \alpha_1 U + S = 0 \quad (6.9)$$

$$-k_B \ln[Z_N(\alpha_1)] + \alpha_1 U + S = 0 \quad (6.10)$$

Recordemos que para un sistema simple a temperatura constante con volúmen y número de partículas constantes la representación termodinámica natural es la de Helmholtz, en la cual la relación fundamental es $F = F(T, V, N)$; la energía libre de Helmholtz satisface la relación: $F - U + TS = 0$. La comparación con la Ec.(6.10) nos sugiere la identificación $\alpha_1 = -1/T$ y

$$F(T, V, N) = -k_B T \ln[Z_N(T)] \quad (6.11)$$

La función partición toma entonces la forma

$$Z_N(T) = e^{-\beta F} = \text{Tr} \left(e^{-\beta \hat{H}} \right) \quad (6.12)$$

donde

$$\beta \equiv \frac{1}{k_B T} \quad (6.13)$$

El operador densidad resulta entonces

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{\text{Tr} \left(e^{-\beta \hat{H}} \right)} \quad (6.14)$$

El ensemble definido por la densidad de probabilidad (6.14) se conoce como **ensemble canónico**. Veremos en breve la consistencia de esta elección. Notemos que la energía libre F se expresa en términos de la función partición. Dado que la forma funcional $F = F(T, V, N)$ es una relación fundamental, vemos entonces que toda la información termodinámica esta contenida en la función partición.

La derivación en el caso clásico es enteramente análoga a la anterior (queda como ejercicio para el lector) y da como resultado

$$\rho(p, q) = \frac{e^{-\beta H(p, q)}}{Z_N(T)} \quad (6.15)$$

con

$$Z_N(T) = e^{-\beta F} = \frac{1}{h^{3N}} \int e^{-\beta H(p, q)} dp dq \quad (6.16)$$

donde la integral se extiende ahora a todo el espacio de las fases correspondiente a un sistema con V y N fijos.

6.2. Densidad de estados

Consideremos la integral en el espacio de las fases de una función arbitraria f que depende de (p, q) a través del Hamiltoniano

$$I = \frac{1}{h^{3N}} \int f[H(p, q)] dp dq$$

La misma puede escribirse como

$$I = \int_0^\infty f(E) g(E) dE$$

donde

$$g(E) = \int_{H(p, q)=E} dp dq / h^{3N} \quad (6.17)$$

se conoce como **densidad de estados**; $g(E) dE$ nos dá el número de estados con energía entre E y $E + dE$. Si en particular elegimos $f(H) = \Theta(E - H)$, donde $\Theta(x)$ es la función escalón de Heaviside, tenemos

$$I = \Sigma(E) / h^{3N} = \int_0^E g(E') dE'$$

de donde

$$g(E) = \frac{1}{h^{3N}} \frac{\partial \Sigma(E)}{\partial E} = \frac{\omega(E)}{h^{3N}} = e^{S_m(E)/k_B} \quad (6.18)$$

donde $S_m(E)$ es la entropía microcanónica correspondiente a una energía E . Tenemos por lo tanto que

$$Z_N(T) = \int_0^\infty e^{-\beta E} g(E) dE \quad (6.19)$$

En el caso cuántico podemos escribir

$$Z_N(T) = \sum_n e^{-\beta E_n} \quad (6.20)$$

donde el índice n simboliza un conjunto completo de números cuánticos, esto es, la suma corre sobre todos los posibles **autoestados** del Hamiltoniano del sistema y E_n son los correspondientes autovalores. Para un sistema finito el espectro es discreto y podemos escribir

$$Z_N(T) = \sum_E g'(E) e^{-\beta E} \quad (6.21)$$

donde, a diferencia de la suma en la Ec.(6.20), esta suma corre sobre todos los autovalores **diferentes** E del Hamiltoniano y $g'(E)$ es la degeneración del autovalor E . En muchos casos se observa que, en el límite termodinámico, el espectro se vuelve denso, de manera que la suma (6.21) converge a una integral del tipo (6.19). Así, podemos asumir esa expresión en general. Veamos un ejemplo.

Consideremos una partícula libre encerrada en una caja cúbica de dimensión L , de manera que $V = L^3$. Supongamos que las paredes son perfectamente rígidas, de manera que la función de onda debe anularse en las mismas. El Hamiltoniano de este sistema es

$$\hat{H} = \frac{\hat{p}^2}{2m}$$

donde $\hat{p} \equiv (\hat{p}_x, \hat{p}_y, \hat{p}_z)$ es el operador momento lineal de la partícula. Sea $|\vec{k}\rangle$ un autoestado ortonormalizado del operador vector de onda $\hat{k} = \hat{p}/\hbar$, con $\langle \vec{k} | \vec{k}' \rangle = \delta_{\vec{k}, \vec{k}'}$. En la representación coordenada tenemos

$$\langle \vec{r} | \vec{k} \rangle = \psi_{n_x, n_y, n_z}(x, y, z) = A \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) \sin\left(\frac{n_z \pi z}{L}\right) \quad (6.22)$$

donde A es una constante de normalización y n_x, n_y, n_z son enteros $1, 2, \dots, \infty$. Los autovalores permitidos del operador vector de onda \hat{k} son

$$\vec{k} = \frac{n_x \pi}{L} \mathbf{i} + \frac{n_y \pi}{L} \mathbf{j} + \frac{n_z \pi}{L} \mathbf{k}$$

donde \mathbf{i}, \mathbf{j} y \mathbf{k} son vectores unitarios en las direcciones x, y, z respectivamente.

Podemos calcular entonces la función partición de una partícula¹ $Z_1(T)$. La misma viene dada por

$$Z_1(T) = \text{Tr}_1 e^{-\beta \hat{p}^2 / 2m} = \sum_{\vec{k}} \langle \vec{k} | e^{-\beta \hbar^2 \vec{k}^2 / 2m} | \vec{k}' \rangle = \sum_{k_x=\pi/L}^{\infty} \sum_{k_y=\pi/L}^{\infty} \sum_{k_z=\pi/L}^{\infty} e^{-\beta E(\vec{k})} \quad (6.23)$$

donde los autovalores de la energía son

$$E(\vec{k}) = \frac{\hbar^2 |\vec{k}|^2}{2m} \quad (6.24)$$

Cada una de las sumas en la Ec.(6.23) corre sobre variables que cambian en una cantidad $\Delta k_i = \pi/L$ con $i = x, y, z$. Podemos reescribir

$$Z_1(T) = \frac{V}{\pi^3} \sum_{k_x=\pi/L}^{\infty} \sum_{k_y=\pi/L}^{\infty} \sum_{k_z=\pi/L}^{\infty} e^{-\beta E(\vec{k})} \Delta k_y \Delta k_x \Delta k_z \quad (6.25)$$

En el límite de volúmen infinito $L \rightarrow \infty$, $\Delta k_i \rightarrow 0$ y las sumas de la ecuación anterior convergen a una integral de Riemann, de manera que para $L \gg 1$ podemos aproximar

¹Esto puede parecer un poco extraño, ya que en general estamos interesados en el comportamiento de $\sim 10^{23}$ partículas. Sin embargo, mas allá del ejercicio teórico que representa esta cálculo, veremos mas adelante que la función partición de un gas ideal puede expresarse en términos de Z_1 .

$$\begin{aligned}
Z_1(T) &\sim \frac{V}{\pi^3} \int_0^\infty dk_x \int_0^\infty dk_y \int_0^\infty dk_z e^{-\beta E(k)} \\
&= \frac{V}{(2\pi)^3} \int_{-\infty}^\infty dk_x \int_{-\infty}^\infty dk_y \int_{-\infty}^\infty dk_z e^{-\beta E(k)} \quad (6.26)
\end{aligned}$$

$$= \frac{4\pi V}{(2\pi)^3} \int_0^\infty k^2 e^{-\beta E(k)} dk \quad (6.27)$$

donde $k = |\vec{k}|$. Haciendo el cambio de variables $k = \sqrt{2mE}/\hbar$ obtenemos finalmente

$$Z_1(T) = \int_0^\infty g(E) e^{-\beta E} dE \quad (6.28)$$

donde la densidad de estados $g(E)$ resulta:

$$g(E) = 2\pi \left(\frac{V}{h^3} \right) (2m)^{3/2} \sqrt{E} \quad (6.29)$$

Esta expresión puede interpretarse fácilmente como sigue. Los estados accesibles del sistema se encuentran distribuidos en los nodos una grilla cúbica en el espacio (k_x, k_y, k_z) . El número de estados con energía entre E y $E + dE$ corresponde al número de nodos en dicho espacio ubicados en un octante de la cáscara encerrada entre las esferas de radio R y $R + dR$ con $R = \sqrt{2mE}/\hbar$. Podemos estimar dicho número dividiendo el volúmen de la cáscara por el volúmen asociado a cada nodo $(\pi/L)^3$. El volúmen de una esfera de radio $R = \sqrt{2mE}/\hbar$ es

$$\Omega(E) = \frac{4\pi}{3} \frac{(2m)^{3/2}}{\hbar^3} E^{3/2}$$

y el volúmen de la cáscara

$$\Omega(E + dE) - \Omega(E) = \frac{d\Omega(E)}{dE} dE = 2\pi \frac{(2m)^{3/2}}{\hbar^3} E^{1/2} dE$$

Dividiendo por $(2\pi/L)^3$ (el factor $2^3 = 8$ viene porque consideramos un octante de la cáscara) se obtiene la Ec.(6.29).

Finalmente, podemos integrar y obtener la función partición

$$Z_1(T) = \frac{V}{\lambda_T^3} \quad (6.30)$$

donde

$$\lambda_T = \frac{h}{\sqrt{2\pi m k_B T}} \quad (6.31)$$

se conoce como *longitud de onda térmica* (esta cantidad tiene dimensiones de longitud). Veremos que constituye una magnitud importante en la descripción de los gases ideales, ya que puede verse que la misma es del orden de la longitud de onda de de Broglie para una partícula de masa m con energía $k_B T$.

Notemos que, de la Ec.(6.26), haciendo el cambio de variables $p = \hbar k$, obtenemos

$$Z_1(T) = \frac{V}{h^3} \int_{-\infty}^\infty e^{-\beta(p_x^2 + p_y^2 + p_z^2)/2m} dp_x dp_y dp_z \quad (6.32)$$

que corresponde exactamente a la función partición de una partícula clásica.

6.2.1. Ejercicios

1. Muestre que la densidad de estados y la función de partición de una partícula en una caja con condiciones de contorno periódicas resultan idénticas a las obtenidas con paredes reflectantes en el límite $V \gg 1$.
2. Calcule la densidad de estados para una partícula y la función de partición $Z_1(T)$ en una y dos dimensiones.

6.3. El gas ideal clásico en el ensemble canónico

La función de partición para el gas ideal clásico viene dada por

$$Z_N(T) = \frac{1}{h^{3N}} \int dq \int dp \exp \left[-\beta \sum_{i=1}^{3N} \frac{1}{2m} p_i^2 \right] = \left(\frac{V}{h^3} \right)^N \prod_{i=1}^{3N} \int_{-\infty}^{\infty} e^{-\beta p_i^2 / 2m} dp_i \quad (6.33)$$

Notemos que

$$Z_N(T) = \left(\frac{V}{h^3} \right)^N \left\{ \int_{-\infty}^{\infty} e^{-\beta(p_x^2 + p_y^2 + p_z^2) / 2m} dp_x dp_y dp_z \right\}^N = Z_1^N(T) = \left(\frac{V}{\lambda_T} \right)^N \quad (6.34)$$

donde hemos usado la Ec.(6.32). De las Ecs.(6.30) y (6.31) tenemos entonces

$$F = -k_B T N \ln Z_1(T) = -k_B T N \ln \left[\frac{V}{h^3} (2\pi m k_B T)^{3/2} \right] \quad (6.35)$$

Notemos que si multiplicamos V y N por un factor λ , $F(T, \lambda V, \lambda N) \neq \lambda F(T, V, N)$. Este es el mismo problema que encontramos en el ensemble microcanónico y la solución es la misma. Para descontar el conteo de múltiples configuraciones equivalentes de partículas indistinguibles, debemos dividir la densidad de estados por $N!$. De la Ec.(6.19), esto resulta equivalente a reemplazar $Z_N(T) \rightarrow Z_N(T)/N!$, con lo cual

$$F = -k_B T N \ln \left[\frac{V}{N h^3} (2\pi m k_B T)^{3/2} \right] - k_B T N \quad (6.36)$$

La entropía resulta

$$S = -\frac{\partial F}{\partial T} = k_B N \left\{ \ln \left[\frac{V}{N h^3} (2\pi m k_B T)^{3/2} \right] + \frac{5}{2} \right\} \quad (6.37)$$

y la energía interna

$$U = F + TS = \frac{3}{2} N k_B T \quad (6.38)$$

Reemplazando la Ec.(6.38) en la Ec.(6.37) recuperamos la fórmula de Sakkur y Tetrode

$$S(U, V, N) = N k_B \ln \left[\frac{V}{N} \left(\frac{U}{N} \right)^{3/2} \right] + \frac{3}{2} N k_B \left(\frac{5}{3} + \ln \frac{4\pi m}{3h^2} \right)$$

esto es, exactamente el mismo resultado obtenido en el ensemble microcanónico.

6.4. Algunas propiedades generales del ensemble canónico

Como hemos visto hasta ahora, el formalismo general puede ser desarrollado en forma completamente análoga en los casos clásico y cuántico. Para evitar de aquí en más duplicar los cálculos, vamos a adoptar la siguiente nomenclatura:

$$Z_N(T) = \text{Tr} e^{-\beta H} \quad (6.39)$$

donde el símbolo Tr va a denotar “suma sobre todos los microestados compatibles con los vínculos del sistema”, esto es, una auténtica traza de operadores definidos en un espacio de Hilbert apropiado en el caso cuántico o una integral en el espacio de las fases $\int \dots dpdq/h^{3N}$ en el clásico. Los símbolos ρ y H podrán denotar tanto operadores como funciones de (p, q) . Salvo que lo indiquemos explícitamente, todos los cálculos con esta notación serán aplicables tanto al caso cuántico como al clásico. Así, el promedio de cualquier observable en el ensemble canónico \hat{O} puede expresarse como

$$\langle \hat{O} \rangle = \frac{1}{Z_N} \text{Tr} [\hat{O} e^{-\beta H}] \quad (6.40)$$

Por ejemplo, la energía interna $U = \langle H \rangle$ resulta

$$U = \langle H \rangle = \frac{1}{Z_N} \text{Tr} [H e^{-\beta H}], \quad (6.41)$$

la cual puede ser escrita como

$$U = -\frac{\partial \ln Z_N}{\partial \beta} = \frac{\partial(\beta F)}{\partial \beta} \quad (6.42)$$

Otras ecuaciones de estado (por ejemplo, para la presión o la entropía) se obtienen directamente a partir de derivar la relación fundamental $F = -k_B T \ln Z_N$. Sin embargo, el formalismo de la Mecánica Estadística no se limita a una receta para calcular relaciones fundamentales. Podemos también obtener diversos tipos de información microscópica. Consideremos, por ejemplo, el calor específico (en el caso de un sistema simple sería el calor específico a volumen constante):

$$c = \frac{1}{N} \frac{\partial U}{\partial T} = -\frac{k_B \beta^2}{N} \frac{\partial U}{\partial \beta} \quad (6.43)$$

Derivando la Ec.(6.41) y usando la definición (6.39)

$$\begin{aligned} \frac{\partial U}{\partial \beta} &= -\frac{1}{Z_N^2} \frac{\partial Z_N}{\partial \beta} \text{Tr} [H e^{-\beta H}] - \frac{1}{Z_N} \text{Tr} [H^2 e^{-\beta H}] \\ &= -\frac{1}{Z_N} \frac{\partial \ln Z_N}{\partial \beta} \text{Tr} [H e^{-\beta H}] - \langle H^2 \rangle \\ &= \langle H \rangle^2 - \langle H^2 \rangle \end{aligned}$$

Y así

$$c = \frac{k_B \beta^2}{N} \left(\langle H^2 \rangle - \langle H \rangle^2 \right) = \frac{k_B \beta^2}{N} \langle (H - \langle H \rangle)^2 \rangle \quad (6.44)$$

El calor específico es proporcional a las *fluctuaciones de energía del sistema*. Recordemos que el calor específico nos da la respuesta del sistema (cantidad de calor absorbida o entregada) ante un cambio infinitesimal de temperatura. Vemos entonces que la respuesta térmica del sistema esta regulada por las fluctuaciones de energía. Como veremos a lo largo del curso, esta es una característica general: la respuesta del sistema ante perturbaciones externas pequeñas está regulada por las fluctuaciones

en equilibrio de alguna magnitud relacionada. La relación (6.44) no solo es de gran importancia conceptual, sino también de gran utilidad en los métodos de simulaciones numéricas en Mecánica Estadística.

6.5. Fluctuaciones de energía y equivalencia entre los ensembles canónico y microcanónico.

En el ensemble canónico la energía del sistema no es constante y por lo tanto pueden esperarse fluctuaciones respecto del valor medio $U = \langle H \rangle$. Vamos a analizar las propiedades estadísticas de dichas fluctuaciones.

Definamos la variable aleatoria $\delta U \equiv (H - U)/U$ que nos da la intensidad relativa de la fluctuación. Claramente $\langle \delta U \rangle = 0$. De la Ec.(6.44) tenemos que

$$\langle (\delta U)^2 \rangle = \frac{cN}{k_B \beta^2 U^2} \quad (6.45)$$

Dado que c es una variable intensiva y que $U \propto N$, tenemos que el desvío estándar de las fluctuaciones relativas decae a cero como

$$\sqrt{\langle (\delta U)^2 \rangle} \sim \frac{1}{\sqrt{N}}$$

En el límite termodinámico $N \rightarrow \infty$ las fluctuaciones relativas de la energía van a cero, lo que significa que, si bien pueden esperarse fluctuaciones en términos absolutos, las variaciones de energía esperables son despreciables comparadas con el valor medio. En otras palabras, si bien el sistema puede acceder en principio a cualquier valor de energía, la probabilidad de que adopte un valor significativamente diferente de U tiende a cero en el límite termodinámico. Veamos esto en mayor detalle.

La probabilidad de que el sistema tenga una energía E en el ensemble canónico viene dada por

$$\begin{aligned} P(E) &= \frac{g(E) e^{-\beta E}}{\int_0^\infty e^{-\beta E} g(E) dE} \\ &= \frac{e^{-\beta E + \ln g(E)}}{\int_0^\infty e^{-\beta E + \ln g(E)} dE} \\ &= \frac{e^{\beta[T S_m(E) - E]}}{\int_0^\infty e^{\beta[T S_m(E) - E]} dE} \end{aligned} \quad (6.46)$$

donde $S_m(E)$ es la entropía microcanónica para una energía E . El promedio en ensemble de una función arbitraria de la energía $f(E)$ esta dado por

$$\langle f(E) \rangle = \frac{\int_0^\infty f(E) e^{\beta[T S_m(E) - E]} dE}{\int_0^\infty e^{\beta[T S_m(E) - E]} dE} \quad (6.47)$$

Analicemos el exponente en estas integrales. La entropía microcanónica es una función creciente de E y cóncava (ya que asumimos que el sistema es estable termodinamicamente). Para valores pequeños de E el exponente es creciente, pero tarde o temprano el término lineal $-E$ domina y el mismo se vuelve decreciente. Por lo tanto, el exponente presenta un máximo para algún valor de $E = \bar{E}$, el cual satisface las condiciones

$$T \left. \frac{\partial S_m(E)}{\partial E} \right|_{E=\bar{E}} = 1 \quad (6.48)$$

$$\left. \frac{\partial^2 S_m(E)}{\partial E^2} \right|_{E=\bar{E}} < 0 \quad (6.49)$$

La primera condición nos dice que \bar{E} corresponde a la energía interna que obtendríamos en el ensemble microcanónico correspondiente a una temperatura T y por lo tanto es proporcional a N . Así, en el límite termodinámico las integrales están dominadas por la contribución en las inmediaciones del máximo de dicho exponente, ya que la función exponencial decae muy rápido al alejarnos de él. Si desarrollamos el exponente en potencias de $E - \bar{E}$:

$$[T S_m(E) - E] = [T S_m(\bar{E}) - \bar{E}] - \frac{1}{2}A(E - \bar{E})^2 + \dots$$

vemos que $P(E)$ tiende a una distribución gaussiana

$$P(E) = \frac{e^{-\beta A(E-\bar{E})^2/2}}{\int_0^\infty e^{-\beta A(E-\bar{E})^2/2} dE} \quad (6.50)$$

donde

$$A = -T \left. \frac{\partial^2 S_m(E)}{\partial E^2} \right|_{E=\bar{E}}$$

De aquí, promediando E resulta inmediato que $U = \bar{E}$. Cambiando de variable en la Ec.(6.50), vemos que las fluctuaciones δU tienen una distribución gaussiana

$$P(\delta U) \propto e^{-(\delta U)^2/2\sigma^2} \quad (6.51)$$

donde la variancia

$$\sigma^2 = \langle (\delta U)^2 \rangle = \frac{1}{\beta A U^2} = -\frac{k_B}{U^2} \left[\left. \frac{\partial^2 S_m(E)}{\partial E^2} \right|_{E=U} \right]^{-1}$$

De la Ec.(6.48) tenemos que

$$\left. \frac{\partial^2 S_m(E)}{\partial E^2} \right|_{E=U} = -\frac{1}{T^2} \frac{\partial T}{\partial U} = -\frac{1}{T^2} \left(\frac{\partial U}{\partial T} \right)^{-1} = -\frac{1}{cNT^2} \quad (6.52)$$

de donde recuperamos la Ec.(6.45).

Finalmente notemos que $U - T S_m = F_m$ es por definición la energía libre de Helmholtz microcanónica. Por otra parte, la energía libre canónica la hemos definido como $Z_N(T) = e^{-\beta F}$. A partir de los resultados anteriores, tenemos que

$$Z_N(T) = e^{-\beta F} = e^{-\beta F_m} \int_0^\infty e^{-\frac{(E-U)^2}{2k_B cNT^2}} dE \quad (6.53)$$

$$\int_0^\infty e^{-\frac{(E-U)^2}{2k_B cNT^2}} dE = \int_{-U}^\infty e^{-\frac{x^2}{2k_B cNT^2}} dx \approx \int_{-\infty}^\infty e^{-\frac{x^2}{2k_B cNT^2}} dx = \sqrt{2\pi k_B cNT^2}$$

y así

$$F \approx F_m - \frac{1}{2}k_B T \ln(N) \quad (6.54)$$

Dado que $F_m \sim N$, el último término es despreciable en el límite termodinámico.

Resumiendo: hemos visto que en el límite termodinámico las fluctuaciones de energía respecto del valor de equilibrio U tienen una distribución de probabilidad gaussiana cuyo ancho tiende a cero como $1/\sqrt{N}$; para $N \rightarrow \infty$ la distribución gaussiana tiende a una delta de Dirac, con lo cual los microestados con probabilidad no nula tienen todos prácticamente la misma energía. En dicho límite, las funciones termodinámicas (entropía, energía libre) obtenidas en ambos ensembles dan el mismo resultado, a menos de una corrección despreciable (crece como $\ln N$, mientras que las funciones termodinámicas crecen como N).

Podemos entender mejor lo anterior como sigue. En el ensemble canónico la probabilidad de un dado microestado es $\rho \propto e^{-\beta H}$, con lo cual decae exponencialmente rápido con la energía del estado, ya que la misma crece como N . Sin embargo, cuando consideramos la probabilidad de medir una dada energía E , debemos multiplicar el factor anterior por el número de microestados con dicha energía, esto es, por la densidad de estados. La densidad de estados nos da una medida del tamaño de la superficie energía, la cual a su vez crece exponencialmente con N . Así, para alguna energía intermedia estos dos efectos se compensan, dando lugar a una probabilidad finita. Para una dada temperatura, la energía donde eso ocurre corresponde al mínimo de la energía libre microcanónica $F = E - TS_m(E)$.

6.6. Sistemas de partículas indistinguibles

Como ya mencionamos, del principio de incerteza resulta que partículas idénticas en Mecánica Cuántica son indistinguibles, a menos que exista una situación particular que limite sus posiciones espaciales (por ejemplo, átomos que están asociados a los sitios de un cristal), permitiendo identificarlas de manera individual.

La propiedad de indistinguibilidad está relacionada con el hecho de que el Hamiltoniano de N partículas, así como otros observables, resultan invariantes ante permutaciones de las variables dinámicas asociadas a diferentes partículas, es decir, ante una reenumeración de las mismas. Por ejemplo, en situaciones bastante generales podemos describir un sistema de partículas a través de un Hamiltoniano que contiene un término de energía cinética y un potencial de interacción de pares

$$\hat{H}_N = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} + \sum_{i < j=1}^{N(N-1)/2} V(\hat{q}_i, \hat{q}_j, \hat{s}_i, \hat{s}_j) \quad (6.55)$$

donde \hat{p}_i , \hat{q}_i y \hat{s}_i son los operadores momento, posición y spin de la partícula i -ésima y donde la energía potencial V resulta invariante ante el intercambio simultáneo de las variables dinámicas de cualquier par de partículas i y j .

Sabemos que por cada operación de simetría en Mecánica Cuántica existe un operador asociado, el cual conmuta con \hat{H} y por lo tanto puede ser diagonalizado simultáneamente. Existirá por lo tanto un conjunto de operadores asociados a la operación de cambiar la enumeración de las partículas, el cual particiona el espacio de Hilbert en diferentes subespacios cada uno caracterizado por un conjunto diferente de autovalores de dichos operadores. Veremos a continuación que existen únicamente dos autovalores posibles asociados a la operación de permutación de partículas. Dichos autovalores constituyen cantidades conservadas y por lo tanto imponen un nuevo vínculo a los microestados accesibles de un sistema. En otras palabras, para un dado sistema de partículas indistinguibles, la matriz densidad de probabilidad solo tendrá elementos no nulos en uno de dichos subespacios.

6.6.1. Simetría de la función de onda de partículas indistinguibles

Supongamos por simplicidad partículas sin spin. El estado de un sistema de N partículas está descrito por una función de onda $\Psi(q_1, \dots, q_N)$, donde q_i denota el vector posición de la partícula i -ésima.

Cambios en la enumeración de las partículas pueden describirse definiendo los operadores permutación de pares \hat{P}_{ik} , los cuales intercambian las coordenadas q_i y q_j en la función de onda

$$\hat{P}_{ik} \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) = \Psi(q_1, \dots, q_k, \dots, q_i, \dots, q_N) \quad (6.56)$$

Si el Hamiltoniano es invariante ante el cambio en la enumeración de todas las partículas, se cumple que

$$[\hat{H}, \hat{P}_{ik}] = 0 \quad \forall \quad i, k = 1, \dots, N \quad \text{con } i \neq k$$

Las autofunciones de \hat{P}_{ik} tienen que satisfacer

$$\begin{aligned} \hat{P}_{ik} \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) &= \lambda \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) \\ \Psi(q_1, \dots, q_k, \dots, q_i, \dots, q_N) &= \lambda \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) \end{aligned} \quad (6.57)$$

donde λ es un autovalor de \hat{P}_{ik} . Si volvemos a aplicar el operador \hat{P}_{ik} en la Ec.(6.57) encontramos

$$\hat{P}_{ik}^2 \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) = \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N) = \lambda^2 \Psi(q_1, \dots, q_i, \dots, q_k, \dots, q_N)$$

esto es, los autovalores de \hat{P}_{ik} son tales que $\lambda^2 = 1$; por lo tanto, solo hay dos posibilidades: $\lambda = +1$, con lo cual la autofunción es *simétrica* respecto al intercambio de las coordenadas i y j , o $\lambda = -1$ en cuyo caso la función de onda es *antisimétrica*.

Si el Hamiltoniano conmuta con todos los operadores de permutación, entonces sus autofunciones pueden ser construidas de manera que sean totalmente simétricas (esto es, simétricas ante cualquier permutación de coordenadas) o totalmente antisimétricas. Una generalización del operador permutación de pares es el operador \hat{P} , que genera una permutación arbitraria de índices

$$\hat{P} \Psi(q_1, q_2, \dots, q_N) = \Psi(q_{P_1}, q_{P_2}, \dots, q_{P_N}) \quad (6.58)$$

donde P_1, \dots, P_N es una permutación de los números $1, \dots, N$ (en realidad, tenemos un operador diferente para cada permutación). Cualquiera de dichas permutaciones se obtiene mediante una secuencia de permutaciones de pares de índices, de manera que el operador \hat{P} resulta de la aplicación sucesiva de un cierto conjunto de operadores \hat{P}_{ik} . Por lo tanto, las autofunciones de \hat{P} también son las funciones simétricas o antisimétricas. Si el número de permutaciones de pares necesarias para construir P_1, \dots, P_N es par (impar) diremos que la permutación es par (impar).

Si uno comienza con una autofunción del Hamiltoniano arbitraria $\Psi(q_1, \dots, q_N)$ sin ninguna simetría de paridad, uno puede construir autofunciones completamente simétricas o antisimétricas vía

$$\Psi^S(q_1, \dots, q_N) = B_S \sum_P \hat{P} \Psi(q_1, \dots, q_N) \quad (6.59)$$

$$\Psi^A(q_1, \dots, q_N) = B_A \sum_P (-1)^P \hat{P} \Psi(q_1, \dots, q_N) \quad (6.60)$$

donde B_S, B_A son constantes de normalización y las sumas corren sobre todas las posibles permutaciones de los índices $1, \dots, N$. El signo $(-1)^P$ es $+1$ si la permutación es par y -1 si la permutación es impar. Por ejemplo, para una función de onda de tres partículas tenemos

$$\begin{aligned} \Psi^S(q_1, q_2, q_3) &= B_S [\Psi(q_1, q_2, q_3) + \Psi(q_2, q_1, q_3) + \Psi(q_1, q_3, q_2) + \Psi(q_3, q_2, q_1) \\ &\quad + \Psi(q_2, q_3, q_1) + \Psi(q_3, q_1, q_2)] \\ \Psi^A(q_1, q_2, q_3) &= B_A [\Psi(q_1, q_2, q_3) - \Psi(q_2, q_1, q_3) - \Psi(q_1, q_3, q_2) - \Psi(q_3, q_2, q_1) \\ &\quad + \Psi(q_2, q_3, q_1) + \Psi(q_3, q_1, q_2)] \end{aligned}$$

Es un hecho experimental que los sistemas de partículas en la naturaleza solo pueden ser descritos por funciones de onda de simetría bien definida. Mas aún, se observa que partículas de una misma especie (electrones, muones, fotones, quarks, etc) solo pueden presentar un único tipo de simetría ante intercambio. En otras palabras, solo existen dos tipos de partículas. Aquellas descritas por funciones de onda *simétricas* se denominan **bosones** en honor al físico indú Satyendranath Bose, quien en 1924 propuso esta propiedad para los fotones. Poco tiempo despues Albert Einstein aplico el mismo concepto a un gas ideal monoatómico. Partículas descritas por funciones de onda *antisimétricas* se denominan **fermiones** en honor al físico italiano Enrico Fermi.

Se observa también que estas propiedades están directamente relacionadas al *spin* de las partículas. Partículas con spin semientero son siempre fermiones (ej., electrones), mientras que partículas con spin entero (ej., fotones) son siempre bosones. La relación entre el spin y la simetría de intercambio de las partículas es un hecho aún no bien comprendido y objeto de debate, de manera que lo asumiremos como un postulado.

Estas propiedades tienen una profunda influencia en la descripción mecánico-estadística ya que los microestados accesibles del sistema deben hallarse exclusivamente en subespacios de estados simétricos o antisimétricos, dependiendo del tipo de partícula. Como veremos, esto lleva a comportamientos radicalmente diferentes. Se dice entonces que los bosones obedecen la **estadística de Bose-Einstein**, mientras que los fermiones obedecen la **estadística de Fermi-Dirac**.

6.6.2. Función de onda de muchas partículas

Para avanzar mas en la teoría debemos entonces elegir una base de autoestados del Hamiltoniano, sobre la cual podamos construir conjuntos de estados simétricos o antisimétricos. Esto resulta particularmente sencillo en el caso de partículas no interactuantes, esto es, cuando el Hamiltoniano puede expresarse como la suma de operadores de una partícula:

$$\hat{H}(\hat{q}_1, \dots, \hat{q}_N, \hat{p}_1, \dots, \hat{p}_N) = \sum_{i=1}^N \hat{h}(\hat{q}_i, \hat{p}_i)$$

Si tenemos resuelto el problema de autovalores de una partícula

$$\hat{h}\phi_k(q) = \epsilon_k \phi_k(q)$$

donde k denota un conjunto completo de números cuánticos, podemos construir un autoestado de \hat{H} como

$$\Psi_{k_1, \dots, k_N}^E(q_1, \dots, q_N) = \prod_{i=1}^N \phi_{k_i}(q_i) \quad (6.61)$$

correspondiente a un autovalor de \hat{H}

$$E = \sum_{i=1}^N \epsilon_{k_i}$$

La función de onda (6.61) puede ser expresada mas claramente en la notación de Dirac. Sea $|k\rangle$ el vector de estado de una partícula correspondiente al autovalor k , esto es, $\phi_{k_i}(q_i) = \langle q_i | k_i \rangle$. El vector de estado de N partículas es el producto directo de los vectores de estado de una partícula

$$|k_1, \dots, k_N\rangle = |k_1\rangle |k_2\rangle \dots |k_N\rangle \quad (6.62)$$

La Ec.(6.62) significa que la partícula 1 se encuentra en el estado k_1 , la partícula 2 en el estado k_2 , etc. El vector hermitiano conjugado (“bra”) resulta

$$\langle k_1, \dots, k_N | = \langle k_N | \langle k_{N-1} | \dots \langle k_1 | \quad (6.63)$$

Si los estados de una partícula esta ortonormalizados $\langle k | k' \rangle = \delta_{k,k'}$ y forman un conjunto completo

$$\sum_k |k\rangle \langle k| = 1$$

las mismas propiedades se cumplen para los autoestados (6.62) en el espacio suma directa:

$$\begin{aligned} \langle k'_1, \dots, k'_N | k_1, \dots, k_N \rangle &= \langle k'_N | \langle k'_{N-1} | \dots \langle k'_1 | k_1 \rangle | k_2 \rangle \dots | k_N \rangle \\ &= \langle k'_1 | k_1 \rangle \dots \langle k'_N | k_N \rangle \\ &= \delta_{k_1, k'_1} \dots \delta_{k_N, k'_N} \end{aligned} \quad (6.64)$$

$$\sum_{k_1, \dots, k_N} |k_1, \dots, k_N\rangle \langle k_1, \dots, k_N| = 1 \quad (6.65)$$

En ese caso, cualquier función de onda arbitraria (aún para partículas interactuantes) puede ser desarrollada en términos de los estados $|k_1, \dots, k_N\rangle$. En la representación coordinada, la función de onda resulta

$$\Psi_{k_1, \dots, k_N}^E(q_1, \dots, q_N) = \langle q_1, \dots, q_N | k_1, \dots, k_N \rangle = \phi_{k_1}(q_1) \phi_{k_2}(q_2) \dots \phi_{k_N}(q_N) \quad (6.66)$$

Esta función de onda no tiene simetría de paridad definida, ya que el intercambio de dos coordenadas (o equivalentemente de dos números cuánticos) lleva a un nuevo autoestado completamente diferente. Autoestados simétricos y antisimétricos pueden construirse a partir de las Ecs.(6.59)-(6.60):

$$\Psi_{k_1, \dots, k_N}^{E,S}(q_1, \dots, q_N) = B_S \sum_P \hat{P} \phi_{k_1}(q_1) \dots \phi_{k_N}(q_N) \quad (6.67)$$

$$\Psi_{k_1, \dots, k_N}^{E,A}(q_1, \dots, q_N) = B_A \sum_P (-1)^P \hat{P} \phi_{k_1}(q_1) \dots \phi_{k_N}(q_N) \quad (6.68)$$

Aquí las sumas corren sobre todas las permutaciones P_1, \dots, P_N de los argumentos q_1, \dots, q_N de las autofunciones de una partícula $\phi_{k_n}(q_n)$. Sin embargo, es claro que dá exactamente lo mismo permutar los índices de las coordenadas o los índices de los números cuánticos. Podemos entonces expresar

$$|k_1, \dots, k_N\rangle^S = B_S \sum_P \hat{P} |k_1, \dots, k_N\rangle \quad (6.69)$$

$$|k_1, \dots, k_N\rangle^A = B_A \sum_P (-1)^P \hat{P} |k_1, \dots, k_N\rangle \quad (6.70)$$

donde ahora las sumas corren sobre permutaciones de los números cuánticos.

La función de onda antisimétrica puede expresarse como el determinante de una matriz

$$\Psi_{k_1, \dots, k_N}^{E,A}(q_1, \dots, q_N) = B_A \det \begin{pmatrix} \phi_{k_1}(q_1) & \dots & \phi_{k_1}(q_N) \\ \vdots & & \vdots \\ \phi_{k_N}(q_1) & \dots & \phi_{k_N}(q_N) \end{pmatrix} \quad (6.71)$$

(recordemos la propiedad de los determinantes de cambiar de signo ante una permutación de dos cualesquiera de sus filas o columnas, lo cual en este caso corresponde a permutar dos coordenadas o dos números cuánticos). De aquí resulta inmediato que la función de onda antisimétrica se anula si dos partículas están en el mismo estado cuántico, ya que en ese caso la matriz tiene dos filas idénticas. Esto es precisamente el enunciado del **Principio de Exclusión de Pauli**: dos fermiones idénticos no pueden ocupar simultáneamente el mismo estado cuántico. La probabilidad de que esto último ocurra (dada por el módulo al cuadrado de la función de onda) es cero.

Las funciones de onda antisimétricas por lo tanto resultan la combinación de productos de todas las posibles permutaciones de N índices diferentes y la normalización resulta

$${}^A \langle k_1, \dots, k_N | k_1, \dots, k_N \rangle^A = B_A^2 N! = 1$$

Así, arribamos a la expresión

$$\Psi_{k_1, \dots, k_N}^{E,A}(q_1, \dots, q_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \phi_{k_1}(q_1) & \cdots & \phi_{k_1}(q_N) \\ \vdots & & \vdots \\ \phi_{k_N}(q_1) & \cdots & \phi_{k_N}(q_N) \end{pmatrix} \quad (6.72)$$

conocida como **determinante de Slater**.

