



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO

Grado en Física

**EL PRINCIPIO DE MÁXIMA ENTROPÍA EN FÍSICA
ESTADÍSTICA: UNA REVISIÓN CON EJEMPLOS
ILUSTRATIVOS**

Autor: Óscar Bernardo Rodríguez

**Tutores: Luis Fernando Hevia de los Mozos,
José Carlos Cobos Hernández**

2023

Resumen

En este trabajo se ha realizado una revisión crítica del Principio de Máxima Entropía y sus aplicaciones a la Física Estadística, buscando, además, ejemplos ilustrativos, con el objetivo de incorporar las innovaciones realizadas a futuras clases de la asignatura “Física Estadística”, obligatoria del Grado en Física y del doble Grado en Física y Matemáticas de la Universidad de Valladolid.

La Física Estadística es una rama de la Física que tiene como finalidad establecer una conexión entre la descripción microscópica de la realidad, que proporcionan otras materias de la Física (Mecánica Clásica, Mecánica Cuántica, Electromagnetismo, Elasticidad...), con la descripción macroscópica del mundo que proporciona la Termodinámica (que no precisa en absoluto de una descripción en términos de variables microscópicas).

Para comprender la asignatura, es necesario interiorizar una serie de conceptos que están repartidos en muchas fuentes distintas. El objetivo de este trabajo es recoger esas nociones y conceptos necesarios para la Física Estadística mediante la búsqueda de bibliografía específica, lecturas, y artículos. Tras recopilar la información necesaria, y una revisión cuidadosa por parte de los tutores de este trabajo (J. C. Cobos y L. F. Hevia), se han introducido estos conceptos por medio de unos ejemplos sencillos y didácticos, que pretenden ilustrar, para un nivel de Tercer Curso del Grado, conceptos como el Principio de Máxima Entropía y su utilidad, los estados de equilibrio termodinámico, el límite termodinámico de los sistemas, y los conjuntos o colectividades estadísticas.

En concreto, se revisan en cada capítulo los siguientes aspectos:

Capítulo 1: Se hace una introducción a dos magnitudes: la entropía termodinámica y la entropía estadística. Esta última se define para una distribución de probabilidades, y será relevante al aplicar el Principio de Máxima Entropía (PME) en los conjuntos estadísticos.

Capítulo 2: Se ilustra la metodología de trabajo con el PME en un problema sencillo: un juego de reparto de fichas en cuatro colores. Se resolverá de forma exacta y utilizando el PME, para luego comparar ambos resultados.

Capítulo 3: Se justifica el PME como un método de razonamiento adecuado para inferir probabilidades desconocidas a partir de la información conocida de un problema.

Capítulo 4: Se hace una breve introducción a los postulados de Callen, al formalismo termodinámico entrópico y a las transformadas de Legendre de la Entropía.

Capítulo 5: Mediante ejemplos sencillos, se ilustra el comportamiento de sistemas formados por muchas entidades elementales independientes.

Capítulo 6: Se hace un estudio breve de los conjuntos microcanónico y canónico, y se presentan sus ecuaciones fundamentales. Se expone también el protocolo de trabajo típico para resolver problemas de Física Estadística, y se resuelve un problema modelo con gran detalle. El apéndice B de este trabajo complementa el capítulo 6, pues contiene la aplicación completa del PME a cada conjunto estadístico, y cómo se obtienen las ecuaciones resumidas en el capítulo. Dicho apéndice contiene, además, otros dos conjuntos de interés: el gran canónico y el isotermodinámico-isobárico.

Índice

Capítulo 1: Introducción. Los distintos conceptos de entropía ..1

1.1	Entropía termodinámica	2
1.1.1	El Segundo Principio desde un punto de vista inductivo	2
1.1.2	El Segundo Principio desde un punto de vista deductivo	4
1.2	Entropía estadística	4
1.2.1	La entropía de Boltzmann-Planck	4
1.2.2	La entropía estadística de Gibbs-von Neumann-Shannon	5
1.2.3	Relación con la Teoría de la Información. Concepto de autoinformación	6
1.2.4	Entropía estadística adimensional	7

Capítulo 2: El Principio de Máxima Entropía. Ejemplos de propósito general8

2.1	Aproximación inicial a la entropía estadística	8
2.2	Un juego de reparto: apuestas en cuatro colores	9
2.2.1	Resolución exacta del problema	9
2.3	Puntos de vista en Teoría de la Probabilidad	10
2.4	El Principio de Máxima Entropía (PME o MaxEnt)	11
2.4.1	Resolución del juego de reparto con el Principio de Máxima Entropía	12

Capítulo 3: Justificación y significado del Principio de Máxima Entropía..... 16

3.1	Comportamiento asintótico de las permutaciones	16
3.1.1	Permutaciones sin repetición: aproximación de Stirling	16
3.1.2	Permutaciones con repetición	16
3.2	Justificación del Principio de Máxima Entropía	17
3.2.1	Demostración del teorema de Jaynes	18
3.3	Remarcas importantes sobre el Principio de Máxima Entropía	19

Capítulo 4: Formalismo termodinámico 20

4.1	El equilibrio termodinámico	20
4.2	Variables termodinámicas y trabajo de configuración	20
4.3	Postulados de Callen. Crítica y limitaciones	21
4.4	La representación entrópica	22
4.5	Transformadas de Legendre de la entropía	23
4.5.1	La función de Massieu	24
4.5.2	La función de Kramers-Landau (potencial gran canónico)	24

4.5.3	La función de Planck.....	24
Capítulo 5: Sistemas con muchas entidades elementales independientes: el límite termodinámico		
25		
5.1	Gas ideal en un recipiente aislado, parte I.....	25
5.2	Gas ideal en un recipiente aislado, parte II. Analogía del desierto	28
5.3	Propiedades de un sistema “ideal” en el límite termodinámico	29
5.3.1	Estrechez de la distribución de probabilidad de las medias muestrales	29
5.3.2	Estrechez de la distribución de probabilidad de las variables extensivas del sistema completo	29
5.3.3	Direccionalidad de los procesos en un sistema aislado.....	30
5.4	Vuelta al juego de reparto. Justificación de la convergencia hacia las probabilidades inferidas mediante el PME.....	30
Capítulo 6: Los conjuntos estadísticos. Aplicación y problemas modelo		
31		
6.1	El conjunto microcanónico.....	31
6.1.1	Inferencia de probabilidades	31
6.1.2	Ecuaciones de estado y funciones respuesta.....	32
6.2	El conjunto canónico	32
6.2.1	Inferencia de probabilidades	33
6.2.2	Ecuaciones de estado, funciones respuesta y fluctuaciones.....	33
6.3	Protocolo de trabajo en Física Estadística	34
6.4	Problema modelo.....	34
Conclusiones		40
Bibliografía		41
Apéndice A: Complementos al Capítulo 2		
43		
A.1.	Tablas de tabulación directa del juego de reparto.....	43
A.1.1.	$N = 5$	43
A.1.2.	$N = 15$	43
A.1.3.	$N = 20$	44
A.2.	Comentario sobre la maximización de la entropía.....	45
Apéndice B: Desarrollo de los conjuntos estadísticos.....		
46		
B.1.	El conjunto microcanónico.....	46
B.1.1.	Inferencia de probabilidades	47
B.1.2.	Ecuaciones de estado	48
B.1.3.	Funciones respuesta.....	49

B.2. El conjunto canónico	49
B.2.1. Inferencia de probabilidades	50
B.2.2. Ecuaciones de estado	53
B.2.3. Funciones respuesta y fluctuaciones de energía.....	54
B.3. El conjunto gran canónico	56
B.3.1. Inferencia de probabilidades	56
B.3.2. Ecuaciones de estado.....	58
B.3.3. Funciones respuesta y fluctuaciones	59
B.4. El conjunto isoterma-isobárico.....	59
B.4.1. Inferencia de probabilidades	60
B.4.2. Ecuaciones de estado.....	62
B.4.3. Funciones respuesta y fluctuaciones	62
B.5. Resumen de los diferentes conjuntos estadísticos	64
Apéndice C: El método del término máximo	65
Apéndice D: Las probabilidades a priori. Los antecedentes históricos del PME.	68
D.1. Introducción. Las probabilidades a priori.....	68
D.2. El argumento de Maxwell.....	68
D.3. El argumento de la distribución de Boltzmann	69

Índice de tablas

Tabla 2.1. Diferentes apuestas posibles en el juego de reparto propuesto	9
Tabla 2.2. El juego de reparto: resultados exactos por tabulación directa, y resultados con MaxEnt.	14
Tabla 3.1. Error cometido con la aproximación de Stirling. Cuanto mayor es N, menor es el error cometido.	16
Tabla 5.1. Correspondencias en la analogía del desierto.....	28

Índice de figuras

Figura 2.1. Tendencia de las probabilidades exactas a las obtenidas por MaxEnt.....	15
Figura 5.1. Un gas encerrado en una caja	25
Figura 5.2. Forma de la distribución.....	26
Figura 5.3. Gas encerrado en dos mitades.....	28
Figura 6.1. Dos niveles de energía.....	35
Figura 6.2. Gráfica de $-\beta\varepsilon(p_1)$	37
Figura 6.3. Gráficas de la dependencia de la EME con variables de interés.....	38
Figura 6.4. Gráficas de la dependencia de la actividad absoluta con dos variables de interés.....	38

Capítulo 1: Introducción. Los distintos conceptos de entropía

La Física es la ciencia experimental más fundamental, cuyo doble objetivo es conocer para intentar comprender la realidad que nos rodea. Como seres humanos que somos, interaccionamos con nuestro entorno, y tratamos de entenderlo asociando ideas y conceptos a aquello que somos capaces de percibir con nuestros sentidos (Einstein afirmó que: “*la física es el estudio de las percepciones sensoriales impersonales*”). A partir de la experiencia somos capaces de idear modelos de realidad y teorías que son capaces de describir (de peor o mejor forma) los sucesos que observamos.

La entropía es una magnitud termodinámica (función de estado, además) que puede resultar un tanto extraña cuando se enseña a los estudiantes de Física, pues la primera impresión que se tiene de ella es que no se parece en nada a otras magnitudes físicas, como pueden ser por ejemplo la masa, la longitud o la fuerza, de las que se tienen conceptos, nociones e ideas que parecen en principio más claras. Y en nuestro caso, aún peor, se confunden habitualmente los conceptos de temperatura y calor, que son los preconceptos en que está basada su definición.

¿Cómo encaja el concepto de entropía en la realidad que nos rodea y qué significado físico tiene? En este trabajo trataremos de dar respuesta a esta pregunta, introduciendo el Segundo Principio de la Termodinámica y exponiendo las definiciones y propiedades de dos magnitudes –en principio– diferentes, pero que deben estar íntimamente relacionadas si debe cumplirse que la descripción empírica experimental coincida con la explicación teórica deductiva:

- La entropía termodinámica, que es una propiedad macroscópica de los sistemas (en equilibrio termodinámico, o en estados generalizados de no-equilibrio), y no depende de ningún modelo de realidad. La Termodinámica es una ciencia macroscópica que no necesita para nada una descripción en términos de variables microscópicas. Eso fue lo que la hizo insensible a los cambios que generaron en el resto de las ramas de la Física tanto la revolución relativista como la cuántica.
- La entropía estadística, que es una función matemática característica de una distribución de probabilidad. Aunque existen diversas definiciones de entropía estadística, dependiendo del valor de autoinformación utilizado, e incluso generalizando a propiedades no-extensivas, nosotros utilizaremos la así llamada entropía de Gibbs-von Neumann-Shannon, ya que es el límite común de muchas de ellas bajo determinadas condiciones físicas razonables, siendo la que se aplica a la descripción estadística de los sistemas macroscópicos si nuestro modelo de realidad parte de la idea de que conocemos el valor medio (primer momento) de las variables características “microscópicas” que lo describen.

Obviamente, las ecuaciones teóricas que genera el formalismo estadístico para describir los sistemas reales deben coincidir con las ecuaciones empíricas macroscópicas que se utilizan en Termodinámica. En resumen, aceptamos el postulado fundamental de la Física Estadística que postula que “la realidad macroscópica no es más que el promedio de las variables microscópicas que caracterizan a los sistemas reales”.

1.1 Entropía termodinámica

1.1.1 El Segundo Principio desde un punto de vista inductivo

La necesidad del Segundo Principio de la Termodinámica surge del hecho de que existen procesos de un sistema que, aun siendo compatibles con el Primer Principio de la Termodinámica, es decir, cumpliendo la conservación de la energía, nunca se observan en la naturaleza. Según Biel (1998), “*aun sin salirse de la limitación impuesta por el Primer Principio (la energía no se crea ni se destruye, sólo se conserva) no toda la energía en forma de calor puede transformarse siempre en otra clase de energía (ni siquiera en un experimento ideal) y no pueden invertirse todos los procesos de transformación de energía*”.

Los siguientes ejemplos sencillos dan cuenta de la direccionalidad natural (asimetría espacial y temporal) de ciertos procesos:

- Ejemplo 1. Un bloque de hielo se sitúa en una habitación a temperatura ambiente (mayor que la que posee el hielo en equilibrio con agua líquida). Su evolución natural es calentarse, fundirse y pasar totalmente al estado líquido. Se produce un flujo de energía en forma de calor entre el cristal de agua y su entorno, de tal forma que es el aire el que calienta al hielo. Sin embargo, nunca se ha observado que el agua líquida forme un bloque de hielo a temperatura ambiente, espontáneamente, calentando el aire a su alrededor. Ambos procesos conservan la energía y cumplen con el Primer Principio de la Termodinámica, pero nadie ha observado el segundo proceso.
- Ejemplo 2. Un recipiente adiabático e impermeable se encuentra dividido en dos mitades por un separador (ya sea una puerta, una válvula que las conecta, etc.). Una mitad está llena de aire, y la otra está vacía. Al retirar el separador, el aire se expande ocupando todo el volumen de la cámara, y sin embargo nunca se ha observado que, espontáneamente, el aire vuelva a ocupar sólo una mitad del recipiente.

Los procesos que se han mencionado en los ejemplos 1 y 2 se realizan espontáneamente en una determinada dirección y no en la contraria. Así pues, el Primer Principio de la Termodinámica no es suficiente para explicar el sentido de evolución de los sistemas y se precisa de un nuevo principio que complementa y completa al primero.

Mientras el trabajo W puede convertirse íntegramente en calor Q de forma continua (por medio de procesos cíclicos del sistema activo de una máquina térmica), la transformación continua del calor en trabajo no puede producirse al 100%. La transformación continua de calor en trabajo exige una cierta compensación, que se puede dar de dos formas diferentes:

- Se produce la cesión de calor a un sistema con temperatura más baja (foco frío) que la del sistema del que se absorbe el calor (foco caliente).
- La transformación de calor en trabajo implica una modificación del estado del sistema (proceso no cíclico).

A lo largo del siglo XIX, diferentes físicos enunciaron el Segundo Principio de la Termodinámica en base al estudio de las máquinas térmicas. Estos enunciados se encuentran recogidos en el Tomo I del libro “*Termodinámica*” de Tejerina (1976):

Enunciado de Carnot (1824): “*Para que una máquina térmica cíclica produzca trabajo a expensas del calor que recibe de una fuente caliente, es necesario que energía en forma de calor pase a otra fuente más fría.*”

Enunciado de Kelvin-Planck (1852): “*Es imposible construir una máquina térmica que funcionando cíclicamente convierta en trabajo toda la energía que recibe en forma de calor de*

una fuente térmica.” La máquina térmica que describe este enunciado es el *móvil perpetuo de segunda especie*, negando su existencia.

Enunciado de Clausius (1850): “*Es imposible construir una máquina térmica cíclica cuyo único resultado sea pasar energía en forma de calor de una fuente fría a otra caliente.*”

Aunque históricamente el Segundo Principio de la Termodinámica surge del estudio de las máquinas térmicas, no es exclusivo de sistemas que formen parte de ellas. Este razonamiento inductivo con máquinas térmicas permite introducir dos nuevas variables para todo sistema termodinámico:

- La **temperatura termodinámica**, T , definida de forma independiente de las sustancias y propiedades termométricas empíricas usadas para su definición y medida experimental.
- La **entropía termodinámica**, S , variable extensiva (función homogénea de grado 1 de las variables extensivas naturales de las que depende $[N, U, V]$ –capítulo 4–), cuyo valor no puede disminuir en un proceso adiabático ($Q = 0$).

En particular, los sistemas aislados están aislados adiabáticamente, por lo que la variación de entropía en tales sistemas no puede ser negativa.

Al conjunto formado por un sistema y su entorno se le suele llamar **universo termodinámico**, que es **un sistema finito y aislado, pero que no es el universo en el sentido cosmológico**. Como la entropía es una variable extensiva, la entropía de un universo termodinámico será la suma de las entropías del sistema y de su entorno. Además, como un universo termodinámico es un sistema aislado:

$$\Delta S_{\text{universo}} = \Delta S_{\text{sistema}} + \Delta S_{\text{entorno}} \geq 0 \quad (1.1)$$

Es decir, la entropía del universo termodinámico no puede disminuir. Esta desigualdad permite clasificar los procesos que se dan en un universo termodinámico:

- Se dice que un proceso es reversible cuando es posible invertir exactamente los intercambios de energía del sistema y del entorno para devolverlos a su estado inicial. Para estos procesos (en la práctica irrealizables) $\Delta S_{\text{universo}} = 0$.
- Se dice que un proceso es irreversible en caso contrario. Para estos procesos, $\Delta S_{\text{universo}} > 0$

La variación de entropía entre dos estados de equilibrio en un sistema cerrado puede calcularse mediante la igualdad de Clausius:

$$dS = \frac{\delta Q_{\text{rev}}}{T} \quad (1.2)$$

donde δQ_{rev} es el calor intercambiado en un proceso infinitesimal reversible que conecte los estados inicial y final.

Dado que los procesos adiabáticos deben también estar caracterizados por una variación de entropía no negativa, se puede analizar el carácter reversible o irreversible del proceso adiabático de un sistema sin necesidad de analizar la variación de entropía de su entorno:

- Principio de aumento de entropía en procesos adiabáticos. En un proceso adiabático la entropía no puede disminuir:

$$\Delta S_{\text{adiabático}} \geq 0 \quad \rightarrow \quad \begin{cases} \Delta S_{\text{adiabático}} = 0 & \Leftrightarrow \text{Proceso adiabático reversible} \\ \Delta S_{\text{adiabático}} > 0 & \Leftrightarrow \text{Proceso adiabático irreversible} \end{cases}$$

1.1.2 El Segundo Principio desde un punto de vista deductivo

Al igual que se hace con las formulaciones modernas de la Mecánica, el Segundo Principio de la Termodinámica puede enunciarse postulando directamente la existencia y las propiedades del equilibrio termodinámico y de la entropía (Callen, 1985). En esta formulación, el inverso de la temperatura se define como la variable canónica conjugada de la energía (capítulo 4). Esta variable (el inverso de la temperatura) es, en realidad, mucho más adecuada que la propia temperatura, que evita discusiones superfluas acerca de la inaccesibilidad del “cero absoluto de temperatura” y permite introducir las temperaturas negativas de manera natural y sin discontinuidades que distorsionan la interpretación del significado de las mismas.

Continuaremos esta exposición en el capítulo 4. Allí, enunciaremos y comentaremos críticamente los postulados de Callen.