En el caso de las funciones de onda simétricas los números cuánticos pueden repetirse. Esto es, un número arbitrario de bosones pueden ocupar un mismo estado de una partícula. La normalización en este caso viene dada por

$$\begin{aligned} {}^S \langle k_1, \dots, k_N | k_1, \dots, k_N \rangle^S &= B_S^2 \sum_P \sum_{P'} \langle k_{P'_1}, \dots, k_{P'_N} | k_{P_1}, \dots, k_{P_N} \rangle \\ &= B_S^2 N! \sum_P \langle k_1, \dots, k_N | k_{P_1}, \dots, k_{P_N} \rangle \\ &= B_S^2 N! \sum_P \delta_{k_1, k_{P_1}} \cdots \delta_{k_N, k_{P_N}} = 1 \end{aligned} \quad (6.73)$$

Dado que ahora puede haber números cuánticos repetidos, el número de términos no nulos en la última suma va a depender de cuantos índices se repitan. Podemos describir esto de manera general, introduciendo los números de ocupación $n_k = 0, 1, \dots, N$ con $\sum_k n_k = N$, los cuales nos dicen cuantas partículas tenemos en cada estado individual para una dada función de onda simétrica. Así, resulta

$$B_S^2 N! \prod_k n_k! = 1 \quad (6.74)$$

Es posible demostrar (queda como ejercicio para el lector) que las autofunciones

$$|k_1, \dots, k_N\rangle^S = \frac{1}{(N! \prod_k n_k!)^{1/2}} \sum_P \hat{P} |k_1, \dots, k_N\rangle \quad (6.75)$$

$$|k_1, \dots, k_N\rangle^A = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \hat{P} |k_1, \dots, k_N\rangle \quad (6.76)$$

forman un conjunto ortonormal completo en los subespacios de funciones simétricas y antisimétricas respectivamente. En otras palabras, estas funciones constituyen bases de dichos subespacios y por lo tanto cualquier función de onda arbitraria de N bosones o fermiones (aún en el caso de partículas interactuantes) puede ser desarrollada en términos de las funciones (6.75) o (6.76) respectivamente. Sin embargo, es importante tener en claro como se indexan los diferentes estados en cada una de

eses bases. Cada conjunto de índices k_1, \dots, k_N designa una *única* función de base, no importando el orden de los mismos: una permutación de índices solo produce un cambio de signo en el caso antisimétrico y deja totalmente inalterada la función en el caso simétrico. En el caso de las funciones fermiónicas, los índices nunca se repiten. Entonces lo que caracteriza una de estas funciones es cuantas partículas tenemos en cada estado. Esto puede ser descrito de manera mas clara a través de los números de ocupación n_k . La especificación de los números de ocupación para todos los estados posibles k de una partícula, con la restricción $\sum_k n_k = N$ determina de manera unívoca un autoestado simetrizado. En el caso de fermiones, tenemos la restricción adicional de que los números de ocupación solo pueden adoptar los valores $n_k = 0, 1$.

La (anti)simetrización de los estados tiene también consecuencias para las posibles cantidades observables del sistema. Para partículas indistinguibles no tiene sentido calcular valores de expectación cuánticos de observables que de alguna manera identifiquen individualmente las partículas. Por ejemplo, ya no será posible calcular la densidad de probabilidad de encontrar la partícula Nro.2 en la posición q . Solo podemos calcular la probabilidad de encontrar *alguna* de las N partículas en q . Esto significa que *todos los observables \hat{O} de un sistema de partículas indistinguibles tienen que ser invariantes ante permutaciones de las partículas:*

$$[\hat{O}, \hat{P}] = 0$$

Esto vale en particular para el operador densidad de probabilidad $\hat{\rho}$.

6.6.3. Función partición de partículas indistinguibles

La evaluación de la función partición $Z_N = \text{Tr} \exp(-\beta \hat{H})$ para partículas indistinguibles debe entonces realizarse en el subespacio de las funciones simétricas (6.75), en el caso de bosones, o antisimétricas (6.76) en el caso de fermiones. Para ello podemos usar las bases de autoestados (6.75) o (6.76).

Para evaluar la traza debemos sumar todos los elementos de matriz diagonales de $\hat{\rho}$ en dichos subespacios. Una forma de enumerar todos los posibles estados en un subespacio es sumar sobre todos los valores posibles de estados de una partícula k_1, \dots, k_N . No obstante, dado que $\hat{\rho}$ es invariante ante el intercambio de partículas, al hacer el cálculo de este modo se repetirán términos que corresponden a un único estado (anti)simetrizado. Debemos entonces descontar esta degeneración. Para el caso de N bosones, para cada conjunto particular de índices k_1, \dots, k_N esta repetición ocurrirá $N!/n_{k_1}!n_{k_2}!\dots$ veces. Tenemos entonces:

$$\begin{aligned} \text{Tr} e^{-\beta \hat{H}} &= \frac{1}{N!} \sum_{k_1} \dots \sum_{k_N} \left(\prod_{k'} n_{k'}! \right)^S \langle k_1, \dots, k_N | e^{-\beta \hat{H}} | k_1, \dots, k_N \rangle^S \\ &= \left(\frac{1}{N!} \right)^2 \sum_{k_1} \dots \sum_{k_N}^{(+)} \langle k_1, \dots, k_N | e^{-\beta \hat{H}} | k_1, \dots, k_N \rangle^{(+)} \\ &= \frac{1}{N!} \sum_{k_1} \dots \sum_{k_N} \langle k_1, \dots, k_N | e^{-\beta \hat{H}} | k_1, \dots, k_N \rangle^{(+)} \end{aligned} \quad (6.77)$$

donde el “bra” en la última expresión es un estado no simetrizado y

$$|k_1, \dots, k_N\rangle^{(+)} \equiv \sum_P \hat{P} |k_1, \dots, k_N\rangle$$

es un estado simetrizado no-normalizado. En forma semejante, para el caso de N fermiones tenemos:

$$\begin{aligned}
\text{Tr } e^{-\beta\hat{H}} &= \frac{1}{N!} \sum_{k_1} \cdots \sum_{k_N}^A \langle k_1, \dots, k_N | e^{-\beta\hat{H}} | k_1, \dots, k_N \rangle^A \\
&= \frac{1}{N!} \sum_{k_1} \cdots \sum_{k_N} \langle k_1, \dots, k_N | e^{-\beta\hat{H}} | k_1, \dots, k_N \rangle^{(-)} \quad (6.78)
\end{aligned}$$

donde

$$|k_1, \dots, k_N \rangle^{(-)} \equiv \sum_P (-1)^P \hat{P} |k_1, \dots, k_N \rangle$$

Consideremos el ejemplo no realista de la función partición $Z_3(T)$ de tres partículas indistinguibles (bosones o fermiones), para el caso de un gas ideal en una caja de volumen $V = L^3$ con paredes perfectamente reflectantes. Como vimos en la sección 6.2, los autoestados de una partícula $\langle \vec{r} | \vec{k} \rangle$ vienen dados por la Ec.(6.22) con

$$\vec{k} = \frac{n_x \pi}{L} \mathbf{i} + \frac{n_y \pi}{L} \mathbf{j} + \frac{n_z \pi}{L} \mathbf{k}$$

y con n_x, n_y, n_z enteros positivos. El Hamiltoniano de 3 partículas viene dado por $\hat{H}_3 = \hat{p}_1^2/2m + \hat{p}_2^2/2m + \hat{p}_3^2/2m$ y de las Ecs.(6.77)-(6.78) tenemos que

$$Z_3(T) = \text{Tr}_3 \left(e^{-\beta\hat{H}_3} \right) = \frac{1}{3!} \sum_{\vec{k}_1} \sum_{\vec{k}_2} \sum_{\vec{k}_3} \langle \vec{k}_1, \vec{k}_2, \vec{k}_3 | e^{-\beta\hat{H}_3} | \vec{k}_1, \vec{k}_2, \vec{k}_3 \rangle^{(\pm)} \quad (6.79)$$

donde

$$| \vec{k}_1, \vec{k}_2, \vec{k}_3 \rangle^{(\pm)} = | \vec{k}_1, \vec{k}_2, \vec{k}_3 \rangle \pm | \vec{k}_1, \vec{k}_3, \vec{k}_2 \rangle \pm | \vec{k}_2, \vec{k}_1, \vec{k}_3 \rangle \pm | \vec{k}_3, \vec{k}_2, \vec{k}_1 \rangle + | \vec{k}_3, \vec{k}_1, \vec{k}_2 \rangle + | \vec{k}_2, \vec{k}_3, \vec{k}_1 \rangle \quad (6.80)$$

El primer término de esta ecuación contribuye a $Z_3(T)$ con

$$\sum_{\vec{k}_1} \sum_{\vec{k}_2} \sum_{\vec{k}_3} \langle \vec{k}_1, \vec{k}_2, \vec{k}_3 | e^{-\beta\hat{H}_3} | \vec{k}_1, \vec{k}_2, \vec{k}_3 \rangle = \left(\sum_{\vec{k}} \langle \vec{k} | e^{-\beta\hat{p}^2/2m} | \vec{k} \rangle \right)^3 = \left[\sum_{\vec{k}} e^{-\beta E(\vec{k})} \right]^3 = [Z_1(T)]^3,$$

el segundo término con

$$\begin{aligned}
\sum_{\vec{k}_1} \sum_{\vec{k}_2} \sum_{\vec{k}_3} \langle \vec{k}_1, \vec{k}_2, \vec{k}_3 | e^{-\beta\hat{H}_3} | \vec{k}_1, \vec{k}_3, \vec{k}_2 \rangle &= \left(\sum_{\vec{k}_1} e^{-\beta E(\vec{k}_1)} \right) \left(\sum_{\vec{k}_2} \sum_{\vec{k}_3} e^{-\beta(E(\vec{k}_2)+E(\vec{k}_3))} \delta_{\vec{k}_2, \vec{k}_3} \right) \\
&= \left(\sum_{\vec{k}_1} e^{-\beta E(\vec{k}_1)} \right) \left(\sum_{\vec{k}} e^{-2\beta E(\vec{k})} \right) \\
&= Z_1(T) Z_1 \left(\frac{T}{2} \right)
\end{aligned}$$

y así sucesivamente. Tenemos entonces

$$Z_3(T) = \frac{1}{3!} \left[(Z_1(T))^3 \pm 3 Z_1(T) Z_1 \left(\frac{T}{2} \right) + 2 Z_1 \left(\frac{T}{3} \right) \right] \quad (6.81)$$

Si recordamos que $Z_1(T) = V/\lambda_T^3$ con

$$\lambda_T = \frac{h}{\sqrt{2\pi m k_B T}}$$

resulta

$$\begin{aligned} Z_3(T) &= \frac{1}{3!} \left[\left(\frac{V}{\lambda_T^3} \right)^3 \pm \frac{3}{2^{3/2}} \left(\frac{V}{\lambda_T^3} \right)^2 + \frac{2}{3^{3/2}} \left(\frac{V}{\lambda_T^3} \right) \right] \\ &= \frac{1}{3!} \left(\frac{V}{\lambda_T^3} \right)^3 \left[1 \pm \frac{3}{2^{3/2}} \left(\frac{\lambda_T^3}{V} \right) + \frac{2}{3^{3/2}} \left(\frac{\lambda_T^3}{V} \right)^2 \right] \end{aligned} \quad (6.82)$$

El límite clásico corresponde a altas temperaturas y/o bajas densidades (V grande) $\lambda_T^3/V \ll 1$, en el cual obtenemos

$$Z_3(T) \sim \frac{1}{3!} \left(\frac{V}{\lambda_T^3} \right)^3 + \mathcal{O} \left(\frac{3 \lambda_T^3}{V} \right)$$

Para un gas de N partículas idénticas en el límite clásico tenemos

$$Z_N(T) \sim \frac{1}{N!} (Z_1(T))^N + \mathcal{O} \left(\frac{N \lambda_T^3}{V} \right) \quad (6.83)$$

Z_1^N es la función partición para un conjunto de N partículas distinguibles. En el límite clásico, los términos permutados son despreciables y el único remanente de la indistinguibilidad es el factor $1/N!$, que es precisamente el conteo correcto de Boltzmann. En este límite no existe diferencia entre el comportamiento de bosones y fermiones. A bajas temperaturas y/o altas densidades, los términos permutados cobran relevancia y ambos tipos de partículas exhiben comportamientos diferentes. Desafortunadamente, la evaluación de Z_N por el método anterior es sumamente complicada. En el capítulo siguiente veremos que este cálculo puede realizarse más fácilmente utilizando otro ensemble diferente.

Ciertos efectos cuánticos aún pueden tomarse en cuenta en una descripción semiclásica de un gas de partículas con grados internos de libertad. Este sería el caso de un gas de moléculas no interactuantes que poseen grados internos de libertad que pueden absorber energía, tales como movimientos rotacionales y vibracionales de la molécula, grados de libertad electrónicos, nucleares, etc. Supongamos por simplicidad que tomamos en cuenta solo los movimientos rotacionales y vibracionales. Para tal sistema el Hamiltoniano de la molécula i -ésima puede escribirse como $\hat{H}_i = \hat{p}_i^2/2m + \hat{H}_{i(rot)} + \hat{H}_{i(vib)}$, donde estamos asumiendo que los diferentes grados de libertad están desacoplados entre sí. La función partición viene dada por

$$Z_N(T) = \text{Tr}_N \exp \left[-\beta \sum_{i=1}^N \left(\frac{\hat{p}_i^2}{2m} + \hat{H}_{i(rot)} + \hat{H}_{i(vib)} \right) \right]$$

Si los diferentes operadores que aparecen en esta expresión *conmutan todos entre sí*, en el límite semiclásico, resulta

$$Z_N(T) \approx \frac{1}{N!} \left(Z_{1(tr)} \times Z_{1(rot)} \times Z_{1(vib)} \right)^N$$

donde $Z_{1(tr)} = \text{Tr}_1 \left(e^{-\beta \hat{p}^2/2m} \right)$, $Z_{1(rot)} = \text{Tr}_1 \left(e^{-\beta \hat{H}_{(rot)}} \right)$ y $Z_{1(vib)} = \text{Tr}_1 \left(e^{-\beta \hat{H}_{(vib)}} \right)$. La operación Tr_1 debe ser llevada a cabo sobre un conjunto completo de estados de una partícula apropiado para el operador asociado.

6.7. Fluidos clásicos no ideales

Fluidos normales son bien descriptos por la estadística clásica, ya que en general a las temperaturas y densidades a las cuales los efectos cuánticos se vuelven relevantes los mismos solidifican.

El tipo de interacción mas común entre partículas involucra solamente las coordenadas, de manera que el Hamiltoniano puede expresarse como la suma de un término de energía cinética y un potencial de interacción $U_p(q_1, \dots, q_N)$. La función partición, por lo tanto, puede factorizarse como

$$Z_N(T) = Z_{GI} Q V^{-N}$$

donde Z_{GI} es la función partición del gas ideal (la cual incluye el conteaje correcto de Boltzmann) y

$$Q = \int dq_1 \cdots dq_N e^{-\beta U_p(q_1, \dots, q_N)} \quad (6.84)$$

es llamada a veces función partición configuracional, donde las integrales son sobre el volumen V . El valor medio de cualquier función que depende solamente de las coordenadas de las partículas resulta entonces

$$\langle f(q_1, \dots, q_N) \rangle = \frac{1}{Q} \int dq_1 \cdots dq_N f(q_1, \dots, q_N) e^{-\beta U_p(q_1, \dots, q_N)} \quad (6.85)$$

Al analizar las propiedades de un fluido debemos tomar en cuenta las simetrías del sistema. En primer lugar, el potencial $U_p(q_1, \dots, q_N)$ debe ser simétrico ante permutaciones de partículas, ya que, aunque distinguibles en la aproximación clásica, son todas idénticas. En segundo lugar, en una fase fluida (gas ó liquido) a diferencia de lo que ocurre en un sólido, tenemos invariancia traslacional ya que no existen posiciones privilegiadas. Esto implica que el potencial de interacción debe ser invariante ante un desplazamiento constante de las coordenadas de todas las partículas y por lo tanto debe depender de las distancias relativas entre las partículas y no de sus posiciones absolutas². Además de ser homogéneo (invariancia traslacional) un fluido es isotrópico y por lo tanto el potencial debe ser invariante ante rotaciones del sistema de coordenadas.

Asumiendo las partículas puntuales, la densidad en un volumen diferencial centrado en r puede escribirse como³

$$\rho_p(r) = \sum_{i=1}^N \langle \delta(r - q_i) \rangle = \frac{1}{Q} \int dq_1 \cdots dq_N \sum_{i=1}^N \delta(r - q_i) e^{-\beta U_p(q_1, \dots, q_N)} \quad (6.86)$$

Debido a la invariancia ante permutaciones de partículas, cada uno de los términos en la suma de la ecuación anterior debe ser igual a los demás. Por lo tanto, podemos escribir

$$\rho_p(r) = \frac{N}{Q} \int dq_2 \cdots dq_N \int dq_1 \delta(r - q_1) e^{-\beta U_p(q_1, \dots, q_N)}$$

Además, dado que U_p depende de las distancias relativas $|q_i - q_j|$, podemos hacer un cambio de variables $q'_i = q_i - q_1$ para $i = 2, \dots, N$, con lo cual

$$\rho_p(r) = \frac{N}{Q} \int dq'_2 \cdots dq'_N e^{-\beta U(q'_2, \dots, q'_N)} \int dq_1 \delta(r - q_1) = \frac{N}{Q} \int dq'_2 \cdots dq'_N e^{-\beta U_p(q'_2, \dots, q'_N)} \quad (6.87)$$

En forma semejante

²Estamos considerando un volumen macroscópico y por lo tanto despreciamos los efectos de borde en el límite termodinámico.

³La coordenada r aquí representa una variable macroscópica, es decir, que varía en cantidades mucho mayores que las longitudes microscópicas características.

$$Q = \int dq'_2 \dots dq'_N e^{-\beta U_p(q'_2, \dots, q'_N)} \int dq_1 = V \int dq'_2 \dots dq'_N e^{-\beta U_p(q'_2, \dots, q'_N)} \quad (6.88)$$

de donde $\rho_p(r) = N/V$ para todo r dentro del volúmen V . La densidad de partículas en un fluido es siempre constante en cualquier punto, independientemente de que las partículas interactúen (en una fase sólida cristalina esta invariancia se rompe, ya que las partículas se localizan en posiciones fijas en el espacio).

Para un gas ideal las posiciones de todas las partículas son totalmente independientes entre sí. En un fluido real las interacciones inducen correlaciones entre las posiciones de las diferentes partículas. Una manera de describir estas correlaciones es a través de la **función de distribución de pares** $g(r)$, que se define como sigue. Sea $\mathbf{r}_{ij} \equiv q_i - q_j$ el vector distancia relativa entre las partículas i y j . Entonces:

$$g(\mathbf{r}) \equiv \frac{2V}{N(N-1)} \sum_{(i,j)} \langle \delta(\mathbf{r} - \mathbf{r}_{ij}) \rangle \quad (6.89)$$

donde la notación (i, j) indica que cada par se cuenta una sola vez. La función $g(\mathbf{r})$ nos da el número medio de pares de partículas ubicadas a una distancia \mathbf{r} . En primer lugar notemos que, por la simetría ante permutaciones de partículas del potencial, los promedios en la Ec.(6.89) son todos iguales. Además, por la isotropía del potencial, g no puede depender de la orientación del vector \mathbf{r} y solo puede depender de su módulo $r = |\mathbf{r}|$. Así

$$g(r) = V \langle \delta(\mathbf{r} - \mathbf{r}_{12}) \rangle \quad (6.90)$$

donde la elección del par 1-2 es arbitraria. En el caso de un gas ideal el potencial $U = 0$. Un cálculo directo da que $\langle \delta(\mathbf{r} - \mathbf{r}_{12}) \rangle = 1/V$ y $g(r) = 1$. Esto significa que todas las distancias entre partículas son equiprobables. Para el caso interactuante tenemos

$$\begin{aligned} g(r) &= \frac{V}{Q} \int dq'_2 \delta(\mathbf{r} - q'_2) \int dq'_3 \dots dq'_N e^{-\beta U_p(q'_2, \dots, q'_N)} \int dq_1 \\ &= \frac{V^2}{Q} \int dq'_3 \dots dq'_N e^{-\beta U_p(r, q'_3, \dots, q'_N)} \end{aligned} \quad (6.91)$$

El caso mas frecuente es aquel en el cual la energía potencial es una suma de potenciales de pares, los cuales dependen solamente de la distancia entre partículas, esto es

$$U_p = \sum_{(i,j)} u(r_{ij}) \quad (6.92)$$

Tomando en cuenta la simetría ante permutaciones de partículas del potencial, tenemos entonces

$$\begin{aligned} \langle U_p \rangle &= \frac{N(N-1)}{2} \langle u(r_{12}) \rangle \\ &= \frac{N(N-1)}{2Q} \int dq'_2 u(q'_2) \int dq'_3 \dots dq'_N e^{-\beta U_p(q'_2, \dots, q'_N)} \int dq_1 \\ &= \frac{V N(N-1)}{2Q} \int d^3r u(r) \int dq'_3 \dots dq'_N e^{-\beta U_p(r, \dots, q'_N)} \end{aligned} \quad (6.93)$$

Usando la Ec.(6.91) y tomando en cuenta la energía cinética obtenemos finalmente la ecuación de estado

$$U = \langle H \rangle = \frac{3}{2} N k_B T + \frac{N(N-1)}{2V} \int d^3r u(r) g(r) \quad (6.94)$$

donde la integral es sobre todo el volúmen V (no debemos olvidar que $g(r)$ depende también de T). La ecuación de estado para la presión también puede expresarse en función de $g(r)$:

$$P = -\frac{\partial F}{\partial V} = k_B T \frac{\partial \ln Z_N}{\partial V} \quad (6.95)$$

$$= k_B T \frac{\partial}{\partial V} \ln \left(\frac{Z_{GI} Q}{V^N} \right) = k_B T \rho_p + k_B T \frac{V^N}{Q} \frac{\partial}{\partial V} \left(\frac{Q}{V^N} \right) \quad (6.96)$$

Para evaluar la última derivada consideremos que el sistema se encuentra en una caja cúbica de volúmen $V = L^3$. Podemos entonces calcular Q integrando en coordenadas cartesianas, donde cada una de las $3N$ coordenadas se integra entre 0 y L . Llamemos entonces \vec{q} al vector que contiene las $3N$ coordenadas. Si hacemos el cambio de variable $\vec{x} = \vec{q}/L$ obtenemos

$$Q = V^N \int_0^1 dx_1 \cdots \int_0^1 dx_{3N} e^{-\beta U_p(L\vec{x})}$$

y así

$$\frac{\partial}{\partial V} \left(\frac{Q}{V^N} \right) = \frac{1}{3V^{2/3}} \frac{\partial}{\partial L} \left(\frac{Q}{V^N} \right) = -\frac{\beta}{3V^{2/3}} \int_0^1 dx_1 \cdots \int_0^1 dx_{3N} \frac{\partial U_p(L\vec{x})}{\partial L} e^{-\beta U_p(L\vec{x})}$$

$$\frac{\partial U_p(L\vec{x})}{\partial L} = \sum_{(i,j)} \frac{du(r_{ij})}{dr} (r_{ij}/L)$$

Transformando nuevamente a las variables q y usando la simetría ante permutaciones de U_p obtenemos

$$\begin{aligned} \frac{\partial}{\partial V} \left(\frac{Q}{V^N} \right) &= -\beta \frac{N(N-1)}{6V^{N+1}} \int dq_1 \cdots \int dq_N r_{12} \frac{du(r_{12})}{dr} e^{-\beta U_p(q_1, \dots, q_N)} \\ &= -\beta \frac{N(N-1)}{6V^N} \int d^3r r \frac{du(r)}{dr} \int dq'_3 \cdots \int dq'_N e^{-\beta U_p(r, q'_3, \dots, q'_N)} \\ &= -\beta \frac{N(N-1)Q}{6V^{N+2}} \int d^3r r \frac{du(r)}{dr} g(r) \end{aligned} \quad (6.97)$$

Reemplazando en la Ec.(6.96), en el límite termodinámico obtenemos finalmente la ecuación de estado para un fluido normal (con interacciones de pares)

$$P = k_B T \rho_p \left[1 - \frac{2\pi\rho_p}{3k_B T} \int_0^\infty r^3 \frac{du(r)}{dr} g(r) dr \right] \quad (6.98)$$

la cual se conoce como ecuación de estado del virial. Vemos que las ecuaciones de estado quedan completamente determinadas si conocemos el potencial de interacción y la función $g(r)$.

La función $g(r)$ puede obtenerse mediante experimentos de scattering. En los mismos un haz de partículas de prueba (rayos X, neutrones o electrones) con momento $\hbar\vec{k}'$ se hace incidir sobre el material (podemos representarlas por ondas planas). Las partículas incidentes interactúan con las moléculas del material, las cuales actúan como centros dispersores. Así, el número de partículas dispersadas con momento $\hbar\vec{k}''$ depende de la distribución de las moléculas en el material y es descrita por la sección diferencial de scattering, la cual en una primera aproximación resulta proporcional a

$$I(\vec{k}) = \left\langle \left| \sum_{j=1}^N e^{i\vec{k} \cdot \vec{q}_j} \right|^2 \right\rangle \quad (6.99)$$

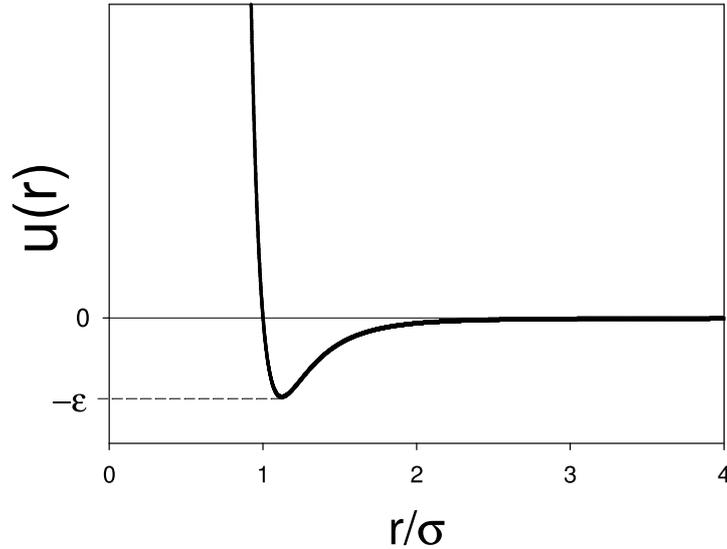


Figura 6.1: Potencial de Lennard-Jones.

donde $\vec{k} = \vec{k}'' - \vec{k}'$ y el promedio se toma en ensemble. $I(\vec{k})$ se conoce como la función de estructura. Tenemos

$$I(\vec{k}) = \left\langle \sum_{j=1}^N \sum_{l=1}^N e^{i\vec{k} \cdot (\vec{q}_j - \vec{q}_l)} \right\rangle = N + \left\langle \sum_{j \neq l} e^{i\vec{k} \cdot (\vec{q}_j - \vec{q}_l)} \right\rangle \quad (6.100)$$

Dado que la cantidad a promediar solo depende de las coordenadas, tenemos

$$\begin{aligned} I(\vec{k}) &= N + \frac{1}{Q} \sum_{j \neq l} \int dq_1 \cdots dq_N e^{i\vec{k} \cdot (\vec{q}_j - \vec{q}_l)} e^{-\beta U(q_1, \dots, q_N)} \\ &= N + \frac{N(N-1)}{Q} \int dq_1 \cdots dq_N e^{i\vec{k} \cdot (\vec{q}_2 - \vec{q}_1)} e^{-\beta U(q_1, \dots, q_N)} \\ &= N + \frac{VN(N-1)}{Q} \int dq'_2 \cdots dq'_N e^{i\vec{k} \cdot \vec{q}'_2} e^{-\beta U(q'_2, \dots, q'_N)} \end{aligned} \quad (6.101)$$

Comparando con la Ec.(6.91) obtenemos

$$I(\vec{k}) = N + \frac{N(N-1)}{V} \int d^3r e^{i\vec{k} \cdot \vec{r}} g(r) \quad (6.102)$$

esto es, la función de estructura está directamente relacionada a la transformada de Fourier de la función de distribución de pares. Esto permite tanto obtener ecuaciones de estado de fluidos de manera empírica, como contrastar predicciones teóricas con resultados experimentales.

Para la mayoría de los fluidos normales el potencial de interacción de pares es repulsivo a muy cortas distancias (interacción de carozo duro) y atractivo a distancias cortas. Un ejemplo típico es el potencial semi-empírico de Lennard-Jones

$$u(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (6.103)$$

el cual se muestra en la Fig.6.1. σ tiene unidades de longitud y representa el diámetro de carozo duro (distancia mínima a la cual pueden aproximarse dos partículas), ya que $u(r)$ se anula en $r = \sigma$ y $u(r)$ crece rápidamente para $r < \sigma$. Para distancias $r > \sigma$ la interacción es atractiva con un mínimo en

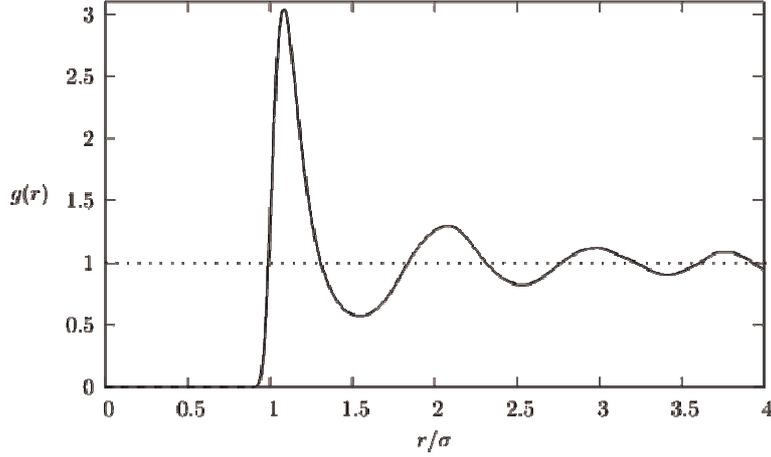


Figura 6.2: Función distribución de pares para el potencial de Lennard-Jones, calculada mediante simulaciones numéricas para $k_B T/\epsilon = 0,71$ y $\rho_p = 0,844$.

$r = \sigma 2^{1/6}$ y tiende rápidamente a cero como $1/r^6$ para $r \gg \sigma$, al igual que una interacción de Van der Waals. La intensidad de la interacción a una distancia $r = 2,5\sigma$ es inferior al 1% del valor en el mínimo. La función $g(r)$ para potenciales de este tipo tiene la forma característica que se muestra esquemáticamente en la Fig.6.2, la cual puede interpretarse como sigue. Si nos paramos en una molécula dada, $\rho_p g(r)$ nos dá la densidad promedio de partículas que observaríamos a una distancia r . Así, $g(r) = 0$ para $r \leq \sigma$, ya que la probabilidad de encontrar dos partículas a una distancia menor que esta es cero. $g(r)$ muestra típicamente picos espaciados a una distancia levemente mayor que σ . El primer pico en general es el dominante, debido a la interacción atractiva a cortas distancias. Los demás picos reflejan la competencia entre interacciones atractivas y repulsiva entre las diferentes partículas. A distancias grandes, esta correlación se pierde y $g(r)$ converge asintóticamente a $g = 1$, aproximando el comportamiento de un gas ideal. La distancia y distribución de los picos brinda información acerca de las correlaciones espaciales entre partículas. Cuanto mas diluido el gas, menos picos se observan en $g(r)$. Estos comportamientos son radicalmente diferentes al que se observa en el caso de un sólido cristalino. En primer lugar, en un sólido $g(\vec{r})$ no es isotrópica. Pero además, la misma exhibe picos espaciados regularmente correspondientes a los sitios del cristal, no importando la distancia.

En el caso de fluidos debilmente interactuantes a bajas densidades es posible obtener una aproximación para la ecuación de estado como sigue. Sea la función

$$f(r) = e^{-\beta u(r)} - 1 \quad (6.104)$$

Denominando $f_{ij} \equiv f(r_{ij})$ podemos escribir la función partición configuracional como

$$Q = \int dq_1 \cdots dq_N \prod_{(i,j)} e^{-\beta u(r_{ij})} = \int dq_1 \cdots dq_N \prod_{(i,j)} [1 + f_{ij}] \quad (6.105)$$

Desarrollando los productos tenemos

$$\begin{aligned} Q &= \int dq_1 \cdots dq_N \left[1 + \sum_{(i,j)} f_{ij} + \frac{1}{2} \sum_{(i,j)} \sum_{(k,l) \neq (i,j)} f_{ij} f_{kl} + \cdots \right] \\ &= V^N + V^{N-2} \sum_{(i,j)} \int dq_i dq_j f_{ij} + \cdots \end{aligned}$$

$$\begin{aligned}
&= V^N + V^{N-1} \frac{N(N-1)}{2} \int d^3r f(r) + \dots \\
&= V^N \left[1 - \frac{N(N-1)}{V} B(T) + \dots \right]
\end{aligned} \tag{6.106}$$

donde

$$B(T) = -\frac{1}{2} \int d^3r \left[e^{-\beta u(r)} - 1 \right] \tag{6.107}$$

se conoce como **segundo coeficiente del virial**. Si despreciamos los términos superiores en el desarrollo anterior tenemos la aproximación

$$\frac{Q}{V^N} \approx 1 - \frac{N^2}{V} B(T)$$

y de la Ec.(6.96)

$$P = k_B T \rho_p + k_B T \frac{\frac{N^2}{V^2} B(T)}{1 - \frac{N^2}{V} B(T)}$$

y para $N^2 B(T)/V \ll 1$ obtenemos finalmente

$$P \approx k_B T \rho_p [1 + \rho_p B(T)] \tag{6.108}$$

De hecho, puede demostrarse que, para potenciales de muy corto alcance, la ecuación de estado admite un desarrollo en potencias de la densidad de partículas

$$P = k_B T [\rho_p + \rho_p^2 B(T) + \rho_p^3 C(T) + \dots] \tag{6.109}$$

el cual se conoce como desarrollo del virial. Vemos entonces que la aproximación (6.108) resulta válida a muy bajas densidades. Si integramos la Ec.(6.107) por partes, para un potencial tipo Lennard-Jones obtenemos

$$B(T) = -\frac{2\pi\beta}{3} \int_0^\infty r^3 \frac{du(r)}{dr} e^{-\beta u(r)} dr \tag{6.110}$$

Reemplazando en la Ec.(6.108) y comparando con la Ec.(6.98) vemos que truncar el desarrollo del virial al segundo orden equivale a aproximar $g(r)$ por

$$g(r) = e^{-\beta u(r)} \tag{6.111}$$

En la Fig.6.3 vemos la aproximación a segundo orden del virial para el potencial de Lennard-Jones. La misma da cuenta solamente del primer pico. Para describir los demás picos resulta necesario tomar en cuenta ordenes superiores, a menos que el gas sea muy diluido.

6.8. Calor específico de los sólidos.

Vamos a revisar el problema de las vibraciones en un sólido cristalino en mas detalle, ahora utilizando el ensemble canónico. Consideremos un sistema de partículas que interactúan a través un potencial de pares, como el que vimos en la sección anterior Ec.(6.92). A temperaturas suficientemente bajas el sistema solidifica, porque existe una configuración ordenada de partículas que minimiza la energía (6.92). Esto no significa que las partículas queden fijas en los sitios del cristal, ya que por efecto térmico, las mismas presentarán pequeñas oscilaciones en torno de dichos sitios (oscilaciones de gran amplitud fundirían el cristal y ocurren a temperaturas mayores). En

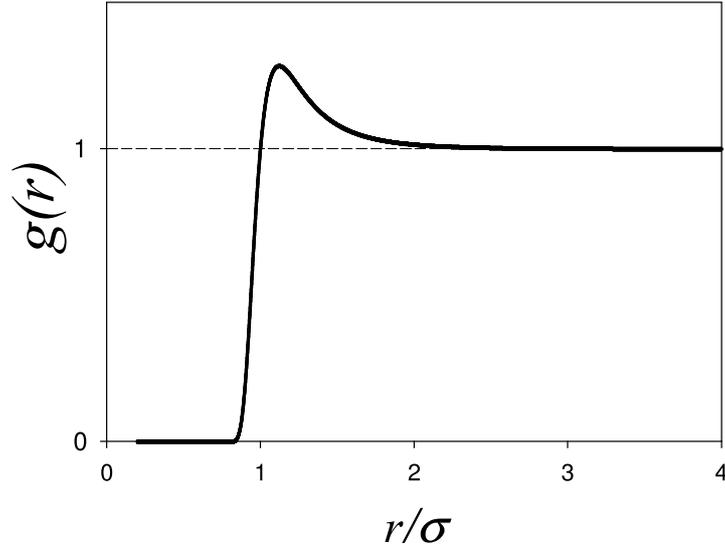


Figura 6.3: Aproximación a segundo orden del virial de la función distribución de pares para el potencial de Lennard-Jones a $k_B T/\epsilon = 1$.

esas condiciones, el potencial (6.92) puede ser aproximado por un potencial armónico (podemos desarrollar el mismo en Taylor hasta el segundo orden), de donde los movimientos colectivos de las partículas pueden ser descompuestos en modos normales. Dichos modos corresponden a ondas de sonido en el cristal (transversales y longitudinales) con una distribución de frecuencias que depende de la estructura cristalina particular.

Tomemos, por ejemplo el caso de un conjunto de N partículas restringidas a moverse en una dimensión, las cuales interactúan a través de un potencial de Lennard-Jones (6.103). Debido al carácter de corto alcance de las interacciones, resulta evidente que el mínimo de energía potencial ocurre si las partículas se encuentran equiespaciadas a una distancia $a \approx \sigma 2^{1/6}$ (la distancia del mínimo de $u(r)$). A temperatura cero (esto es, en ausencia de energía cinética) este será el estado de equilibrio. El mismo corresponde a partículas localizadas a lo largo de una cadena, en las posiciones $x_n = na$ con $n = 1, 2, \dots, N$. a se conoce como parámetro de red. Por simplicidad vamos a asumir condiciones de contorno periódicas $x_{N+1} = x_1$, con lo cual la cadena se transforma en un anillo. Esto no es para nada restrictivo, ya que en el límite termodinámico las propiedades deberían ser independientes de las condiciones de contorno. A bajas temperaturas las partículas tendrán cierta energía cinética y por lo tanto podrán efectuar pequeños desplazamientos en torno de las posiciones de equilibrio. En estas circunstancias, cada partícula está básicamente ligada a un sitio del anillo y por lo tanto se vuelven distinguibles.

Definamos el desplazamiento de la partícula n -ésima respecto de la posición de equilibrio x_n como $y_n \equiv q_n - x_n$. Para bajas temperaturas los desplazamientos serán pequeños y por lo tanto podemos aproximar el potencial de interacción entre cada par de partículas mediante un potencial cuadrático (desarrollo en Taylor hasta segundo orden en torno de $r = a$). Así, a menos de una constante el Hamiltoniano clásico puede ser escrito como

$$H = \sum_{n=1}^N \frac{p_n^2}{2m} + \frac{1}{2} K \sum_{n=1}^N (y_{n+1} - y_n)^2 \quad (6.112)$$

donde $y_{N+1} = y_1$ y $K = d^2u(r)/dr^2|_{r=a} > 0$. De la mecánica clásica, sabemos que para un Hamiltoniano cuadrático como el anterior existe una transformación canónica a un conjunto de variables (modos normales) (P_k, Q_k) , tal que en términos de dichas variables el mismo resulta

$$H = \sum_{k=1}^N \left[\frac{1}{2m} P_k^2 + \frac{1}{2} m \omega_k^2 Q_k^2 \right] \quad (6.113)$$

Estas nuevas coordenadas describen movimientos oscilatorios colectivos (modos normales) que corresponden a ondas longitudinales en la cadena. Las frecuencias de oscilación pueden obtenerse notando que los modos normales corresponden a combinaciones lineales de soluciones de la forma

$$y_n \propto e^{i(kna - \omega_k t)} \quad (6.114)$$

Las condiciones de contorno periódicas implican que $y_{n+N} = y_n$ y por lo tanto

$$e^{ikNa} = 1$$

de donde los vectores de onda permitidos son

$$k = \frac{2\pi m}{aN}$$

con $m = 0, \pm 1 \pm 2, \dots, \pm(N-1)/2, N/2$; otros valores de m repiten alguna de las soluciones anteriores, ya que si cambiamos k en $2\pi/a$ la solución permanece invariante. Así, tenemos exactamente N soluciones independientes, cuyos valores de k se encuentran en el intervalo $\pm\pi/a$, el cual se conoce como 1ra Zona de Brillouin.