1.2 Entropía estadística

En Física Estadística, existen dos formas diferentes de introducir la entropía (existen muchas más, en función de la autoinformación que utilicemos, pero no lo estudiaremos aquí). Por un lado se tiene la entropía de Boltzmann–Planck, y por otro lado la entropía de Gibbs–von Neumann–Shannon, siendo esta última mucho más general y aquella un caso particular de ésta. Las estudiaremos en los siguientes apartados.

1.2.1 La entropía de Boltzmann-Planck

Boltzmann fue uno de los primeros físicos en relacionar la entropía y la probabilidad. Definió la entropía como una medida del número de posibles microestados que puede adoptar un sistema aislado en equilibrio termodinámico, consistente con sus propiedades macroscópicas; es decir, con el macroestado del sistema.

Por ejemplo, a nivel microscópico, un gas estará formado por un gran número de átomos o de moléculas (hipótesis atómica) que se mueven libremente e interaccionan entre sí y con las paredes del recipiente. Desde el punto de vista de la Mecánica Clásica, un microestado sería una descripción de las posiciones y momentos de todas las partículas del gas (espacio de las fases). Al ser un gran número de partículas (del orden de 10^{23}), existe un número abismalmente grande de microestados, es decir, de diferentes configuraciones posibles. Sin embargo, su comportamiento colectivo da lugar a propiedades macroscópicas bien definidas, denominadas variables termodinámicas, como son el volumen, la presión y la temperatura.

M. Planck (1914), en las páginas 119-120 de su libro “*The Theory of Heat Radiation*”, comenta: “*La conexión logarítmica entre entropía y probabilidad fue establecida por primera vez por L. Boltzmann en su teoría cinética de gases. [...] A diferencia de [Boltzmann], nosotros asignamos un cierto valor absoluto a la entropía*”. Con ello, Planck está haciendo referencia a la relación que expone en ese mismo libro entre la entropía S y el número de microestados Ω **de un sistema aislado** (que está caracterizado por unas ligaduras macroscópicas conocidas como son el número de partículas, la energía interna y los desplazamientos generalizados que dan cuenta del trabajo mecánico de configuración – el volumen, por ejemplo –, que permanecen fijas):

$$S = k_{\text{B}} \ln \Omega \quad (1.3)$$

donde k_{B} es la constante de Boltzmann, que da dimensiones a la entropía. Como el número de microestados Ω es un número entero positivo y mayor que 1 (no se puede tener medio microestado), la entropía así definida será nula cuando se tenga un único microestado, y positiva en el resto de los casos.

No obstante, la entropía de Boltzmann-Planck (1.3), grabada sobre la tumba de L. Boltzmann en el cementerio central de Viena, resulta ser un caso particular de la entropía de J. W. Gibbs (1902), también utilizada por J. von Neumann (1932) en sus “*Mathematical Foundations of Quantum Mechanics*”, para explicar el proceso de medida en Mecánica Cuántica; y por C. E. Shannon (1948), al desarrollar “*A Mathematical Theory of Communication*” (razón por la cual también se la conoce como entropía de información). Dicha entropía de Gibbs–Von Neumann–Shannon es la que llamaremos **entropía estadística**, magnitud central de este trabajo.

1.2.2 La entropía estadística de Gibbs-von Neumann-Shannon

La entropía estadística de Gibbs–von Neumann–Shannon nos permitirá inferir las probabilidades asociadas a microestados que dependan de ligaduras macroscópicas que describen sistemas no-aislados, que intercambian con su entorno materia y energía, y se estructuran microscópicamente de acuerdo con su entropía. Esta magnitud es mucho más adecuada a la hora de discutir la conexión entre las propiedades físicas macroscópicas y las microscópicas de las que son consecuencia, y se escribe en términos de las distribuciones de probabilidad asociadas al conjunto de posibles microestados en los que puede presentarse el sistema.

Dada la distribución de probabilidades $\{p_1, p_2, \dots, p_\Omega\}$ asociada a los Ω sucesos **elementales** de una variable aleatoria discreta (que, en particular, son mutuamente excluyentes, de forma que si ocurre un suceso elemental no ocurren los otros), donde p_i es la probabilidad correspondiente al suceso elemental i -ésimo; definimos la **entropía estadística** S como:

$$S(p_1, \dots, p_\Omega) = -k \sum_{i=1}^{\Omega} p_i \ln p_i = \sum_{i=1}^{\Omega} p_i \left(k \ln \frac{1}{p_i} \right) \quad (1.4)$$

donde k es una constante de proporcionalidad que además introduce las unidades en las que se mide la entropía (en Física Estadística, el valor de la constante es precisamente la constante de Boltzmann, de forma que $k = k_B = 1.380649 \cdot 10^{-23}$ J/K). En esta definición, se admite que cuando $p_i = 0$ el término $p_i \ln p_i = 0$, lo que equivale a admitir una extensión por continuidad, ya que $\lim_{p_i \rightarrow 0^+} (p_i \ln p_i) = 0$.

En el caso de sucesos elementales equiprobables, la entropía de Gibbs–von Neumann–Shannon se reduce a la entropía de Boltzmann-Planck. Este hecho se puede demostrar fácilmente.

Consideremos una variable aleatoria discreta con Ω sucesos elementales y equiprobables, de forma que tiene asociada una distribución de probabilidades $\left\{ \frac{1}{\Omega}, \frac{1}{\Omega}, \dots, \frac{1}{\Omega} \right\}$. Llevando estas condiciones a la entropía estadística (1.4):

$$S = -k \sum_{i=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} = -k \frac{\Omega}{\Omega} \ln \frac{1}{\Omega} = -k \ln \Omega^{-1} = k \ln \Omega$$

obteniéndose la expresión de Boltzmann–Planck para la entropía. Este hecho ya permite ver la equiprobabilidad de los microestados accesibles a un sistema aislado, pero volveremos sobre ello en el capítulo 6.

J. von Neumann (1932) escribió la fórmula correspondiente a la entropía estadística (1.4) en términos del operador densidad ρ en Mecánica Cuántica (generalización del concepto de estado cuántico de un sistema que incluye mezclas estadísticas de estados cuánticos):

$$S = -k_B \operatorname{tr}(\rho \ln \rho) \quad (1.5)$$

1.2.3 Relación con la Teoría de la Información. Concepto de autoinformación

En la Teoría de la Comunicación de Shannon, el esquema fundamental de funcionamiento a caracterizar consiste en estudiar qué ocurre con la señal que envía un emisor (mensaje) a un receptor por medio de un canal de comunicación, y cómo puede afectarle el ruido que exista en dicho canal de comunicación.

Shannon (1948) demostró en su artículo “*A Mathematical Theory of Communication*” que se puede obtener una estimación precisa del contenido de información que puede transmitirse con una señal. Se trata de una idea muy importante, pues nos permite caracterizar inequívocamente cuánta información podemos transmitir sin problemas.

- Si una señal porta un mensaje que tiene una probabilidad muy alta de ocurrir, entonces no ganamos mucha información cuando decodificamos dicho mensaje.
- Por otro lado, si una señal porta un mensaje que tiene una baja probabilidad de ocurrir, cuando nos llega ese mensaje ganamos una cantidad muy significativa de información.

Shannon cuantificó esto tomando el logaritmo en base 2 de la probabilidad de que ocurra un cierto mensaje. Entonces, si denotamos como I la información contenida en un mensaje, también conocida como **autoinformación**, y la probabilidad de que ocurra dicho mensaje como p , tenemos:

$$I = -\log_2 p \quad (1.6)$$

Shannon (1948) definió la autoinformación de forma que cumpliera los siguientes axiomas:

1. Un acontecimiento con una probabilidad de ocurrir del 100% no es sorprendente y no aporta ninguna información. Si $p = 1$, entonces $I = 0$.
2. Cuanto menos probable es un suceso, más sorprendente es, y más información aporta. Si $p = 0$ entonces $I = +\infty$.
3. Si se miden por separado dos sucesos independientes, la cantidad total de información es la suma de las autoinformaciones de los sucesos individuales.

La unidad en la que se mide la autoinformación (1.6) es el “*bit*” (*contracción de “binary digit”*). *John Tukey, 1947, Laboratorios Bell*, aunque también se le conoce históricamente como el “*shannon*”. La definición histórica de bit es:

$$\log_2 2 = 1 \text{ bit}$$

El signo negativo en la ecuación (1.6) asegura que la información contenida en un mensaje sea positiva (pues $p < 1$, para el caso de esta definición asociada a Shannon –no siempre es así y hay probabilidades extendidas mayores que 1 o que son negativas–), y que cuanto menos probable sea el mensaje, mayor sea la información contenida.

Resumimos el uso de los logaritmos para caracterizar la información contenida en una señal afirmando lo siguiente:

- Un mensaje transmitido que es poco probable, tiene asociada una baja probabilidad, y por lo tanto tiene un gran contenido de información.
- Un mensaje transmitido que es muy probable, tiene asociada una alta probabilidad, y por lo tanto tiene un pequeño contenido de información.

De forma más general, puede medirse la autoinformación en unidades arbitrarias:

$$I = -k \log p \quad (1.7)$$

donde si $k = 1 / \log 2$ recuperamos la autoinformación en bits (1.6).

Dada una distribución discreta de probabilidad $\{p_1, p_2, \dots, p_\Omega\}$, el valor medio (primer momento de la distribución de probabilidad) de una variable aleatoria discreta X , que toma los valores $\{x_1, x_2, \dots, x_\Omega\}$ en los distintos sucesos elementales, se calcula mediante la ecuación:

$$\langle X \rangle = \sum_{i=1}^{\Omega} p_i x_i \quad (1.8)$$

La entropía de información se define, entonces, como el valor medio de la autoinformación:

$$\langle I \rangle = \sum_i p_i I_i = \sum_i p_i (-k \ln p_i) = -k \sum_i p_i \ln p_i \quad (1.9)$$

que coincide con la ecuación (1.4).