Las ecuaciones de movimiento tienen la forma

$$m \frac{d^2 y_n}{dt^2} = -K(2y_n - y_{n+1} - y_{n-1}) \quad n = 1, \dots, N \quad (6.115)$$

Si reemplazamos la solución (6.114) en las ecuaciones de movimiento obtenemos

$$\begin{aligned} -m\omega_k^2 e^{i(kna - \omega_k t)} &= -K[2 - e^{ika} - e^{-ika}] e^{i(kna - \omega_k t)} \\ &= -2K(1 - \cos ka) e^{i(kna - \omega_k t)} \end{aligned}$$

de donde obtenemos la relación de dispersión

$$\omega_k = \sqrt{\frac{2K(1 - \cos ka)}{m}} = 2\sqrt{\frac{K}{m}} |\sin(ka/2)| \quad (6.116)$$

la cual se muestra en la Fig.6.4. Las soluciones (6.114) describen ondas longitudinales que se propagan a lo largo de la cadena con velocidad de fase $c_s = \omega_k/k$ y velocidad de grupo $v = \partial\omega/\partial k$. Para valores pequeños de k (longitudes de onda grandes) la relación de dispersión es aproximadamente lineal y la velocidad de grupo es aproximadamente independiente de k e igual a la velocidad de fase $c_s \approx a\sqrt{K/m}$. Esta es la velocidad del sonido en el sólido.

El desarrollo anterior puede extenderse fácilmente al caso tridimensional⁴. En este caso los átomos pueden ordenarse en diferentes estructuras cristalinas. El caso mas simple corresponde a una estructura cúbica simple, en la cual las tres coordenadas atómicas de los sitios de mínima energía se encuentran equiespaciadas una distancia a : $x_n = n_x a$, $y_n = n_y a$, $z_n = n_z a$, con $n_x, n_y, n_z = 1, \dots, L$. El número total de átomos es $N = L^3$. Imponiendo condiciones de contorno periódicas en las tres direcciones, los vectores de onda permitidos para los modos normales tienen componentes

⁴Es importante recalcar que el desarrollo anterior no intenta formular un modelo de solidificación en una dimensión, sino simplemente resolver el problema teórico de las oscilaciones en una dimensión para luego extenderlo al caso tridimensional. Por motivos que exceden el presente curso, puede verse que no es posible estabilizar un cristal a cualquier temperatura finita en dimensiones menores que 3.

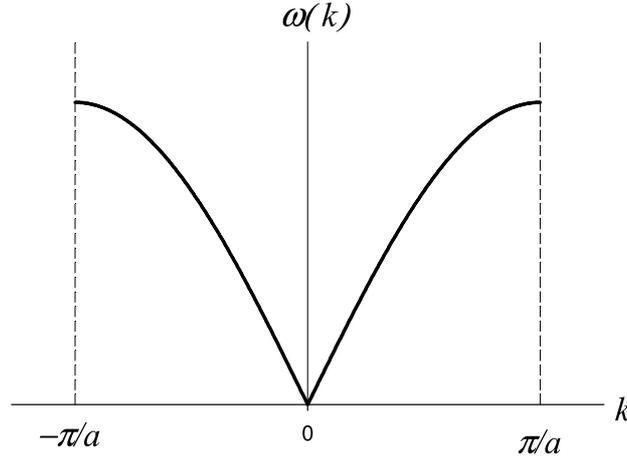


Figura 6.4: Relación de dispersión para una cadena lineal armónica.

$$k_i = \frac{2\pi m_i}{aL} \quad \text{con } i = x, y, z \quad \text{y} \quad m_i = 0, \pm 1, \pm 2, \dots \pm (L-1)/2, L/2 \quad (6.117)$$

totalizando N vectores de onda. Cada componente toma valores en el intervalo $\pm\pi/aL$. Esto define un cubo de lado $2\pi/aL$ en el espacio recíproco (el espacio de los vectores de onda), es cual constituye la 1ra zona de Brillouin en el caso 3D. En tres dimensiones es posible tener ondas longitudinales y transversales. Esto da lugar, por cada vector de onda, a tres soluciones diferentes, cada una con una relación de dispersión $\omega_s(\vec{k})$ con $s = 1, 2, 3$. Para valores pequeños de k estas relaciones de dispersión son aproximadamente lineales en k en forma semejante al caso unidimensional, dando lugar a dos velocidades del sonido diferentes, una para ondas longitudinales c_l y otra para ondas transversales c_t (los dos modos transversales corresponden a dos polarizaciones diferentes y por lo tanto tienen la misma velocidad del sonido). El Hamiltoniano puede escribirse entonces como una superposición de $3N$ osciladores armónicos independientes, que difieren en sus frecuencias:

$$H = \sum_{k=1}^N \sum_s \left[\frac{1}{2m} P_{\vec{k},s}^2 + \frac{1}{2} m \omega_s(\vec{k})^2 Q_{\vec{k},s}^2 \right] \quad (6.118)$$

Podemos ahora volver al problema del calor específico de este sistema. Como vimos en la sección 5.2.3, el calor específico clásico de un sistema de osciladores armónicos independientes es totalmente independiente de las frecuencias de los osciladores (este resultado puede reobtenerse fácilmente en el ensamble canónico). Por lo tanto, el desarrollo anterior en modos normales no nos aporta nada nuevo. Debemos entonces considerar la versión cuantificada del Hamiltoniano (6.118), esto es

$$H = \sum_{\vec{k}} \sum_s (n_{\vec{k},s} + 1/2) \hbar \omega_s(\vec{k}) \quad (6.119)$$

donde $n_{\vec{k},s} = 0, 1, \dots$. Los autoestados de este Hamiltoniano son el producto directo de los $3N$ autoestados de cada oscilador armónico. En esta base el operador densidad es diagonal y podemos escribir

$$\begin{aligned} Z_N(T) &= \text{Tr}_N \exp \left[-\beta \sum_{\vec{k}} \sum_s (n_{\vec{k},s} + 1/2) \hbar \omega_s(\vec{k}) \right] \\ &= \sum_{\{n_{\vec{k},s}\}} \exp \left[-\beta \sum_{\vec{k}} \sum_s (n_{\vec{k},s} + 1/2) \hbar \omega_s(\vec{k}) \right] \end{aligned} \quad (6.120)$$

donde $\{n_{\vec{k},s}\}$ denota todas las posibles combinaciones de valores enteros para los números cuánticos $n_{\vec{k},s}$; en otras palabras, tenemos $3N$ sumas que van de cero a infinito. Así

$$\begin{aligned} Z_N(T) &= \prod_{\vec{k},s} \left\{ \sum_{n_{\vec{k},s}=0}^{\infty} \exp \left[-\beta(n_{\vec{k},s} + 1/2) \hbar\omega_s(\vec{k}) \right] \right\} \\ &= \prod_{\vec{k},s} \frac{e^{-\beta\hbar\omega_s(\vec{k})/2}}{1 - e^{-\beta\hbar\omega_s(\vec{k})}} \end{aligned} \quad (6.121)$$

y

$$\ln Z_N(T) = - \sum_{\vec{k},s} \frac{\beta\hbar\omega_s(\vec{k})}{2} - \sum_{\vec{k},s} \ln \left[1 - e^{-\beta\hbar\omega_s(\vec{k})} \right] \quad (6.122)$$

$$U = - \frac{\partial \ln Z_N(T)}{\partial \beta} = \sum_{\vec{k},s} \frac{\hbar\omega_s(\vec{k})}{2} + \sum_{\vec{k},s} \frac{\hbar\omega_s(\vec{k})}{e^{\beta\hbar\omega_s(\vec{k})} - 1} \quad (6.123)$$

Los vectores de onda en las sumas anteriores están espaciados a una distancia $\Delta k_i = 2\pi/aL$. Así, en el límite termodinámico $L \rightarrow \infty$ podemos aproximar las sumas por integrales. La energía por partícula $u = U/N$ resulta

$$u = \frac{a^3}{(2\pi)^3} \sum_s \int_{-\pi/a}^{\pi/a} dk_x \int_{-\pi/a}^{\pi/a} dk_y \int_{-\pi/a}^{\pi/a} dk_z \left[\frac{\hbar\omega_s(\vec{k})}{2} + \frac{\hbar\omega_s(\vec{k})}{e^{\beta\hbar\omega_s(\vec{k})} - 1} \right] \quad (6.124)$$

y el calor específico

$$c = \frac{\partial u}{\partial T} = - \frac{1}{k_B T^2} \frac{\partial u}{\partial \beta} = \frac{a^3 \hbar^2}{k_B T^2 (2\pi)^3} \sum_s \int_{-\pi/a}^{\pi/a} dk_x \int_{-\pi/a}^{\pi/a} dk_y \int_{-\pi/a}^{\pi/a} dk_z \frac{\omega_s^2(\vec{k}) e^{\beta\hbar\omega_s(\vec{k})}}{(e^{\beta\hbar\omega_s(\vec{k})} - 1)^2} \quad (6.125)$$

En general no es fácil resolver la integral anterior y se hace necesario introducir alguna aproximación. El modelo de Einstein consiste en asumir que todas las frecuencias son iguales. No obstante, vimos que esta aproximación tan drástica no describe correctamente el comportamiento a muy bajas temperaturas. A continuación veremos un modelo más refinado introducido por Debye.

6.8.1. El modelo de Debye.

El error principal introducido en el modelo de Einstein consiste en no considerar los modos de bajas frecuencias (ondas de sonido), los cuales son particularmente relevantes a bajas temperaturas. La aproximación de Debye consiste en reemplazar las tres relaciones de dispersión por tres relaciones acústicas idénticas $\omega(\vec{k}) = c_s |\vec{k}|$, donde c_s es una velocidad del sonido promedio. Además, para simplificar los cálculos se reemplaza la integral en la primera zona de Brillouin (un cubo) por una integral en una esfera de radio k_D . El radio k_D se elige de manera tal que el número de estados en la esfera sea igual a N . Si consideramos que cada estado ocupa en el espacio recíproco un volumen $(2\pi/aL)^3$ tenemos que

$$N = \frac{4}{3} \pi k_D^3 / (8\pi^3/a^3 N) \Rightarrow k_D^3 = \frac{6\pi^2}{a^3}$$

En la aproximación de Debye, el calor específico resulta

$$c = \frac{3 a^3 \hbar^2 c_s^2}{2 k_B T^2 \pi^2} \int_0^{k_D} \frac{k^4 e^{\beta \hbar c_s k}}{(e^{\beta \hbar c_s k} - 1)^2} dk = \frac{9 k_B}{(\beta \hbar \omega_D)^3} \int_0^{\Theta_D/T} \frac{x^4 e^x}{(e^x - 1)^2} dx \quad (6.126)$$

donde $\omega_D = c_s k_D$ es la frecuencia de Debye y $\Theta_D = \hbar \omega_D / k_B$ es la temperatura de Debye. Para altas temperaturas $\Theta_D/T \ll 1$ podemos expandir en potencias de x el integrando en la Ec.(6.126), con lo cual nos queda $\sim x^2$. Así, un cálculo directo muestra que $c \rightarrow 3k_B$, con lo cual recuperamos la ley de Dulong y Petit. Para bajas temperaturas $\Theta_D/T \gg 1$ podemos reemplazar el límite superior de integración por $+\infty$, ya que el integrando decae exponencialmente para $x \gg 1$. De Tablas

$$\int_0^{\infty} \frac{x^4 e^x}{(e^x - 1)^2} dx = \frac{4\pi^4}{15}$$

y así

$$c \approx \frac{12k_B\pi^4}{5} \left(\frac{T}{\Theta_D} \right)^3 \quad (6.127)$$

que reproduce el comportamiento observado experimentalmente.

Capítulo 7

El Ensemble Gran Canónico

A pesar de que los ensembles microcanónico y canónico dan resultados equivalentes (en el límite termodinámico), puede argumentarse que conceptualmente el ensemble canónico brinda una representación mas cercana a las situaciones físicas reales. En los experimentos resulta imposible aislar completamente un sistema. Mas aún, uno nunca mide directamente la energía total de un sistema macroscópico: solo accedemos a ella a través de mediciones indirectas. Por el contrario, la temperatura es un parámetro facilmente medible y controlable. La equivalencia entre ensembles nos permite entonces desarrollar una intuición física, en relación mas directa con la situación física real, sin tener que pasar por una construcción teórica relativamente artificial.

En la misma línea argumental, podemos considerar poco físico especificar el número de partículas de un sistema macroscópico, ya que resulta imposible en la práctica conocer esta cantidad con precisión. Lo único que podemos determinar en un experimento es el número promedio de partículas, cantidad a la cual accedemos también de manera indirecta. Esta es una de las motivaciones para introducir un nuevo ensemble, el **ensemble gran canónico**, en el cual el sistema puede tener un número arbitrario de partículas, siendo únicamente su promedio especificado por las condiciones externas. La situación es enteramente análoga a la del ensemble canónico, en el cual la energía media del sistema está determinada por la temperatura del reservorio con el cual se encuentra en contacto.

Como veremos, los resultados obtenidos en este nuevo ensemble resultan equivalentes a los obtenidos en el ensemble canónico y microcanónico en el límite termodinámico. Así, mas allá de cualquier interpretación física, cabe resaltar que el uso de uno u otro ensemble en aplicaciones concretas de la teoría depende mas de conveniencias de cálculo que de consideraciones físicas. Como veremos, ciertos cálculos muy complicados en un ensemble pueden resultar casi triviales en otro. Así, la elección del ensemble apropiado puede resultar crucial a la hora de aplicar la teoría a problemas particulares.

7.1. Formulación general: relación con la termodinámica

Supongamos entonces un sistema que puede intercambiar tanto partículas como energía con el medio ambiente en el cual se encuentra inserto. El medio ambiente actúa entonces a la vez como reservorio de temperatura y de partículas. Vamos a suponer que el volumen del sistema se mantiene constante, de manera que pensamos que el sistema se encuentra separado del reservorio por paredes rígidas pero permeables¹.

Si el sistema se encuentra en equilibrio con el reservorio, el número de partículas no será mas constante, pero podemos asumir que tanto su energía media como su densidad media, esto es, el número medio de partículas, permanecen constantes. Los estados accesibles del sistema son en este

¹Que existan o no realmente este tipo de paredes no afecta los resultados de la teoría. En ultima instancia, podemos considerarlas como una idealización conveniente.

caso los autoestados de la energía para una partícula, para dos partículas, etc.. De esta manera, el espacio de Hilbert estará constituido por la suma directa de los espacios de Hilbert de una, dos, tres, etc. partículas. Dichos subespacios pueden ser caracterizados definiendo un operador número de partículas \hat{N} , tal que cualquier estado con un número bien definido de partículas N es autoestado de \hat{N} con autovalor N .

Vamos a *asumir* que el operador Hamiltoniano \hat{H} no mezcla estados de subespacios con diferentes números de partículas, es decir, que conmuta con \hat{N} . Esto significa la matriz que lo representa en una base de autoestados de \hat{N} tendrá una estructura diagonal en bloques $\hat{H}_0, \hat{H}_1, \text{etc.}$, donde \hat{H}_N es el hamiltoniano de N partículas.

Sea $|E_l^N, N\rangle$ es autoestado de \hat{H}_N con autovalor E_l^N entonces

$$\hat{H} |E_l^N, N\rangle = E_l^N |E_l^N, N\rangle$$

y

$$\hat{N} |E_l^N, N\rangle = N |E_l^N, N\rangle$$

donde el índice l corre sobre todos los autoestados del Hamiltoniano \hat{H}_N . Si \hat{H} conmuta con \hat{N} , entonces $\hat{\rho}$ también lo hace y por lo tanto también tiene una estructura diagonal en bloques.

El número máximo de partículas puede estar acotado o no, dependiendo del problema. En el caso de partículas con un volumen atómico finito, existirá un número máximo N_{max} que podremos alojar en un volumen V . En ese caso, los elementos de matriz de $\hat{\rho}$ deberán ser cero en cualquier subespacio con $N > N_{max}$. En el caso de partículas puntuales $N_{max} = \infty$. Asumamos por simplicidad este último caso. Si trabajamos en la base de autoestados de la energía $|E_l^N, N\rangle$, el operador densidad será diagonal y sus autovalores serán

$$\rho_l^N = \langle E_l^N | \hat{\rho} | E_l^N \rangle$$

(de aquí en más omitiremos el doble subíndice para simplificar la notación). Tenemos entonces que $\sum_l \rho_l^N$ nos da la probabilidad de que el sistema contenga N partículas.

Para encontrar $\hat{\rho}$ aplicamos el principio variacional de Gibbs, esto es, vamos a maximizar la entropía

$$S = -k_B \text{Tr} (\hat{\rho} \ln \hat{\rho}) = -k_B \sum_{N=0}^{\infty} \sum_l \rho_l^N \ln \rho_l^N \quad (7.1)$$

sujeta a los vínculos de normalización

$$\text{Tr} (\hat{\rho}) = \sum_{N=0}^{\infty} \sum_l \rho_l^N = 1, \quad (7.2)$$

energía media

$$U \equiv \langle \hat{H} \rangle = \text{Tr} (\hat{H} \hat{\rho}) = \sum_{N=0}^{\infty} \sum_l E_l^N \rho_l^N = cte \quad (7.3)$$

y número medio de partículas

$$\langle N \rangle = \text{Tr} (\hat{N} \hat{\rho}) = \sum_{N=0}^{\infty} N \sum_l \rho_l^N = cte \quad (7.4)$$

Introduciendo multiplicadores de Lagrange para cada uno de los vínculos anteriores, la variación de la entropía (7.1) al producir variaciones independientes $\delta\rho_i^N$ en los autovalores de $\hat{\rho}$ viene dada por

$$\delta \left[\sum_{N=0}^{\infty} \sum_l \left\{ \alpha_0 \rho_l^N + \alpha_1 E_l^N \rho_l^N + \alpha_2 N \rho_l^N - k_B \rho_l^N \ln \rho_l^N \right\} \right] = 0 \quad (7.5)$$

$$\sum_{N=0}^{\infty} \sum_l \delta \rho_l^N \left\{ (\alpha_0 - k_B) + \alpha_1 E_l^N + \alpha_2 N - k_B \ln \rho_l^N \right\} = 0 \quad (7.6)$$

de donde

$$-k_B \ln \rho_l^N + (\alpha_0 - k_B) + \alpha_1 E_l^N + \alpha_2 N = 0 \quad (7.7)$$

Dado que esto vale para cada uno de los autovalores de $\hat{\rho}$, \hat{H} y \hat{N} tenemos la igualdad entre operadores

$$-k_B \ln(\hat{\rho}) + (\alpha_0 - k_B)\hat{I} + \alpha_1\hat{H} + \alpha_2\hat{N} = 0 \quad (7.8)$$

de donde

$$\hat{\rho} = e^{(\alpha_0/k_B - 1)} \exp\left(\frac{\alpha_1}{k_B}\hat{H} + \frac{\alpha_2}{k_B}\hat{N}\right) \quad (7.9)$$

A fin de identificar los multiplicadores de Lagrange multiplicamos primero la ecuación anterior por $\hat{\rho}$ y tomamos traza, con lo cual la condición de normalización nos da que

$$\mathcal{Z} \equiv e^{(1 - \alpha_0/k_B)} = \text{Tr} \exp\left(\frac{\alpha_1}{k_B}\hat{H} + \frac{\alpha_2}{k_B}\hat{N}\right) \quad (7.10)$$

donde la ecuación anterior *define* la *función gran partición* \mathcal{Z} . Para determinar los multiplicadores α_1 y α_2 multiplicamos la Ec.(7.8) por $\hat{\rho}$ y tomamos traza, de donde

$$S + \alpha_1 U + \alpha_2 \langle N \rangle - k_B \ln \mathcal{Z} = 0 \quad (7.11)$$

Si multiplicamos la Ec. anterior por T y comparamos con la transformada de Legendre de la energía con respecto a S y a $\langle N \rangle$:

$$TS + \alpha_1 T U + \alpha_2 T \langle N \rangle - k_B T \ln \mathcal{Z} = 0 \quad (7.12)$$

$$\Omega = U - TS - \mu \langle N \rangle \quad (7.13)$$

donde μ es el potencial químico, vemos que podemos identificar $\alpha_1 = -1/T$, $\alpha_2 = \mu/T$ y

$$\Omega(\mu, T) = -k_B T \ln \mathcal{Z}(\mu, T) \quad (7.14)$$

Ω se conoce como *el potencial gran canónico* y la relación anterior nos da la ecuación fundamental para el ensemble gran canónico. Con estas identificaciones tenemos finalmente las expresiones:

$$\mathcal{Z}(T, V, \mu) = e^{-\beta\Omega(T, V, \mu)} = \text{Tr} e^{-\beta(\hat{H} - \mu\hat{N})} \quad (7.15)$$

y

$$\hat{\rho} = \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H} - \mu\hat{N})} \quad (7.16)$$

donde hemos incluido ahora explícitamente el volumen del sistema.

En el caso clásico, la derivación es enteramente análoga. En este caso definimos una densidad de probabilidad $\rho(p, q, N)$, la cual nos da la probabilidad de que el sistema contenga N partículas con momentos y coordenadas (p, q) . Podemos pensarla como una familia discreta de funciones densidad indexadas por N , tal que densidad N -ésima es función de $6N$ variables reales, sujetas a la condición de normalización

$$\sum_{N=0}^{\infty} \int \rho(p, q, N) dp dq = 1$$

donde cada integral se extiende sobre todo el espacio de las fases de N partículas.

La densidad que maximiza la entropía de Gibbs toma la forma

$$\rho(p, q, N) = \frac{1}{\mathcal{Z}} e^{-\beta(H^N(p,q) - \mu N)}$$

donde

$$\mathcal{Z} = \sum_{N=0}^{\infty} \frac{1}{B_N h^{3N}} \int e^{-\beta(H^N(p,q) - \mu N)} dp dq$$

con $B_N = 1$ para partículas distinguibles y $B_N = N!$ para partículas indistinguibles. Vemos que, al igual que en el ensemble canónico, todas las expresiones formales tienen la misma forma en el caso cuántico que en el clásico si adoptamos en este último la notación

$$\text{Tr} \equiv \sum_{N=0}^{\infty} \frac{1}{B_N h^{3N}} \int \dots dp dq$$

Salvo indicación contraria, todo lo que sigue es válido tanto para el caso cuántico como para el clásico.

Para completar la identificación de la expresión (7.14) con la transformada de Legendre de la energía debemos verificar que la misma sea consistente con las relaciones termodinámicas derivadas. Para ello calculemos a partir de la Ec.(7.14) la derivada

$$-\frac{\partial \Omega}{\partial T} = k_B \ln(\mathcal{Z}) + k_B T \frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial T} = k_B \ln(\mathcal{Z}) - \frac{1}{T \mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \beta} \quad (7.17)$$

donde hemos usado que

$$\frac{\partial}{\partial T} = -\frac{1}{k_B T^2} \frac{\partial}{\partial \beta}$$

De la definición de \mathcal{Z} tenemos que

$$\frac{\partial \mathcal{Z}}{\partial \beta} = -\text{Tr} \left[e^{-\beta(H - \mu N)} (H - \mu N) \right] \quad (7.18)$$

de donde

$$-\frac{\partial \Omega}{\partial T} = k_B \ln(\mathcal{Z}) + \frac{1}{T} U - \frac{\mu}{T} \langle N \rangle \quad (7.19)$$

Pero teníamos de la expresión variacional que

$$k_B \ln(\mathcal{Z}) = -\frac{1}{T} U + \frac{\mu}{T} \langle N \rangle + S$$

de donde verificamos la relación

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V, \mu} \quad (7.20)$$

Con un cálculo semejante podemos verificar también que

$$\langle N \rangle = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T, V} \quad (7.21)$$

Notemos que $\Omega \sim V$ para $V \gg 1$, es decir, es una función extensiva del volumen. Así, $\langle N \rangle \sim V$.

Por otra parte tenemos las siguientes relaciones para la energía interna a partir de la Ec.(7.18):

$$U(T, V, \mu) = - \frac{\partial \ln \mathcal{Z}}{\partial \beta} + \mu \langle N \rangle = \frac{\partial(\beta \Omega)}{\partial \beta} + \mu \langle N \rangle \quad (7.22)$$

Además, usando la ecuación de Euler $U = TS - PV + \mu N$ tenemos que

$$\Omega = -PV \quad \rightarrow \quad P = -\frac{\Omega}{V} \quad (7.23)$$

Finalmente, si calculamos la función partición en la base de autoestados de la energía obtenemos la expresión general

$$\begin{aligned} \mathcal{Z} &= \text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} = \sum_{N=0}^{\infty} \sum_l e^{-\beta(E_l^N - \mu N)} \\ &= \sum_{N=0}^{\infty} e^{\beta \mu N} \sum_l e^{-\beta E_l^N} \\ &= \sum_{N=0}^{\infty} z^N Z_N(T) \end{aligned} \quad (7.24)$$

donde $Z_N(T)$ es la función partición canónica de N partículas y

$$z \equiv e^{\beta \mu} \quad (7.25)$$

se conoce como la **fugacidad**.

7.2. Fluctuaciones en el número de partículas: equivalencia de ensembles

En el ensamble gran canónico la energía media y el número medio de partículas son constantes para T y μ fijos. No obstante, dado que el sistema puede intercambiar energía y partículas con el medio externo, existirán fluctuaciones con respecto a estos valores medios. Es importante entonces conocer la magnitud de estas fluctuaciones. A fin de calcular las fluctuaciones en el número medio de partículas podemos escribir la condición de normalización de la siguiente manera

$$\text{Tr} \rho = \text{Tr} e^{\beta(\Omega(T, \mu) - H + \mu N)} = 1 \quad (7.26)$$

Derivando respecto de μ tenemos

$$\text{Tr} \left[\left(\beta \frac{\partial \Omega}{\partial \mu} + \beta N \right) e^{\beta(\Omega(T, \mu) - H + \mu N)} \right] = 0 \quad (7.27)$$

$$\beta \frac{\partial \Omega}{\partial \mu} + \beta \text{Tr} \left[N e^{\beta(\Omega(T, \mu) - H + \mu N)} \right] = 0 \quad (7.28)$$

Derivando nuevamente respecto de μ tenemos

$$\beta \frac{\partial^2 \Omega}{\partial \mu^2} + \beta \text{Tr} \left[N \beta \left(\frac{\partial \Omega}{\partial \mu} + N \right) e^{\beta(\Omega(T, \mu) - H + \mu N)} \right] = 0 \quad (7.29)$$

$$\beta \frac{\partial^2 \Omega}{\partial \mu^2} + \beta^2 \frac{\partial \Omega}{\partial \mu} \text{Tr} [N \rho] + \beta^2 \text{Tr} [N^2 \rho] = 0 \quad (7.30)$$

Usando la Ec.(7.21) tenemos

$$\langle N^2 \rangle - \langle N \rangle^2 = -k_B T \frac{\partial^2 \Omega}{\partial \mu^2} = k_B T \frac{\partial \langle N \rangle}{\partial \mu} \quad (7.31)$$

de donde la desviación fraccional en el número de partículas se comporta como

$$\frac{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}{\langle N \rangle} \sim V^{-1/2}$$

A medida que el volúmen (y por lo tanto el número de partículas) se incrementa el tamaño de las fluctuaciones se vuelve cada vez mas pequeño comparado con la magnitud del número medio de partículas. Con un cálculo semejante al utilizado en el ensamble canónico puede mostrarse tambien que

$$\frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} \sim V^{-1/2}$$

Vemos que en el límite termodinámico la mayoría de los microestados probables tienen el mismo número de partículas y la misma energía, con lo cual los tres ensembles resultan equivalentes.

Finalmente, notemos la similitud en la estructura del operador densidad en los ensembles canónico y gran canónico, en relación al principio variacional de Gibbs. El operador resulta siempre proporcional a la exponencial de una combinación lineal de observables, uno por cada vínculo macroscópico que imponemos al sistema. Los valores medios de dichos observables corresponden a variables termodinámicas extensivas. Los coeficientes de dicha combinación lineal están asociados a los multiplicadores de Lagrange respectivos y resultan iguales al parámetro termodinámico intensivo asociado al observable en la representación entropía: $1/T$ en el caso de la energía y $-\mu/T$ en el caso del número de partículas. La generalización de este procedimiento permite derivar diferentes tipos de ensembles, apropiados para cada problema particular, los cuales resultan todos equivalentes en el límite termodinámico. Un ejemplo es el llamado **ensemble de las presiones**, que corresponde a la situación física de una sistema en contacto con un reservorio de temperatura y presión. En este caso, el número de partículas es constante, pero la energía y el volúmen pueden fluctuar. Los vínculos externos son entonces la energía media y el volúmen medio. Mediante una derivación semejante a la utilizada para el ensemble gran canónico, se encuentra que la función gran partición en el caso del ensemble de las presiones viene dada por

$$\Upsilon(P, T, N) = \int_0^\infty \left(\text{Tr} e^{-\beta(H+PV)} \right) dV = \int_0^\infty e^{-\beta PV} Z_N(T, V) dV$$

donde $Z_N(T, V)$ es la función partición canónica para N partículas en un volúmen V y

$$G(T, P, N) = -k_B T \ln \Upsilon(P, T, N)$$

es la energía libre de Gibbs. La derivación se deja como ejercicio para el lector.

7.3. Gases Ideales Cuánticos

El tratamiento cuántico correcto de un gas ideal requiere tomar en consideración la indistinguibilidad de las partículas. Esto significa que los microestados accesibles del sistema deben corresponder exclusivamente a funciones de onda simétricas, en el caso de bosones o antisimétricas, en el caso de fermiones. Como vimos en la sección 6.6 el cálculo de la función partición canónica sobre una base de estados totalmente simétricos o antisimétricos es bastante complicada. Por contraposición, veremos que este cálculo es extremadamente sencillo en el ensemble gran canónico. Este es un buen ejemplo de como la equivalencia entre ensembles nos permite este tipo de flexibilidad.

Consideremos entonces un sistema de partículas no interactuantes y supongamos que conocemos los autovectores y autovalores de una partícula

$$H_1 |k\rangle = \epsilon_k |k\rangle$$

donde k representa un conjunto completo de números cuánticos para una partícula. Como vimos en la sección 6.6, con los mismos podemos construir una base de estados simetrizados $|k_1, k_2, \dots, k_N\rangle^{S,A}$ de N partículas. Así, podemos construir una base para cada subespacio de N partículas y la suma directa de todas esas bases nos da una base para un sistema con un número arbitrario de partículas. Ahora bien, como también vimos en la sección 6.6, el orden de los índices en dichos estados es completamente irrelevante. Así, al sumar sobre todos los posibles valores de los índices para calcular la función partición, debemos descontar la repetición de términos. Por otra parte, la especificación de los números de ocupación n_k con $\sum_k n_k = N$ determina unívocamente cada estado simetrizado de N partículas. Así, si indexamos los autoestados simetrizados mediante los números de ocupación no necesitamos preocuparnos por la repetición de términos. Esto resulta particularmente útil en el caso del ensemble gran canónico, ya que ahora podemos levantar la restricción $\sum_k n_k = N$. Vamos entonces a denotar por

$$|n_{k_1}, n_{k_2}, \dots\rangle^{S,A}$$

un estado simetrizado con n_{k_1} partículas en el nivel ϵ_{k_1} , n_{k_2} partículas en el nivel ϵ_{k_2} , etc, donde los índices k_1, k_2, \dots ahora corren sobre **todos** los posibles números cuanticos **diferentes** de una partícula. Por ejemplo, el estado con $n_{k_i} = 1$ y todos los demás números de ocupación iguales a cero corresponde al estado de una partícula $|k_i\rangle$; el estado *simétrico* con $n_{k_i} = 2$ y todos los demás números de ocupación iguales a cero corresponde al estado de dos partículas $|k_i\rangle |k_i\rangle$; el estado *antisimétrico* con $n_{k_i} = 1$, $n_{k_j} = 1$ y todos los demás números de ocupación iguales a cero corresponde al estado de dos partículas

$$\frac{1}{\sqrt{2}} [|k_i\rangle |k_j\rangle - |k_j\rangle |k_i\rangle]$$

y así sucesivamente. Dado que cada uno de estos estados está unívocamente determinado por el conjunto de valores de los números de ocupación, podemos expresar

$$\mathcal{Z} = \sum_{n_{k_1}=0}^{\infty} \sum_{n_{k_2}=0}^{\infty} \dots \sum_{n_{k_\infty}=0}^{\infty} {}^{S,A} \langle n_{k_1}, n_{k_2}, \dots | e^{-\beta(\hat{H}-\mu\hat{N})} | n_{k_1}, n_{k_2}, \dots \rangle^{S,A} \quad (7.32)$$

Claramente, estos estados son autoestados de \hat{N} con autovalores $N = \sum_k n_k$. Además

$$\hat{H} |n_{k_1}, n_{k_2}, \dots\rangle^{S,A} = \left(\sum_k n_k \epsilon_k \right) |n_{k_1}, n_{k_2}, \dots\rangle^{S,A}$$

y así

$$e^{-\beta(\hat{H}-\mu\hat{N})} |n_{k_1}, n_{k_2}, \dots\rangle^{S,A} = \exp\left(-\beta \sum_k (\epsilon_k - \mu) n_k\right) |n_{k_1}, n_{k_2}, \dots\rangle^{S,A}$$

En el caso de partículas de Bose-Einstein (BE) tenemos entonces que

$$\mathcal{Z}_{BE}(T, V, \mu) = \sum_{n_{k_1}=0}^{\infty} \sum_{n_{k_2}=0}^{\infty} \cdots \sum_{n_{k_\infty}=0}^{\infty} \exp\left\{-\beta \sum_k n_k (\epsilon_k - \mu)\right\} \quad (7.33)$$

En el caso de partículas de Fermi-Dirac (FD) las funciones de onda no nulas son aquellas para las cuales todos los números de ocupación adoptan los valores cero o uno. La función gran partición para el gas de FD tiene por lo tanto la forma:

$$\mathcal{Z}_{FD}(T, V, \mu) = \sum_{n_{k_1}=0}^1 \sum_{n_{k_2}=0}^1 \cdots \sum_{n_{k_\infty}=0}^1 \exp\left\{-\beta \sum_k n_k (\epsilon_k - \mu)\right\} \quad (7.34)$$

Finalmente, resulta ilustrativo considerar que ocurriría si tuvieramos un gas cuántico de partículas idénticas *distinguibles*. En este caso no tendríamos ninguna restricción en los valores posibles de los números de ocupación, pero tenemos que tomar en cuenta que, dado un conjunto de números de ocupación fijos $\{n_k\}$, el intercambio de dos partículas en dos autoestados arbitrarios diferentes k_i y k_j , con números de ocupación n_{k_i} y n_{k_j} corresponde a un nuevo estado, *diferente del anterior*, sin alterar los números de ocupación y por lo tanto con el mismo factor exponencial. Así, para cada conjunto de valores $\{n_k\}$, debemos multiplicar dicho factor por el número de combinaciones de partículas entre estados diferentes. La función gran partición para un gas de partículas distinguibles tiene entonces la forma

$$\mathcal{Z}_{dist}(T, V, \mu) = \sum_{n_{k_1}=0}^{\infty} \sum_{n_{k_2}=0}^{\infty} \cdots \sum_{n_{k_\infty}=0}^{\infty} \frac{N!}{n_{k_1}! \cdots n_{k_j}! \cdots n_{k_\infty}!} \exp\left\{-\beta \sum_k n_k (\epsilon_k - \mu)\right\} \quad (7.35)$$

donde el factor multinomial nos da el número de posibles combinaciones de partículas distinguibles correspondientes a un dado conjunto de números de ocupación $\{n_k\}$ tales que $\sum_k n_k = N$.

Como veremos en el próximo capítulo, a temperaturas altas el número medio de bosones en cualquier estado k es muy pequeño, lo que significa que los únicos microestados que contribuyen significativamente a la función gran partición son aquellos para los cuales todos los números de ocupación toman valores cero o uno. Así, el comportamiento a altas temperaturas de ambos tipos de partículas (bosones o fermiones) es el mismo. Las mismas consideraciones valen en el caso de un gas de partículas distinguibles. En esa situación, si comparamos las Eqs.(7.33) y (7.34) con la Eq.(7.35) vemos que la única diferencia reside en el factor $N!$ en la última, ya que $n_k! = 1$ para la mayoría de los números de ocupación. Así, el comportamiento de los gases ideales cuánticos a altas temperaturas corresponde al de un gas de partículas distinguibles con la corrección de eliminar el factor $N!$. Este es precisamente el conteaje correcto de Boltzmann.

Un sistema de partículas cuya estadística esta descrito por la función gran partición

$$\mathcal{Z}_{MB}(T, V, \mu) = \sum_{n_{k_1}=0}^{\infty} \sum_{n_{k_2}=0}^{\infty} \cdots \sum_{n_{k_\infty}=0}^{\infty} \frac{1}{n_{k_1}! \cdots n_{k_j}! \cdots n_{k_\infty}!} \exp\left\{-\beta \sum_k n_k (\epsilon_k - \mu)\right\} \quad (7.36)$$

se conoce como *gas de Maxwell-Boltzmann* (MB) y el mismo describe el comportamiento a altas temperaturas de todos los gases ideales. Comparando con la Ec.(7.24) es facil ver que

$$\mathcal{Z}_{MB}(T, V, \mu) = \sum_{N=0}^{\infty} z^N Z_N(T)/N! \quad (7.37)$$

donde $Z_N(T)$ es la función partición canónica de N partículas distinguibles.

7.4. Gas ideal de Maxwell-Boltzmann

El gas de Maxwell-Boltzmann corresponde un sistema idealizado que describe el comportamiento de altas temperaturas de los gases ideales cuánticos. Resulta un ejercicio interesante derivar una vez más las ecuaciones del gas ideal a partir de este modelo usando el ensamble gran canónico. Precisamente la eliminación del factor $N!$ de la Ec.(7.36) vuelve este problema sumamente fácil de resolver, ya que la función gran partición se factoriza de la forma:

$$\mathcal{Z}_{MB}(T, V, \mu) = \prod_k \left(\sum_{n_k=0}^{\infty} \frac{1}{n_k!} e^{-\beta n_k (\varepsilon_k - \mu)} \right) = \prod_k \left(\exp \left[e^{-\beta (\varepsilon_k - \mu)} \right] \right) \quad (7.38)$$

donde hemos usado el desarrollo en serie de Taylor de la función exponencial $e^x = \sum_{n=0}^{\infty} (1/n!) x^n$. El potencial gran canónico resulta entonces

$$\begin{aligned} \Omega_{MB}(T, V, \mu) &= -k_B T \ln \mathcal{Z}_{MB}(T, V, \mu) \\ &= -k_B T \sum_k e^{-\beta (\varepsilon_k - \mu)} \\ &= -k_B T z \sum_k e^{-\beta \varepsilon_k} \end{aligned} \quad (7.39)$$

donde z es la fugacidad. El número medio de partículas viene dado por

$$\langle N \rangle = - \left(\frac{\partial \Omega_{MB}}{\partial \mu} \right)_{T, V} = \sum_k z e^{-\beta \varepsilon_k} \quad (7.40)$$

Tenemos además que

$$\langle N \rangle = \sum_k \langle n_k \rangle \quad (7.41)$$

de donde

$$\langle n_k \rangle = z e^{-\beta \varepsilon_k} \quad (7.42)$$

es el número medio de partículas en el autoestado k . Este resultado puede obtenerse directamente calculando $(1/\mathcal{Z}_{MB}) \text{Tr } n_k \exp(-\beta(H - \mu N))$.

Notemos que la suma sobre k en la Ec.(7.39) corresponde exactamente a la función partición canónica de 1 partícula y así

$$\Omega_{MB}(T, V, \mu) = -k_B T z Z_1(T) \quad (7.43)$$

La Ec.(7.43) es general para partículas no interactuantes con cualquier espectro. Vamos a restringirnos por simplicidad al caso de partículas únicamente con grados de libertad traslacionales. Veremos al final de la sección la generalización al caso que incluye otros grados de libertad, tales como rotacionales, spin, etc.. En este caso la energía es puramente cinética. Vamos a asumir además que el gas se encuentra contenido en una caja cúbica de lado L y volumen $V = L^3$ con paredes

perfectamente reflectantes. Como vimos en la sección 6.2, la función partición de una partícula en este caso es $Z_1 = V/\lambda_T^3$, donde

$$\lambda_T \equiv \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{1/2}$$

es la longitud de onda térmica. Así, la relación fundamental para el gas de Maxwell-Boltzmann en el ensemble gran canónico es

$$\Omega_{MB}(T, V, \mu) = -k_B T \frac{V z}{\lambda_T^3} = -k_B T V z \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \quad (7.44)$$

De la Ec.(7.40) tenemos

$$\langle N \rangle = z Z_1(T) = \frac{V z}{\lambda_T^3} = -\beta \Omega_{MB} \quad (7.45)$$

De aquí podemos obtener el potencial químico en función de la densidad de partículas y de T

$$\mu = k_B T \ln \left(\frac{\langle N \rangle \lambda_T^3}{V} \right) \quad (7.46)$$

La entropía del gas de Maxwell-Boltzmann viene dada por

$$S = - \left(\frac{\partial \Omega_{MB}}{\partial T} \right)_{V, \mu} = \left(\frac{5k_B}{2} - \frac{\mu}{T} \right) \frac{V z}{\lambda_T^3}. \quad (7.47)$$

Usando las Ecs.(7.45) y (7.46) podemos expresar S en función de T y $\langle N \rangle$:

$$S = \frac{5k_B \langle N \rangle}{2} - k_B \langle N \rangle \ln \left(\frac{\langle N \rangle}{V} \lambda_T^3 \right). \quad (7.48)$$

Usando la Ec.(7.23) tenemos que

$$P = - \frac{\Omega_{MB}}{V} = \frac{k_B T z}{\lambda_T^3} = \frac{\langle N \rangle k_B T}{V} \quad (7.49)$$

que es el resultado esperado para un gas ideal clásico. De las Ecs.(7.22) y (7.49) tenemos que la energía interna viene dada por

$$U = - \frac{\partial \langle N \rangle}{\partial \beta} + \mu \langle N \rangle. \quad (7.50)$$

Usando la Ec.(7.45) obtenemos entonces el resultado esperado

$$U = \frac{3}{2} \langle N \rangle k_B T = \frac{3}{2} P V \quad (7.51)$$

Veremos que esta última igualdad es válida para todos los gases ideales. Finalmente, reemplazando la ecuación anterior en la expresión (7.48), se reobtiene la fórmula de Sackur y Tetrode.

Supongamos ahora que tenemos un gas de moléculas no interactuantes. Sea $\zeta(T)$ la función partición canónica de una molécula, correspondiente a los grados internos de libertad (rotacionales, vibracionales, etc, todos excepto la energía cinética), los cuales asumimos desacoplados de la energía cinética. De la Ec.(7.37) tenemos, en la aproximación semiclásica, que

$$\mathcal{Z} = \sum_{N=0}^{\infty} (z \zeta(T))^N (Z_1(T))^N / N! \quad (7.52)$$

donde $Z_1(T)$ es la función partición canónica de una partícula libre puntual. Notemos que los grados internos de libertad pueden ser absorbidos en la fugacidad, esto es, este sistema puede pensarse como un gas ideal con un potencial químico efectivo $\mu'(T)$, tal que $z \rightarrow z' = z\zeta(T)$, ó $\mu' = \mu + k_B T \ln \zeta(T)$. Así, el potencial gran canónico resulta

$$\Omega(T, V, \mu) = -k_B T z \zeta(T) Z_1(T) \quad (7.53)$$

Ejercicios

1. Obtenga la Ec.(7.43) directamente de la expresión (7.37).
2. Muestre que para el sistema descrito por la Ec.(7.53) se cumple que

$$PV = \langle N \rangle k_B T$$

$$U = \langle N \rangle k_B T \left[\frac{3}{2} + T \frac{\partial \ln \zeta(T)}{\partial T} \right]$$

$$\mu = k_B T \ln \left(\frac{\langle N \rangle \lambda_T^3}{V \zeta(T)} \right) \quad (7.54)$$

7.5. Adsorción en superficies

Supongamos que tenemos en contacto un material sólido con un gas (o mas genéricamente, un fluido), en equilibrio a presión y temperatura constantes. Consideremos la interfase entre ambos materiales (superficie del sólido). A diferencia de los átomos (o moléculas) en el volúmen, los átomos en la superficie encuentran desbalanceadas sus interacciones con otros átomos del sólido, pudiendo por lo tanto atraer moléculas del gas. Se define como **adsorción** al fenómeno en el cual moléculas del **adsorbato** (el gas) se ligan a la superficie del sólido. El término adsorción debe distinguirse del término *absorción*, el cual describe la situación en la cual las moléculas del gas penetran en el volúmen, llenando poros del sólido. En la adsorción las moléculas del adsorbato quedan ligadas a sitios fijos de la superficie e incluso puede ocurrir **desorción**, el proceso inverso. En equilibrio termodinámico, se adsorverán tantas moléculas como se desorverán y la concentración de adsorbato en la superficie del sólido permanecerá constante. El resultado final de este proceso es la formación de una película líquida o gaseosa en la superficie del material. Estos procesos son de enorme importancia industrial en la actualidad, siendo algunos de los adsorbentes mas comunes el carbón activado y la sílica gel.

El proceso de adsorción es en general complejo, pero podemos tener una idea general del comportamiento termodinámico característico a través de un modelo idealizado introducido por Irving Langmuir en 1916. El mismo contituye un buen ejemplo de aplicación del ensemble gran canónico clásico. El modelo se basa en suponer (a) que las moléculas del adsorbato se adsorven en sitios fijos de la superficie, los cuales denominaremos sitios de adsorción o trampas, (b) que cada sitios puede adsorver a lo sumo una molécula, (c) que las moléculas adsorvidas no interactúan entre sí y (d) que el gas se comporta como un gas ideal.

Dado que la superficie del sólido (sistema de trampas) se encuentra en equilibrio con el gas, podemos pensar que este último actúa como un reservorio de partículas para la primera. El procedimiento consiste entonces en calcular las ecuaciones de estado para el sistema de trampas y luego imponer las condiciones de equilibrio termodinámico con el gas.

Supongamos entonces que tenemos M sitios de adsorción y sea $-\gamma$ la energía de ligadura de una molécula (γ es la energía necesaria para evaporar una molécula del adsorbato ligada a la

superficie). Supongamos además que una molécula tiene grados internos de libertad. Sea $\zeta(T)$ la función partición canónica para los grados internos de libertad de una molécula. La función gran partición para el sistema de trampas viene dada entonces por

$$\mathcal{Z}_M = \sum_{N=0}^M z^N Z_N(T)$$

donde Z_N es la función partición canónica para N moléculas adsorvidas. El conteaje correcto de Boltzmann no se incluye en este caso, ya que las trampas son distinguibles. Dado que las moléculas adsorvidas son independientes, tenemos que

$$Z_N(T) = g(N) (Z_1(T))^N = g(N) \left(e^{\beta\gamma\zeta(T)} \right)^N$$

donde $g(N)$ es el número de maneras de acomodar N moléculas en M trampas. Así

$$\mathcal{Z}_M = \sum_{N=0}^M \frac{M!}{N!(M-N)!} \left(z e^{\beta\gamma\zeta(T)} \right)^N = \left(1 + z e^{\beta\gamma\zeta(T)} \right)^M \quad (7.55)$$

Alternativamente, podemos derivar la ecuación anterior como sigue. Como ya vimos, los grados internos de libertad pueden ser absorbidos en la fugacidad, reemplazando $z \rightarrow z' = z \zeta(T)$. Cada trampa puede pensarse como una sistema de dos estados: con molécula adsorbida o vacía, con energías $-\gamma$ y cero respectivamente. Así, podemos asociar a cada trampa un número de ocupación $n_i = 0, 1$, tal que cero y uno representan los estados de la trampa i -ésima con molécula y sin molécula adsorbida respectivamente. El Hamiltoniano de este sistema puede entonces escribirse como $H = -\gamma \sum_{i=1}^M n_i$ y $N = \sum_{i=1}^M n_i$. La función gran partición puede escribirse como

$$\mathcal{Z}_M = \text{Tr} e^{-\beta(H - \mu' N)} = \sum_{n_1=0,1} \dots \sum_{n_M=0,1} e^{\sum_i n_i \beta(\gamma + \mu')} = \left(1 + e^{\beta(\gamma + \mu')} \right)^M$$

de donde resulta la Ec.(7.55). El número medio de partículas adsorvidas viene dado por

$$\langle N \rangle = z \frac{\partial \ln \mathcal{Z}_M}{\partial z} = M \frac{z e^{\beta\gamma\zeta(T)}}{1 + z e^{\beta\gamma\zeta(T)}}$$

Se define el *cubrimiento* θ de la superficie como la fracción de moléculas adsorvidas. Así

$$\theta = \frac{z e^{\beta\gamma\zeta(T)}}{1 + z e^{\beta\gamma\zeta(T)}} \quad (7.56)$$

La condición de equilibrio termodinámico entre el sólido y el gas es la igualdad entre sus potenciales químicos. De la Ec.(7.54) tenemos que

$$z \zeta(T) = \frac{\langle N \rangle \lambda_T^3}{V} = \frac{P \lambda_T^3}{k_B T}$$

y por lo tanto

$$\theta = \frac{P}{P_0 + P} \quad (7.57)$$

donde

$$P_0 = \frac{k_B T}{\lambda_T^3} e^{-\beta\gamma}$$

La Ec.(7.57) se conoce como *isoterma de Langmuir*. Para $P/P_0 \ll 1$, el cubrimiento se comporta como $\theta \sim P/P_0$ y $\theta \rightarrow 1$ para $P/P_0 \gg 1$.

Capítulo 8

Gases ideales de Bose-Einstein

La función gran partición (7.33) para el gas ideal de Bose-Einstein puede escribirse como

$$\mathcal{Z}_{BE}(T, V, \mu) = \prod_k \left(\sum_{n_k=0}^{\infty} e^{-\beta n_k (\varepsilon_k - \mu)} \right) = \prod_k \left(\frac{1}{1 - e^{-\beta(\varepsilon_k - \mu)}} \right) \quad (8.1)$$

donde hemos usado que $\sum_{n=0}^{\infty} a^n = 1/(1-a)$; como veremos enseguida, $\varepsilon_k - \mu \geq 0$, con lo cual resulta válido usar este resultado. El potencial gran canónico resulta entonces

$$\Omega_{BE}(T, V, \mu) = -k_B T \ln \mathcal{Z}_{BE}(T, V, \mu) = k_B T \sum_k \ln \left(1 - e^{-\beta(\varepsilon_k - \mu)} \right) \quad (8.2)$$

de donde podemos calcular

$$\langle N \rangle = - \left(\frac{\partial \Omega_{BE}}{\partial \mu} \right)_{T, V} = \sum_k \left(\frac{e^{-\beta(\varepsilon_k - \mu)}}{1 - e^{-\beta(\varepsilon_k - \mu)}} \right) = \sum_k \left(\frac{1}{e^{\beta(\varepsilon_k - \mu)} - 1} \right) \quad (8.3)$$

Recordando que

$$\langle N \rangle = \sum_k \langle n_k \rangle \quad (8.4)$$

tenemos entonces que el número medio de partículas en el estado k es

$$\langle n_k \rangle = \left(\frac{1}{e^{\beta(\varepsilon_k - \mu)} - 1} \right) = \left(\frac{1}{z^{-1} e^{\beta \varepsilon_k} - 1} \right) \quad (8.5)$$

Dado que $\langle n_k \rangle \geq 0$ tenemos que $e^{\beta(\varepsilon_k - \mu)} \geq 1$ y por lo tanto $\varepsilon_k - \mu \geq 0 \forall k$.

8.1. La condensación de Bose-Einstein

Supongamos por simplicidad que tenemos partículas libres en una caja de volúmen $V = L^3$ con condiciones de contorno periódicas, esto es, las funciones de onda de una partícula deben satisfacer $\phi_k(x + L, y + L, z + L) = \phi_k(x, y, z)$. Las autofunciones en estas condiciones son ondas planas normalizadas de la forma:

$$\phi_{\vec{k}}(\vec{r}) = \frac{1}{L^{3/2}} e^{i\vec{k} \cdot \vec{r}}$$

donde los autovalores del operador vector de onda vienen dados por

$$k_i = \frac{2\pi}{L} n_i \quad i = x, y, z$$

con $n_i = 0 \pm 1, \pm 2 \dots$ y donde los autovalores de la energía son $\varepsilon(\vec{k}) = \hbar^2 k^2 / 2m$. La suma sobre autoestados corresponde en este caso a

$$\sum_k = \sum_{n_x=0,\pm 1,\dots} \sum_{n_y=0,\pm 1,\dots} \sum_{n_z=0,\pm 1,\dots} \quad (8.6)$$

Reescribiendo la suma (8.6) como

$$\sum_k \dots = \frac{V}{(2\pi)^3} \sum_{k_x=0,\pm 2\pi/L,\dots} \sum_{k_y=0,\pm 2\pi/L,\dots} \sum_{k_z=0,\pm 2\pi/L,\dots} \dots \Delta k_x \Delta k_y \Delta k_z \quad (8.7)$$

donde $\Delta k_i = 2\pi/L$, podemos aproximar en el límite termodinámico $L \rightarrow \infty$:

$$\sum_k \dots \sim \frac{V}{(2\pi)^3} \int \dots d^3 k = \frac{4\pi V}{(2\pi)^3} \int_0^\infty \dots k^2 dk \quad (8.8)$$

donde la última igualdad vale solo si el integrando es función únicamente de $k = |\vec{k}|$.

En este sistema el mínimo valor del espectro de una partícula corresponde al estado con $\vec{k} = 0$, $\varepsilon_0 = 0$. Esto implica que $\mu \leq 0$ y por lo tanto $0 \leq z \leq 1$. Reemplazando las sumas por integrales (en el límite termodinámico) en la Ec.(8.3) tenemos

$$\langle N \rangle = \frac{4\pi V}{(2\pi)^3} \int_0^\infty \frac{k^2 z}{e^{\beta \hbar^2 k^2 / 2m} - z} dk = \frac{4V}{\sqrt{\pi}} \left(\frac{mk_B T}{2\pi \hbar^2} \right)^{3/2} \int_0^\infty x^2 \left(\frac{z}{e^{x^2} - z} \right) dx \quad (8.9)$$

de donde la densidad de partículas $\rho \equiv \langle N \rangle / V$ resulta

$$\rho = \frac{1}{\lambda_T^3} g_{3/2}(z) \quad (8.10)$$

La función $g_{3/2}(z)$ se define como

$$g_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 \left(\frac{z}{e^{x^2} - z} \right) dx = \sum_{k=1}^\infty \frac{z^k}{k^{3/2}} \quad (8.11)$$

la cual es un caso particular de la familia de funciones $g_n(z)$ definidas como

$$g_n(z) = \sum_{k=1}^\infty \frac{z^k}{k^n}. \quad (8.12)$$

Estas funciones se han sido sumamente estudiadas y se encuentran tabuladas en la literatura. La función (8.11) es acotada y monótona creciente de z para todo $0 \leq z \leq 1$, y toma los valores límites:

$$g_{3/2}(0) = 0$$

$$g_{3/2}(1) = \sum_{k=1}^\infty \frac{1}{k^{3/2}} = \zeta(3/2) = 2,612 \dots$$

donde $\zeta(x)$ es la función zeta de Riemann. La derivada de $g_{3/2}(z)$ diverge para $z \rightarrow 1$ y para valores pequeños de z vemos, del desarrollo en serie (8.11), que $g_{3/2}(z) \sim z$ (ver Fig.8.1).

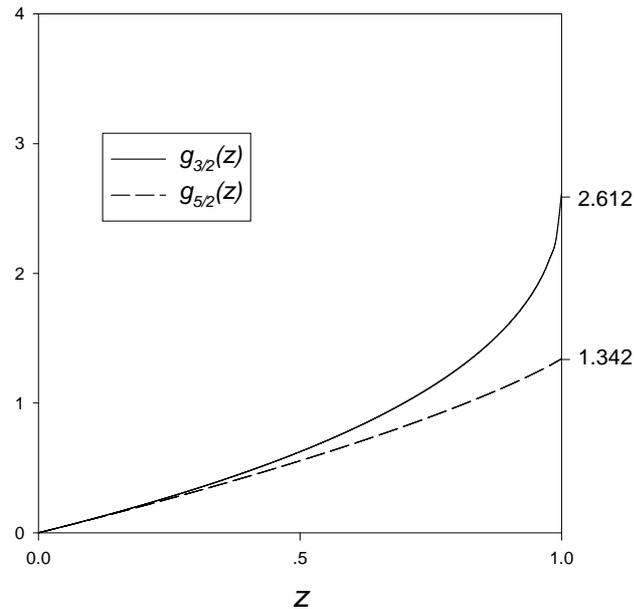


Figura 8.1: Funciones $g_{3/2}$ y $g_{5/2}$.

La ecuación

$$\lambda_T^3 \rho = g_{3/2}(z)$$

es una ecuación implícita para z en función de ρ y T . Pero el producto $\lambda_T^3 \rho$ puede tomar cualquier valor arbitrariamente grande para temperaturas suficientemente bajas y/o altas densidades. De manera que para $\lambda_T^3 \rho > 2,612$ no existe solución a esta ecuación, ya que z no puede ser mayor que uno. Evidentemente esta es una región singular de las funciones termodinámicas, lo cual nos está señalando la existencia de una transición de fase, y claramente para $\lambda_T^3 \rho > 2,612$ algo está incorrecto en nuestros cálculos previos.

Podemos darnos cuenta de donde está el problema si revisamos con más cuidado las Ecs.(8.3) y (8.5). El término en la suma (8.3) correspondiente al estado fundamental $\vec{k} = 0$, esto es, el número de medio de partículas $\langle n_0 \rangle$ en el estado con $\varepsilon_0 = 0$ es divergente cuando $z \rightarrow 1$ (o $\mu \rightarrow 0$):

$$\langle n_0 \rangle = \left(\frac{z}{1-z} \right)^{z \rightarrow 1} \infty. \quad (8.13)$$

Esto significa que el estado fundamental ¹ tiene una ocupación macroscópica cuando $z \rightarrow 1$ en el límite termodinámico, en el cual el número medio de partículas también diverge, ya que exigimos que la densidad sea finita. Vemos entonces que en el cálculo anterior hemos perdido este efecto. Como es posible que aparezca una singularidad en las ecuaciones termodinámicas? La respuesta está en el límite termodinámico y veremos que el error en los cálculos estuvo en la manera en que tomamos este límite.

Vamos entonces a recalcular $\langle N \rangle$ en el límite $V \rightarrow \infty$, pero aislando el término divergente de la suma (8.3), esto es, vamos a aproximar

¹El estado fundamental $\varepsilon = 0$ no necesita ser el estado con $\vec{p} = 0$. Este valor surge en particular para las condiciones de contorno periódicas. En general vamos a tener este fenómeno cualquiera sea el estado fundamental, el cual podemos asumir siempre $\varepsilon = 0$, ya que el cero de la energía está indefinido para las soluciones de la Ec. de Schrödinger. Así, si $\varepsilon_0 \neq 0$ podemos restar a todos los niveles el valor ε_0 y redefinir el potencial químico $\mu' = \mu - \varepsilon_0$.

$$\langle N \rangle = \frac{z}{1-z} + \frac{4\pi V}{(2\pi)^3} \int_{2\pi/L}^{\infty} \frac{k^2 z}{e^{\beta\hbar^2 k^2/2m} - z} dk = \frac{z}{1-z} + \frac{4V}{\lambda_T^3 \sqrt{\pi}} \int_{\lambda_T \sqrt{\pi}/L}^{\infty} x^2 \left(\frac{z}{e^{x^2} - z} \right) dx \quad (8.14)$$

De la misma manera podemos reescribir el potencial gran canónico como

$$\Omega_{BE} = k_B T \ln(1-z) + \frac{4\pi k_B T V}{(2\pi)^3} \int_{2\pi/L}^{\infty} k^2 \ln(1 - ze^{-\beta\hbar^2 k^2/2m}) dk \quad (8.15)$$

$$= k_B T \ln(1-z) + \frac{4k_B T V}{\lambda_T^3 \sqrt{\pi}} \int_{\lambda_T \sqrt{\pi}/L}^{\infty} x^2 \ln(1 - ze^{-x^2}) dx. \quad (8.16)$$

Las integrales (8.14) y (8.16) excluyen una esfera de radio $2\pi\hbar/L$ alrededor del origen, lo cual corresponde a sumar a partir de los primeros estados excitados con $k = 2\pi k/L$. Podemos entonces escribir las ecuaciones de estado de la siguiente manera:

$$\rho = \frac{\langle N \rangle}{V} = \frac{1}{V} \frac{z}{1-z} + \frac{1}{\lambda_T^3} g_{3/2}(z) - I_\rho \left(z, \frac{\lambda_T \sqrt{\pi}}{L} \right) \quad (8.17)$$

$$P = -\frac{\Omega_{BE}}{V} = -\frac{k_B T}{V} \ln(1-z) + \frac{k_B T}{\lambda_T^3} g_{5/2}(z) - I_p \left(z, \frac{\lambda_T \sqrt{\pi}}{L} \right) \quad (8.18)$$

donde

$$g_{5/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^{\infty} x^2 \ln(1 - ze^{-x^2}) dx = \sum_{k=1}^{\infty} \frac{z^k}{k^{5/2}} \quad (8.19)$$

$$I_\rho(z, a) = \frac{4}{\lambda_T^3 \sqrt{\pi}} \int_0^a x^2 \left(\frac{z}{e^{x^2} - z} \right) dx \quad (8.20)$$

$$I_p(z, a) = -\frac{4k_B T}{\lambda_T^3 \sqrt{\pi}} \int_0^a x^2 \ln(1 - ze^{-x^2}) dx. \quad (8.21)$$

La función $g_{5/2}(z)$ es también monótona creciente y toma los valores $g_{5/2}(0) = 0$ y $g_{5/2}(1) = \zeta(5/2) = 1,342\dots$ (ver Fig.8.1).

Las integrales (8.20) y (8.21) se anulan en el límite $L \rightarrow \infty$ para todo valor de z , excepto tal vez para $z = 1$, en el cual los respectivos integrandos son singulares. Desarrollando en serie de potencias el término exponencial en ambos integrandos es fácil ver que

$$\lim_{a \rightarrow 0} I_\rho(1, a) = \lim_{a \rightarrow 0} I_p(1, a) = 0. \quad (8.22)$$

Analícemos entonces la nueva ecuación para la densidad

$$\rho = \frac{1}{V} \frac{z}{1-z} + \frac{1}{\lambda_T^3} g_{3/2}(z) \quad (8.23)$$

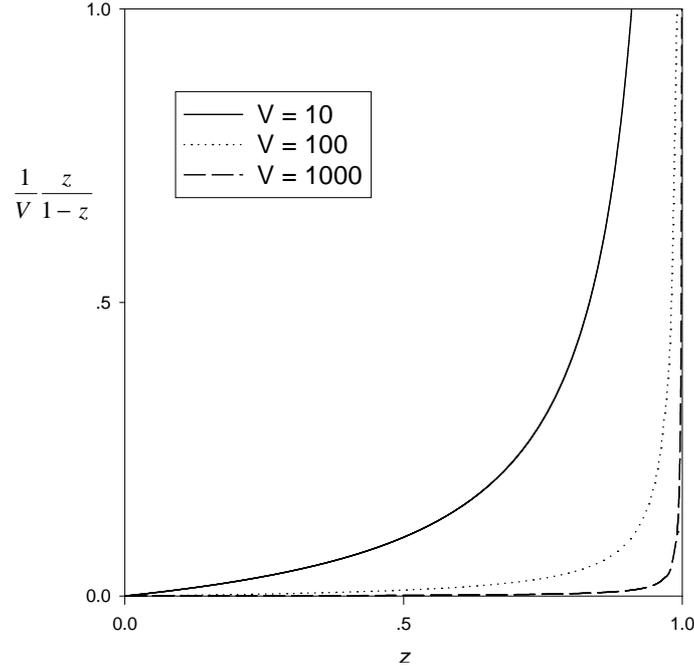


Figura 8.2: Comportamiento del primer término de la Ec.(8.23) s para diferentes valores de V .

El comportamiento del primer término de la Ec.(8.23) se muestra en la Fig.8.2 para diferentes valores de V . Tenemos entonces que resolver la ecuación implícita para z :

$$\lambda_T^3 \rho = \frac{\lambda_T^3}{V} \frac{z}{1-z} + g_{3/2}(z) \quad (8.24)$$

la cual puede resolverse gráficamente para cualquier par de valores de ρ y T como se muestra en al Fig.8.3a. Vemos que para V finito la divergencia del lado derecho de la Ec.(8.24) cuando $z \rightarrow 1$ impide que la raíz de dicha ecuación alcance el valor $z = 1$ para cualquier valor finito de T y ρ . Para $T \rightarrow 0$ y $\rho \rightarrow \infty$ tenemos que $z \rightarrow 1$ y por lo tanto $\langle n_0 \rangle \rightarrow \infty$, lo cual es esperable, ya que en esas condiciones todas las partículas estarán en el estado fundamental ². La solución de z en función de $\lambda_T^3 \rho$ para un gas contenido en un volumen V *finito* se muestra en la Fig.8.3b.

Supongamos ahora V finito, pero *grande*: $V \gg 1$. Para $\lambda_T^3 \rho < 2,612$ las soluciones de la Ec.(8.10) tienden a las de la Ec.(8.23) cuando $V \rightarrow \infty$. Para $\lambda_T^3 \rho \geq 2,612$ tenemos que las raíces de la Ec.(8.24) serán cercanas a $z = 1$. Así, podemos aproximar

$$\lambda_T^3 \rho \approx \frac{\lambda_T^3}{V} \frac{z}{1-z} + g_{3/2}(1)$$

de donde podemos despejar

$$z(V) \approx \frac{\rho_0 V}{1 + \rho_0 V} = \frac{1}{1 + \frac{1}{\rho_0 V}} \sim 1 - \frac{1}{\rho_0 V} \quad (8.25)$$

donde ρ_0 es una cantidad de que no depende de V . Así, vemos que todas las soluciones para $\lambda_T^3 \rho \geq 2,612$ tienden a $z = 1$ cuando $V \rightarrow \infty$. La fugacidad del gas ideal de Bose-Einstein, en el límite termodinámico, es por lo tanto

²Este límite tiene que ser tomado conjuntamente para que $z \rightarrow 1$. Si tomamos el límite $T \rightarrow 0$ manteniendo ρ finita, el primer término del lado derecho de la Ec.(8.24) también diverge, provocando que z tienda a un valor *finito* tal que $\langle n_0 \rangle \rightarrow \langle N \rangle = \rho V$.

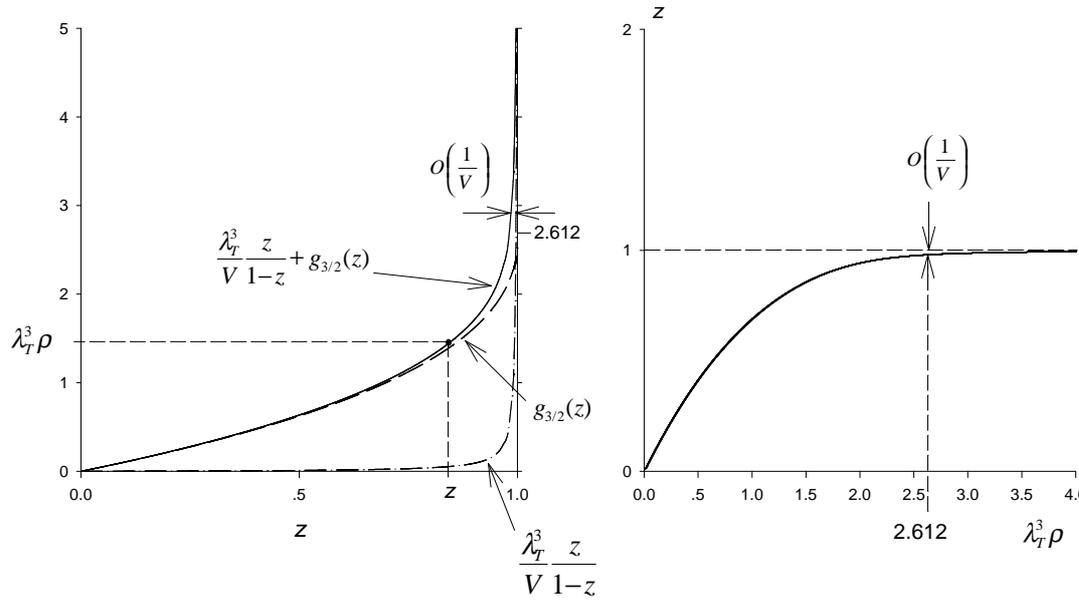


Figura 8.3: (a) Solución gráfica de la Ec.(8.24). (b) Fugacidad de un gas ideal de Bose-Einstein contenido en un volumen finito V .

$$z = \begin{cases} 1 & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \text{la raíz de } \lambda_T^3 \rho = g_{3/2}(z) & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.26)$$

y se muestra en la Fig.8.4.

De la Ec.(8.25) tenemos además que

$$\lim_{V \rightarrow \infty} \left(\frac{1}{V} \frac{z(V)}{1 - z(V)} \right) = \lim_{V \rightarrow \infty} \frac{\langle n_0 \rangle}{V} = \begin{cases} \rho_0 & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ 0 & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.27)$$

y por lo tanto de la Ec.(8.23) tenemos que la densidad media de partículas tiene satisfase

$$\rho = \frac{1}{v} = \begin{cases} \rho_0 + \frac{1}{\lambda_T^3} g_{3/2}(1) & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \frac{1}{\lambda_T^3} g_{3/2}(z) & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.28)$$

donde z viene dado por (8.26) y v es el volumen por partícula.

Vemos entonces que para $\lambda_T^3 \rho \geq g_{3/2}(1)$ un número *macroscópico* de partículas ocupan el estado fundamental ε_0 . Este fenómeno se conoce como **condensación de Bose-Einstein** y comienza a ocurrir cuando $z \rightarrow 1$ (ó, equivalentemente, cuando $\mu \rightarrow 0$), es decir, cuando la densidad ρ y la temperatura T son tales que

$$\lambda_T^3 \rho = g_{3/2}(1). \quad (8.29)$$

La Ec.(8.29) nos permite obtener la temperatura crítica T_c de la transición en función de la densidad:

$$\lambda_{T_c}^3 = \frac{g_{3/2}(1)}{\rho} \quad (8.30)$$

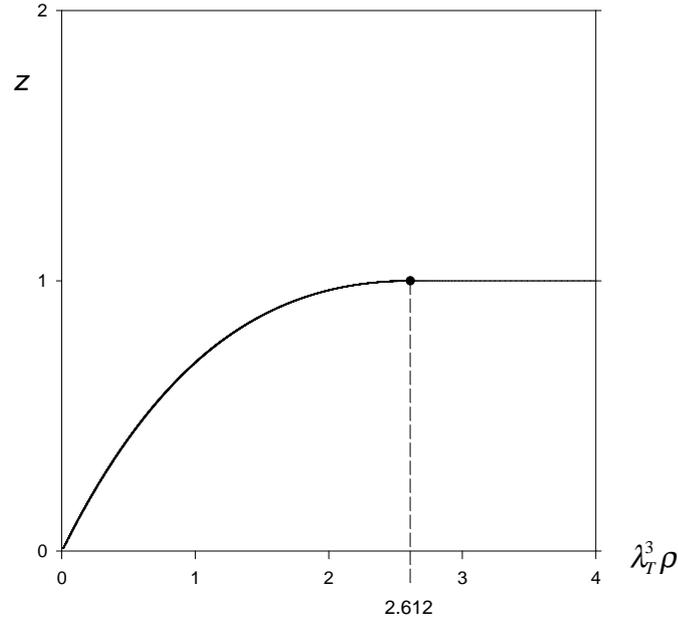


Figura 8.4: Fugacidad de un gas ideal de Bose-Einstein en el límite termodinámico.

de donde

$$T_c = \left(\frac{2\pi\hbar^2}{mk_B} \right) \left(\frac{\rho}{g_{3/2}(1)} \right)^{2/3} \quad (8.31)$$

También podemos invertir la ecuación (8.31) y expresar el volumen específico crítico en función de T :

$$v_c = \left(\frac{2\pi\hbar^2}{mk_B} \right)^{3/2} \frac{g_{3/2}(1)}{T^{3/2}} \quad (8.32)$$

De la Ec.(8.28) podemos obtener la fracción de partículas en el estado fundamental para la región de condensación $z = 1$, manteniendo fija la densidad y variando la temperatura:

$$1 = \frac{\rho_0}{\rho} + \frac{1}{\rho\lambda_T^3} g_{3/2}(1) \quad (8.33)$$

$$\frac{\langle n_0 \rangle}{\langle N \rangle} = \frac{\rho_0}{\rho} = 1 - \frac{1}{\rho\lambda_T^3} g_{3/2}(1) = 1 - \frac{\lambda_{T_c}^3}{\lambda_T^3} = 1 - \left(\frac{T}{T_c} \right)^{3/2} \quad (8.34)$$

Así, podemos definir un parámetro de orden η para la transición como

$$\eta \equiv \frac{\langle n_0 \rangle}{\langle N \rangle} = \begin{cases} 1 - \left(\frac{T}{T_c} \right)^{3/2} & \text{si } T \leq T_c \\ 0 & \text{si } T > T_c \end{cases} \quad (8.35)$$

el cual se muestra en la Fig.8.5. Para $T \sim T_c$ tenemos que

$$\eta \sim \frac{T_c - T}{T_c} \quad (8.36)$$

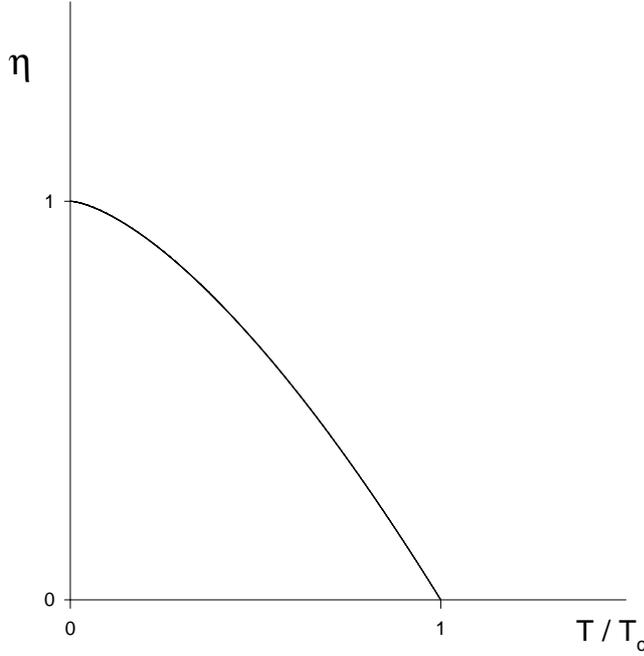


Figura 8.5: Parámetro de orden $\eta = \langle n_0 \rangle / \langle N \rangle$ vs. temperatura reducida.

Esto nos sugiere que la transición es de segundo orden. No obstante, no es posible encuadrar esta transición dentro de la clasificación usual de primer y segundo orden. En particular, el calor específico no es divergente en el punto crítico, si bien presenta un comportamiento anómalo. Mas aún, veremos que desde cierto punto de vista podría interpretarse como una transición de primer orden.

Notemos que si mantenemos fija la temperatura y variamos la densidad, podemos expresar también la fracción de partículas en el estado fundamental a partir de las Ecs.(8.28) y (8.34) como

$$\frac{\langle n_0 \rangle}{\langle N \rangle} = 1 - \frac{\rho_c}{\rho} = 1 - \frac{v}{v_c} \quad (8.37)$$

Vamos a calcular entonces las restantes funciones termodinámicas. Consideremos la expresión (8.18) para la presión. Es evidente que $\lim_{V \rightarrow \infty} (1/V) \ln(1-z) = 0$ para todo $z \neq 1$. Además, de la Ec.(8.25) tenemos que

$$\lim_{V \rightarrow \infty} \left(\frac{1}{V} \ln(1-z(V)) \right) = 0 \quad (8.38)$$

Así, reemplazando las Ecs(8.22), (8.26) y (8.38) en (8.18) tenemos que

$$P = \begin{cases} \frac{k_B T}{\lambda_T^3} g_{5/2}(1) & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \frac{k_B T}{\lambda_T^3} g_{5/2}(z) & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.39)$$

Notemos que la presión es independiente de la densidad para $\lambda_T^3 \rho \geq g_{3/2}(1)$.

Analicemos ahora las isotermas del gas de Bose-Einstein en el espacio de parámetros (P, v) . A temperatura constante tenemos un punto de transición $P = P_c(v_c)$ que se obtiene de tomar $z = 1$ en la Ec.(8.39) y expresar T en función de v a través de la Ec.(8.32). Variando T se define entonces una línea de transición

$$P_c(v_c) = \frac{2\pi\hbar^2 g_{5/2}(1)}{m(g_{3/2}(1))^{5/3}} \frac{1}{v_c^{5/3}} \quad (8.40)$$

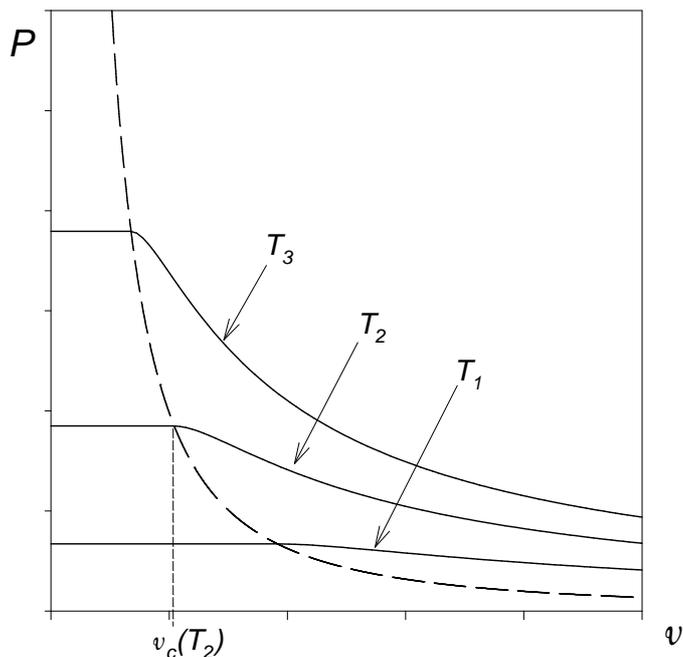


Figura 8.6: Gráfico esquemático de las isothermas de gas ideal de Bose-Einstein para tres temperaturas diferentes $T_1 < T_2 < T_3$. La línea de trazos corresponde a la curva $P_c(v_c)$.