De esta forma establecemos que **la entropía estadística es una medida de la aleatoriedad o falta de información asociada a una distribución de probabilidad.**

1.2.4 Entropía estadística adimensional

En ocasiones conviene trabajar con una entropía estadística adimensional $\sigma = S / k$ en la que se omite la constante k que da las unidades a la entropía estadística:

$$\boxed{\sigma(p_1, \dots, p_\Omega) = \frac{S(p_1, \dots, p_\Omega)}{k} = -\sum_{i=1}^{\Omega} p_i \ln p_i} \quad (1.10)$$

Así debería haberse definido la entropía, en su momento, con lo que se hubiesen evitado muchas malinterpretaciones que se hacen de la magnitud entropía, la magnitud temperatura se mediría en unidades de energía, y no habría hecho falta introducir la constante de Boltzmann.

Capítulo 2: El Principio de Máxima Entropía. Ejemplos de propósito general

En este segundo capítulo se exponen ejemplos para ilustrar el significado de la entropía estadística y la metodología de su uso para inferir probabilidades desconocidas cuando hay poca información disponible. Dicha metodología se apoya en el llamado Principio de Máxima Entropía (Jaynes, 1957), cuyas condiciones de validez también se ilustrarán con un ejemplo.

2.1 Aproximación inicial a la entropía estadística

Para obtener un mayor conocimiento y familiaridad con esta magnitud, vamos a estudiar algunos ejemplos sencillos en los que se aplica y su interpretación en términos de información (Lee, 1988; Baixauli y Cobos, 1996).

Ejemplo 1. Una moneda

Cuando lanzamos al aire una moneda ideal no trucada, la probabilidad de obtener cara (suceso 1) es $p_1 = 0.5$; y la de obtener cruz (suceso 2) es $p_2 = 0.5$. Tenemos así una distribución de probabilidad de la forma $\{p_1, p_2\} = \{0.5, 0.5\}$. La entropía estadística para este sistema será la siguiente:

$$S_1 = -k(p_1 \ln p_1 + p_2 \ln p_2) = -k[0.5 \ln(0.5) + 0.5 \ln(0.5)] = -k \ln 0.5 = k \ln 2 = 0.69315k$$

Ejemplo 2. Un cruce de cuatro caminos

Un turista, que está buscando el camino para llegar al museo de la ciudad, llega a un cruce y no sabe por qué calle continuar. ¿Cuál será la entropía estadística asociada a este caso? Denotaremos las cuatro direcciones posibles del cruce como E, O, N y S . Hay entonces cuatro posibilidades (cuatro sucesos). La distribución de probabilidades será $\{p_E, p_O, p_N, p_S\} = \{1/4, 1/4, 1/4, 1/4\}$ y la entropía estadística para este ejemplo es:

$$S_2 = -k(0.25 \ln 0.25 + 0.25 \ln 0.25 + 0.25 \ln 0.25 + 0.25 \ln 0.25) = k \ln 4 = 1.38629k$$

Pero si algún peatón le indicase el camino, la incertidumbre desaparecería. Por lo tanto, la cantidad $1.38629k$ puede ser interpretada por ejemplo como el valor de la información “*ir hacia el oeste*”. Vemos así que la entropía estadística mide la aleatoriedad inherente a nuestro conocimiento de la realidad, estando estrechamente relacionada con esa incertidumbre.

Ejemplo 3. Un dado ideal

Por último, si tiramos un dado ideal no trucado, la probabilidad de obtener cada una de las seis caras es $1/6$. En este caso, la distribución de probabilidad asociada será $\{p_1, p_2, p_3, p_4, p_5, p_6\} = \{1/6, 1/6, 1/6, 1/6, 1/6, 1/6\}$, y su entropía es:

$$S_3 = k \ln 6 = 1.79176k$$

Al comparar los dos primeros ejemplos observamos que la probabilidad en el primero es $p_i = 0.5$ y la entropía es $S_1 = 0.69315k$, mientras que en el segundo $p_i = 0.25$ y la entropía se duplica hasta $S_2 = 1.38629k$. Así pues, en el caso de sucesos equiprobables, cuanto menor sea la probabilidad mayor es la incertidumbre. La entropía es mayor cuando nuestro desconocimiento (en este caso el número de opciones) es mayor, algo que se ajusta al concepto intuitivo de

información. En el caso de la moneda hay dos posibilidades y $S_1 = k \ln 2$, en el cruce de caminos hay cuatro posibilidades y $S_2 = k \ln 4$. Por último, en el dado tenemos seis posibilidades y la entropía es $S_3 = k \ln 6$. Si se generaliza esto a Ω sucesos elementales y equiprobables de una variable aleatoria discreta, se tiene $S = k \ln \Omega$, que es la entropía de Boltzmann-Planck (1.3).

2.2 Un juego de reparto: apuestas en cuatro colores

En la sección anterior se ha podido asociar una magnitud a cada una de las distribuciones de probabilidad de los ejemplos utilizados: una moneda, un cruce de caminos y un dado ideal. Esta magnitud es la entropía estadística (1.4), que da cuenta de la libertad de aparición de todos los posibles sucesos, o visto de otra forma, de la incertidumbre acerca del resultado concreto que aparecerá en cada caso. Ilustraremos la metodología de trabajo con el Principio de Máxima Entropía con el siguiente ejemplo sencillo, adaptado (Hevia, 2023) a partir de otro presentado en la monografía de Lee (1988):

Un juego de reparto: En este juego especial, disponemos de 10 fichas idénticas por partida a repartir entre los cuatro posibles colores que hay en una mesa: rojo, negro, amarillo y verde. Apostar una ficha a un color tiene los siguientes costes:

- Rojo: $d_1 = 4\text{€}$
- Negro: $d_2 = 2\text{€}$
- Amarillo: $d_3 = 1\text{€}$
- Verde: $d_4 = 0\text{€}$

En cada partida hay que gastar obligatoriamente 10 euros en total repartiendo las 10 fichas entre los cuatro colores. *¿Cuál será la probabilidad de que un color reciba una ficha?*

2.2.1 Resolución exacta del problema

En primer lugar, resolveremos este problema utilizando el método de tabulación directa.

Número de apuesta (j)	Rojo: n_{1j} $d_1 = 4\text{€}$	Negro: n_{2j} $d_2 = 2\text{€}$	Amarillo: n_{3j} $d_3 = 1\text{€}$	Verde: n_{4j} $d_4 = 0\text{€}$	Degeneración (Ω_j)
1	0	0	10	0	1
2	0	1	8	1	90
3	0	2	6	2	1260
4	0	3	4	3	4200
5	0	4	2	4	3150
6	0	5	0	5	252
7	1	0	6	3	840
8	1	1	4	4	6300
9	1	2	2	5	7560
10	1	3	0	6	840
11	2	0	2	6	1260
12	2	1	0	7	360
Total fichas a un color: F_i	18780	53370	79270	109710	$\Omega = \sum_{j=1}^{12} \Omega_j = 26113$
Probabilidad	0.071918	0.204381	0.303565	0.420136	$N\Omega = 261130$

Tabla 2.1. Diferentes apuestas posibles en el juego de reparto propuesto

El objetivo del método de tabulación directa es obtener todas las combinaciones posibles de apuestas permitidas teniendo en cuenta la información del enunciado: en cada apuesta se han de gastar 10 euros ($D = 10 \text{ €}$) repartiendo 10 fichas ($N = 10$) entre los cuatro colores.

En la Tabla 2.1 se muestran las diferentes configuraciones posibles del juego. En cada fila se ha escrito el número n_{ij} de fichas al color i en la apuesta j . El número total de sucesos elementales es Ω , que son todas las formas en las que se puede realizar el reparto.

Permutaciones con repetición. Sea un conjunto de N elementos en el que se tiene un elemento x_1 repetido n_1 veces, un elemento x_2 repetido n_2 veces, ... , y un elemento x_s repetido n_s veces. Sus permutaciones con repetición son las diferentes formas de ordenar sus N elementos:

$$PR_N^{n_1, \dots, n_s} = \frac{N!}{n_1! n_2! \dots n_s!} \quad (2.1)$$

Las degeneraciones incluidas en la tabla se deben a las diferentes permutaciones con repetición de las fichas en cada una de las posibles apuestas. Por ejemplo, en la última apuesta de la tabla (caso 12) hay $PR_{10}^{2,1,0,7}$ formas de seleccionar 2 fichas rojas, 1 negra, 0 amarillas y 7 verdes. El número total de posibilidades de seleccionarlás será:

$$PR_{10}^{2,1,0,7} = \Omega_{12} = \frac{10!}{(n_{1,12}!)(n_{2,12}!)(n_{3,12}!)(n_{4,12}!)} = \frac{10!}{(2!)(1!)(0!)(7!)} = 360$$

La frecuencia absoluta F_i con la que aparece en el espacio muestral la ficha correspondiente al color i se calcula sumando el número de fichas de ese color en todas las combinaciones posibles:

$$F_i = \sum_{j=1}^{12} n_{ij} \cdot \Omega_j$$

y el número total de fichas apostadas a todos los colores es $\sum_{i=1}^4 F_i = N\Omega = 261130$.

Para conocer las probabilidades, se divide el número total de fichas F_i apostadas al color i entre el número total de fichas apostadas $N\Omega$. Así, se obtienen las probabilidades de que un determinado color reciba una ficha: $p_i = F_i / (N\Omega)$

$p_{1(\text{rojo})} = \frac{18780}{261130} = 0.071918$	$p_{2(\text{negro})} = \frac{53370}{261130} = 0.204381$
$p_{3(\text{amarillo})} = \frac{79270}{261130} = 0.303565$	$p_{4(\text{verde})} = \frac{109710}{261130} = 0.420136$

De esta forma se ha resuelto el problema planteado. En principio se puede aplicar este procedimiento en cualquier problema de reparto. Sin embargo, cuando el número de fichas se incrementa hasta millones, y el número de colores posibles es igual de elevado, el método de tabulación directa no resulta práctico en absoluto. Mostramos esto en el **Apéndice A.1**, en el que hemos resuelto el juego de reparto para $N = 5, 15$ y 20 fichas, con $D = 5, 15$ y 20 € respectivamente. Si nos fijamos, la tabla para $N = 20$ ya es demasiado extensa (Hevia, 2023), y el método no resultará nada práctico para valores mayores de N .

Es por ello por lo que se propone utilizar el **Principio de Máxima Entropía** para la resolución de este tipo de problemas.

2.3 Puntos de vista en Teoría de la Probabilidad

El problema de la asignación de probabilidades en los casos en los que hay muy poca información disponible es tan antiguo como la Teoría de la Probabilidad. Para analizarlo en profundidad, se debe reconocer el hecho de que la Teoría de la Probabilidad se ha desarrollado

en dos direcciones diferentes en cuanto a sus fundamentos (Jaynes, 1957; Jaynes, 1978; Jaynes, 2003; Hevia, 2014; Hevia, 2023):

- El **punto de vista frecuentista** considera la probabilidad de un suceso como una **propiedad del propio suceso**, verificable en experimentos mediante la observación de **frecuencias relativas**. Cuando un frecuentista calcula una distribución de probabilidad, cree que hace predicciones que son, en principio, verificables al detalle, como las que hace la Mecánica Clásica. Una distribución de probabilidad es buena desde el punto de vista frecuentista cuando **representa correctamente las fluctuaciones** observables de una variable aleatoria.
- Por otro lado, el **punto de vista bayesiano** mira las probabilidades como **expresiones formales de la ignorancia humana**; la probabilidad de un suceso indica simplemente la **esperanza de que un suceso ocurra, basada en cualquier información disponible**. Para el bayesiano, el fin de la Teoría de la Probabilidad es servir de herramienta para sacar conclusiones plausibles en casos en los que no hay suficiente información para obtener conclusiones seguras; por tanto, no se espera una verificación detallada. Una distribución de probabilidad es buena desde el punto de vista bayesiano cuando **representa correctamente la información** que se tiene del valor de una variable aleatoria.

Aunque las teorías basadas en los puntos de vista frecuentista y bayesiano son matemáticamente idénticas, los conceptos que subyacen no lo son. En los problemas estadísticos que se presentan en Física, ambos puntos de vista son necesarios.

2.4 El Principio de Máxima Entropía (PME o MaxEnt)

En todo problema que requiera de la aplicación de la Teoría de la Probabilidad, se ha de realizar un primer paso importante: asignar unos valores iniciales de probabilidades a los diferentes sucesos elementales de un experimento aleatorio (Jaynes, 1978). Esta asignación inicial de probabilidades fue descrita y enunciada por J. Bernoulli en 1713 con el nombre de **Principio de Indiferencia**: en ausencia de información que indique lo contrario, los Ω sucesos elementales que conforman el espacio muestral de un experimento aleatorio han de ser equiprobables, de forma que a cada uno de ellos se le asigna la probabilidad $1/\Omega$. El **Principio de Máxima Entropía** consiste en una generalización del Principio de Indiferencia (Jaynes, 1957; Jaynes, 1978; Jaynes, 2003; Hevia, 2023):

El problema a resolver es el siguiente: partiendo de una determinada información sobre un experimento aleatorio, se quiere conocer la distribución de probabilidades más adecuada que se debe asignar al conjunto de sucesos elementales del experimento, de forma que dicha distribución sea compatible con la información conocida.

El Principio de Máxima Entropía establece que la distribución de probabilidad más adecuada es aquella que, teniendo en cuenta la información disponible, maximiza la entropía estadística.

La información disponible se introduce en el problema en forma de unas ciertas ligaduras entre las probabilidades de la distribución.

Con el Principio de Máxima Entropía se asume el punto de vista bayesiano, pues es un método que permite inferir probabilidades desconocidas a partir de la información disponible. En ocasiones nos referiremos al Principio de Máxima Entropía como “PME” o como “MaxEnt” a modo de abreviación.

2.4.1 Resolución del juego de reparto con el Principio de Máxima Entropía

El objetivo es resolver el problema del juego de reparto anterior utilizando el Principio de Máxima Entropía. En el problema se tienen cuatro posibles sucesos correspondientes a los cuatro colores. Cada color está caracterizado por su coste d_i y su probabilidad p_i . Sea p_1 la probabilidad de poner una ficha en el color rojo, p_2 en el negro, p_3 en el amarillo y p_4 en el verde, la entropía estadística asociada a esta distribución de probabilidades será:

$$S = -k \sum_{i=1}^4 p_i \ln p_i \quad (2.2)$$

La entropía de la distribución está sujeta a las siguientes ligaduras:

$$\sum_{i=1}^4 p_i = 1 \quad (2.3)$$

$$\sum_{i=1}^4 p_i d_i = \langle d \rangle \quad (2.4)$$

La ligadura (2.3) se corresponde con la condición de que la suma de las probabilidades de todos los sucesos elementales del experimento ha de ser la unidad. La ligadura (2.4) introduce la información sobre el coste medio por ficha y partida, denotado como $\langle d \rangle$. En el juego de reparto, la información correspondiente al coste medio por ficha se obtiene dividiendo el dinero total entre el número de fichas:

$$\langle d \rangle = \frac{10 \text{ € por partida}}{10 \text{ fichas por partida}} = 1 \text{ €}$$

Una vez conocidas las ligaduras del problema, se puede aplicar el Principio de Máxima Entropía. La distribución de probabilidad más adecuada es aquella que maximiza la entropía estadística sujeta a la información conocida. Para maximizar la entropía estadística sujeta a las ligaduras, se emplea el método de los multiplicadores de Lagrange (O'Connell y Haile, 2005):

Multiplicadores de Lagrange. El método de los multiplicadores de Lagrange es un procedimiento para encontrar los puntos críticos de funciones de varias variables sujetas a unas ciertas ligaduras o restricciones en su dominio. Sea una función de varias variables $f(x_1, x_2, \dots, x_n)$, y sean s ligaduras $g_k(x_1, x_2, \dots, x_n) = 0$, $k = 1, 2, \dots, s$. El problema de encontrar los puntos críticos de la función $f(x_1, x_2, \dots, x_n)$ sujeta a las ligaduras $g_k(x_1, x_2, \dots, x_n) = 0$ es equivalente a encontrar los puntos críticos de otra función, llamada función de Lagrange, que no está sujeta a ninguna ligadura y se construye de la siguiente forma:

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_s) = f(x_1, \dots, x_n) + \sum_{k=1}^s \lambda_k g_k(x_1, \dots, x_n) \quad (2.5)$$

donde los λ_k son unas nuevas variables escalares denominadas multiplicadores de Lagrange. Este método reduce el problema inicial de n variables con s ligaduras, a uno de $n + s$ variables sin ligaduras.

En el juego de reparto, la función de Lagrange con los multiplicadores de Lagrange λ_1 y λ_2 se construye de la siguiente manera:

$$L(p_1, \dots, p_4, \lambda_1, \lambda_2) = -k \sum_{i=1}^4 p_i \ln p_i + \lambda_1 \left(\sum_{i=1}^4 p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^4 p_i d_i - \langle d \rangle \right) \quad (2.6)$$

Las condiciones de extremo condicionado para $S(p_1, \dots, p_4)$ con las ligaduras dadas se obtienen calculando las derivadas parciales de $L(p_1, \dots, p_4, \lambda_1, \lambda_2)$ respecto a sus parámetros, y después igualándolas a cero:

$$\left(\frac{\partial L}{\partial p_i}\right)_{\lambda_1, \lambda_2, p_j \neq p_i} = 0 \quad \left(\frac{\partial L}{\partial \lambda_1}\right)_{\lambda_2, p_i} = 0 \quad \left(\frac{\partial L}{\partial \lambda_2}\right)_{\lambda_1, p_i} = 0 \quad (2.7)$$

Las derivadas con respecto a los multiplicadores de Lagrange λ_1 y λ_2 no aportan información adicional, pues se obtienen de nuevo las ecuaciones de las ligaduras impuestas a la entropía:

$$\frac{\partial L}{\partial \lambda_1} = \sum_{i=1}^4 p_i - 1 = 0 \quad \rightarrow \quad \sum_{i=1}^4 p_i = 1 \quad \frac{\partial L}{\partial \lambda_2} = \sum_{i=1}^4 p_i d_i - \langle d \rangle = 0 \quad \rightarrow \quad \sum_{i=1}^4 p_i d_i = \langle d \rangle$$

De las derivadas parciales con respecto a las p_i se tiene lo siguiente:

$$\frac{\partial L}{\partial p_i} = -k(\ln p_i + 1) + \lambda_1 + \lambda_2 d_i = 0 \quad (2.8)$$

Despejando de la ecuación (2.8) se obtiene la siguiente expresión para las probabilidades p_i :

$$p_i = \exp\left(\frac{\lambda_1}{k} + \frac{\lambda_2 d_i}{k} - 1\right) \quad (2.9)$$

Las constantes a determinar en el problema son λ_1 y λ_2 . Se pueden definir dos nuevas variables Z y β de la siguiente forma:

$$Z^{-1} = \exp\left(\frac{\lambda_1}{k} - 1\right) \quad \beta = -\frac{\lambda_2}{k} \quad (2.10)$$

Sustituyendo las nuevas variables (2.10) en la expresión para las probabilidades (2.9):

$$\boxed{p_i = \frac{e^{-\beta d_i}}{Z}} \quad (2.11)$$

Una vez conocida la expresión para las probabilidades, llevamos la expresión (2.11) a las ligaduras (2.3) y (2.4):

$$\sum_{i=1}^4 p_i = 1 = \sum_{i=1}^4 \frac{e^{-\beta d_i}}{Z} = \frac{1}{Z} \sum_{i=1}^4 e^{-\beta d_i} \quad \rightarrow \quad \boxed{Z = \sum_i e^{-\beta d_i}} \quad (2.12)$$

$$\sum_{i=1}^4 p_i d_i = \langle d \rangle = \sum_{i=1}^4 \frac{d_i e^{-\beta d_i}}{Z} = \frac{1}{Z} \sum_{i=1}^4 d_i e^{-\beta d_i} \quad \rightarrow \quad \boxed{Z \langle d \rangle = \sum_i d_i e^{-\beta d_i}} \quad (2.13)$$

La resolución del sistema de dos ecuaciones formado por las expresiones (2.12) y (2.13) nos proporciona los valores de Z y β .