En la Fig.8.6 podemos ver la forma general de las isothermas. Para $v < v_c(T)$ tenemos que $P = P_c = cte$. Esto recuerda fuertemente las isothermas de la transición usual gas-líquido en la región de coexistencia. Mas aún, dado que el estado condensado consiste en una mezcla de un conjunto macroscópico de partículas en el estado fundamental y otro conjunto distribuido en el resto de los niveles (la fracción de partículas en cualquier otro nivel individual tiende cero en el límite termodinámico). De esta manera podemos interpretar dicho estado como una coexistencia entre una fase gaseosa y una fase “condensada”, si bien esta última correspondería a una condensación en el espacio de los momentos. A partir de esta consideraciones suele interpretarse a veces la condensación de BE como una transición de primer orden. Si aceptamos esta interpretación tenemos que preguntarnos entonces cual es el volúmen específico de la fase condensada. Comparando la Fig.8.6 con el diagrama de consistencia de la transición gas-líquido todo indica que el volúmen específico del condensado es cero y el del gas es $v = v_c$. Pero esto esto significa que el condensado tiene densidad infinita! Sin embargo este aspecto no-físico es consistente con el hecho de que estamos considerando un gas de partículas no-interactuantes, lo cual implica que las partículas pueden acercarse entre ellas a distancias arbitrariamente pequeñas. Mas aún, veremos que esta interpretación, mas allá del aspecto no físico expuesto, es bastante consistente.

Consideremos por ejemplo la presión crítica en función de la temperatura, esto es

$$P_c(T) = \frac{k_B T}{\lambda_T^3} g_{5/2}(1) \propto (k_B T)^{5/2} \quad (8.41)$$

De acuerdo a la interpretación anterior, esta función, la cual se muestra en la Fig.8.7, correspondería a la curva de presión de vapor de la transición, es decir, a una curva de coexistencia. Sin embargo notemos que nunca podemos atravesar la curva. Nunca podemos tener una fase condensada pura, excepto a $T = 0$. Derivando la Ec.(8.41) tenemos

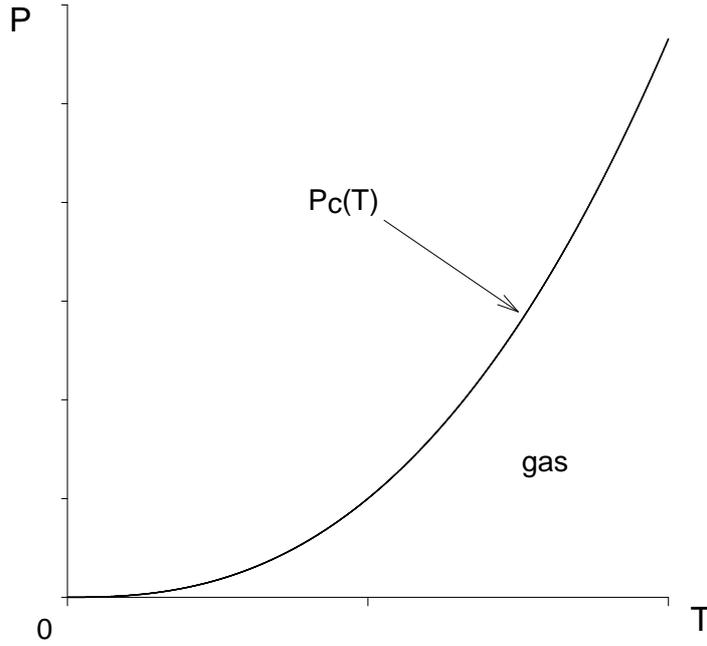


Figura 8.7: Línea de coexistencia $P_c(T)$.

$$\frac{dP_c(T)}{dT} = \frac{5}{2} \frac{k_B g_{5/2}(1)}{\lambda_T^3} = \frac{1}{Tv_c} \left(\frac{5}{2} k_B T \frac{g_{5/2}(1)}{g_{3/2}(1)} \right) \quad (8.42)$$

La Ec.(8.42) correspondería la ecuación de Clausius-Clapeyron para la transición. Dado que $\Delta v = v_c$, esto implica un calor latente de la forma

$$l = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{5}{2} k_B T. \quad (8.43)$$

Calculemos ahora la entropía por unidad de volúmen. Tenemos que

$$s = \lim_{V \rightarrow \infty} -\frac{1}{V} \left(\frac{\partial \Omega_{BE}}{\partial T} \right)_{V,\mu} = \lim_{V \rightarrow \infty} \left(\frac{\partial P}{\partial T} \right)_{V,\mu} \quad (8.44)$$

Derivando entonces la Ec.(8.39) y usando la propiedad

$$\frac{dg_n(z)}{dz} = \frac{1}{z} g_{n-1}(z) \quad (8.45)$$

tenemos que

$$s = \begin{cases} \frac{5}{2} \frac{k_B}{\lambda_T^3} g_{5/2}(1) & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \frac{5}{2} \frac{k_B}{\lambda_T^3} g_{5/2}(z) - k_B \rho \ln(z) & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.46)$$

Vemos que $s = 0$ a $T = 0$, de acuerdo con la tercera ley de la termodinámica. Hemos visto además que la fase condensada pura solo existe a $T = 0$. De aquí concluimos que el condensado tiene entropía cero, lo cual es consistente con una fase en la cual todas las partículas se encuentran en el mismo estado cuántico. Así, a temperatura finita, la contribución a la entropía en la región de coexistencia viene dada exclusivamente por la fracción de partículas en la fase gaseosa. Si llamamos s' a la entropía por partícula del sistema y s'_g a la entropía por partícula *en la fase gaseosa*, tenemos que

$$s' = \left(1 - \frac{\langle n_0 \rangle}{\langle N \rangle}\right) s'_g = \frac{v}{v_c} s'_g \quad (8.47)$$

donde hemos usado la Ec.(8.37). De la Ec.(8.46) tenemos que en la región de coexistencia

$$s' = \lim_{V \rightarrow \infty} \frac{S}{\langle N \rangle} = v s = \frac{5 k_B}{2 \lambda_T^3} g_{5/2}(1) v \quad (8.48)$$

Comparando las dos últimas ecuaciones, y teniendo en cuenta que la entropía de la fase condensada es cero, tenemos que la diferencia de entropía entre la fase gaseosa y la condensada es

$$\Delta s = s'_g = \frac{5 k_B}{2 \lambda_T^3} g_{5/2}(1) v_c = \frac{5 g_{5/2}(1)}{2 g_{3/2}(1)} k_B \quad (8.49)$$

Comparando con la Ec.(8.43) vemos que

$$l = T \Delta s \quad (8.50)$$

lo cual es consistente con la interpretación de una transición de primer orden. Finalmente, es posible ver que la compresibilidad isotérmica $\kappa_T \rightarrow \infty$ cuando $v \rightarrow v_c$. Dado que las isothermas son planas en la región de coexistencia, $\kappa_T = \infty$ en toda la región de coexistencia, lo cual es consistente con la imagen de un fluido infinitamente compresible.

Mas alla de cualquier interpretación, es claro que los aspectos no físicos del gas ideal de Bose-Einstein son un resultado de despreciar las interacciones entre partículas. Mas aún, el efecto mas interesante, esto es, la condensación, aparece cuando

$$\frac{\lambda_T}{v^{1/3}} = \left(g_{3/2}(1)\right)^{1/3} \approx 1,377$$

esto es, cuando la longitud de onda de deBroglie promedio es del orden de la distancia media entre partículas. En esta situación las interacciones entre partículas en general no pueden ser despreciadas. Modelos mas realistas que toman en cuenta interacciones repulsivas entre partículas a cortas distancias muestran que la condensación persiste, pero los efectos no físicos (isothermas planas) desaparecen. En este caso la transición de fase es claramente de segundo orden.

Asi, las predicciones mas importantes del gas ideal de Bose-Einstein son: (a) Es posible tener una transición de fase como resultado exclusivamente de la estadística, a diferencia de otras transiciones de fase que son un resultado exclusivamente de las interacciones entre partículas; (b) la fase de bajas temperaturas corresponde a un **estado cuántico coherente macroscópico**. Esto último implica que propiedades cuánticas, las cuales usualmente solo son observables a escalas microscópicas, puedan ser observadas a escalas macroscópicas.

A partir de la expresión (8.46) podemos calcular el calor específico a densidad constante mediante la expresión

$$c_\rho = T \left(\frac{\partial s}{\partial T} \right)_\rho \quad (8.51)$$

Notemos que para calcular c_ρ tenemos que mantener ρ constante en lugar de μ . Asi, vamos a tener que calcular

$$\left(\frac{\partial z}{\partial T} \right)_\rho = - \frac{3 z g_{3/2}(z)}{2 T g_{1/2}(z)} \quad (8.52)$$

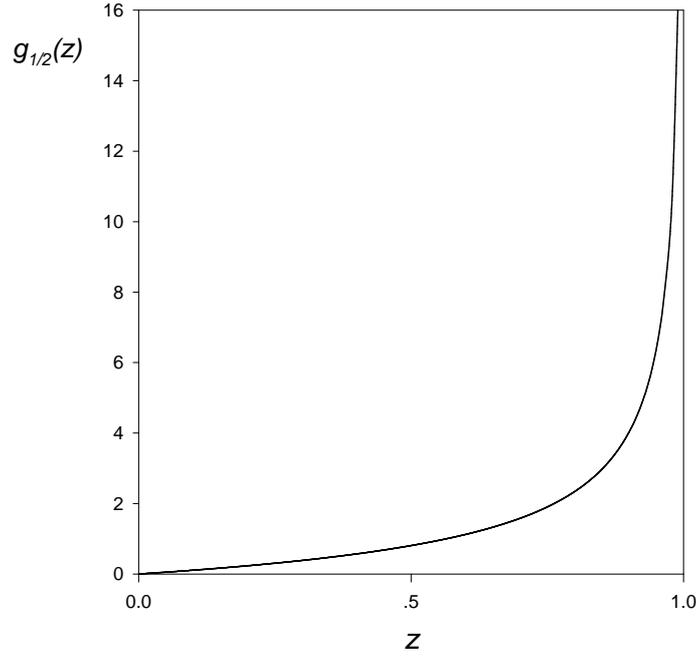


Figura 8.8:

la cual se obtiene derivando la ecuación $\lambda_T^3 \rho = g_{3/2}(z)$ con respecto a T manteniendo $\rho = cte$; la función $g_{1/2}(z)$ se muestra en la Fig.8.8. Usando esta expresión obtenemos

$$c_\rho = \begin{cases} \frac{15}{4} \frac{k_B}{\lambda_T^3} g_{5/2}(1) & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \frac{15}{4} \frac{k_B}{\lambda_T^3} g_{5/2}(z) - k_B \rho \frac{9}{4} \frac{g_{3/2}(z)}{g_{1/2}(z)} & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.53)$$

El calor específico en función de T se muestra en la Fig.8.9. Notemos que $g_{1/2}(z) \rightarrow \infty$ para $z \rightarrow 1$ y por lo tanto c_ρ es continua en el punto crítico, mientras que su derivada es discontinua. A altas temperaturas c_ρ se hace constante, tal como se espera en un gas clásico. Vemos que tanto c_ρ como s tienden a cero cuando $T \rightarrow 0$ de acuerdo con la tercera ley de la termodinámica, con una dependencia en T de la forma $T^{3/2}$.

Finalmente, la energía interna por unidad de volumen puede calcularse a partir de la Ec.(7.22) obteniéndose:

$$u = \lim_{V \rightarrow \infty} \frac{U}{V} = \begin{cases} \frac{3}{2} \frac{k_B T}{\lambda_T^3} g_{5/2}(1) & \text{si } \lambda_T^3 \rho \geq g_{3/2}(1) \\ \frac{3}{2} \frac{k_B T}{\lambda_T^3} g_{5/2}(z) & \text{si } \lambda_T^3 \rho < g_{3/2}(1) \end{cases} \quad (8.54)$$

Notemos que $u = \frac{3}{2} P$, al igual que en el gas de Maxwell-Boltzmann, aún en la región de condensación.

A altas temperaturas o bajas densidades, esto es cuando

$$\frac{\lambda_T}{v^{1/3}} \ll \left(g_{3/2}(1) \right)^{1/3}$$

tenemos que $z \rightarrow 0$ y $g_{5/2}(z) \approx g_{3/2}(z) \approx g_{1/2}(z) \sim z$. De la Ec.(8.28) obtenemos para la densidad

$$\rho \approx \frac{z}{\lambda_T^3} \quad (8.55)$$

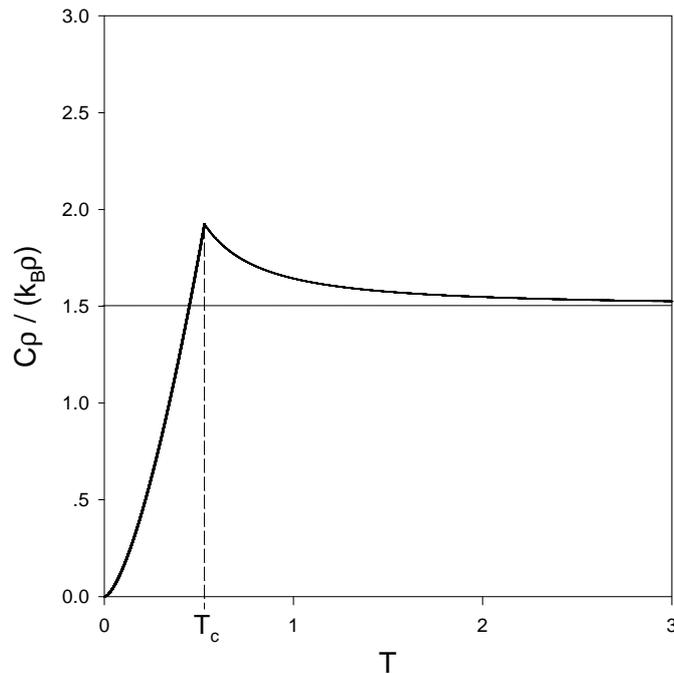


Figura 8.9: Calor específico a densidad constante en función de la temperatura para el gas ideal de Bose-Einstein.

mientras que de la Ec.(8.39) obtenemos para la presión

$$P \approx \frac{k_B T z}{\lambda_T^3} = \rho k_B T = \frac{\langle N \rangle k_B T}{V} \quad (8.56)$$

y de la Ec.(8.53) tenemos que

$$c_\rho \approx \frac{15}{4} \frac{k_B z}{\lambda_T^3} - \frac{9}{4} k_B \rho = \frac{3}{2} \rho k_B \quad (8.57)$$

Vemos que a altas temperaturas y/o bajas densidades el gas de Bose-Einstein se comporta como un gas ideal clásico, esto es, los efectos de la estadística se vuelven despreciables.

8.2. Radiación de cuerpo negro

Consideremos las propiedades de equilibrio de una cavidad de volumen V vacía de materia, a temperatura T . Un sistema de este tipo se conoce como *cuerpo negro* y puede producirse experimentalmente evacuando una cavidad en cualquier material y colocando la misma en contacto con un reservorio a temperatura T . Los átomos en las paredes de la cavidad van a emitir y absorber permanentemente radiación electromagnética, de manera que en equilibrio habrá en la cavidad radiación de diversas frecuencias con diferentes intensidades. En efecto, se observa que la distribución de energía en función de la frecuencia resulta independiente del tiempo y de las características de la cavidad. En otras palabras, las propiedades del campo electromagnético en el interior de la cavidad son independientes del tiempo y de la historia de la muestra y por lo tanto este sistema se encuentra en equilibrio termodinámico.

Si bien hemos desarrollado el formalismo de la mecánica estadística para sistemas de partículas, no existe en principio impedimento alguno para extenderlo al campo electromagnético. Lo primero que debemos preguntarnos es cuáles son los microestados accesibles del campo electromagnético. Ya hemos tratado un sistema semejante en el tratamiento cuántico de un sistema de partículas en una

caja, esto es ondas de materia. Al igual que en dicho caso, los microestados del presente sistema en equilibrio corresponden a ondas estacionarias, esto es, a los modos del campo electromagnético en la cavidad. Sin embargo, para que podamos aplicar el formalismo de la mecánica estadística debemos formular dichas soluciones de las ecuaciones de Maxwell en términos de variables canónicas. En otras palabras, debemos expresar las ecuaciones del campo en términos de coordenadas y momentos generalizados que satisfagan las ecuaciones de Hamilton.

Si el volumen de la cavidad es suficientemente grande, puede verse a partir de consideraciones termodinámicas, que las propiedades termodinámicas de la radiación en la cavidad son independientes de la naturaleza de la cavidad (forma, composición química, etc). De esta manera, podemos elegir para el campo de radiación las condiciones de contorno que nos resulten mas convenientes. Supongamos entonces que tenemos un cubo de lado L con condiciones de contorno periódicas.

Consideremos la energía del campo electromagnético en el interior de la cavidad:

$$\mathcal{H} = \frac{1}{8\pi} \int_V (\vec{E}^2 + \vec{H}^2) dV \quad (8.58)$$

donde los campos \vec{E} y \vec{H} son funciones de la posición y del tiempo. Los mismos obedecen la ecuaciones de Maxwell en ausencia de fuentes:

$$\nabla \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t} \quad (8.59)$$

$$\nabla \cdot \vec{E} = 0 \quad (8.60)$$

$$\nabla \times \vec{H} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t} \quad (8.61)$$

$$\nabla \cdot \vec{H} = 0 \quad (8.62)$$

Si expresamos los campos en términos del potencial vector \vec{A} y del potencial escalar ϕ como

$$\vec{H} = \nabla \times \vec{A} \quad (8.63)$$

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \nabla \phi \quad (8.64)$$

las ecuaciones (8.59) y (8.62) se satisfacen automáticamente. Notemos sin embargo que los potenciales se encuentran definidos a menos de una transformación de gauge, esto es, si introducimos nuevos potenciales

$$\vec{A} \rightarrow \vec{A} - \nabla \psi \quad \phi \rightarrow \phi + \frac{1}{c} \frac{\partial \psi}{\partial t}$$

donde ψ es una función arbitraria, obtenemos exactamente los mismos campos \vec{E} y \vec{H} . Esta arbitrariedad en la definición de \vec{A} nos permite imponer la condición de transversalidad

$$\nabla \cdot \vec{A} = 0. \quad (8.65)$$

conocida como gauge de Coulomb. Dado que no hay cargas en la cavidad podemos tomar $\phi = 0$. Así, la ecuación (8.60) se satisface automáticamente y la Ec.(8.61) nos dá la ecuación de ondas para el potencial vector

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = 0 \quad (8.66)$$

La solución general de la Ec.(8.66) para la caja con condiciones de contorno periódicas tiene la forma

$$\vec{A} = \sum_{\vec{k}} \left[\vec{a}_{\vec{k}} \exp \left(i\omega(\vec{k})t + i\vec{k} \cdot \vec{r} \right) + c.c. \right] \quad (8.67)$$

donde las componentes del vector de onda vienen dadas por $k_i = 2\pi l_i/L$ con $i = x, y, z$ y $l_i = 0, \pm 1, \pm 2$; la frecuencia satisface la relación de dispersión

$$\omega(\vec{k}) = c|\vec{k}| \quad (8.68)$$

y *c.c.* indica el complejo conjugado del término anterior. Los coeficientes $\vec{a}_{\vec{k}}$ se determinan a partir de las condiciones iniciales. La condición de transversalidad (8.65) implica que

$$\vec{k} \cdot \vec{a}_{\vec{k}} = 0. \quad (8.69)$$

Los campos \vec{E} y \vec{H} vienen dados entonces por

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} = \sum_{\vec{k}} \left[-ik \vec{a}_{\vec{k}} \exp \left(i\omega(\vec{k})t + i\vec{k} \cdot \vec{r} \right) + c.c. \right] \quad (8.70)$$

$$\vec{H} = \nabla \times \vec{A} = \sum_{\vec{k}} \left[i(\vec{k} \times \vec{a}_{\vec{k}}) \exp \left(i\omega(\vec{k})t + i\vec{k} \cdot \vec{r} \right) + c.c. \right] \quad (8.71)$$

esto es, tenemos un desarrollo en ondas planas de los campos. Usando la propiedad

$$\int_V \exp \left[i(\vec{k} - \vec{k}') \cdot \vec{r} \right] dV = V \delta_{\vec{k}, \vec{k}'} \quad (8.72)$$

es fácil demostrar que

$$\int_V E^2 dV = \sum_{\vec{k}} \left[-Vk^2 \vec{a}_{\vec{k}} \cdot \vec{a}_{-\vec{k}} \exp \left(2i\omega(\vec{k})t \right) - Vk^2 \vec{a}_{\vec{k}}^* \cdot \vec{a}_{-\vec{k}}^* \exp \left(-2i\omega(\vec{k})t \right) + 2Vk^2 \vec{a}_{\vec{k}} \cdot \vec{a}_{\vec{k}}^* \right] \quad (8.73)$$

De la misma manera usando las propiedades del producto vectorial mixto y la transversalidad de los campos, se puede demostrar que

$$\int_V H^2 dV = \sum_{\vec{k}} \left[Vk^2 \vec{a}_{\vec{k}} \cdot \vec{a}_{-\vec{k}} \exp \left(2i\omega(\vec{k})t \right) + Vk^2 \vec{a}_{\vec{k}}^* \cdot \vec{a}_{-\vec{k}}^* \exp \left(-2i\omega(\vec{k})t \right) + 2Vk^2 \vec{a}_{\vec{k}} \cdot \vec{a}_{\vec{k}}^* \right] \quad (8.74)$$

Por lo tanto tenemos que

$$\mathcal{H} = \frac{1}{8\pi} \int_V \left(\vec{E}^2 + \vec{H}^2 \right) dV = \frac{V}{2\pi} \sum_{\vec{k}} k^2 \vec{a}_{\vec{k}} \cdot \vec{a}_{\vec{k}}^* \quad (8.75)$$

8.2.1. Solución clásica

Notemos que la expresión (8.75) para la energía no es un Hamiltoniano, ya que la misma no se encuentra expresada en términos de variables canónicas, esto es, coordenadas y momentos generalizados que satisfagan las ecuaciones de Hamilton. Así, no podemos aplicar en la presente forma el formalismo de la mecánica estadística clásica.

Definimos entonces las coordenadas generalizadas:

$$\vec{Q}_{\vec{k}}(t) = \alpha \left[\vec{a}_{\vec{k}} \exp \left(i\omega(\vec{k})t \right) + \vec{a}_{\vec{k}}^* \exp \left(-i\omega(\vec{k})t \right) \right] \quad (8.76)$$

donde α es una constante real a determinar, y los momentos generalizados

$$\vec{P}_{\vec{k}}(t) = \frac{d}{dt} \vec{Q}_{\vec{k}}(t) = \alpha \left[i\omega(\vec{k}) \vec{a}_{\vec{k}} \exp(i\omega(\vec{k})t) - i\omega(\vec{k}) \vec{a}_{\vec{k}}^* \exp(-i\omega(\vec{k})t) \right] \quad (8.77)$$

De las Ecs.(8.76) y (8.77) podemos expresar

$$\vec{a}_{\vec{k}} = \frac{1}{2\alpha} e^{-i\omega(\vec{k})t} \left(\vec{Q}_{\vec{k}} - \frac{i}{\omega(\vec{k})} \vec{P}_{\vec{k}} \right) \quad (8.78)$$

Reemplazando la Ec.(8.78) en (8.75) tenemos que

$$\mathcal{H} = \frac{V}{8\pi\alpha^2 c^2} \sum_{\vec{k}} \left[\vec{P}_{\vec{k}}^2 + \omega(\vec{k})^2 \vec{Q}_{\vec{k}}^2 \right] \quad (8.79)$$

Es facil ver que eligiendo

$$\alpha = \sqrt{\frac{V}{4\pi c^2}} \quad (8.80)$$

las coordenadas y momentos generalizados (8.76) y (8.77) satisfacen las ecuaciones de Hamilton:

$$\frac{d}{dt} \vec{Q}_{\vec{k}} = \frac{\partial \mathcal{H}}{\partial \vec{P}_{\vec{k}}} \quad \frac{d}{dt} \vec{P}_{\vec{k}} = -\frac{\partial \mathcal{H}}{\partial \vec{Q}_{\vec{k}}} \quad (8.81)$$

Ahora si, al disponer de una forma canónica podemos aplicar el formalismo de la mecánica estadística clásica. Es importante notar que la transversalidad de los campos impone que

$$\vec{Q}_{\vec{k}} \cdot \vec{k} = \vec{P}_{\vec{k}} \cdot \vec{k} = 0 \quad (8.82)$$

esto es, los vectores $\vec{Q}_{\vec{k}}$ y $\vec{P}_{\vec{k}}$ son normales a la dirección de propagación de la onda dada por el vector \vec{k} , y por lo tanto el Hamiltoniano (8.79) puede escribirse como

$$\mathcal{H} = \frac{1}{2} \sum_{\vec{k},j} \left[P_{\vec{k},j}^2 + \omega(\vec{k})^2 Q_{\vec{k},j}^2 \right] = \sum_{\vec{k},j} H_{\vec{k},j} \quad (8.83)$$

donde el índice $j = 1, 2$ corresponde a las dos diferentes direcciones de polarización lineal de la onda y $H_{\vec{k},j}$ es el Hamiltoniano de un oscilador armónico unidimensional de frecuencia $\omega(\vec{k}) = ck$. Dado que estamos a temperatura constante, podemos trabajar en el ensamble canónico. La función partición tiene entonces la forma

$$Z = \prod_{\vec{k},j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dQ_{\vec{k},j} dP_{\vec{k},j} e^{-\beta H_{\vec{k},j}} \quad (8.84)$$

Usando

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (8.85)$$

tenemos que

$$\ln Z = \sum_{\vec{k},j} \ln \left(\frac{2\pi}{\beta kc} \right) \quad (8.86)$$

de donde obtenemos la energía interna:

$$\frac{U}{V} = -\frac{1}{V} \frac{\partial}{\partial \beta} \ln Z = \frac{1}{V} \sum_{\vec{k},j} \frac{1}{\beta} = 2 \frac{1}{(2\pi)^3} \frac{1}{\beta} \int d^3k \rightarrow \infty \quad (8.87)$$

Esta divergencia se conoce como *catástrofe ultravioleta*. Sea $u(\nu)$ la densidad espectral de energía, esto es, la cantidad de energía por unidad de volumen del campo electromagnético con frecuencia entre ν y $\nu + d\nu$. Tenemos entonces que

$$\frac{U}{V} = \int_0^\infty u(\nu) d\nu \quad (8.88)$$

De la Ec.(8.87) podemos escribir, al menos formalmente,

$$\frac{U}{V} = \frac{1}{4\pi^3} k_B T \int_0^\infty 4\pi k^2 dk = \frac{8\pi}{c^3} k_B T \int_0^\infty \nu^2 d\nu \quad (8.89)$$

donde

$$\nu = \frac{1}{2\pi} \omega(\vec{k}) = \frac{ck}{2\pi} \quad (8.90)$$

Comparando las ecuaciones anteriores obtenemos:

$$u(\nu) = \frac{8\pi}{c^3} k_B T \nu^2 \quad (8.91)$$

Esta ecuación se conoce como *ley de Rayleigh-Jeans*.

8.2.2. Solución cuántica: la ley de radiación de Plank

Lo que acabamos de ver fue el comienzo de la mecánica cuántica. La densidad espectral de energía había sido medida con gran precisión a fines del siglo XIX y, excepto a bajas frecuencias, resultaba bien diferente a la ley de Rayleigh-Jeans. Esto implicaba que algo estaba equivocado en (a) la mecánica estadística ó (b) el electromagnetismo clásico. En 1900 Planck asumió que la falla estaba en el electromagnetismo, dando comienzo a la mecánica cuántica.

La cuantificación del Hamiltoniano (8.83) da como resultado un Hamiltoniano de la forma

$$\mathcal{H} = \sum_{\vec{k},j} \hbar\omega(\vec{k}) n_{\vec{k},j} \quad (8.92)$$

donde $j = 1, 2$ corresponde a las dos polarizaciones y $n_{\vec{k},j} = 0, 1, \dots$. La función de partición canónica puede entonces escribirse como

$$Z = \prod_{\vec{k},j} Z_{\vec{k},j} \quad (8.93)$$

(notemos que el número de osciladores en este caso es infinito, aún en el caso de un volumen finito) donde

$$Z_{\vec{k},j} = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(\vec{k})n} = \frac{1}{1 - e^{-\beta\hbar\omega(\vec{k})}} \quad (8.94)$$

Así

$$\ln Z = \sum_{\vec{k},j} \ln Z_{\vec{k},j} = -2 \sum_{\vec{k},j} \ln [1 - e^{-\beta\hbar\omega(\vec{k})}] \quad (8.95)$$

y la energía interna resulta

$$U = -\frac{\partial}{\partial\beta} \ln Z = 2\hbar c \sum_{\vec{k}} \frac{k}{e^{\beta\hbar ck} - 1} \quad (8.96)$$

En el límite termodinámico, reemplazando

$$\sum_{\vec{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3k$$

obtenemos:

$$U(T, V) = \frac{V\hbar c}{\pi^2} \int_0^\infty \frac{k^3}{e^{\beta\hbar\omega(\vec{k})} - 1} dk \quad (8.97)$$

De la Ec.(8.97) tenemos

$$\frac{U}{V} = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3}{e^{\beta h\nu} - 1} d\nu \quad (8.98)$$

de donde obtenemos la *ley de radiación de Planck*:

$$u(\nu, T) = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{\beta h\nu} - 1} \quad (8.99)$$

En el límite de bajas frecuencias $\nu \rightarrow 0$ la Ec.(8.99) reproduce la ley de Rayleigh-Jeans

$$u(\nu) \sim \frac{8\pi}{c^3} k_B T \nu^2, \quad (8.100)$$

sin embargo, la integral (8.98) es finita. De hecho, haciendo el cambio de variable $x = \beta h\nu$ obtenemos la *ley de Stefan-Boltzmann*:

$$\frac{U}{V} = \frac{8\pi}{(hc)^3} (k_B T)^4 \int_0^\infty \frac{x^3}{e^x - 1} dx = \sigma T^4 \quad (8.101)$$

El calor específico a volumen constante es por lo tanto

$$c_v = 4\sigma T^3 \quad (8.102)$$

que se anula para $T \rightarrow 0$ de acuerdo con la tercera ley de la termodinámica.

8.2.3. El gas de fotones

Los resultados anteriores pueden ser derivados a partir de una interpretación bastante mas interesante. En 1928 P. Dirac suirió que el campo electromagnético podía ser cuantizado tratando las variables canónicas clásicas $Q_{\vec{k},j}$ y $P_{\vec{k},j}$ como operadores que obedecen las relaciones de conmutación

$$[Q_{\vec{k},j}, P_{\vec{k}',j'}] = i\hbar \delta_{\vec{k},\vec{k}'} \delta_{j,j'}$$

$$[Q_{\vec{k},j}, Q_{\vec{k}',j'}] = [P_{\vec{k},j}, P_{\vec{k}',j'}] = 0$$

Dirac también introdujo los operadores

$$a_{\vec{k},j} = \frac{\omega(\vec{k})}{2\hbar} Q_{\vec{k},j} + \frac{i}{\omega(\vec{k})} P_{\vec{k},j} \quad (8.103)$$

y sus hermitianos conjugados

$$a_{\vec{k},j}^\dagger = \frac{\omega(\vec{k})}{2\hbar} Q_{\vec{k},j} - \frac{i}{\omega(\vec{k})} P_{\vec{k},j} \quad (8.104)$$

los cuales satisfacen las relaciones de conmutación

$$\left[a_{\vec{k},j}^-, a_{\vec{k}',j'}^\dagger \right] = \delta_{\vec{k},\vec{k}'} \delta_{j,j'} \quad (8.105)$$

$$\left[a_{\vec{k},j}^\dagger, a_{\vec{k}',j'}^\dagger \right] = \left[a_{\vec{k},j}^-, a_{\vec{k}',j'}^- \right] = 0 \quad (8.106)$$

Reemplazando entonces en el Hamiltoniano (8.83) se obtiene el Hamiltoniano cuantizado:

$$H = \sum_{\vec{k},j} \hbar\omega(\vec{k}) \left(a_{\vec{k},j}^\dagger a_{\vec{k},j} + \frac{1}{2} \right) \quad (8.107)$$

el cual incluye el término de punto cero (irrelevante para la Mecánica Estadística, ya que solo representa una constante aditiva en la energía). Es posible ver a partir de las relaciones de conmutación (8.105)-(8.106) que los operadores $a_{\vec{k},j}^\dagger, a_{\vec{k},j}^-$ (hermitianos) tienen autovalores enteros $n_{\vec{k},j} = 0, 1, 2, \dots$ (comparar con Eq.(8.92)). Mas aún, un análisis detallado muestra que los autoestados simultáneos del conjunto anterior de operadores corresponde exactamente a los autoestados simetrizados de N partículas independientes con energías de una partícula $\hbar\omega(\vec{k})$ y $N = \sum_{\vec{k},j} n_{\vec{k},j}$. Los números cuánticos $n_{\vec{k},j}$ asociados a cada oscilador armónico pueden por lo tanto interpretarse como números de ocupación correspondientes a los estados de una partícula con energía $\hbar\omega(\vec{k})$.

Así, el campo electromagnético puede interpretarse como compuesto por cierto tipo de partículas, a las cuales se denomina **fotones**, las cuales obedecen la estadística de Bose-Einstein. Estas partículas tienen momento lineal $\vec{p} = \hbar\vec{k}$ y se mueven a la velocidad de la luz c , lo cual requiere que tengan masa en reposo nula. De esta manera, la energía resulta consistente con la expresión relativista $\varepsilon = \sqrt{c^2p^2 + m^2c^4} = cp$. Los fotones son partículas de spin 1 (a pesar de que por la condición de transversalidad solo pueden asumir dos valores, correspondientes a las diferentes polarizaciones). Otra característica que surge de la teoría es que el número de fotones no se conserva, ya que los mismos pueden ser emitidos o absorbidos por átomos. Así, el número de fotones no está definido, ni siquiera en valor medio, lo cual implica que el potencial químico sea $\mu = 0$. De esta manera, la función gran partición viene dada por la Ec.(8.2) tomando $\mu = 0$:

$$\Omega(T, V) = 2k_B T \sum_{\vec{k}} \ln \left(1 - e^{-\beta\hbar\omega(\vec{k})} \right) \quad (8.108)$$

El número medio de fotones con momento $\hbar\vec{k}$, independientemente de la polarización, es

$$\langle n_{\vec{k}} \rangle = \frac{2}{e^{\beta\hbar\omega(\vec{k})} - 1} \quad (8.109)$$

donde el factor 2 proviene de las dos posibles polarizaciones. La energía interna puede obtenerse directamente de

$$U(T, V) = \sum_{\vec{k}} \hbar\omega(\vec{k}) \langle n_{\vec{k}} \rangle \quad (8.110)$$

Reemplazando la Ec.(8.109) en la ecuación anterior reobtenemos la expresión (8.96). También podemos calcular la presión de radiación en la cavidad a partir de

$$P = -\frac{\Omega}{V} = -\frac{k_B T}{\pi^2} \int_0^\infty k^2 \ln \left(1 - e^{-\beta\hbar ck} \right) dk \quad (8.111)$$

Integrando por partes la Ec.(8.111) y comparando con la Ec.(8.97) obtenemos la ecuación de estado

$$P = \frac{U}{3V} \quad (8.112)$$

Notemos que en las funciones termodinámicas no aparece ninguna singularidad, esto es, el gas de fotones no presenta condensación de Bose. Físicamente esto se debe a que el número de fotones no se conserva, lo cual hace que los fotones desaparezcan en lugar de condensar.

8.3. Calor específico de los sólidos revisado: el gas de fonones

Si comparamos el Hamiltoniano (8.83) para el campo electromagnético en una cavidad con el Hamiltoniano (6.119) para las oscilaciones de un cristal, vemos que la situación es entéramente análoga. En este último caso, la cuantificación del campo de vibraciones del cristal resulta equivalente a un gas de bosones llamados **fonones**. Así, el calor específico de los sólidos puede ser derivado en el ensemble gran canónico, considerando un gas de bosones con potencial químico cero (el número de fonones tampoco se conserva) y las relaciones de dispersión apropiadas (ver capítulo 6). Fonones y fotones son llamados *cuasi partículas*.