En el caso concreto del juego de reparto planteado, los costes son:

$$\langle d \rangle = 1 \text{ €} \quad d_1 = 4 \text{ €} \quad d_2 = 2 \text{ €} \quad d_3 = 1 \text{ €} \quad d_4 = 0 \text{ €}$$

y las ecuaciones a resolver son las siguientes:

$$\left. \begin{array}{l} Z = 1 + e^{-\beta} + e^{-2\beta} + e^{-4\beta} \\ Z = e^{-\beta} + 2e^{-2\beta} + 4e^{-4\beta} \end{array} \right\} \rightarrow x = e^{-\beta} \rightarrow \left\{ \begin{array}{l} Z = 1 + x + x^2 + x^4 \\ Z = x + 2x^2 + 4x^4 \end{array} \right. \quad (2.14)$$

En ellas, se ha introducido el cambio de variable $x = e^{-\beta}$. Sustituyendo una en otra:

$$3x^4 + x^2 - 1 = 0 \quad \rightarrow \quad x = \sqrt{\frac{-1 + \sqrt{13}}{6}} = 0.658983 \quad (2.15)$$

Se trata de una ecuación bicuadrada cuya única solución real positiva es $x = 0.658983$. De esta forma, se obtienen los siguientes valores para Z y β , con los que se pueden calcular las probabilidades de cada uno de los colores:

$$\boxed{\beta = -\ln x = 0.417058} \quad \boxed{Z = 1 + x + x^2 + x^4 = 2.281822} \quad p_i = \frac{e^{-\beta d_i}}{Z} = \frac{x^{d_i}}{Z} \quad (2.16)$$

Las probabilidades que se inferen al aplicar el Principio de Máxima Entropía son las siguientes:

$$\boxed{p_{1(rojo)} = 0.082645} \quad \boxed{p_{2(negro)} = 0.190312} \quad \boxed{p_{3(amarillo)} = 0.288797} \quad \boxed{p_{4(verde)} = 0.438246}$$

Si se comparan estos valores con los obtenidos mediante tabulación directa para $N = 10$ fichas y $D = 10 \text{ €}$, el acuerdo es considerable.

En la Tabla 2.2 se comparan las probabilidades obtenidas con el PME con las obtenidas por tabulación directa. Las seis primeras filas se corresponden con la resolución por tabulación directa para $\langle d \rangle = 1\text{€}$ / ficha según el número de fichas N (se han incluido otros casos más, con 5, 15, 20, 30 y 40 fichas, todos con el mismo valor medio del coste por ficha):

N / fichas	D / €	$p_1 = p(\text{rojo})$	$p_2 = p(\text{negro})$	$p_3 = p(\text{amarillo})$	$p_4 = p(\text{verde})$
5	5	0.056338	0.225352	0.323944	0.394366
10	10	0.071918	0.204381	0.303565	0.420136
15	15	0.075680	0.199331	0.298617	0.426372
20	20	0.077491	0.196954	0.296129	0.429426
30	30	0.079254	0.194661	0.293662	0.432423
40	40	0.080118	0.193545	0.292436	0.433900
MaxEnt con $\langle d \rangle = 1\text{€}$ / ficha		<u>0.082645</u>	<u>0.190312</u>	<u>0.288797</u>	<u>0.438246</u>

Tabla 2.2. El juego de reparto: resultados exactos por tabulación directa, y resultados con MaxEnt.

La tendencia de las probabilidades se muestra en la Figura 2.1. Observamos que, al aumentar el número de fichas, los resultados exactos obtenidos por tabulación directa se aproximan cada vez más a los resultados que predice el PME. Este análisis sugiere la siguiente conclusión:

Las probabilidades calculadas por tabulación directa tienden a las probabilidades inferidas con el Principio de Máxima Entropía al aumentar el número de fichas apostadas.

Las razones de esta tendencia se explicarán en la sección 5.4 a partir de lo expuesto en el capítulo 3 y en las secciones 5.1–5.3.

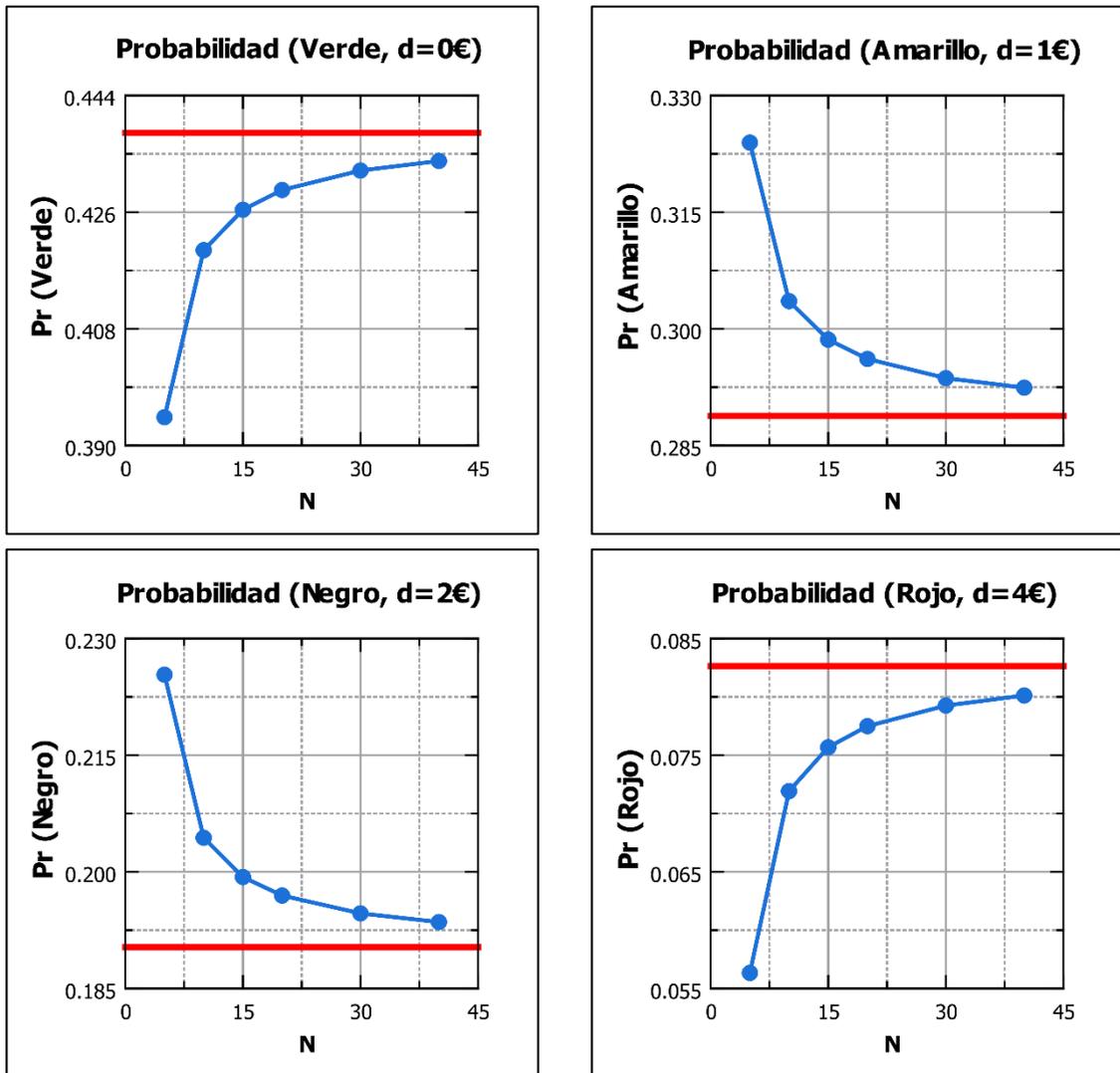


Figura 2.1. Tendencia de las probabilidades exactas a las obtenidas por MaxEnt

En las gráficas se muestra la probabilidad calculada de que los 4 colores reciban una ficha para $\langle d \rangle = 1\text{€}/\text{ficha}$. Los puntos representan los resultados del método de tabulación directa para distintos valores del número de fichas N . Las líneas horizontales continuas son los resultados obtenidos a partir del Principio de Máxima Entropía.

Capítulo 3: Justificación y significado del Principio de Máxima Entropía

Desde el punto de vista frecuentista, una distribución de probabilidades necesita ser verificada experimentalmente en términos de frecuencias relativas. La frecuencia relativa de un suceso es el cociente entre el número de resultados favorables a dicho suceso y el número total de resultados. Nos proponemos demostrar que las probabilidades obtenidas del Principio de Máxima Entropía tienen una conexión directa con la obtención de dichas probabilidades como frecuencias relativas en la repetición de un experimento en condiciones idénticas. Veremos que dicha conexión es una consecuencia matemática directa de las propiedades de las permutaciones cuando el número de elementos a permutar es muy grande.