Capítulo 9

Gas ideal de Fermi-Dirac

Los fermiones son partículas de spin semi-entero. Supongamos el caso mas simple de spin $1/2$, esto es, partículas para las cuales $S_z = \pm\hbar/2$, estados que vamos a denotar por $\sigma = \uparrow, \downarrow$. Denotemos por l a cada estado asociado a los números cuánticos $l = (\vec{k}, \sigma)$. En ausencia de campos electromagnéticos, la energía no depende de σ . Así, cada estado $\varepsilon_{\vec{k}}$ podrá alojar hasta dos partículas con spines opuestos, sin violar el principio de exclusión. De esta manera, la función gran partición (7.34) para el gas ideal de Fermi-Dirac puede escribirse como

$$\mathcal{Z}_{FD}(T, V, \mu) = \prod_{\vec{k}} \left(\sum_{n_{\vec{k},\uparrow}=0}^1 e^{-\beta n_{\vec{k},\uparrow}(\varepsilon_{\vec{k}} - \mu)} \sum_{n_{\vec{k},\downarrow}=0}^1 e^{-\beta n_{\vec{k},\downarrow}(\varepsilon_{\vec{k}} - \mu)} \right) = \prod_{\vec{k}} \left(1 + e^{-\beta(\varepsilon_{\vec{k}} - \mu)} \right)^2 \quad (9.1)$$

Para un gas de fermiones de spin s genérico, la potencia 2 de la expresión anterior se reemplaza por la degeneración $g = 2s + 1$. Así

$$\mathcal{Z}_{FD}(T, V, \mu) = \prod_{\vec{k}} \left(1 + e^{-\beta(\varepsilon_{\vec{k}} - \mu)} \right)^g \quad (9.2)$$

El potencial gran canónico resulta entonces

$$\Omega_{FD}(T, V, \mu) = -k_B T \ln \mathcal{Z}_{FD}(T, V, \mu) = -k_B T g \sum_{\vec{k}} \ln \left(1 + e^{-\beta(\varepsilon_{\vec{k}} - \mu)} \right) \quad (9.3)$$

de donde podemos calcular

$$\langle N \rangle = - \left(\frac{\partial \Omega_{FD}}{\partial \mu} \right)_{T,V} = \sum_{\vec{k}} \left(\frac{g e^{-\beta(\varepsilon_{\vec{k}} - \mu)}}{1 + e^{-\beta(\varepsilon_{\vec{k}} - \mu)}} \right) = \sum_{\vec{k}} \left(\frac{g}{e^{\beta(\varepsilon_{\vec{k}} - \mu)} + 1} \right) = \sum_{\vec{k}} \langle n_{\vec{k}} \rangle \quad (9.4)$$

de donde

$$\langle n_{\vec{k}} \rangle = \left(\frac{g}{e^{\beta(\varepsilon_{\vec{k}} - \mu)} + 1} \right) = \left(\frac{gz}{e^{\beta\varepsilon_{\vec{k}}} + z} \right) \quad (9.5)$$

Para el gas ideal de Fermi-Dirac el número medio de partículas nunca es divergente. El potencial químico puede tomar cualquier valor $-\infty < \mu < \infty$ y por lo tanto $0 \leq z < \infty$. Calculemos las funciones termodinámicas. Reemplazando las sumas por integrales en el límite termodinámico tenemos

$$P = -\frac{\Omega_{FD}}{V} = \frac{4\pi k_B T g}{(2\pi)^3} \int_0^\infty k^2 \ln(1 + z e^{-\beta\hbar^2 k^2/2m}) dk \quad (9.6)$$

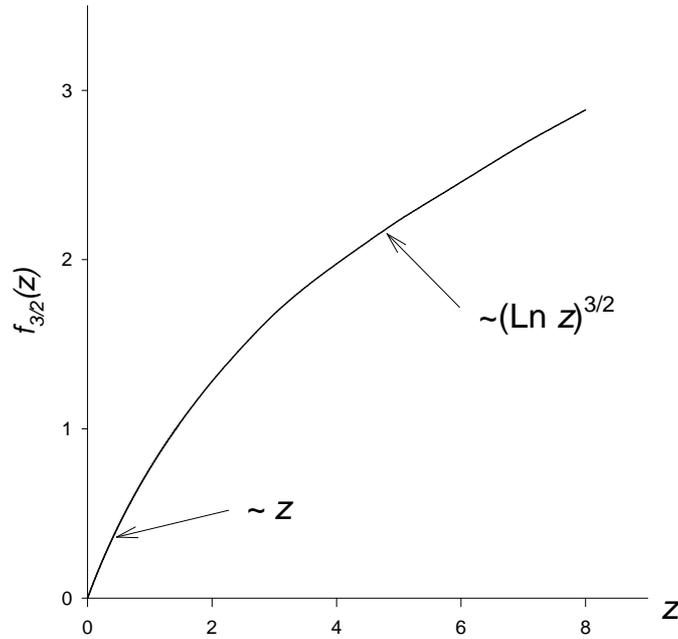


Figura 9.1: Funcion $f_{3/2}(z)$.

$$\langle N \rangle = \frac{4\pi V g}{(2\pi)^3} \int_0^\infty k^2 \left(\frac{z}{e^{\beta \hbar^2 k^2 / 2m} + z} \right) dk \quad (9.7)$$

Haciendo el cambio de variable usual $x = \hbar k \sqrt{\beta / 2m}$ tenemos que

$$P = -\frac{\Omega_{FD}}{V} = \frac{k_B T g}{\lambda_T^3} f_{5/2}(z) \quad (9.8)$$

$$\langle N \rangle = \frac{gV}{\lambda_T^3} f_{3/2}(z) \quad (9.9)$$

donde

$$f_{5/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 \ln [1 + z e^{-x^2}] = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{z^k}{k^{5/2}} \quad (9.10)$$

$$f_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 \left(\frac{z}{e^{x^2} + z} \right) dx = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{z^k}{k^{3/2}} \quad (9.11)$$

La funcion $f_{3/2}$ se muestra en la Fig.9.1.

Las ecuaciones de estado del gas de Fermi-Dirac se determinan eliminando z en función de ρ y T de la ecuación:

$$\lambda_T^3 \rho = g f_{3/2}(z) \quad (9.12)$$

Los distintos comportamientos de interés ocurren para valores de $z \ll 1$ y $z \gg 1$. Para valores pequeños de z tenemos del desarrollo en serie (9.11)

$$f_{3/2}(z) = z - \frac{z^2}{2^{3/2}} + \dots \quad (9.13)$$

Para valores grandes de z podemos obtener un desarrollo asintótico de la siguiente manera. Llamemos

$$\nu \equiv \beta\mu = \ln z.$$

Entonces, usando el cambio de variable $y = x^2$,

$$f_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 \left(\frac{1}{e^{x^2-\nu} + 1} \right) dx = \frac{2}{\sqrt{\pi}} \int_0^\infty \sqrt{y} \left(\frac{1}{e^{y-\nu} + 1} \right) dy = \frac{4}{3\sqrt{\pi}} \int_0^\infty \frac{y^{3/2} e^{y-\nu}}{(e^{y-\nu} + 1)^2} dy \quad (9.14)$$

donde en el último paso hemos integrado por partes. La función

$$\frac{e^{y-\nu}}{(e^{y-\nu} + 1)^2}$$

tiene un pico simétrico en $y = \nu$ y decae rápidamente para valores de $y \neq \nu$. Así, podemos desarrollar $y^{3/2}$ en serie de Taylor alrededor de $y = \nu$:

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}} \int_0^\infty dy \frac{e^{y-\nu}}{(e^{y-\nu} + 1)^2} \left[\nu^{3/2} + \frac{3}{2} \nu^{1/2} (y - \nu) + \frac{3}{8} \nu^{-1/2} (y - \nu)^2 + \dots \right] \quad (9.15)$$

$$= \frac{4}{3\sqrt{\pi}} \int_{-\nu}^\infty dt \frac{e^t}{(e^t + 1)^2} \left[\nu^{3/2} + \frac{3}{2} \nu^{1/2} t + \frac{3}{8} \nu^{-1/2} t^2 + \dots \right] \quad (9.16)$$

Escribimos entonces

$$\int_{-\nu}^\infty = \int_{-\infty}^\infty - \int_{-\infty}^{-\nu}$$

La segunda integral es de orden $e^{-\nu}$. Así

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}} \int_{-\infty}^\infty dt \frac{e^t}{(e^t + 1)^2} \left[\nu^{3/2} + \frac{3}{2} \nu^{1/2} t + \frac{3}{8} \nu^{-1/2} t^2 + \dots \right] + \mathcal{O}(e^{-\nu}) \quad (9.17)$$

$$= \frac{4}{3\sqrt{\pi}} \left(I_0 \nu^{3/2} + \frac{3}{2} I_1 \nu^{1/2} + \frac{3}{8} I_2 \nu^{-1/2} + \dots \right) + \mathcal{O}(e^{-\nu}) \quad (9.18)$$

donde

$$I_n \equiv \int_{-\infty}^\infty dt \frac{t^n e^t}{(e^t + 1)^2} \quad (9.19)$$

El término

$$\frac{e^t}{(e^t + 1)^2}$$

es una función par de t . Así, para n impar tenemos que $I_n = 0$. Para $n = 0$ tenemos

$$I_0 = -2 \int_0^\infty dt \frac{d}{dt} \frac{1}{(e^t + 1)} = 1. \quad (9.20)$$

Para $n > 0$ par tenemos

$$I_n = -2 \left[\frac{\partial}{\partial \lambda} \int_0^\infty dt \frac{t^{n-1}}{(e^{\lambda t} + 1)} \right]_{\lambda=1} = -2 \left[\frac{\partial}{\partial \lambda} \lambda^{-n} \right]_{\lambda=1} \int_0^\infty du \frac{u^{n-1}}{(e^u + 1)} \quad (9.21)$$

$$= 2n \int_0^\infty du \frac{u^{n-1}}{(e^u + 1)} = (n-1)!(2n)(1-2^{1-n})\zeta(n) \quad (9.22)$$

donde $\zeta(n)$ es la función zeta de Riemann ($\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$, $\zeta(6) = \pi^6/945$). Por lo tanto

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}} \left[(\ln z)^{3/2} + \frac{\pi^2}{8} (\ln z)^{-1/2} + \dots \right] + \mathcal{O}(z^{-1}) \quad (9.23)$$

$$= \frac{4}{3\sqrt{\pi}} \left[\left(\frac{\mu}{k_B T} \right)^{3/2} + \frac{\pi^2}{8} \left(\frac{\mu}{k_B T} \right)^{-1/2} + \dots \right] + \mathcal{O}(e^{-\beta\mu}) \quad (9.24)$$

9.0.1. Bajas temperaturas/Altas densidades ($\lambda_T^3/v \gg 1$)

Este es el régimen en el cual los efectos cuánticos, en particular el principio de exclusión de Pauli, se vuelven relevantes. A bajas temperaturas tendremos que

$$\frac{\lambda_T^3 \rho}{g} = \frac{4}{3\sqrt{\pi}} \left[\left(\frac{\mu}{k_B T} \right)^{3/2} + \frac{\pi^2}{8} \left(\frac{\mu}{k_B T} \right)^{-1/2} \right] + \mathcal{O} \left[\left(\frac{T}{\mu} \right)^{5/2} \right] \quad (9.25)$$

y recordemos que

$$\lambda_T \equiv \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{1/2}.$$

Vemos que cuando $T \rightarrow 0$ el potencial químico tiende a un valor constante $\mu = \varepsilon_F$, que define la **energía de Fermi**:

$$\varepsilon_F \equiv \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{g} \rho \right)^{2/3} \quad (9.26)$$

Para entender el significado físico de ε_F examinemos el comportamiento del número medio $\langle n_{\vec{k}} \rangle$ de partículas con momento $\vec{p} = \hbar\vec{k}$ en las cercanías del cero absoluto:

$$\langle n_{\vec{k}} \rangle \approx \frac{g}{e^{\beta(\varepsilon_{\vec{k}} - \varepsilon_F)} + 1} \quad (9.27)$$

En el límite $T \rightarrow 0$ ($\beta \rightarrow \infty$) la exponencial en el denominador diverge para todo valor $\varepsilon_{\vec{k}} > \varepsilon_F$ y se anula para todo valor $\varepsilon_{\vec{k}} < \varepsilon_F$. Así, tenemos que

$$\langle n_{\vec{k}} \rangle_{T=0} = \begin{cases} g & (\varepsilon_{\vec{k}} < \varepsilon_F) \\ 0 & (\varepsilon_{\vec{k}} > \varepsilon_F) \end{cases} \quad (9.28)$$

El significado físico de esta expresión es claro. El principio de exclusión de Pauli no permite más de g partículas en cada autoestado de \hat{p} . En el estado fundamental del gas las partículas ocupan los estados de menor energía posible, y por lo tanto van llenando los niveles inferiores hasta alcanzar un valor finito ε_F . Así ε_F es simplemente el menor autovalor de energía de una partícula por debajo del cual hay $\langle N \rangle_{T=0}/g$ niveles ocupados. En el espacio de los momentos las partículas ocupan una esfera de radio $p_F = \sqrt{2m\varepsilon_F}$, llamado **momento de Fermi**, cuya superficie se denomina

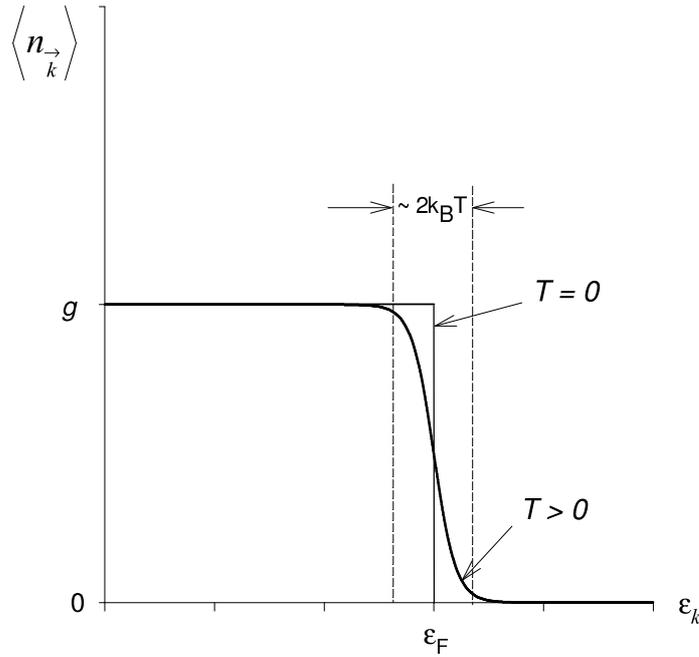


Figura 9.2: Valor medio de los números de ocupación en el gas ideal de Fermi, para una dirección \vec{k} particular en el espacio recíproco.

superficie de Fermi. Con esta interpretación podemos calcular directamente la energía de Fermi, sin tomar ningún límite. Cada estado en el espacio de momentos ocupa un volumen $(2\pi\hbar)^3/V$. Así una esfera de radio p_F contendrá

$$\frac{V}{(2\pi\hbar)^3} \frac{4}{3} \pi p_F^3 = \frac{\langle N \rangle}{g}$$

estados, de donde recuperamos la expresión (9.26). A temperaturas suficientemente bajas $\langle n_{\vec{k}} \rangle$ tiene el comportamiento que se muestra en la Fig.9.2. Únicamente una pequeña fracción de las partículas con energías del orden de $k_B T$ por debajo de ϵ_F son excitadas por encima de ϵ_F , mientras que las restantes permanecen con sus energías fijas. Por lo tanto únicamente las partículas con energías en torno de ϵ_F podrán alterar su estado ante una perturbación externa, como por ejemplo un campo eléctrico.

Para obtener las funciones termodinámicas del gas a bajas temperaturas tenemos que considerar el siguiente término en el desarrollo (9.25), el cual puede escribirse como

$$\begin{aligned} \epsilon_F^{3/2} &= \mu^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 + \dots \right] \\ &= \mu^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\epsilon_F} \right)^2 \left(\frac{\epsilon_F}{\mu} \right)^2 + \dots \right] \\ &= \mu^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\epsilon_F} \right)^2 \right] + \mathcal{O} \left[\left(\frac{k_B T}{\epsilon_F} \right)^4 \right] \end{aligned} \quad (9.29)$$

$$\frac{\mu}{\epsilon_F} \approx \frac{1}{\left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\epsilon_F} \right)^2 \right]^{2/3}} = 1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 + \dots \quad (9.30)$$

$$\mu = \frac{\nu}{\beta} = \varepsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] \quad (9.31)$$

El parámetro de la expansión es $k_B T / \varepsilon_F$. Si definimos la **temperatura de Fermi** como

$$k_B T_F \equiv \varepsilon_F \quad (9.32)$$

entonces bajas temperaturas y altas densidades significa que $T \ll T_F^1$. En este dominio se dice que el gas se encuentra *degenerado*, dado que todas las partículas tienden a estar en el estado de menor energía posible.

La energía interna viene dada por

$$U = \sum_{\vec{k}} \langle n_{\vec{k}} \rangle \varepsilon_{\vec{k}} = \frac{V}{(2\pi)^3} \frac{2\pi\hbar^2}{m} \int_0^\infty k^4 \langle n_{\vec{k}} \rangle dk \quad (9.33)$$

Integrando por partes

$$U = \frac{V\hbar^2}{4\pi^2 m} \int_0^\infty \frac{k^5}{5} \left(-\frac{\partial}{\partial k} \langle n_{\vec{k}} \rangle \right) dk = \frac{V\beta g \hbar^4}{20\pi^2 m^2} \int_0^\infty k^6 \frac{e^{\beta\varepsilon_{\vec{k}} - \nu}}{(e^{\beta\varepsilon_{\vec{k}} - \nu} + 1)^2} dk \quad (9.34)$$

Usando el cambio de variable $y = \beta\hbar^2 k^2 / 2m$ obtenemos

$$U = \frac{Vg}{20\pi^2 \hbar^3 m} (2mk_B T)^{5/2} \int_0^\infty y^{5/2} \frac{e^{y-\nu}}{(e^{y-\nu} + 1)^2} dy \quad (9.35)$$

Siguiendo el mismo procedimiento que usamos para obtener el desarrollo de $f_{3/2}$, desarrollamos $y^{5/2}$ en serie de Taylor alrededor de $y = \nu$:

$$U = \frac{Vg}{20\pi^2 \hbar^3 m} (2mk_B T)^{5/2} \int_{-\infty}^\infty dt \frac{e^t}{(e^t + 1)^2} \left[\nu^{5/2} + \frac{5}{2} \nu^{3/2} t + \frac{15}{8} \nu^{1/2} t^2 + \dots \right] \quad (9.36)$$

$$= \frac{Vg}{20\pi^2 \hbar^3 m} (2mk_B T)^{5/2} \left[\nu^{5/2} + \frac{5}{8} \pi^2 \nu^{1/2} + \dots \right] \quad (9.37)$$

Reemplazando la expresión (9.31) para ν y conservando términos hasta orden cuadrático en $k_B T / \varepsilon_F$ tenemos

$$U = \frac{Vg}{20\pi^2 \hbar^3 m} (2m\varepsilon_F)^{5/2} \left[\left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right)^{5/2} + \frac{5}{8} \pi^2 \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] \quad (9.38)$$

$$= \frac{3}{5} \langle N \rangle \varepsilon_F \left[1 + \frac{5}{12} \pi^2 \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] \quad (9.39)$$

De esta última expresión podemos calcular la forma asintótica del calor específico a bajas temperaturas:

$$C_v = \frac{1}{V} \left(\frac{\partial U}{\partial T} \right)_{V, \langle N \rangle} = \frac{\rho \pi^2}{2} \frac{k_B^2 T}{\varepsilon_F} + \dots \quad (9.40)$$

Vemos que el calor específico tiende a cero linealmente cuando $T \rightarrow 0$ y por lo tanto satisface la tercera ley de la termodinámica.

¹De la Ec.(9.26) es fácil ver que $(T/T_F)^{3/2} \propto \lambda_T^3 / v$, donde el coeficiente de proporcionalidad es de orden uno para cualquier valor razonable de g .

Finalmente, integrando por partes dos veces la Ec.(9.10) obtenemos para $f_{5/2}(z)$ la expresión:

$$f_{5/2}(z) = \frac{8}{15\pi} \int_0^\infty y^{5/2} \frac{e^{y-\nu}}{(e^{y-\nu} + 1)^2} dy \quad (9.41)$$

de donde usando el mismo procedimiento que anteriormente llegamos a la expresión

$$P = \frac{2}{5} \rho \varepsilon_F \left[1 + \frac{5}{12} \pi^2 \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right] = \frac{2}{3} \frac{U}{V} \quad (9.42)$$

Esto nos muestra que aún a $T = 0$ se necesitan paredes externas para contener un gas ideal de Fermi, ya que la presión no se anula (a diferencia del gas ideal clásico y el gas de Bose). Esta es una manifestación del principio de exclusión de Pauli, que solo permite un número g de partículas con momento cero. Todas las restantes partículas tienen momento finito y dan lugar a la presión del punto cero.

9.0.2. Altas temperaturas/Bajas densidades ($\lambda_T^3/v \ll 1$)

Usando el desarrollo (9.13) en la Ec.(9.12) tenemos

$$\frac{\lambda_T^3}{gv} = z - \frac{z^2}{2^{3/2}} + \dots \quad (9.43)$$

Esta ecuación puede invertirse desarrollando z en una serie de potencias en λ_T^3/v ; reemplazando en la ecuación anterior y comparando iguales potencias se obtiene que:

$$z = \frac{\lambda_T^3}{gv} + \frac{1}{2^{3/2}} \left(\frac{\lambda_T^3}{gv} \right)^2 + \dots \quad (9.44)$$

Para partículas sin spin $g = 1$, el primer término de esta ecuación nos da la ecuación de estado correspondiente para el gas de Maxwell-Boltzmann (7.45).

Tenemos además de la Ec.(9.8)

$$\frac{P}{k_B T} = \frac{g}{\lambda_T^3} \left(z - \frac{z^2}{2^{5/2}} + \dots \right) = \frac{gz}{\lambda_T^3} \left(1 - \frac{z}{2^{5/2}} + \dots \right) \quad (9.45)$$

y usando la Ec.(9.44) tenemos

$$\frac{Pv}{k_B T} = 1 + \frac{1}{2^{5/2}} \frac{\lambda_T^3}{gv} + \dots \quad (9.46)$$

Esto es, tenemos la ecuación de estado del gas ideal más correcciones. Vemos así que el gas de Fermi-Dirac se comporta también como el gas de Maxwell-Boltzmann a altas temperaturas y bajas densidades, tal como fue anticipado.

9.0.3. Ejemplo: calor específico electrónico de los metales

En un sólido metálico puede considerarse en una primera aproximación que los electrones de valencia se desacoplan de los iones, los cuales se localizan formando la estructura cristalina. Dichos electrones (llamados electrones de conducción) se comportan entonces como un gas de fermiones cargados negativamente, los cuales se mueven en un potencial periódico creado por los iones. Si bien podríamos esperar que la interacción coulombiana entre electrones juegue un papel importante, el efecto de la carga positiva de los iones sumado al carácter de largo alcance de la interacción coulombiana apantalla de manera efectiva el potencial que ve cada electron; así, en una primera aproximación podemos considerar los electrones como partículas independientes que se mueven en

un potencial periódico generado por los iones. En un segundo nivel de aproximación, podemos despreciar el potencial periódico, con lo cual nos resta un gas ideal de fermiones en ausencia de potencial externo. Si bien esta última aproximación, conocida como Teoría de Sommerfeld, es inadecuada para describir las propiedades de conducción eléctrica en metales, resulta una aproximación sorprendentemente buena para describir las propiedades térmicas en equilibrio de gran cantidad de metales. La aplicación entonces de los resultados de las secciones anteriores resulta inmediata.

Analicemos primero bajo qué condición debemos analizar el problema. Supongamos el caso más simple de un electrón de conducción por átomo. El volumen específico resulta entonces $v = a^3$, donde a es el parámetro de red, el cual toma un valor típico $a \sim 5\text{Å}$. Dado que la masa del electrón es $m \approx 9 \times 10^{-28}$ g y recordando que $\hbar \approx 10^{-27}$ erg.seg, a temperatura ambiente $T \sim 300$ K esto nos da una longitud de onda térmica $\lambda_T \sim 4 \times 10^{-7}$ cm. Así, $\lambda_T^3/v \sim 1000$, con lo cual aún a temperaturas relativamente altas el gas de electrones se encuentra fuertemente degenerado. En particular, esto implica que la contribución de los electrones (9.40) al calor específico es en general mucho menor al valor clásico $3/2 \rho k_B$ ($T \ll T_F$). Dado que la temperatura de Debye para la mayoría de los metales está entre 100 y 400 K, el calor específico de los mismos a temperatura ambiente está principalmente determinado por las vibraciones de los iones del cristal.

A temperaturas suficientemente bajas, no obstante, la contribución electrónica lineal en la temperatura puede ser importante comparada con el término $\propto T^3$ del modelo de Debye. Así

$$C_v \approx AT + BT^3$$

Experimentalmente suele representarse $C/T = A + BT^2$ vs. T^2 , encontrándose que los datos presentan un acuerdo excelente con una línea recta.

Capítulo 10

Mecánica Estadística y Termodinámica de sistemas magnéticos

Consideremos un sistema de N partículas con momento magnético, cuya componente en la dirección de un campo magnético externo \vec{B} viene dada por μ_i , $i = 1, \dots, N$. El Hamiltoniano del sistema viene dado entonces por

$$H = H_0 - B \sum_{i=1}^N \mu_i \quad (10.1)$$

donde $B = |\vec{B}|$ y H_0 es el Hamiltoniano del sistema en ausencia de campo externo, el cual incluye la energía cinética e interacciones entre partículas. Vamos a analizar la mecánica estadística de este sistema en el ensamble canónico. Tenemos que

$$\rho = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z} \exp \left[-\beta \left(H_0 - B \sum_{i=1}^N \mu_i \right) \right] \quad (10.2)$$

con

$$Z = \text{Tr} e^{-\beta H} = \text{Tr} \exp \left[-\beta \left(H_0 - B \sum_{i=1}^N \mu_i \right) \right] \quad (10.3)$$

Sea

$$M \equiv \left\langle \sum_{i=1}^N \mu_i \right\rangle = \frac{1}{Z} \text{Tr} \left[\left(\sum_{i=1}^N \mu_i \right) e^{-\beta H} \right] \quad (10.4)$$

la magnetización total del sistema. La entropía del sistema viene dada por

$$S = -k_B \text{Tr} \rho \ln \rho \quad (10.5)$$

reemplazando la expresión (10.2) tenemos que

$$S = \frac{1}{T} \langle H \rangle + k_B \ln Z \quad (10.6)$$

de donde podemos definir la función

$$F(T, B, V, N) \equiv -k_B T \ln Z(T, B, V, N) = \langle H \rangle - TS \quad (10.7)$$

Para un sistema de partículas *no magnéticas* la función (10.7) corresponde a la *energía libre de Helmholtz*, esto es, a la transformada de Legendre de la energía con respecto a la entropía. A que potencial termodinámico (energía libre) corresponde F en el caso magnético? Supongamos por simplicidad un sólido magnético (el caso mas frecuente), en cuyo caso las variables termodinámicas relevantes son T , B y N (el volúmen es proporcional al número de partículas). Esto es, $F = F(T, B, N)$, con lo cual F resulta el análogo a la energía libre de Gibbs en un fluido, esto es, la transformada de Legendre de la energía interna respecto de la entropía y la variable extensiva magnética relevante, la magnetización. No obstante, si definimos la energía interna como $U = \langle H \rangle$, resulta que $F = U - TS$. Esta forma indicaría que F es la transformada de Legendre de la energía respecto solo de la entropía, es decir, al análogo del potencial de Helmholtz en un fluido. Una formulación consistente con el formalismo termodinámico puede obtenerse definiendo

$$U \equiv \langle H_0 \rangle \quad (10.8)$$

Con esta definición tenemos que

$$F(T, B, N) = U - TS - BM \quad (10.9)$$

que tiene la forma de una transformada de Legendre de U con respecto a S y M . Veamos si esto es consistente. Sabemos que

$$S = - \left(\frac{\partial F}{\partial T} \right)_{B, N} \quad (10.10)$$

ya que esta ecuación es la misma que para el caso de una gas. Si nuestra interpretación es correcta debe cumplirse tambien que

$$M = - \left(\frac{\partial F}{\partial B} \right)_{T, N} \quad (10.11)$$

De la Ec.(10.7) tenemos que

$$- \left(\frac{\partial F}{\partial B} \right)_{T, V, N} = k_B T \left(\frac{\partial \ln Z}{\partial B} \right)_{T, V, N} \quad (10.12)$$

$$= \frac{1}{Z} \text{Tr} \left\{ \sum_{i=1}^N \mu_i e^{[-\beta(H_0 - B \sum_{i=1}^N \mu_i)]} \right\} \quad (10.13)$$

$$= \left\langle \sum_{i=1}^N \mu_i \right\rangle \quad (10.14)$$

de donde (10.11) es correcta. De hecho, si consideramos un sistema descrito por el Hamiltoniano H_0 en contacto con un reservorio de temperatura y de campo magnético (esto es, añadimos el vínculo de magnetización media constante), la aplicación del principio variacional de Gibbs nos lleva a la expresión (10.3) para la función partición, donde el multiplicador de Lagrange asociado al nuevo vínculo resulta proporcional a B . Esto significa que de hecho estamos trabajando en un ensamble diferente del canónico. Sin embargo, resulta habitual considerar que el Hamiltoniano del sistema es (10.1) y trabajarlo en el ensamble canónico. Los resultados son totalmente independientes de la intrerpretación. En lo que sigue adoptaremos este último punto de vista, pero conviene tener conocimiento de esta diferencia sutil, ya que en circunstancias muy particulares puede llevar a errores.

En el caso de trabajar en el ensamble gran canónico la energía libre que se obtiene resulta la transformada de Legendre de la energía interna con respecto a S , M y $\langle N \rangle$.

Finalmente la susceptibilidad magnética a campo nulo se define como

$$\chi_0 \equiv \frac{1}{V} \left(\frac{\partial M}{\partial B} \right)_{T,V,N} \Big|_{B=0} \quad (10.15)$$

Así, para campos débiles tenemos que

$$M/V \approx \chi_0 B.$$

De las Ecs.(10.3) y (10.4) tenemos que

$$\left(\frac{\partial M}{\partial B} \right)_{T,V,N} = -\frac{\beta}{Z^2} \text{Tr} \left[\left(\sum_{i=1}^N \mu_i \right) e^{-\beta H} \right] \frac{\partial Z}{\partial B} + \frac{\beta}{Z} \text{Tr} \left[\left(\sum_{i=1}^N \mu_i \right)^2 e^{-\beta H} \right] \quad (10.16)$$

$$= \beta \left\langle \left(\sum_{i=1}^N \mu_i \right)^2 \right\rangle - \frac{\beta}{Z^2} \left(\text{Tr} \left[\left(\sum_{i=1}^N \mu_i \right) e^{-\beta H} \right] \right)^2 \quad (10.17)$$

$$= \beta \left[\left\langle \left(\sum_{i=1}^N \mu_i \right)^2 \right\rangle - M^2 \right] \quad (10.18)$$

Así

$$\chi_0 = \frac{1}{k_B T V} \left[\left\langle \left(\sum_{i=1}^N \mu_i \right)^2 \right\rangle - M^2 \right]_{B=0} \quad (10.19)$$

Este resultado es un caso particular de lo que se conoce como *Teorema de Fluctuación-Disipación*, el cual establece que la respuesta de una cantidad macroscópica a una perturbación externa pequeña (esto es, χ_0 como respuesta al campo externo B) es proporcional a las fluctuaciones de dicha cantidad en equilibrio en ausencia de la perturbación. Este resultado es análogo a la relación (6.44) entre el calor específico y las fluctuaciones de energía. La expresión anterior, por otra parte, es de gran utilidad en el cálculo de χ_0 mediante simulaciones numéricas.

Capítulo 11

Magnetismo en medios materiales

Las propiedades magnéticas de la materia macroscópica están determinadas casi exclusivamente por los momentos magnéticos de los electrones de los átomos constituyentes. Los momentos magnéticos nucleares son del orden de mil veces más débiles que los de los electrones y por lo tanto pueden despreciarse. Así, los diferentes comportamientos magnéticos que se observan en la materia están estrechamente relacionados con la estadística de Fermi y por lo tanto con el principio de exclusión de Pauli. En muchos materiales estos momentos magnéticos presentan una fuerte interacción entre ellos. Esto da origen a diferentes tipos de propiedades magnéticas, siendo las más comunes el ferromagnetismo y el antiferromagnetismo.

Por otra parte, aún en el caso de espines débilmente interactuantes, en presencia de campos magnéticos externos los electrones presentan dos efectos diferentes: (a) movimiento en órbitas cuantizadas normales al campo magnético, las cuales generan acoplamientos entre los momentos angulares orbitales y el campo; (b) los espines electrónicos tienden a alinearse con el campo externo. El primer efecto da lugar al *diamagnetismo* (esto es, sistemas para los cuales $\chi < 0$) mientras que el segundo da lugar al *paramagnetismo*. Vamos a estudiar estos efectos por separado, comenzando con el *diamagnetismo* y fenómenos relacionados.

11.1. Diamagnetismo de Landau

El fenómeno de diamagnetismo está relacionado con la interacción entre el campo externo y el momento magnético asociado con el momento angular **orbital** de los electrones. Clásicamente, podríamos pensar que el campo magnético externo induce corrientes circulares, las cuales en virtud de la ley de Lenz generan momentos magnéticos opuestos al campo. Sin embargo, veremos que esto es falso. El fenómeno es netamente cuántico. Podemos ver esto a través de un modelo ultrasimplificado, en el cual vamos a ignorar el spin de los electrones así como también la interacción entre ellos; esto daría un modelo ideal de diamagnetismo en metales. El Hamiltoniano de una partícula de carga $-e$ y masa m en presencia de un campo magnético \vec{B} está dado por la expresión:

$$H = \frac{1}{2m} \left(\vec{p} + \frac{e}{c} \vec{A} \right)^2 \quad (11.1)$$

donde \vec{A} es el potencial vector asociado con el campo y c es la velocidad de la luz.

En el contexto de la mecánica clásica la presencia del campo externo en este caso no tiene ninguna influencia sobre las propiedades termodinámicas. Podemos ver fácilmente esto si escribimos la función de partición para una partícula en el espacio de las fases clásico:

$$Z_1 = \int d^3r \int d^3p \exp \left[-\frac{\beta}{2m} \left(\vec{p} + \frac{e}{c} \vec{A} \right)^2 \right] \quad (11.2)$$

La transformación $\vec{p} \rightarrow \vec{p} - (e/c)\vec{A}$ es canónica y por lo tanto tiene determinante uno. Así la función de partición es equivalente a la del gas ideal clásico y por lo tanto no hay diamagnetismo. Físicamente, esto se debe a que la fuerza de Lorentz producida por el campo externo es siempre centrípeta y por lo tanto no produce trabajo mecánico. Así, la energía de un electrón es independiente de la presencia del campo. Dado que estamos sumando sobre todos los posibles valores de los momentos, estamos considerando corrientes circulares en ambos sentidos de rotación alrededor del campo magnético con igual peso estadístico y por lo tanto los efectos se cancelan.

El diamagnetismo es un fenómeno puramente cuántico. Necesitamos entonces encontrar las soluciones de la Ec. de Schrödinger independiente del tiempo correspondiente al Hamiltoniano (11.1).

Vamos a considerar un campo magnético externo uniforme \vec{B} en la dirección del eje z . La invariancia de gauge nos da libertad de elegir de distintas maneras el potencial vector. Vamos a elegir:

$$A_x = -By \quad A_y = A_z = 0 \quad (11.3)$$

El Hamiltoniano (11.1) toma entonces la forma

$$H = \frac{1}{2m} \{ [p_x - (eB/c)y]^2 + p_y^2 + p_z^2 \} \quad (11.4)$$

Podemos resolver la ecuación de Schrödinger proponiendo para la función de onda la siguiente forma:

$$\psi(x, y, z) = e^{i(k_x x + k_z z)} f(y) \quad (11.5)$$

Tenemos que

$$H e^{i(k_x x + k_z z)} f(y) = \frac{1}{2m} \{ [\hbar k_x - (eB/c)y]^2 + p_y^2 + \hbar^2 k_z^2 \} e^{i(k_x x + k_z z)} f(y). \quad (11.6)$$

Así, la ecuación de Schrödinger $H\psi = \epsilon\psi$ nos queda

$$\left[\frac{1}{2m} p_y^2 + \frac{1}{2} m \omega_0^2 (y - y_0)^2 \right] f(y) = \epsilon' f(y) \quad (11.7)$$

donde $\epsilon' = \epsilon - \hbar^2 k_z^2 / 2m$, $\omega_0 = eB/mc$ e $y_0 = (\hbar c/eB)k_x$. Vemos que $f(y)$ satisface la ecuación de un oscilador armónico cuya frecuencia ω_0 es igual a la frecuencia de Larmor clásica, esto es, la frecuencia de rotación de una partícula clásica cargada en un campo magnético uniforme. Los autovalores de energía son entonces

$$\epsilon(k_z, n) = \frac{\hbar^2 k_z^2}{2m} + \hbar \omega_0 \left(n + \frac{1}{2} \right) \quad (11.8)$$

donde $n = 0, 1, \dots$. Estos autovalores se conocen como **niveles de Landau**. Si consideramos que el gas está contenido en una caja cúbica de lado L con condiciones de contorno periódicas entonces $k_{x,z} = 2\pi n_{x,z}/L$, $n_{x,z} = 0, \pm 1, \pm 2, \dots$. Notemos que los niveles de Landau no dependen de k_x y por lo tanto tendrán una degeneración igual al número de autovalores permitidos de k_x . Dado que y_0 depende de k_x los autovalores permitidos de k_x serán aquellos tales que $0 < y_0 < L$, esto es

$$0 < n_x < \frac{eB}{\hbar c} L^2$$

Así, $g = (eB/\hbar c)L^2$ es la degeneración de los niveles de Landau. La dependencia de la degeneración en L^2 nos dice que esta es sumamente alta, aún para campos débiles. De esta manera podemos construir nuevas autofunciones para cada nivel de Landau combinando las soluciones obtenidas para un mismo nivel. En particular, pueden construirse soluciones que son a la vez autoestados de

la componente del momento angular orbital en la dirección del campo. Esto puede verse eligiendo el siguiente gauge en lugar de (11.3):

$$\vec{A} = -\frac{1}{2} (\vec{r} \times \vec{B}) = \frac{1}{2} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} B \quad (11.9)$$

el cual satisface $\nabla \cdot \vec{A} = 0$ (gauge de Coulomb). Es simple ver (queda como ejercicio para el lector) que usando el gauge (11.9) el Hamiltoniano (11.1) adopta la forma

$$H = \frac{\hat{p}^2}{2m} + \omega_0 \hat{L}_z + \frac{m\omega_0^2}{2} (x^2 + y^2) \quad (11.10)$$

donde \hat{L}_z es la componente z del operador momento angular $\vec{L} = \vec{r} \times \hat{p}$. Este Hamiltoniano resulta la suma de tres contribuciones: una componente de energía cinética en la dirección z : $p_z^2/2m$, un acoplamiento Zeeman $\omega_0 \hat{L}_z$ y un oscilador armónico bidimensional de frecuencia ω_0 . Dado que este último tiene simetría azimutal alrededor del eje z , las autofunciones pueden ser elegidas proporcionales a las autofunciones del operador \hat{L}_z . De esta manera, los niveles de Landau pueden interpretarse como resultado de la cuantización de las órbitas clásicas circulares en el plano perpendicular al campo magnético.