3.1 Comportamiento asintótico de las permutaciones

3.1.1 Permutaciones sin repetición: aproximación de Stirling

El factorial de un número N admite un desarrollo denominado serie de Stirling:

$$N! = \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \left[1 + \frac{1}{12N} + O\left(\frac{1}{N}\right)\right] \quad (3.1)$$

donde la notación $O(f(N))$ simboliza términos que tienden a 0 más rápido que $f(N)$ cuando $N \rightarrow \infty$. El correspondiente desarrollo del logaritmo natural del factorial se obtiene tomando el logaritmo natural a ambos lados de la expresión (3.1):

$$\ln(N!) \approx N \ln N - N + \frac{1}{2} \ln(2\pi N) + \frac{1}{12N} + O\left(\frac{1}{N^2}\right) \quad (3.2)$$

Cuando $N \gg \ln N$, solo los dos primeros términos en (3.2) son significativos, de forma que la aproximación de Stirling para el logaritmo de un factorial es

$$\ln(N!) \approx N \ln N - N \quad (3.3)$$

En la Tabla 3.1 se recogen los valores exactos del logaritmo natural de un factorial, y el valor que proporciona la aproximación de Stirling.

N	$\ln(N!)$	$N \ln N - N$	Error relativo
10	15.104413	13.025851	13.761%
100	363.739376	360.517019	0.886%
1000	5912.128178	5907.755279	0.074%
10^{23}	$5.1959457 \cdot 10^{24}$	$5.1959457 \cdot 10^{24}$	$\sim 0\%$

Tabla 3.1. Error cometido con la aproximación de Stirling. Cuanto mayor es N , menor es el error cometido.

3.1.2 Permutaciones con repetición

Utilizando la serie de Stirling para desarrollar las permutaciones con repetición (definidas en la ecuación (2.1)), unas cuentas directas, aunque algo laboriosas, llevan a los siguientes desarrollos:

$$PR_N^{n_1, \dots, n_s} = \exp\{N\sigma(g_1, \dots, g_s)\} (\sqrt{2\pi N})^{s-1} \left(\sqrt{\prod_{i=1, g_i \neq 0}^s \frac{1}{g_i}} \right) \cdot \exp\left\{ \frac{1}{12N} \left(1 - \sum_{i=1, g_i \neq 0}^s \frac{1}{g_i} \right) + O\left(\frac{1}{N^2}\right) \right\} \quad (3.4)$$

$$\frac{1}{N} \ln(PR_N^{n_1, \dots, n_s}) = \sigma(g_1, \dots, g_s) + \frac{1}{2N} \left[(s-1) \ln(2\pi N) - \sum_{i=1, g_i \neq 0}^s \ln g_i \right] + O\left(\frac{1}{N}\right) \quad (3.5)$$

donde σ es la función entropía adimensional definida por la ecuación (1.10) y $g_i = n_i / N$ es la “frecuencia relativa” con la que se repite cada elemento.

3.2 Justificación del Principio de Máxima Entropía

La justificación del Principio de Máxima Entropía se basa en el siguiente teorema (Jaynes, 1978; Jaynes, 2003; Hevia, 2023):

Teorema. Sea un experimento aleatorio, cuyos sucesos elementales son $\{A_1, \dots, A_\Omega\}$. Supongamos que disponemos de una cierta **información** (conjunto de datos objetivos y, por tanto, independientes del observador, acerca de los sucesos, al que corresponden ciertas ligaduras matemáticas entre las frecuencias relativas de dichos sucesos). Con el objeto de obtener experimentalmente las frecuencias relativas de los sucesos, f_i , se repite dicho experimento N veces en condiciones idénticas. Sea $W(f_1, \dots, f_\Omega)$ el número de formas diferentes en el que puede ocurrir una cierta distribución de frecuencias relativas $\{f_1, \dots, f_\Omega\}$.

Suponemos, además, que la información cumple con las tres hipótesis siguientes:

- i. Es **correcta** (esto es, las frecuencias relativas obtenidas se adecúan a la información inicial, pues de lo contrario dicha información sería incorrecta o el experimento se habría realizado de forma incompatible con ella)
- ii. Es **suficiente** (es decir, incluye todas las ligaduras físicas relevantes del experimento).
- iii. Cuando se utiliza dicha información para hallar el máximo condicionado de $W(f_1, \dots, f_\Omega)$ y de $\sigma(f_1, \dots, f_\Omega)$, donde σ es la entropía estadística adimensional, **dicho máximo condicionado es único**.

Entonces, cuando $N \rightarrow \infty$, “la distribución experimental de frecuencias relativas se aproxima muy rápidamente” a la distribución de frecuencias relativas que maximiza $\sigma(f_1, \dots, f_\Omega)$ de acuerdo con la información, en el sentido siguiente:

- a) Para N suficientemente grande, maximizar $W(f_1, \dots, f_\Omega)$ de acuerdo con la información disponible equivale a maximizar la función $\sigma(f_1, \dots, f_\Omega)$ de acuerdo con la misma información.
- b) El número de formas en las que se puede obtener una distribución de frecuencias relativas *diferente* de la del máximo de $W(f_1, \dots, f_\Omega)$ condicionado a la información disminuye exponencialmente en comparación con dicho máximo al aumentar N .

Por tanto, para N suficientemente grande, es de esperar que ocurra la distribución de frecuencias relativas que maximiza $\sigma(f_1, \dots, f_\Omega)$, simplemente porque puede ocurrir de un número enormemente mayor de formas con relación a cualquier otra distribución. La distribución de frecuencias relativas obtenida de maximizar $\sigma(f_1, \dots, f_\Omega)$ es matemáticamente idéntica a la distribución de probabilidades obtenida de maximizar $\sigma(p_1, \dots, p_\Omega)$, **lo que justifica, desde el punto de vista frecuentista, la elección de la distribución de probabilidades obtenidas del Principio de Máxima Entropía.**

3.2.1 Demostración del teorema de Jaynes

Como resultado del experimento, se obtiene n_i veces el suceso elemental A_i , siendo $n_1 + \dots + n_\Omega = N$. Las frecuencias relativas serán, por tanto, $f_i = n_i/N$.

¿De cuántas formas puede ocurrir una distribución de frecuencias relativas? Para contabilizarlas, debemos tener en cuenta que, aunque hay $N!$ posibles ordenaciones de los resultados del experimento, son iguales todas las ordenaciones en las que se permutan entre sí solamente los n_i resultados favorables a cada suceso A_i . En otras palabras, el número de formas, $W(f_1, \dots, f_\Omega)$, en que puede ocurrir dicha distribución de frecuencias es igual a las permutaciones con repetición de N elementos, en donde hay n_1 elementos iguales, ..., n_Ω elementos iguales:

$$W(f_1, \dots, f_\Omega) = \frac{N!}{n_1! \dots n_\Omega!} = \frac{N!}{(Nf_1)! \dots (Nf_\Omega)!} \quad (3.6)$$

a) Para demostrar la primera parte del teorema, estudiaremos la función $[\ln W(f_1, \dots, f_\Omega)] / N$, que es una función monótona creciente de $W(f_1, \dots, f_\Omega)$ (y, por tanto, su máximo condicionado debe ocurrir para los mismos valores de las frecuencias relativas) y veremos que, para N arbitrariamente grande, el valor de esta función es igual a $\sigma(f_1, \dots, f_\Omega)$, donde σ es la entropía estadística adimensional.

De acuerdo con la fórmula asintótica (3.5):

$$\frac{\ln W(f_1, \dots, f_\Omega)}{N} = \sigma(f_1, \dots, f_\Omega) + \frac{1}{2N} \left[(\Omega - 1) \ln(2\pi N) + \sum_{\substack{i=1 \\ f_i \neq 0}}^{\Omega} \ln f_i \right] + O\left(\frac{1}{N}\right). \quad (3.7)$$

Pero, de acuerdo con (3.7):

$$\lim_{N \rightarrow \infty} \left[\frac{\ln W(f_1, \dots, f_\Omega)}{N} \right] = \sigma(f_1, \dots, f_\Omega). \quad (3.8)$$

Por ello, **para un número N suficientemente grande de experimentos realizados, maximizar $W(f_1, \dots, f_\Omega)$ de forma condicionada a la información disponible es equivalente a maximizar la $\sigma(f_1, \dots, f_\Omega)$ de acuerdo con esas mismas ligaduras.**

b) Para demostrar la segunda parte del teorema, compararemos, para N grande, el valor del cociente entre el número de formas en que se puede realizar la distribución de frecuencias relativas del máximo condicionado y el número de formas en que se puede realizar cualquier otra distribución de frecuencias relativas, y veremos que este cociente tiende exponencialmente a infinito con N tendiendo a infinito.