Podemos entonces escribir la función gran partición utilizando el espectro (11.8) como

$$\mathcal{Z} = \prod_l (1 + ze^{-\beta\epsilon_l}) \quad (11.11)$$

donde l denota ahora el conjunto de números cuánticos $l = (k_z, n, n_x)$. Dado que la energía ϵ_l no depende de n_x tenemos que

$$\mathcal{Z} = \prod_{k_z, n} (1 + ze^{-\beta\epsilon(k_z, n)})^g \quad (11.12)$$

$$\ln \mathcal{Z} = g \sum_{n=0}^{\infty} \sum_{k_z=-\infty}^{\infty} \ln (1 + ze^{-\beta\epsilon(k_z, n)}) \quad (11.13)$$

En el límite termodinámico $L \rightarrow \infty$

$$\ln \mathcal{Z} = \frac{gL}{\pi} \sum_{n=0}^{\infty} \int_0^{\infty} \ln (1 + ze^{-\beta\epsilon(k_z, n)}) dk_z \quad (11.14)$$

donde

$$\epsilon(k_z, n) = \frac{\hbar^2 k_z^2}{2m} + \hbar\omega_0 \left(n + \frac{1}{2} \right)$$

y el potencial gran canónico resulta

$$\Omega = -\frac{gLk_B T}{\pi} \sum_{n=0}^{\infty} \int_0^{\infty} \ln (1 + ze^{-\beta\epsilon(k, n)}) dk \quad (11.15)$$

donde hemos eliminado el subíndice z de k por simplicidad. Para el número de partículas tenemos

$$\langle N \rangle = -\left(\frac{\partial \Omega}{\partial \mu} \right)_{T, B, V} = \frac{gL}{\pi} \sum_{n=0}^{\infty} \int_0^{\infty} \frac{1}{(z^{-1} e^{\beta\epsilon(k, n)} + 1)} dk \quad (11.16)$$

11.1.1. Límite de altas temperaturas: diamagnetismo

Para $T \rightarrow \infty$ ($\beta \rightarrow 0$) tenemos que $z \rightarrow 0$, ya que $\langle N \rangle$ tiene que mantenerse finito. Desarrollando entonces la Ec.(11.15) en potencias de z tenemos que

$$\begin{aligned}\Omega(T, B, V) &= -\frac{gLk_B T z}{\pi} \sum_{n=0}^{\infty} e^{-\beta \hbar \omega_0 (n+1/2)} \int_0^{\infty} e^{-\beta \hbar^2 k^2 / 2m} dk \\ &= -\frac{gLk_B T z}{\lambda_T} e^{-\beta \hbar \omega_0 / 2} \sum_{n=0}^{\infty} e^{-\beta \hbar \omega_0 n}\end{aligned}\quad (11.17)$$

donde λ_T es la longitud de onda térmica. Así

$$\Omega(T, B, V) = -\frac{gLk_B T z}{\lambda_T} \frac{e^{-\beta \hbar \omega_0 / 2}}{1 - e^{-\beta \hbar \omega_0}} = -\frac{Vk_B T z}{2\lambda_T} \frac{eB}{hc} \frac{1}{\sinh(\beta \mu_0 B)} \quad (11.18)$$

donde

$$\mu_0 \equiv \frac{\hbar e}{2mc}$$

es el magnetón de Bohr. El número medio de partículas resulta entonces

$$\langle N \rangle = \frac{Vz}{2\lambda_T} \frac{eB}{hc} \frac{1}{\sinh(\beta \mu_0 B)} \quad (11.19)$$

La magnetización resulta

$$\begin{aligned}M &= -\left(\frac{\partial \Omega}{\partial B}\right)_{T, V, \mu} \\ &= \frac{Vk_B T z}{2\lambda_T} \frac{e}{hc} \left\{ \frac{1}{\sinh(\beta \mu_0 B)} - B\beta \mu_0 \frac{\cosh(\beta \mu_0 B)}{\sinh^2(\beta \mu_0 B)} \right\} \\ &= k_B T \frac{\langle N \rangle}{B} \left[1 - \frac{\beta B \mu_0}{\tanh(\beta \mu_0 B)} \right] \\ &= -\mu_0 \langle N \rangle \mathcal{L}(\beta \mu_0 B)\end{aligned}\quad (11.20)$$

donde

$$\mathcal{L}(x) \equiv \cotgh(x) - \frac{1}{x}$$

se conoce como función de Langevin y se muestra en la Fig.11.1. Del desarrollo asintótico

$$\cotgh(x) = \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \dots$$

vemos que $\mathcal{L}(x) \sim x/3$ para $x \ll 1$. Así, a altas temperaturas y campos débiles tenemos que

$$M = -\frac{\langle N \rangle \mu_0^2 B}{3k_B T} \quad (11.21)$$

de donde finalmente la susceptibilidad resulta

$$\chi_0 = -\frac{\mu_0^2 \rho}{3k_B T} \quad (11.22)$$

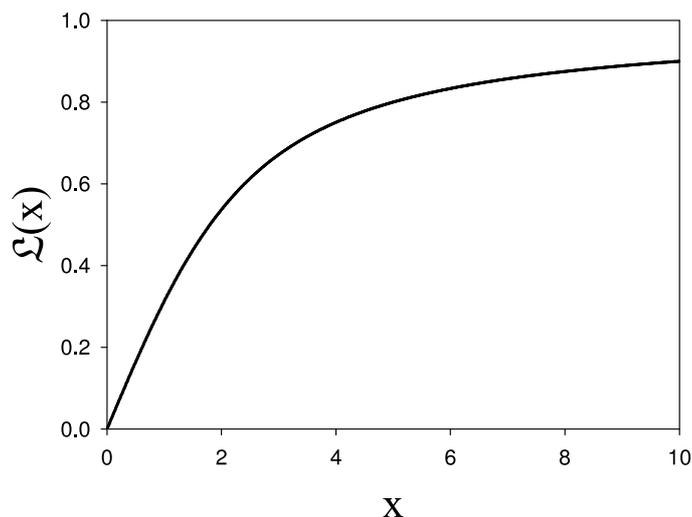


Figura 11.1: Función de Langevin.

esto es, la magnetización resulta opuesta al campo externo. Notemos que el diamagnetismo, aún siendo un efecto de origen cuántico, se manifiesta a altas temperaturas.

Finalmente, debemos notar que la susceptibilidad magnética es la derivada segunda de una energía libre respecto de un parámetro intensivo. El criterio de estabilidad termodinámica implica que la misma debe ser estrictamente positiva. El diamagnetismo parece por lo tanto violar la estabilidad termodinámica. Físicamente, esto se relaciona con el hecho de que el trabajo realizado por la fuente de campo externo sobre el sistema es BdM . Así, una susceptibilidad negativa significa que al aumentar el campo externo el sistema *entrega* energía a la fuente, lo cual es claramente inestable, ya que podemos utilizar esa energía para incrementar aún más el campo. Sin embargo, debemos recordar que BdM representa el trabajo *adicional* realizado por la fuente para establecer el campo dentro de la muestra, respecto del trabajo necesario para establecer el campo en el vacío. Así, en presencia de la muestra la fuente externa debe realizar menos trabajo que en ausencia de la misma, pero el trabajo neto es positivo. Si se incorpora esta contribución a la definición termodinámica de energía del sistema, no existe ninguna violación a la estabilidad termodinámica.

11.1.2. Límite de bajas temperaturas: efecto De Haas - Van Alphen

A bajas temperaturas no hay diamagnetismo, pero aparece otro fenómeno asociado con los niveles de Landau: en ciertos sólidos la magnetización presenta un comportamiento oscilatorio al variar el campo externo. Este fenómeno fue medido por primera vez en 1931 por W. J. De Haas y P. M. Van Alphen. La descripción correcta del efecto involucra la resolución de la Ec. de Schrödinger para una partícula en un potencial periódico en presencia de campo magnético externo, lo que constituye un problema matemático sumamente complicado. Podemos, no obstante, tener una idea cualitativa del fenómeno mediante un modelo simplificado de un gas bidimensional de electrones no interactuantes, en presencia de campo, esto es, nuestro modelo anterior pero despreciando los movimientos en la dirección z . Además, vamos a considerar el caso $k_B T \ll \hbar \omega_0$, de manera que podemos asumir $T = 0$. El problema se reduce entonces al cálculo del estado fundamental.

La idea básica por detrás del fenómeno es la siguiente: la degeneración de los niveles de Landau es independiente del nivel de energía y depende de la intensidad del campo externo. Para un campo suficientemente grande todas las partículas se encontrarán en el nivel de menor energía. Si disminuimos el campo, a partir de cierto valor dicho nivel no podrá alojar todas las partículas

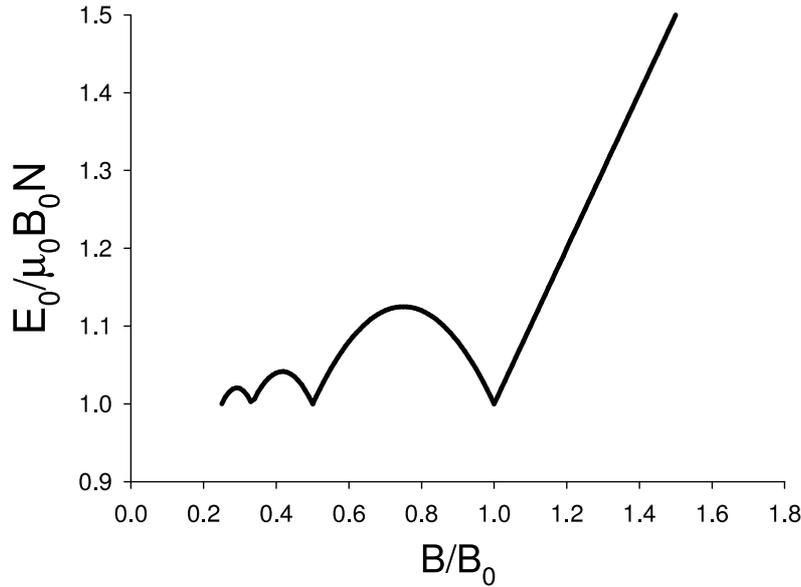


Figura 11.2: Energía a $T = 0$ de un gas de electrones sin spin en función del campo magnético aplicado.

y algunas se verán forzadas a saltar a niveles de mayor energía. Esto producirá cambios en las propiedades magnéticas del sistema.

Podemos escribir los niveles de Landau y su degeneración de la siguiente manera

$$\epsilon(n) = 2\mu_0 B \left(n + \frac{1}{2} \right) \quad (11.23)$$

$$g = NB/B_0 \quad B_0 = \rho hc/e \quad (11.24)$$

donde $\rho = N/L^2$. El campo B_0 es el valor de B por encima del cual un único nivel puede alojar a las N partículas.

Si $B/B_0 > 1$ todas las partículas se encontrarán en el menor nivel de Landau $n = 0$ y la energía del estado fundamental será

$$E_0 = N\mu_0 B \quad (11.25)$$

Si $B/B_0 < 1$ entonces algunas de las partículas ocuparán niveles de mayor energía. Por ejemplo, si $1/2 < B/B_0 < 1$, el nivel $n = 0$ estará lleno y el nivel $n = 1$ parcialmente ocupado. Para $B/B_0 = 1/2$ la degeneración de un nivel será igual a la mitad de partículas. Así, para $B/B_0 < 1/2$ los dos niveles inferiores estarán completos y el nivel con $n = 2$ estará parcialmente ocupado.

Supongamos que B es tal que todos los niveles hasta $n = k$ se encuentran completos y el $n = k + 1$ está parcialmente ocupado. La condición para B es

$$(k + 1)g < N < (k + 2)g \quad (11.26)$$

ó

$$\frac{1}{(k + 2)} < \frac{B}{B_0} < \frac{1}{(k + 1)} \quad (11.27)$$

Para B en este intervalo tenemos

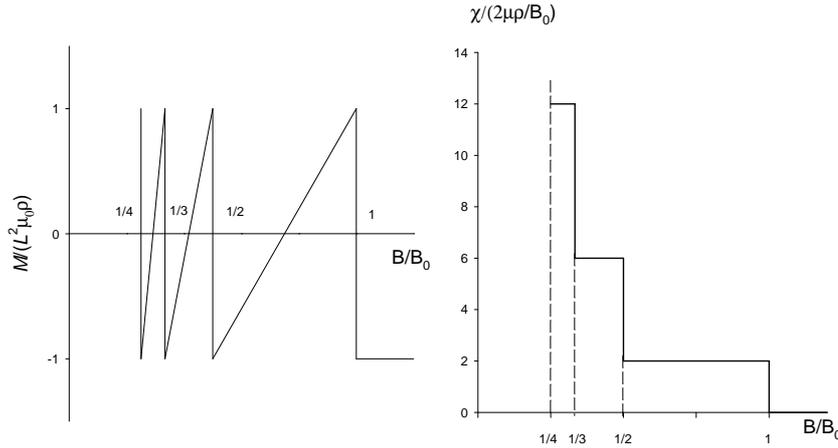


Figura 11.3: Efecto De Haas - Van Alphen.

$$E_0 = g \sum_{i=0}^k \epsilon(i) + [N - (k+1)g] \epsilon(k+1) \quad (11.28)$$

Reemplazando las Ecs.(11.23) y (11.24) tenemos, luego de cierta álgebra,

$$\frac{E_0}{N} = \mu_0 B \left[2k + 3 - (k+1)(k+2) \left(\frac{B}{B_0} \right) \right] \quad (11.29)$$

Introduciendo el parámetro

$$x \equiv \frac{B}{B_0}$$

podemos resumir

$$\frac{E_0(B)}{N} = \begin{cases} \mu_0 B_0 x & x > 1 \\ \mu_0 B_0 x [2k + 3 - (k+1)(k+2)x] & \frac{1}{k+2} < x < \frac{1}{k+1}, \quad k = 0, 1, \dots \end{cases} \quad (11.30)$$

Tenemos así dos efectos que se compensan: al aumentar el campo magnético la energía de cada nivel aumenta, pero por otro lado al aumentar la degeneración los niveles de mayor energía se despueblan. Esto da lugar a un comportamiento oscilatorio de la energía, como se ilustra en la Fig.11.2. La magnetización por unidad de área y la susceptibilidad magnéticas resultan entonces

$$\begin{aligned} \frac{M}{L^2} &= -\rho \left(\frac{\partial(E_0/N)}{\partial B} \right)_{L,N} \\ &= \begin{cases} -\mu_0 \rho & x > 1 \\ \mu_0 \rho [2(k+1)(k+2)x - (2k+3)] & \frac{1}{k+2} < x < \frac{1}{k+1}, \end{cases} \end{aligned} \quad (11.31)$$

$$\begin{aligned}\chi &= \rho \left(\frac{\partial(M/L^2)}{\partial B} \right)_{L,N} \\ &= \begin{cases} 0 & x > 1 \\ 2\mu_0\rho/B_0(k+1)(k+2) & \frac{1}{k+2} < x < \frac{1}{k+1} \end{cases}\end{aligned}\quad (11.32)$$

y se muestran en la Fig.11.3

11.2. Paramagnetismo de Pauli

Consideremos ahora el efecto del acoplamiento directo entre los spines electrónicos y el campo externo, despreciando el acoplamiento del momento angular orbital, esto es, un sistema descrito por el Hamiltoniano:

$$H = \sum_{i=1}^N \left[\frac{1}{2m} p_i^2 - \mu_0 \vec{B} \cdot \vec{\sigma}_i \right] \quad (11.33)$$

donde μ_0 es el magnetón de Bohr y $\vec{\sigma}$ son las matrices de Pauli. Para un único electrón con el campo orientado en la dirección del eje z tenemos

$$H_1 = \frac{1}{2m} p^2 - \mu_0 B \sigma_z \quad (11.34)$$

y por lo tanto el espectro de una partícula es

$$\epsilon(\vec{p}, s) = \frac{1}{2m} p^2 - \mu_0 B s \quad (11.35)$$

donde $s = \pm 1$. El potencial gran canónico resulta entonces:

$$\begin{aligned}\Omega &= -k_B T \sum_{\vec{p}} \sum_{s=\pm 1} \ln \left(1 + z e^{-\beta \epsilon(\vec{p}, s)} \right) \\ &= -k_B T \sum_{\vec{p}} \sum_{s=\pm 1} \ln \left\{ 1 + z \exp \left(-\frac{\beta}{2m} p^2 + \beta \mu_0 B s \right) \right\}\end{aligned}\quad (11.36)$$

Para partículas en una caja de volumen $V = L^3$, $\vec{p} = \hbar \vec{k}$, donde $k_i = 2\pi n_i/L$ y $n_i = 0, \pm 1, \pm 2, \dots$. En el límite termodinámico:

$$\begin{aligned}\Omega &= -\frac{4\pi k_B T V}{(2\pi)^3} \int_0^\infty k^2 \sum_{s=\pm 1} \ln \left\{ 1 + z \exp \left(-\frac{\beta}{2m} \hbar^2 k^2 + \beta \mu_0 B s \right) \right\} dk \\ &= -\frac{k_B T V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \epsilon^{1/2} \sum_{s=\pm 1} \ln \left\{ 1 + z e^{-\beta \epsilon + \beta \mu_0 B s} \right\} d\epsilon\end{aligned}\quad (11.37)$$

donde hemos usado el cambio de variable $\epsilon = \hbar^2 k^2 / 2m$. Podemos entonces escribir

$$\Omega = \Omega_+ + \Omega_- \quad (11.38)$$

donde

$$\Omega_{\pm} = -\frac{k_B T V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \epsilon^{1/2} \left\{ \ln \left\{ 1 + z e^{-\beta \epsilon \pm \beta \mu_0 B} \right\} \right\} d\epsilon \quad (11.39)$$

El número medio de electrones esta dado por

$$\langle N \rangle = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V,B} = \langle N_+ \rangle + \langle N_- \rangle \quad (11.40)$$

donde

$$\langle N_{\pm} \rangle = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\infty} \epsilon^{1/2} \left\{ 1 + z^{-1} e^{\beta \epsilon \mp \beta \mu_0 B} \right\}^{-1} d\epsilon \quad (11.41)$$

es facil ver que

$$M = - \left(\frac{\partial \Omega}{\partial B} \right)_{T,V,\mu} = \mu_0 (\langle N_+ \rangle - \langle N_- \rangle) \quad (11.42)$$

11.2.1. Magnetización en el estado fundamental

Para $\beta \rightarrow \infty$

$$z \rightarrow e^{\beta \epsilon_F}$$

donde ϵ_F es la energía de Fermi, esto es, el potencial químico a $T = 0$. Por ejemplo, el integrando de $\langle N_+ \rangle$ se comporta como

$$\begin{aligned} \frac{1}{1 + z^{-1} \exp(\beta \epsilon - \beta \mu_0 B)} &\sim \frac{1}{1 + \exp[\beta(\epsilon - \mu_0 B - \epsilon_F)]} \\ &\sim \begin{cases} 1 & \epsilon - \mu_0 B - \epsilon_F < 0 \\ 0 & \epsilon - \mu_0 B - \epsilon_F > 0 \end{cases} \end{aligned} \quad (11.43)$$

Así

$$\begin{aligned} \langle N_+ \rangle &= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\epsilon_F + \mu_0 B} \epsilon^{1/2} d\epsilon \\ &= \frac{2}{3} \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (\epsilon_F + \mu_0 B)^{3/2} \end{aligned} \quad (11.44)$$

De la misma manera obtenemos

$$\langle N_- \rangle = \frac{2}{3} \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (\epsilon_F - \mu_0 B)^{3/2} \quad (11.45)$$

Tenemos entonces

$$\langle N \rangle = \frac{1}{6} \frac{V}{\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} [(\epsilon_F + \mu_0 B)^{3/2} + (\epsilon_F - \mu_0 B)^{3/2}] \quad (11.46)$$

$$M = \frac{1}{6} \frac{V}{\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0 [(\epsilon_F + \mu_0 B)^{3/2} - (\epsilon_F - \mu_0 B)^{3/2}] \quad (11.47)$$

Escribiendo

$$\langle N \rangle = \frac{1}{6} \frac{V}{\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon_F^{3/2} \left[(1 + \mu_0 B / \epsilon_F)^{3/2} + (1 - \mu_0 B / \epsilon_F)^{3/2} \right] \quad (11.48)$$

$$M = \frac{1}{6} \frac{V}{\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0 \epsilon_F^{3/2} \left[(1 + \mu_0 B / \epsilon_F)^{3/2} - (1 - \mu_0 B / \epsilon_F)^{3/2} \right] \quad (11.49)$$

podemos desarrollar para campos débiles $\mu_0 B \ll \epsilon_F$:

$$\langle N \rangle = \frac{1}{3} \frac{V}{\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon_F^{3/2} + \hat{O} \left[\left(\frac{\mu_0 B}{\epsilon_F} \right)^2 \right] \quad (11.50)$$

$$M = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0 \epsilon_F^{3/2} \left(\frac{\mu_0 B}{\epsilon_F} \right) + \hat{O} \left[\left(\frac{\mu_0 B}{\epsilon_F} \right)^3 \right] \quad (11.51)$$

Así, tenemos que, para campos débiles

$$M \sim \frac{3}{2} \langle N \rangle \mu_0 \left(\frac{\mu_0 B}{\epsilon_F} \right) \quad (11.52)$$

esto es, la magnetización es directamente proporcional al campo magnético B y por lo tanto la susceptibilidad a campo nulo es constante y positiva:

$$\chi_0 = \frac{1}{V} \left(\frac{\partial M}{\partial B} \right)_{V,N} \Big|_{T=0, B=0} = \frac{3\rho\mu_0^2}{2\epsilon_F} \quad (11.53)$$

Este es uno de los resultados característicos del **paramagnetismo de Pauli**.

11.2.2. Magnetización en el límite degenerado $T \ll T_F$

A temperatura finita la magnetización viene dada por

$$\begin{aligned} M &= \mu_0 (\langle N_+ \rangle - \langle N_- \rangle) \\ &= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0 \int_0^\infty \epsilon^{1/2} \{ f(\epsilon - \mu_0 B) - f(\epsilon + \mu_0 B) \} d\epsilon \end{aligned} \quad (11.54)$$

donde

$$f(x) = \frac{1}{1 + z^{-1} e^{\beta x}} \quad (11.55)$$

es la distribución de Fermi. Vamos a hacer ahora un desarrollo de las funciones $f(x)$ en el integrando de la Ec.(11.54) para campos débiles $B \ll 1$. Conservando el primer orden no nulo en el desarrollo tenemos

$$\begin{aligned} M &\sim -\frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0^2 B \int_0^\infty \epsilon^{1/2} f'(\epsilon) d\epsilon \\ &= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_0^2 B \int_0^\infty \epsilon^{-1/2} f(\epsilon) d\epsilon \end{aligned} \quad (11.56)$$

donde en el último paso hemos integrado por partes. De la misma manera tenemos que

$$\begin{aligned} \langle N \rangle &\sim \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \epsilon^{1/2} \{ f(\epsilon - \mu_0 B) + f(\epsilon + \mu_0 B) \} d\epsilon \\ &= \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \epsilon^{1/2} f(\epsilon) d\epsilon \end{aligned} \quad (11.57)$$

La ecuación anterior puede re-escribirse como

$$\langle N \rangle = 2V\lambda_T^3 f_{3/2}(z) \quad (11.58)$$

donde $f_{3/2}(z)$ viene dada por la Ec.(9.11), esto es, la ecuación correspondiente al caso sin campo externo, de la cual se determina μ en función de T y ρ . Así, para campos débiles y bajas temperaturas $T \ll T_F$ tenemos de las Ecs.(9.31) y (9.26) que

$$\mu = \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 + \dots \right] \quad (11.59)$$

con

$$\epsilon_F = \frac{\hbar^2}{2m} (3\pi^2 \rho)^{2/3} \quad (11.60)$$

Por otra parte, puede obtenerse un desarrollo asintótico de la integral de la Ec.(11.56), utilizando el mismo método que usamos para obtener $f_{3/2}(z)$ a bajas temperaturas, esto es, integrando primero por partes, desarrollando luego el factor $\epsilon^{1/2}$ resultante en serie de Taylor alrededor de $\epsilon = \mu$ e integrando término a término. Obtenemos así

$$M = \frac{3\mu_0^2 B \langle N \rangle}{2\epsilon_F} \left(\frac{\mu}{\epsilon_F} \right)^{1/2} \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\mu} \right)^2 + \dots \right] \quad (11.61)$$

Usando la Ec.(11.59) obtenemos

$$M = \frac{3\mu_0^2 B \langle N \rangle}{2\epsilon_F} \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 + \dots \right] \quad (11.62)$$

de donde podemos obtener la susceptibilidad a campo nulo

$$\chi_0 = \frac{3\mu_0^2 \rho}{2\epsilon_F} \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 + \dots \right] \quad (11.63)$$

que incorpora la primera corrección debida a fluctuaciones térmicas a la Ec.(11.53). Esta fue la expresión obtenida por Pauli en su trabajo pionero, explicando la dependencia débil de la susceptibilidad con la temperatura que ocurre en los metales alcalinos, para los cuales T_F es muy grande.

11.2.3. Límite de altas temperaturas

Para altas temperaturas tenemos que $z \ll 1$. Así

$$f(x) \sim z e^{-\beta x} \quad (11.64)$$

esto es, la distribución de Fermi tiende a la distribución de Maxwell-Boltzmann. Reemplazando en las Ecs.(11.54) y (11.57) tenemos que

$$M = \frac{V\mu_0 z}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sinh(\beta\mu_0 B) \int_0^\infty \epsilon^{1/2} e^{-\beta\epsilon} d\epsilon \quad (11.65)$$

$$\langle N \rangle = \frac{Vz}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \cosh(\beta\mu_0 B) \int_0^\infty \epsilon^{1/2} e^{-\beta\epsilon} d\epsilon \quad (11.66)$$

Combinando estas ecuaciones obtenemos

$$M = \mu_0 \langle N \rangle \tanh\left(\frac{\mu_0 B}{k_B T}\right) \quad (11.67)$$

de donde la susceptibilidad a campo nulo resulta

$$\chi_0 = \frac{\rho \mu_0^2}{k_B T} \quad (11.68)$$

Esta última expresión se conoce como **ley de Curie**, siendo característica de la mayoría de los sistemas paramagnéticos. Estas expresiones son las mismas que se obtienen considerando un sistema de N spines distinguibles en el ensamble canónico. Vemos así que dicho resultado corresponde al límite clásico, en el sentido de la distinguibilidad de las partículas, de un gas de electrones con spin.

11.3. Ferromagnetismo

Uno de los fenómenos más interesantes de la física del estado sólido es el ferromagnetismo, esto es, materiales cristalinos que permanecen magnetizados aún en ausencia de un campo externo. De hecho, históricamente este fue el primer fenómeno magnético en ser descubierto¹ y el que originó el estudio del magnetismo en los albores de la física. Sin embargo, el problema es de tal complejidad que aún hoy existen aspectos que no se comprenden completamente.

Desde un punto de vista microscópico el ferromagnetismo se presenta como un fenómeno cooperativo, en el cual una fracción macroscópica de spines (es decir, momentos magnéticos) se orientan más o menos en la misma dirección por debajo de cierta temperatura crítica (en muchos casos bastante alta). Como ya hemos visto, sistemas de partículas no interactuantes no presentan ferromagnetismo, esto es, la magnetización se anula siempre a campo nulo. La pregunta es entonces, que tipo de interacción es capaz de dar lugar a un ordenamiento espontáneo, coherente, de los momentos magnéticos? Sabemos que existen fuerzas de interacción magnéticas entre dipolos permanentes, las cuales dependen de la orientación relativa de los dipolos y decaen con la distancia. No obstante, los valores típicos de la energía de interacción entre dos dipolos atómicos son sumamente pequeñas (del orden de $10^{-4} eV$, lo cual equivale a temperaturas del orden del grado Kelvin) siendo insuficientes para explicar el ordenamiento a temperaturas razonablemente altas, en muchos casos superiores a la temperatura ambiente. Por otro lado, las energías electrostáticas entre electrones o entre electrones e iones, son suficientemente intensas como para explicar un ordenamiento a las temperaturas observadas. En 1930 Heisenberg, Dirac y otros propusieron un mecanismo que combina la interacción coulombiana entre electrones con el principio de exclusión de Pauli.

11.3.1. Interacciones de intercambio: el modelo de Heisenberg

A fin de ilustrar como una interacción electrostática combinada con el principio de exclusión de Pauli puede dar lugar a una interacción magnética efectiva, vamos a considerar el caso de dos electrones interactuantes en una molécula de hidrógeno. Esto es vamos a analizar las funciones de onda para dos electrones en un potencial que resulta de dos núcleos hidrogenoides localizados a cierta distancia fija entre sí. Podemos pensar este sistema como un “sólido” conformado juntando $N = 2$ átomos con un electrón de valencia cada uno. Si despreciamos las interacciones spin-órbita, el Hamiltoniano H de este sistema no depende del spin. Así, la solución general Ψ de la ecuación estacionaria $H\Psi = E\Psi$ factoriza en una parte puramente orbital $\psi(\vec{r}_1, \vec{r}_2)$, solución de la ecuación

¹La existencia de la magnetita Fe_3O_4 , imanes naturales que se encontraban en la provincia de Magnesia y de ahí su nombre, era conocida por los antiguos griegos. Sin embargo, existen evidencias brujulas a base de magnetita en China entre el 3000 y el 2500 a.C..

$$H\phi = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) \phi + V(\vec{r}_1, \vec{r}_2)\phi = E\phi, \quad (11.69)$$

y una parte dependiente del spin de los electrones; esta última se construye como combinación lineal de los estados

$$|\sigma_1\rangle_1 |\sigma_2\rangle_2$$

donde $|\sigma_i\rangle_i$ es el autoestado de la matriz de Pauli σ_i^z asociada a la partícula i correspondiente al autovalor $\sigma_i = \uparrow, \downarrow$.

El principio de exclusión de Pauli impone que la función de onda total Ψ sea antisimétrica, lo cual puede obtenerse exigiendo que la parte orbital sea simétrica y la de spin antisimétrica o viceversa². Resulta conveniente elegir como función de spin los autoestados del spin total S^2 y su componente S^z , ya que estas serán cantidades conservadas. Dichos estados a la vez tienen paridad definida, de manera que podemos designarlos según su paridad:

$$\chi_A(\sigma_1, \sigma_2) = \frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2) \quad (11.70)$$

la función antisimétrica que corresponde al *singlete* de spin total $S = 0$ y

$$\chi_S(\sigma_1, \sigma_2) = \begin{cases} |\uparrow\rangle_1 |\uparrow\rangle_2 \\ \frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\downarrow\rangle_2 + |\downarrow\rangle_1 |\uparrow\rangle_2) \\ |\downarrow\rangle_1 |\downarrow\rangle_2 \end{cases} \quad (11.71)$$

las funciones simétricas correspondientes al *triplete*, compuesto por los autoestados de spin total $S = 1$ y de S^z con autovalores 1, 0 y -1 respectivamente.

Los dos tipos de funciones de onda globalmente antisimétricas tienen entonces la forma:

$$\psi_s(\vec{r}_1, \vec{r}_2) = \phi_S(\vec{r}_1, \vec{r}_2)\chi_A(\sigma_1, \sigma_2) \quad (11.72)$$

con la parte posicional simétrica y la parte de spin antisimétrica y

$$\psi_t(\vec{r}_1, \vec{r}_2) = \phi_A(\vec{r}_1, \vec{r}_2)\chi_S(\sigma_1, \sigma_2) \quad (11.73)$$

con la situación invertida.

La primera aproximación a este problema consiste en despreciar la interacción electrón-electrón y obtener las autofunciones orbitales de partículas independientes para luego combinarlas de manera de obtener las propiedades de paridad requeridas. Una vez obtenidas las autofunciones orbitales $\phi_{A,S}$, la interacción electrón-electrón puede ser considerada perturbativamente, esto es, a primer orden se considera el valor médio del potencial coulombiano en los diferentes estados orbitales.

Existen distintas aproximaciones para obtener las funciones de onda orbitales $\phi_{A,S}$. Una aproximación particularmente indicada para generalizar el tratamiento a un sólido (es decir, a un sistema de N átomos) es considerar inicialmente los dos átomos separados a una distancia muy grande, cada uno con un electrón en un estado orbital hidrogenoide $\varphi_i(\vec{r})$, $i = 1, 2$, donde el subíndice i indica el núcleo en el cual está centrado el orbital. Si la energía del orbital es ϵ_0 , entonces el estado fundamental de este sistema es $2\epsilon_0$, el cual tiene degeneración cuatro, correspondiente a las diferentes orientaciones relativas de cada spin electrónico. Los estados singlete y triplete están por lo tanto degenerados: $E_s = E_t = 2\epsilon_0$. Al acercarse los átomos se producirá una superposición de los orbitales y las autofunciones del (11.69) ya no corresponderán al producto de los orbitales aislados. Esto dará lugar un desdoblamiento de este nivel, levantándose la degeneración entre la

²Notemos que siempre podemos elegir las autofunciones del Hamiltoniano (11.69) con paridad bien definida, en virtud de la simetría ante intercambio de partículas del potencial $V(\vec{r}_1, \vec{r}_2)$

energía del singlete E_s y la del triplete E_t ; los tres estados del triplete continúan degenerados, ya que la interacción es independiente del spin y los tres estados tienen la misma parte orbital. De hecho es posible demostrar que el estado fundamental exacto del Hamiltoniano (11.69) corresponde a una función orbital simétrica y por lo tanto al estado singlete.

Si el orbital $\varphi(\vec{r})$ es una función muy localizada en el núcleo, la superposición de orbitales localizados en diferentes núcleos será pequeña así como el desdoblamiento del nivel. Así, una posible aproximación de las funciones orbitales consiste en utilizar combinaciones simétricas y antisimétricas de orbitales localizados en diferentes núcleos, esto es

$$\phi_{S,A}(\vec{r}_1, \vec{r}_2) = \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) \pm \varphi_2(\vec{r}_1)\varphi_1(\vec{r}_2) \quad (11.74)$$

Esta elección se conoce como **aproximación de Heitler-London**. Su generalización al caso de N átomos constituye una buena aproximación a los sólidos aislantes. En ausencia de interacción entre los electrones, la aproximación anterior puede ser muy mala, ya que excluye combinaciones de orbitales localizados ambos en el mismo núcleo; en ausencia de repulsión entre electrones esta contribución puede ser importante. Sin embargo, esta última situación elevaría mucho la energía en presencia de repulsión coulombiana y por lo tanto la aproximación de Heitler-London resulta apropiada en este caso.

El desdoblamiento entre niveles $E_s - E_t$ debido a la interacción electrón-electrón puede calcularse entonces de manera perturbativa como

$$E_s - E_t = \left\langle \psi_s \left| \left(\frac{e^2}{r_{12}} \right) \right| \psi_s \right\rangle - \left\langle \psi_t \left| \left(\frac{e^2}{r_{12}} \right) \right| \psi_t \right\rangle \quad (11.75)$$

Reemplazando la Ec.(11.74) y usando que el potencial de interacción es simétrico ante el intercambio $1 \rightleftharpoons 2$ tenemos que

$$E_s - E_t = 4 \left\langle \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) \left| \left(\frac{e^2}{r_{12}} \right) \right| \varphi_2(\vec{r}_1)\varphi_1(\vec{r}_2) \right\rangle \quad (11.76)$$

El segundo término en la ecuación anterior es un elemento de matriz entre dos estados que difieren entre sí solamente a través del intercambio de las coordenadas de los electrones y de ahí que se lo denomina **término de intercambio**.

En las condiciones en que resulta válida la aproximación de Heitler-London, el desdoblamiento del estado fundamental es pequeño comparado con las otras energías de excitación del sistema de dos electrones. Así, las propiedades del sistema, al menos a temperaturas relativamente bajas, estarán dominadas por estos cuatro estados. En el subespacio de estos estados podemos construir un Hamiltoniano efectivo de acuerdo a lo siguiente. Sean $\vec{S}_i = \vec{\sigma}_i/2$, $i = 1, 2$, los operadores de spin asociados con cada electrón, donde $\vec{\sigma}_i$ son matrices de Pauli (hemos tomado $\hbar = 1$). Tenemos que $\vec{S}_i^2 = \frac{1}{2} \left(\frac{1}{2} + 1 \right) = \frac{3}{4}$, de manera que el operador de spin total \vec{S} satisface:

$$\vec{S}^2 = (\vec{S}_1 + \vec{S}_2)^2 = \frac{3}{2} + 2 \vec{S}_1 \cdot \vec{S}_2 \quad (11.77)$$

y por lo tanto los estados singlete y triplete son también autoestados del operador

$$\vec{S}_1 \cdot \vec{S}_2 = \frac{1}{2} \vec{S}^2 - \frac{3}{4} \quad (11.78)$$

Dado que el operador \vec{S}^2 tiene autovalores $S(S+1)$, el operador (11.78) tiene autovalor $-3/4$ en el singlete ($S = 0$) y $1/4$ en el triplete ($S = 1$). En consecuencia, el operador

$$\mathcal{H} = \frac{1}{4}(E_s + 3E_t) - (E_s - E_t)\vec{S}_1 \cdot \vec{S}_2 \quad (11.79)$$

tiene autovalor E_s cuando aplicado al singlete y autovalor E_t cuando aplicado a cualquiera de los estados del triplete. Redefiniendo el cero de energía tenemos el Hamiltoniano efectivo

$$\mathcal{H} = -J\vec{S}_1 \cdot \vec{S}_2 \quad J = E_s - E_t \quad (11.80)$$

que representa una interacción magnética efectiva entre los spines electrónicos, la cual se conoce como **interacción de intercambio**. Esta interacción favorece que los spines esten alineados si $J > 0$ y antialineados si $J < 0$. Esto se corresponde con el hecho de que los spines se encuentran antialineados en el singlete ($E_s < E_t \Rightarrow J < 0$) y alineados en los estados triplete ($E_s > E_t \Rightarrow J > 0$); J se conoce como constante de intercambio. Esta interacción magnética efectiva es resultado combinado de las interacciones coulombianas y las exigencias de antisimetría de las funciones de onda exclusivamente.

Para un cristal regular con N átomos, si bien el tratamiento es mas complicado, se arriba básicamente al mismo resultado. En este caso, el estado fundamental de N átomos muy separados correspondiente a un orbital localizado tiene una degeneración 2^N . El desdoblamiento de estos niveles en el caso en que los iones se encuentran cerca, pero aún lo suficientemente distanciados como para que el desdoblamiento de energía sea pequeño comparado con otras energías de excitación, viene dado por el **Hamiltoniano de Heisenberg**:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j \quad (11.81)$$

donde los operadores de spin se encuentran asociados con los sitios del cristal y la suma $\sum_{\langle i,j \rangle}$ corre sobre sitios primeros vecinos. Esto se debe a que la integral de intercambio involucra superposiciones de funciones de onda localizadas, las cuales son apreciables en general solo entre sitios primeros vecinos.

Lo que es mas notable es que el Hamiltoniano de Heisenberg describe una gran variedad de casos, incluyendo situaciones en las cuales las funciones de onda de átomos individuales esta lejos de ser consideradas orbitales localizados, como es el caso de los metales. De hecho, es posible en ciertas circunstancias derivar un modelo de Heisenberg efectivo en metales a partir de la teoría de bandas³ Dependiendo del tipo de orbitales la constante de intercambio puede adoptar valores positivos o negativos. El primer caso describe un material ferromagnético, cuyo estado fundamental corresponde a todos los spines alineados, en tanto que el segundo describe un sistema **antiferromagnético**, el cual describiremos mas adelante. No obstante, a diferencia del caso $N = 2$ antes descrito, en el caso general no se conoce el espectro completo del Hamiltoniano (11.81).

11.3.2. Modelos derivados del modelo de Heisenberg

El modelo de Heisenberg ferromagnético presenta magnetización espontanea, con una temperatura crítica $T_c > 0$, en cristales tridimensionales. En cristales uni o bidimensionales la magnetización a campo nulo es cero para toda temperatura finita.

La mecánica estadística del modelo de Heisenberg resulta extremadamente complicada de obtener, ya que ni siquiera se conoce de manera exacta el espectro del Hamiltoniano (11.81). No obstante, la fenomenología de la transición de fase ferromagnética puede analizarse utilizando modelos mas sencillos, los cuales pueden derivarse del de Heisenberg.

Las interacciones de intercambio presentes en el modelo de Heisenberg son totalmente isotrópicas. No obstante, en sólidos reales los electrones presentan acoplamientos spin-órbita con las distribuciones de carga de los iones en el cristal. Esta distribuciones de carga tienen a menudo las simetrías del cristal, dando lugar a campos magnéticos efectivos en las direcciones de ejes cristalinos, las cuales introducen anisotropías en las interacciones de intercambio. Estas anisotropías dan

³D. C. Mattis, The Theory of Magnetism made Simple, World scientific, 2006.

lugar a lo que se conoce como "ejes fáciles" de magnetización, esto es, sistemas en los cuales la magnetización espontánea ocurre preferencialmente en la dirección de algún eje particular. Una manera posible de describir estos sistemas es mediante el **modelo de Heisenberg anisotrópico**:

$$\mathcal{H} = - \sum_{\langle i,j \rangle} \left[J_z S_i^z S_j^z + J_{xy} (S_i^x S_j^x + S_i^y S_j^y) \right] \quad (11.82)$$

El límite $J_z \rightarrow 0$ del Hamiltoniano (11.82)

$$\mathcal{H}_{xy} = -J_{xy} \sum_{\langle i,j \rangle} (S_i^x S_j^x + S_i^y S_j^y) \quad (11.83)$$

se conoce como **modelo XY o planar**. En el límite $J_{xy} \rightarrow 0$ tenemos el Hamiltoniano

$$\mathcal{H} = -\frac{J_z}{4} \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z \quad (11.84)$$

donde σ_i^z es una matriz de Pauli asociada con el sitio i . Este Hamiltoniano es diagonal en la base de estados

$$\prod_{i=1}^N |\sigma_i\rangle$$

con autovalores

$$H(\{\sigma_i\}) = -\frac{J_z}{4} \sum_{\langle i,j \rangle} \sigma_i \sigma_j \quad (11.85)$$

donde $\sigma_i = \pm 1$. La función $H(\{\sigma_i\})$ corresponde al Hamiltoniano de **Ising**. A veces este modelo es considerado clásico, en el sentido de estar descrito por variables de estado discretas en lugar de operadores en un espacio de Hilbert. Sin embargo, su origen es netamente cuántico. De hecho el modelo de Ising, a pesar de su extrema simplificación, describe correctamente el comportamiento termodinámico de diversos materiales magnéticos reales. En el capítulo siguiente analizaremos la mecánica estadística de este modelo.

Capítulo 12

El modelo de Ising

El modelo de Ising ¹ es uno de los pocos modelos de partículas interactuantes para el cual se conoce su solución exacta e indudablemente el más simple de todos ellos. Este modelo tiene enorme interés por diversas razones. Por un lado resulta interesante su papel en el desarrollo histórico de la comprensión del ferromagnetismo y de las transiciones de fase, en cuyo proceso representó un papel fundamental. En segundo lugar, el método de solución en una dimensión presentado por Ising y luego extendido a dos dimensiones por Onsager, constituye la base de diversos métodos modernos de cálculo en la física estadística de los fenómenos críticos. Finalmente, hoy se sabe que el modelo de Ising y generalizaciones del mismo sirven para explicar una variedad de fenómenos, no solo físicos (como veremos más adelante), como también de diversas áreas de la biología.

12.1. Modelo de Ising en una dimensión: solución exacta

En una dimensión ($d = 1$) el Hamiltoniano de Ising puede ser escrito de la forma

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - B \sum_{i=1}^N \sigma_i \quad (12.1)$$

donde $\sigma_i = \pm 1$ y vamos a utilizar condiciones de contorno periódicas, esto es

$$\sigma_{i+N} = \sigma_i$$

En una dimensión, las condiciones de contorno periódicas son equivalentes a resolver el problema en un anillo. Así, podemos escribir el Hamiltoniano (12.1) de manera más simétrica:

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - \frac{B}{2} \sum_{i=1}^N (\sigma_i + \sigma_{i+1}) \quad (12.2)$$

La función de partición canónica viene dada por

$$Z_N = \sum_{\{\sigma_i\}} \exp \left[K \sum_{i=1}^N \sigma_i \sigma_{i+1} + \frac{h}{2} \sum_{i=1}^N (\sigma_i + \sigma_{i+1}) \right] \quad (12.3)$$

donde $K \equiv \beta J$ y $h \equiv \beta B$. Notemos que la función de partición puede ser escrita de la forma

$$Z_N = \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_N=\pm 1} \prod_{i=1}^N T(\sigma_i, \sigma_{i+1}) \quad (12.4)$$

¹E. Ising, Z. Phys. **31**, 253 (1925).

donde

$$T(\sigma_i, \sigma_{i+1}) = \exp \left[K\sigma_i\sigma_{i+1} + \frac{h}{2}(\sigma_i + \sigma_{i+1}) \right] \quad (12.5)$$

Esta última expresión puede ser interpretada como un elemento de una matriz 2×2 simétrica, indexada por los valores de las variables de spin $\sigma_i = \pm 1$. Definimos entonces la **matriz de transferencia**:

$$\mathbf{T} = \begin{pmatrix} T(+1, +1) & T(+1, -1) \\ T(-1, +1) & T(-1, -1) \end{pmatrix} = \begin{pmatrix} e^{K+h} & e^{-K} \\ e^{-K} & e^{K-h} \end{pmatrix} \quad (12.6)$$

Así, podemos escribir la función de partición como

$$Z_N = \sum_{\sigma_1=\pm 1} \sum_{\sigma_3=\pm 1} \cdots \sum_{\sigma_N=\pm 1} \left(\sum_{\sigma_2=\pm 1} T(\sigma_1, \sigma_2)T(\sigma_2, \sigma_3) \right) T(\sigma_3, \sigma_4) \cdots T(\sigma_N, \sigma_1) \quad (12.7)$$

$$= \sum_{\sigma_1=\pm 1} \sum_{\sigma_4=\pm 1} \cdots \sum_{\sigma_N=\pm 1} \left(\sum_{\sigma_3=\pm 1} T^2(\sigma_1, \sigma_3)T(\sigma_3, \sigma_4) \right) T(\sigma_4, \sigma_5) \cdots T(\sigma_N, \sigma_1) \quad (12.8)$$

$$\vdots \quad (12.9)$$

$$= \sum_{\sigma_1=\pm 1} T^N(\sigma_1, \sigma_1) \quad (12.10)$$

$$= \text{Tr}(\mathbf{T}^N) \quad (12.11)$$

La matriz \mathbf{T} es simétrica por construcción y por lo tanto sus autovalores son reales. Si λ_{\pm} son los autovalores de \mathbf{T} entonces

$$Z_N = \lambda_+^N + \lambda_-^N \quad (12.12)$$

Los autovalores resultan

$$\lambda_{\pm} = e^K \left[\cosh(h) \pm \sqrt{\cosh^2(h) - 2e^{-2K}\sinh(2K)} \right] \quad (12.13)$$

A campo nulo $h = 0$ tenemos

$$\lambda_+ = 2 \cosh(K) \geq \lambda_- = 2 \sinh(K) > 0$$

y en general $\lambda_+ \geq \lambda_- \geq 0$. Si escribimos

$$Z_N = \lambda_+^N \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \quad (12.14)$$

tenemos entonces el límite termodinámico $N \rightarrow \infty$ para la energía libre por partícula

$$f(T, B) = \lim_{N \rightarrow \infty} \left[-\frac{k_B T}{N} \ln Z_N \right] = -k_B T \ln \lambda_+ \quad (12.15)$$

$$= -J - \frac{1}{\beta} \ln \left\{ \cosh(\beta B) + \sqrt{\cosh^2(\beta B) - 2e^{-2\beta J}\sinh(2\beta J)} \right\} \quad (12.16)$$

que es una función analítica de T y B . Notemos que, a diferencia de los gases, en este caso el límite termodinámico se toma como $N \rightarrow \infty$ en lugar de $V \rightarrow \infty$. En general al tratar con

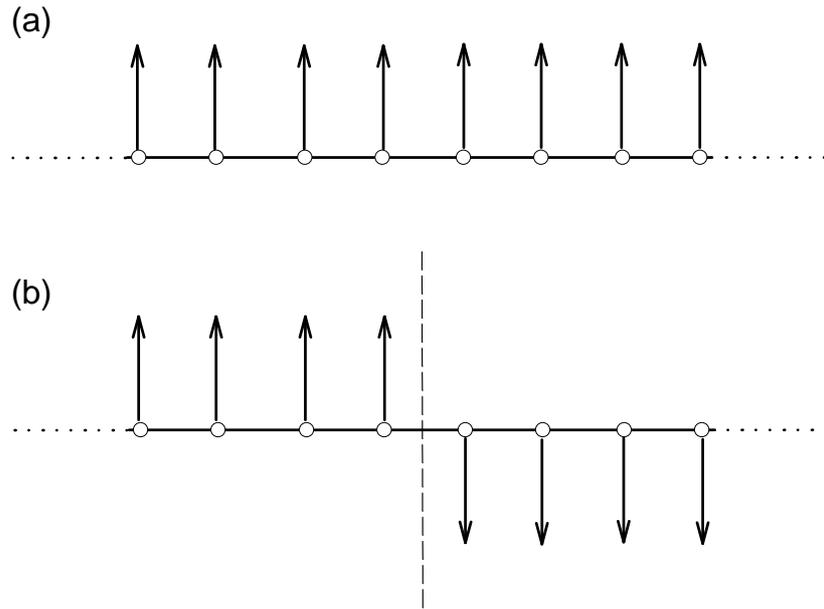


Figura 12.1: (a) Estado fundamental del modelo de Ising unidimensional. (b) Pared de dominio.

modelos en redes V y N son variables dependientes, de manera que solo una de ellas es relevante termodinámicamente. Habitualmente se escoge el número de partículas, esto es, el número de sitios en la red.

La magnetización por spin esta dada por

$$m(T, B) = - \left(\frac{\partial f}{\partial B} \right)_T = \frac{\sinh(\beta B)}{\sqrt{\cosh^2(\beta B) - 2e^{-2\beta J} \sinh(2\beta J)}} \quad (12.17)$$

Vemos que $m \rightarrow 0$ cuando $B \rightarrow 0$, para toda $T \neq 0$, y por lo tanto este sistema no presenta transición de fase a temperatura finita, esto es, se comporta como un paramagneto a toda temperatura. Podemos calcular también la entropía por spin $s(T, B) = -(\partial f / \partial T)_B$, y a mediante esta el calor específico a campo constante

$$C_{B=0} = T \left(\frac{\partial s}{\partial T} \right)_{B=0} = -T \left(\frac{\partial^2 f}{\partial T^2} \right)_{B=0} = \frac{J^2}{k_B T^2} \frac{1}{\cosh^2 \left(\frac{J}{k_B T} \right)} \quad (12.18)$$

que es también una función analítica para toda $T > 0$.

Existe un argumento atribuido a Landau para mostrar que no es posible tener un estado ordenado a temperatura finita (y, por lo tanto, una transición de fase) en un sistema unidimensional con interacciones de corto alcance. Aplicado al modelo de Ising es como sigue.

El estado fundamental del modelo unidimensional, en ausencia de campo, es una cadena con todos los spines alineados, tal como se muestra en la Fig.12.1a. Este es el mínimo absoluto de la energía. Veamos ahora la configuración de spines de la Fig.12.1b. Esta corresponde a dos *dominios magnéticos*² con magnetizaciones opuestas separados por dos *paredes de dominio* (estamos

²Un dominio se define como una región conexas con todos los spines alineados. Una pared de dominio es la interface que separa dos dominios.

suponiendo siempre condiciones de contorno periódicas; la extensión del argumento a una cadena abierta es directa). La diferencia de energía entre esta configuración y el estado fundamental es $\Delta U = 4J$ ($2J$ es el costo energético de cada pared). Por otra parte, el conjunto de las configuraciones con una pared de dominio tiene una entropía muy grande $\Delta S = k_B \ln N$, ya que las paredes pueden ser colocadas en N posiciones distintas de la cadena (anillo). A temperatura finita, la diferencia de energía libre entre el estado fundamental y el estado con dos dominios de magnetización opuesta es

$$\Delta F = 2J - k_B T \ln N \quad (12.19)$$

En el límite termodinámico $N \rightarrow \infty$ la expresión anterior se vuelve negativa, favoreciendo por lo tanto la creación de dominios. Dado que la magnetización promedia a cero en el conjunto de configuraciones dos dominios, el estado ordenado resulta inestable frente a un estado de magnetización nula *a cualquier temperatura finita*. El modelo de Ising solo ordena para $T = 0$, por lo cual se dice que presenta una transición de fase a temperatura cero.

Este argumento puede extenderse a todo tipo de modelos unidimensionales con interacciones de corto alcance.

12.2. Modelo de Ising en dos dimensiones

El modelo bidimensional fue resuelto por primera vez usando la técnica de matriz de transferencia por Onsager ³ con campo nulo y por Yang ⁴ con campo magnético. Posteriormente fue resuelto por diferentes autores usando diferentes técnicas de cálculo ⁵. En cualquiera de ellas, la resolución es muchísimo más complicada que en el caso unidimensional. Nos limitaremos aquí a presentar los resultados principales y discutiremos sus consecuencias.

En el límite termodinámico (en la red cuadrada y a campo nulo) la energía libre por partícula resulta

$$\beta f(T) = -\ln(2 \cosh 2\beta J) - \frac{1}{2\pi} \int_0^\pi d\phi \ln \left[\frac{1}{2} \left(1 + \sqrt{1 - \kappa^2 \sin^2 \phi} \right) \right] \quad (12.20)$$

donde

$$\kappa = \frac{2 \sinh(2\beta J)}{\cosh^2(2\beta J)} \quad (12.21)$$

La energía interna por spin viene dada por

$$u(T) = \frac{\partial(\beta f(T))}{\partial \beta} = -2J \tanh(2\beta J) + \frac{\kappa}{2\pi} \frac{d\kappa}{d\beta} \int_0^\pi d\phi \frac{\sin^2 \phi}{\Delta(1 + \Delta)} \quad (12.22)$$

donde $\Delta(\phi) = \sqrt{1 - \kappa^2 \sin^2 \phi}$. Se verifica fácilmente que

$$\int_0^\pi d\phi \frac{\sin^2 \phi}{\Delta(1 + \Delta)} = \frac{1}{\kappa^2} \int_0^\pi \frac{d\phi}{\Delta(\phi)} - \frac{\pi}{\kappa^2} \quad (12.23)$$

Reemplazando en la Ec.(12.22) obtenemos, con algo de trabajo, que

$$u(T) = -J \coth(2\beta J) \left[1 + \frac{2}{\pi} \kappa' K_1(\kappa) \right] \quad (12.24)$$

³L. Onsager, Phys. Rev. **65**, 117 (1944); B. Kaufman, Phys. Rev. **76**, 1232 (1949); B. Kaufman y L. Onsager, Phys. Rev. **76**, 1244 (1949).

⁴C. N. Yang, Phys. Rev. **85**, 809 (1952).

⁵T. D. Schultz, D. C. Mattis y E. H. Lieb, Rev. Mod. Phys. **36**, 856 (1964).

donde

$$K_1(\kappa) \equiv \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - \kappa^2 \sin^2 \phi}} \quad (12.25)$$

es una integral elíptica completa de primera especie y

$$\kappa' \equiv 2 \tanh^2(2\beta J) - 1$$

$$\kappa^2 + \kappa'^2 = 1$$

Derivando la Ec.(12.24) podemos obtener, con cierto trabajo, el calor específico

$$c(T) = \frac{du(T)}{dT} = \frac{2k_B}{\pi} (\beta J \coth(2\beta J))^2 \left\{ 2K_1(\kappa) - 2E_1(\kappa) - (1 - \kappa') \left[\frac{\pi}{2} \kappa' K_1(\kappa) \right] \right\} \quad (12.26)$$

donde

$$E_1(\kappa) \equiv \int_0^{\pi/2} d\phi \sqrt{1 - \kappa^2 \sin^2 \phi} \quad (12.27)$$

es una integral elíptica completa de segunda especie.

La integral elíptica $K_1(\kappa)$ tiene una singularidad en $\kappa = 1$ ($\kappa' = 0$), en cuyo entorno tenemos que

$$K_1(\kappa) \sim \ln \frac{4}{|\kappa'|}$$

$$E_1(\kappa) \sim 1$$

En dicho punto todas las funciones termodinámicas son no-analíticas. La temperatura crítica resulta entonces de la condición

$$\kappa = \frac{2 \sinh(2\beta_c J)}{\cosh^2(2\beta_c J)} = 1 \quad (12.28)$$

o bien de $\kappa' = 0$

$$2 \tanh^2(2\beta_c J) = 1 \quad (12.29)$$

de donde resulta

$$k_B T_c / J = \frac{2}{\ln(1 + \sqrt{2})}. \quad (12.30)$$

Otra relación satisfecha por T_c es

$$\sinh^2(2\beta_c J) = 1. \quad (12.31)$$

En el entorno de T_c tenemos entonces que

$$c(T) \sim -\frac{2k_B}{\pi} \left(\frac{2J}{k_B T_c} \right)^2 \ln |\kappa'| \quad (12.32)$$

Pero

$$\kappa' = 2 \tanh^2(2\beta J) - 1 = 2 \tanh^2(2\beta J) - 2 \tanh^2(2\beta_c J) \sim D(T - T_c)$$

donde D es una constante. Así

$$c(T) \sim -\frac{2k_B}{\pi} \left(\frac{2J}{k_B T_c} \right)^2 \ln \left| 1 - \frac{T}{T_c} \right| \quad (12.33)$$

esto es, $c(T)$ presenta una singularidad logarítmica, en contraste con la predicción de campo medio de una discontinuidad.

El parámetro de orden, esto es, la magnetización por partícula es

$$m(T) = \begin{cases} 0 & T > T_c \\ \left\{ 1 - [\sinh(2\beta J)]^{-4} \right\}^{1/8} & T \leq T_c \end{cases} \quad (12.34)$$

En el entorno del punto crítico tenemos que

$$m(T) \sim \left[\sinh^4(2\beta J) - \sinh^4(2\beta_c J) \right]^{1/8} \sim A(T_c - T)^{1/8} \quad (12.35)$$

donde A es una constante. El exponente crítico del parámetro de orden resulta entonces $\beta = 1/8$.

En dimensiones $d \leq 3$ no se conoce hasta hoy la solución exacta del modelo Ising en ningún tipo de red. Sin embargo, existen numerosas soluciones aproximadas, así como resultados de simulaciones numéricas, las cuales permiten obtener la temperatura crítica y los exponentes críticos en 3 dimensiones. En particular, el exponente crítico para la magnetización resulta independiente del tipo de red cristalina y es $\beta \approx 1/3$, en concordancia con el exponente crítico para la transición gas-líquido (la independencia de los exponentes críticos de los detalles microscópicos del sistema se conoce como **universalidad**).

12.3. Aproximación de campo medio

Una aproximación no controlada de enorme uso es la que se campo medio y se basa en la teoría del ferromagnetismo de Weiss. Si bien esta aproximación no describe quantitativamente bien los resultados experimentales, es de enorme importancia teórica, debido especialmente a su simplicidad. La misma puede desarrollarse como sigue.

Si consideramos un sistema con invariancia traslacional (red infinita ó finita con condiciones de contorno periódicas), tendremos que la magnetización local $\langle \sigma_i \rangle$ será $\langle \sigma_i \rangle = m$ independiente del sitio particular i , donde m es la magnetización por spin:

$$m = \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle$$

Podemos entonces escribir el Hamiltoniano como sigue

$$H = -J \sum_{\langle i,j \rangle} \left[(\sigma_i - m)(\sigma_j - m) + m(\sigma_i + \sigma_j) - m^2 \right] - B \sum_i \sigma_i \quad (12.36)$$

Sea q el número de primeros vecinos de un sitio (propiedad de la red particular), el cual se conoce como **número de coordinación**. El Hamiltoniano (12.36) puede escribirse como

$$H = \frac{qN}{2} m^2 J - (qJm + B) \sum_i \sigma_i - J \sum_{\langle i,j \rangle} \Delta \sigma_i \Delta \sigma_j \quad (12.37)$$

donde $\Delta\sigma_i \equiv \sigma_i - m$ es una variable aleatoria que representa las fluctuaciones en la magnetización del sitio i . La aproximación de campo medio consiste entonces en despreciar el término de fluctuaciones en el Hamiltoniano (12.37), con lo cual obtenemos

$$H_{cm} = \frac{qN}{2} m^2 J - B_{ef} \sum_i \sigma_i \quad (12.38)$$

donde $B_{ef} \equiv qJm + B$. En esta aproximación entonces, el sistema es descrito por un conjunto de spines independientes en presencia de un campo efectivo generado por la magnetización promedio de sus vecinos mas el campo externo. Este corresponde al campo medio o campo molecular introducido por Weiss. La función partición resulta entonces

$$Z_{cm} = \exp(-\beta q N m^2 J / 2) \left[\sum_{\sigma=\pm 1} e^{\beta B_{ef} \sigma} \right]^N = \exp(-\beta q N m^2 J / 2) [2 \cosh(\beta B_{ef})]^N \quad (12.39)$$

y la energía libre por spin

$$f_{cm} = -\frac{k_B T}{N} \ln Z_{cm} = \frac{1}{2} m^2 J q - k_B T \ln [2 \cosh(\beta B_{ef})] \quad (12.40)$$

Para que la aproximación resulte consistente, al magnetización por spin m debe resultar

$$m = \langle \sigma_i \rangle = \frac{1}{Z_{cm}} \text{Tr} \sigma_i e^{-\beta H_{cm}} = \frac{\sum_{\sigma=\pm 1} \sigma e^{\beta B_{ef} \sigma}}{\sum_{\sigma=\pm 1} e^{\beta B_{ef} \sigma}} = \tanh(\beta B_{ef}) \quad (12.41)$$

Así, arribamos a la ecuación autoconsistente para m :

$$m = \tanh[\beta(qJm + B)] \quad (12.42)$$

la cual se conoce como **ecuación de Curie-Weiss**. La misma puede resolverse gráficamente. Tomemos el caso $B = 0$: $m = \tanh(\beta q J m)$, cuya solución se ilustra en la Fig.12.2. $m = 0$ es siempre solución. Cuando la pendiente de la tangente hiperbólica en el origen es mayor que uno, aparecen dos soluciones simétricas $m = \pm m_0(T)$. La condición para que ocurra esto último es entonces $\beta_c J q = 1$ ó $k_B T_c = q J$. Es posible ver que la solución $m = 0$ representa un máximo de la energía libre f_{cm} para $T < T_c$. Así, $m_0(T)$ corresponde a la magnetización espontánea. Para $B \neq 0$, solo existe una solución estable (mínimo absoluto de la energía libre) para cualquier temperatura. Para temperaturas levemente menores que T_c tendremos que $m_0 \ll 1$. Así, podemos desarrollar la Ec.(12.42) con $B = 0$ como

$$m_0 = \beta q J m_0 + (\beta q J m_0)^3 / 3 + \mathcal{O}(m_0^5)$$

de donde

$$m_0(T) \sim \left[3 \left(1 - \frac{T}{T_c} \right) \right]^{1/2}$$

El exponente crítico para la magnetización en la aproximación de campo medio es por lo tanto $\beta = 1/2$. Notemos que este exponente es totalmente independiente de la dimensión de la red (la dimensión en esta aproximación entra solamente a través del número de coordinación q), y difiere de los valores exactos para $d = 2$ (1/8) y $d = 3$ (1/3). Mas aún, esta aproximación predice una transición de fase incluso para $d = 1$. Puede demostrarse que la aproximación de campo medio para el modelo de Ising se vuelve exacta en dimensiones $d \geq 4$.

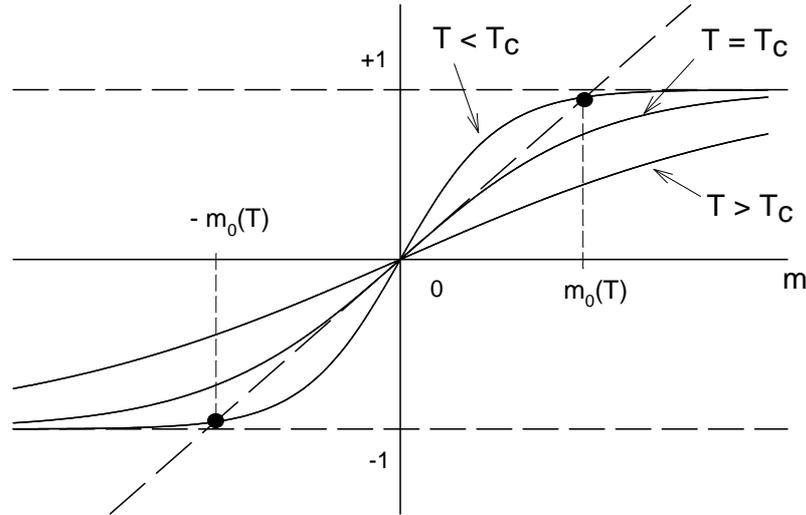


Figura 12.2: Solución gráfica de la ecuación de Curie-Weiss a campo externo nulo $B = 0$.

12.4. El gas de red

Este es un modelo de partículas interactuantes para describir la transición líquido-gas. En este modelo las partículas solo pueden ocupar posiciones discretas en el espacio. De esta manera, podemos asociar a cada posición un sitio o celda en una red de Bravais con N sitios. En otras palabras, dividimos el espacio en celdas ubicadas regularmente, tal que cada celda puede alojar a lo sumo una partícula. El número de sitios N es proporcional al volumen del sistema. Como vimos en el capítulo 6, la función partición canónica puede descomponerse en una parte correspondiente a la energía cinética (gas ideal) y una parte configuracional. En el ensamble gran canónico, la primera puede absorberse dentro del potencial químico. Consideramos entonces solamente la parte configuracional y asumimos un potencial de interacción entre dos partículas del tipo “carozo duro” (es decir, repulsión infinita) a muy cortas distancias y atractiva a distancias intermedias. El potencial de carozo duro impone que en un sitio no puede haber más de una partícula. Si llamamos $u(|\mathbf{r}_i - \mathbf{r}_j|)$ al potencial de interacción entre dos partículas ubicadas en las posiciones \mathbf{r}_i y \mathbf{r}_j , entonces:

$$u(r) = \begin{cases} \infty & r = 0 \\ -\epsilon_0 & r = 1 \\ 0 & \text{otro caso} \end{cases} \quad (12.43)$$

(compare con el potencial de Lennard-Jones visto en el capítulo 6). Asociamos a cada sitio i una variable $n_i = 0, 1$, tal que $n_i = 0$ corresponde a un sitio vacío y $n_i = 1$ a un sitio ocupado. Entonces, dada una configuración de n sitios ocupados $\{n_1, n_2, \dots, n_N\}$, con $\sum_{i=1}^N n_i = n$, la energía del gas será

$$H_g = -\epsilon_0 \sum_{\langle i, j \rangle} n_i n_j \quad (12.44)$$

y la función gran partición

$$\mathcal{Z}_g = \sum_{n=0}^N \sum_{\{n_i\}} e^{-\beta(H_g - \mu n)} \quad (12.45)$$

donde la suma $\sum_{\{n_i\}}$ es sobre todas las configuraciones de $\{n_i\}$ tales que $\sum_{i=1}^N n_i = n$.

Hacemos ahora el cambio de variables

$$n_i = \frac{1}{2}(1 + \sigma_i)$$

con $\sigma_i = \pm 1$. Tenemos que

$$H_g = -\frac{\epsilon_0}{4} \left[\sum_{\langle i,j \rangle} (\sigma_i + \sigma_j) + \sum_{\langle i,j \rangle} \sigma_i \sigma_j \right] \quad (12.46)$$

donde hemos eliminado los términos constantes. Se define el *número de coordinación* de la red como el número de primeros vecinos de un sitio cualquiera. Tenemos entonces que

$$H_g = -\frac{\epsilon_0}{4} \left[q \sum_{i=1}^N \sigma_i + \sum_{\langle i,j \rangle} \sigma_i \sigma_j \right] \quad (12.47)$$

Por otra parte

$$n = \sum_{i=1}^N n_i = \frac{N}{2} + \frac{1}{2} \sum_{i=1}^N \sigma_i \quad (12.48)$$

y así

$$H_g - \mu n = -\frac{\epsilon_0}{4} \sum_{\langle i,j \rangle} \sigma_i \sigma_j - \left(\frac{\mu}{2} + \frac{q\epsilon_0}{4} \right) \sum_{i=1}^N \sigma_i - \frac{\mu N}{2} \quad (12.49)$$

Además

$$\sum_{n=0}^N \sum_{\{n_i\}} = \sum_{\sigma_1=\pm 1} \dots \sum_{\sigma_N=\pm 1}$$

Tenemos por lo tanto que

$$\mathcal{Z}_g = \mathcal{Z}_I e^{\beta\mu N/2} \quad (12.50)$$

donde \mathcal{Z}_I es la función de partición de un modelo de Ising con $J = \epsilon_0/4$ y $B = \mu/2 + qJ$. El potencial gran canónico resulta entonces

$$\Omega_g(T, \mu, V) = -k_B T \ln \mathcal{Z}_g = F_I(T, B = \mu/2 + qJ, N) - \frac{\mu N}{2} \quad (12.51)$$

Si consideramos una red hipercúbica en d dimensiones con parámetro de red unitario tenemos que $V = N$. Así, la presión resulta

$$P = -\frac{\Omega_g}{V} = -f_I(T, B = \mu/2 + qJ) + \frac{\mu}{2} \quad (12.52)$$

donde f_I es la energía libre por spin. De la Ec.(12.48) obtenemos la segunda ecuación de estado:

$$\rho(T, \mu) \equiv \frac{\langle n \rangle}{N} = \frac{1}{2} [1 + m(T, B = \mu/2 + qJ)] \quad (12.53)$$

donde $m(T, B)$ es la magnetización por spin.

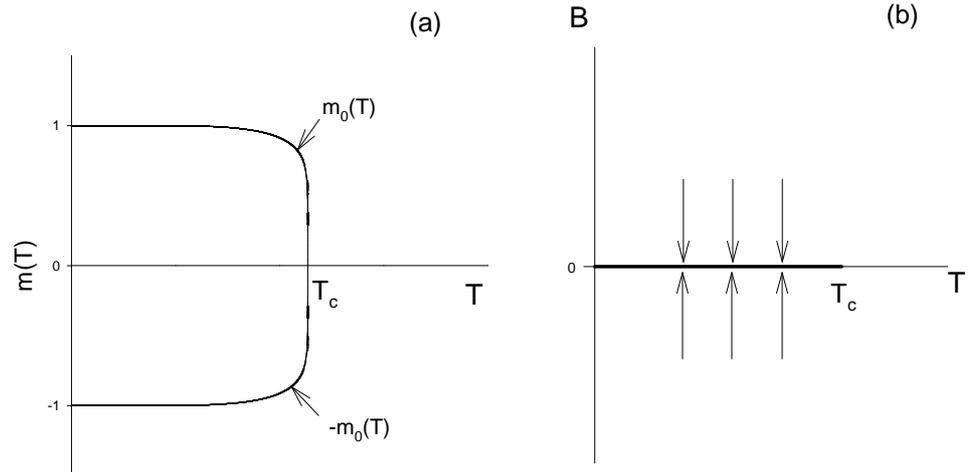


Figura 12.3: (a) Magnetización espontanea del modelo de Ising para $d \geq 2$. (b) Línea de transición de primer orden.

El modelo de Ising en dimensiones $d \geq 2$ presenta una transición de fase de primer orden para $B = 0$ y $T < T_c$, donde el parámetro de orden cambia de manera discontinua de $m_0(T)$ a $-m_0(T)$, siendo $m_0(T) = \lim_{B \rightarrow 0^+} m(B, T)$ (ver Fig.(12.3)). En la versión gas de red, dicha transición ocurre entre dos estados de diferente densidad

$$\rho(T) = \frac{1}{2} [1 \pm m_0(T)]$$

correspondiendo por lo tanto a las fases líquida y gaseosa.

12.4.1. Ejercicios

1. Incorpore la parte de energía cinética (en la aproximación clásica) a las ecuaciones de estado del gas de red.
2. Resuelva numéricamente la ecuación de Curie-Weiss y grafique las curvas de coexistencia para el gas de red en los espacios P vs. T y P vs. $v = 1/\rho$, dentro de la aproximación de campo medio.

12.5. Modelo de Ising antiferromagnético

Consideremos el Hamiltoniano de Ising con campo nulo y $J < 0$. En este caso la configuración ferromagnética con todos los spines alineados resulta el máximo absoluto de la energía. El mínimo absoluto debería corresponder ahora a una configuración en la cual todo par de spines primeros vecinos se encuentran **anti-alineados**. Es esto posible? Depende de la red. Evidentemente es posible en la red unidimensional, como se muestra en la Fig.(12.4). También es posible en la red cuadrada como se muestra en la Fig.(12.5). ¿Qué tienen en común estas redes? En ambos casos podemos dividir la red en dos subredes A y B, tal que todo sitio de una de las subredes tiene como primeros vecinos solamente sitios de la otra subred. En el caso de la cadena cada subred tiene parámetro de

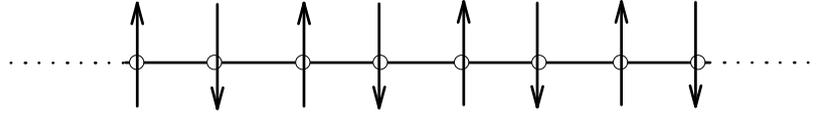


Figura 12.4: Orden antiferromagnético en la cadena lineal.

red 2 (suponiendo el parámetro de la red original 1) y en la red cuadrada cada subred es a su vez una red cuadrada con parámetro de red $\sqrt{2}$. Redes que pueden ser subdivididas de esta manera se dicen **bipartitas**. En estas redes el estado fundamental del Hamiltoniano de Ising corresponde a cada una de las subredes en configuraciones ferromagnéticas con magnetización opuesta. Estas configuraciones se conocen como **estados de Néel ó antiferromagnéticos**. A fin de cuantificar el orden antiferromagnético se define como parámetro de orden la **magnetización staggered (alternante)**:

$$m_s \equiv m_A - m_B \quad (12.54)$$

donde

$$m_A = \left\langle \frac{1}{N} \sum_{i \in A} \sigma_i \right\rangle \quad m_B = \left\langle \frac{1}{N} \sum_{i \in B} \sigma_i \right\rangle \quad (12.55)$$

A temperatura cero $m_s = 1$. A temperatura finita existe orden antiferro en redes con $d \geq 2$, esto es $m_s(T) \neq 0$ por debajo de cierta temperatura crítica T_N llamada temperatura de Néel. Para $T > T_N$ se tiene que $m_s(T) = 0$. El estado antiferromagnético tiene magnetización total nula para toda T .

Consideremos ahora la función partición del modelo de Ising a campo nulo:

$$Z(\beta, J) = \sum_{\{\sigma_i\}} \exp \left[\beta J \sum_{\langle i, j \rangle} \sigma_i \sigma_j \right] \quad (12.56)$$

Introducimos una transformación: $\sigma_i \rightarrow -\sigma_i \forall i \in B$. La función partición permanece invariante, ya que la suma corre sobre todas las configuraciones de spines. Así, tenemos que

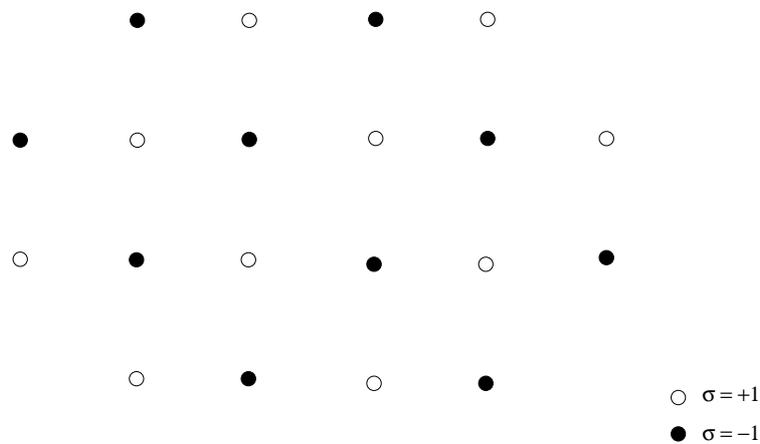


Figura 12.5: Orden antiferromagnético en la red cuadrada.

$$Z(\beta, J) = Z(\beta - J)$$

Por otra parte, la aplicación de esta transformación al cálculo de la magnetización nos da

$$m(\beta, J) = m_s(\beta, -J)$$

Así, las soluciones para el modelo antiferromagnético a campo nulo son completamente simétricas con las del modelo ferromagnético. En particular tendremos que $T_N = T_c$, donde T_c es la temperatura crítica del modelo ferromagnético. Esta simetría se rompe en presencia de campo magnético. En este caso el comportamiento del antiferro es completamente diferente del ferro. En el caso ferromagnético la presencia de campo destruye la transición de fase. En el caso antiferro la transición persiste y el efecto del campo es solamente alterar la temperatura crítica, creando una **línea crítica**, tal como se muestra en la figura 12.6. En este caso el comportamiento de la magnetización m es la de un paramagneto. De hecho el antiferromagnetismo fue propuesto por primera vez por Néel en 1932 como manera de explicar la susceptibilidad paramagnética de metales como el Cr o el Mn, la cual era mucha más grande que la predicha por la teoría de Pauli. En años posteriores pudo ser verificada la existencia del antiferromagnetismo mediante difracción de neutrones. Hoy en día se sabe que el antiferromagnetismo es mucho más frecuente que el ferromagnetismo en la naturaleza.

Finalmente, en redes que no son bipartitas no existe orden antiferromagnético. Tomemos como ejemplo la red triangular y consideremos una celda unitaria. Es imposible hacer que todos los pares estén antialineados. En esta situación el mínimo de energía corresponde a dos espines en una dirección y el tercero en la opuesta. Este estado tiene degeneración 3. En la red completa la situación es parecida, aunque mucho más compleja. Sistema con estas características se dicen **frustrados**, y están caracterizados por una degeneración altísima del estado fundamental (infinita, en el límite termodinámico), sin ningún tipo de orden magnético espacial.

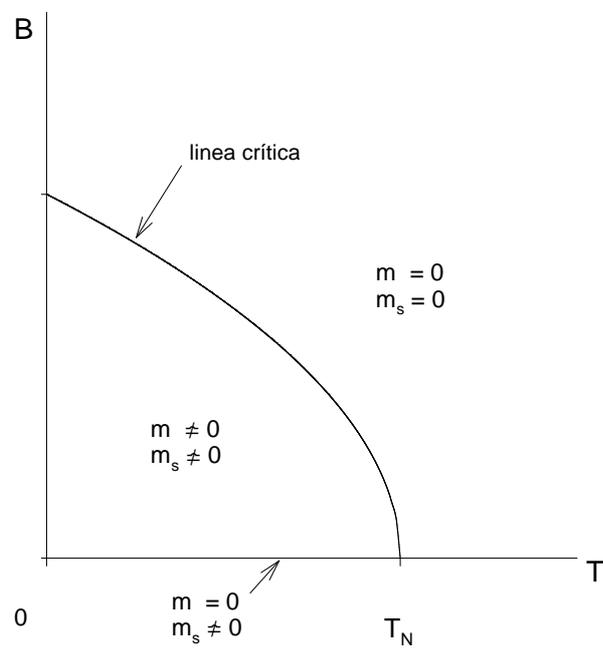


Figura 12.6: Diagrama de fases del modelo de Ising antiferromagnético con campo externo.