Supongamos que hemos encontrado la distribución de frecuencias relativas $\{f_{1m}, \dots, f_{\Omega m}\}$ que da lugar al único máximo condicionado de $W(f_1, \dots, f_\Omega)$, al que denotaremos por W_m . Sea $\{f'_1, \dots, f'_\Omega\}$ otra distribución **arbitraria** de probabilidad, con la única condición de que sea compatible con la información disponible; denotemos $W' = W(f'_1, \dots, f'_\Omega)$ el número de formas en que puede ocurrir esta distribución de frecuencias relativas. De acuerdo con la fórmula asintótica (3.4), el cociente entre W' / W_m se puede escribir:

$$\frac{W'}{W_m} = \exp\{-N \cdot [\sigma(f_{1m}, \dots, f_{sm}) - \sigma(f'_1, \dots, f'_s)]\} \cdot A \cdot \exp\left\{\frac{B}{N} + O\left(\frac{1}{N^2}\right)\right\} \quad (3.9)$$

$$\text{donde } A = \sqrt{\prod_{\substack{i=1 \\ f_{im}, f'_i \neq 0}}^{\Omega} \frac{f_{im}}{f'_i}} \quad \text{y} \quad B = \frac{1}{12} \left(\sum_{\substack{i=1 \\ f_{im} \neq 0}}^{\Omega} \frac{1}{f_{im}} - \sum_{\substack{i=1 \\ f'_i \neq 0}}^{\Omega} \frac{1}{f'_i} \right).$$

Para N arbitrariamente grande, el último factor exponencial es prácticamente igual a la unidad, con lo que la tendencia dominante es el primer factor exponencial decreciente (ya que, por primera parte del teorema y la hipótesis de máximo único, $\sigma(f_{1m}, \dots, f_{sm}) - \sigma(f'_1, \dots, f'_s) > 0$), que disminuye enormemente rápido al aumentar N . Así pues, **el número de formas en las que se puede obtener una distribución de frecuencias relativas diferente de la del máximo de $W(f_1, \dots, f_\Omega)$ disminuye exponencialmente en comparación con dicho máximo a medida que el número de experimentos realizados N tiende a infinito.**

3.3 Remarcas importantes sobre el Principio de Máxima Entropía

A continuación, reproducimos algunas remarcas hechas por el propio Jaynes en varios de sus numerosos escritos sobre el PME (Jaynes, 1978; Jaynes, 2003; Hevia, 2023).

Se debe tener en mente que **el PME no es una teoría física** que permita calcular inequívocamente las probabilidades de los sucesos de un experimento. **El PME es un método de razonamiento** que permite obtener las “mejores” predicciones que pueden hacerse con la información de la que se parte. Si cambia la información de partida, las predicciones cambian.

Debemos insistir en que las conexiones entre las frecuencias relativas y el PME no afirman que las frecuencias obtenidas como probabilidades a partir del PME se observarán con certeza en un experimento; de hecho, ningún principio de la Teoría de la Probabilidad puede predecir con certeza qué ocurrirá en un experimento real. Lo correcto sería decir que **la distribución de frecuencias relativas correspondiente con las probabilidades del PME es, con diferencia abismal, la que puede ocurrir de mayor número de maneras en un experimento real, siempre y cuando las ligaduras físicas que operen en el experimento sean las mismas que las que se supongan en la aplicación del PME.**

La mayor parte de las distribuciones de frecuencia producidas en experimentos reales son de máxima entropía, simplemente porque se pueden realizar de muchas más maneras. Producir una distribución de frecuencias que se aleje apreciablemente de la de máxima entropía requiere la introducción de ligaduras físicas muy efectivas en el experimento. **Cualquier desviación estadísticamente significativa**, aunque no sea muy grande, **de las predicciones del PME constituye una evidencia importante** (y, si persiste, concluyente) **de que hay información que no se ha tenido en cuenta adecuadamente en el cálculo** (es incompleta o errónea).

En la línea de esta última posibilidad, Jaynes (1957) pone **el descubrimiento de la Mecánica Cuántica** como ejemplo de la potencia del método de razonamiento. La inferencia de las probabilidades de los posibles estados mecánicos de un sistema puede hacerse sobre la base de que obedece a la Mecánica Clásica o a la Mecánica Cuántica. En ambos casos, se puede usar el PME como método de razonamiento para obtener las probabilidades de dichos estados, pero sin embargo los resultados obtenidos a partir de la Mecánica Clásica no se corresponden con los experimentos. El fracaso de las predicciones probabilísticas no proviene, por tanto, del PME, sino de haber partido de información errónea (el modelo cuántico, y no el clásico, es el que proporciona correctamente la información sobre los posibles estados del sistema).

Recordemos que la **información** disponible es un conjunto de **datos objetivos** y, por tanto, independientes del observador, acerca de los sucesos, al que corresponden ciertas **ligaduras entre las probabilidades o entre las frecuencias relativas**. Hay que tener en cuenta que cuando las ligaduras no son lineales en las probabilidades, el PME en ocasiones no puede hacer predicciones suficientemente satisfactorias, en el sentido de que puede haber diferentes máximos locales. **Si las ligaduras son lineales en las probabilidades, el máximo es único y el PME hace predicciones satisfactorias en ese sentido.**

Capítulo 4: Formalismo termodinámico

En este capítulo se resumen, de forma sintética, ciertos contenidos de Termodinámica que serán necesarios para comprender la aplicación del Principio de Máxima Entropía a la Física Estadística.

4.1 El equilibrio termodinámico

Nosotros denominaremos en lo que sigue **sistema termodinámico** a todo sistema macroscópico que cumpla las siguientes condiciones (Hevia, 2023): ausencia de fuerzas a larga distancia entre distintas partes del cuerpo, ausencia de inhomogeneidades debidas a campos externos y ausencia de histéresis. Para cualquier caso en el que una de las tres hipótesis anteriores sea violada, las conclusiones que presentaremos serán parcialmente aplicables o no aplicables en absoluto. Hay que mencionar que no existe (por el momento) una teoría termodinámica general ampliamente aceptada para cuando existen fuerzas a larga distancia o histéresis, por lo que el análisis de estos sistemas debe hacerse con cuidado en cada caso particular.

En general, los sistemas termodinámicos se presentan en estados difíciles de describir, salvo en unos casos particulares denominados **estados de equilibrio termodinámico** (o, simplemente, **estados de equilibrio**). Éstos tienen las siguientes características (Tejerina, 1976; Hevia, 2023):

- Son **estados estacionarios**: las variables de estado permanecen constantes en el tiempo con el suficiente grado de aproximación, es decir, con **fluctuaciones suficientemente pequeñas como para poder ser despreciadas**.
- No hay **intercambios netos de materia ni energía** entre sus diferentes partes ni con el entorno.

La condición de presentar fluctuaciones despreciables será retomada más adelante en relación con el límite termodinámico (sección 5.3).

4.2 Variables termodinámicas y trabajo de configuración

Las variables macroscópicas empleadas para la descripción de los estados de equilibrio de un sistema termodinámico reciben el nombre de **variables termodinámicas**. Ahora bien, no todas son independientes. Las variables termodinámicas que describen de forma completa y unívoca el estado de un sistema reciben el nombre de **variables de estado**. El resto de las variables, no empleadas para describir el estado, se denominan **funciones de estado**, pues pueden expresarse en función de las variables de estado (Tejerina, 1976).

Dos clases importantes de variables termodinámicas son las intensivas y las extensivas (Tejerina, 1976; Hevia, 2023):

- Se dice que una variable es **intensiva** si toma el mismo valor en cualquier parte de un sistema homogéneo, independientemente del tamaño o forma de la parte seleccionada. Ejemplos: la densidad, la masa molar, la presión, etc.
- Se dice que una variable es **extensiva (o aditiva)** en un sistema cuando su valor en el sistema es igual a la suma de los valores que toma en cada una de sus partes. Ejemplos: el volumen, la masa, la carga, etc. Si se cumplen las hipótesis establecidas en

la sección anterior para un sistema termodinámico, la energía interna de un sistema es extensiva.

El **trabajo de configuración** elemental (Tejerina, 1976; Hevia, 2023) sobre un sistema termodinámico se puede escribir en la forma:

$$\delta W = \sum_{i=1}^k A_i da_i \quad (4.1)$$

- Las variables A_i representan las **fuerzas generalizadas (externas)** que el entorno ejerce sobre el sistema (presión, tensión superficial, tensión de un hilo...). Son variables **intensivas**.
- Las variables a_i se llaman **desplazamientos generalizados** y son variables del propio sistema distintas de su energía interna, y que además son **extensivas** (volumen, área, longitud de un hilo...).

Por trabajar con unas variables concretas, supondremos en las siguientes secciones que el trabajo de configuración del sistema es hidrostático:

$$\delta W = -PdV \quad (4.2)$$

donde P es la presión externa y V es el volumen del sistema.

4.3 Postulados de Callen. Crítica y limitaciones

Los postulados de Callen (1985) para una formulación deductiva de la Termodinámica son los siguientes:

■ **Postulado I.** “*En los sistemas simples existen estados particulares (denominados estados de equilibrio) que, desde el punto de vista macroscópico, están caracterizados completamente por la energía interna U , el volumen V , y el número de partículas N_1, N_2, \dots, N_r de los componentes químicos del sistema*”.

Las variables mencionadas en el *Postulado I* reciben el nombre de **variables naturales de la entropía**. Notemos que son todas extensivas, y por ello muchas veces se hace referencia a ellas como “las variables extensivas” o “los parámetros extensivos”.

■ **Postulado II.** “*Existe una función de estado (llamada entropía, S) de los parámetros extensivos de cualquier sistema compuesto. La entropía $S = S(U, V, N_1, \dots, N_r)$ está definida para todos los estados de equilibrio del sistema, y tiene la siguiente propiedad: los valores que toman los parámetros extensivos, en ausencia de ligaduras internas, son aquellos que maximizan la entropía respecto de los estados de equilibrio*”.

El *Postulado II* está inspirado en la Física Estadística (en las propiedades de la entropía estadística de un sistema aislado, que coincide con la entropía de Boltzmann-Planck, como se verá en el capítulo 6) y no puede obtenerse mediante la formulación inductiva de máquinas térmicas. En la formulación deductiva, es este postulado el que contiene la información de la direccionalidad de los procesos, pues permite hallar el estado final de equilibrio al que llegará un sistema termodinámico aislado al suprimir una de sus ligaduras internas (lo que Callen llama el “problema básico de la Termodinámica”).

■ **Postulado III.** “La entropía de un sistema compuesto es aditiva respecto a sus subsistemas constituyentes. La entropía es una función continua, diferenciable y monótona creciente de la energía”.

Del *Postulado III* se derivan inmediatamente varias consecuencias matemáticas. La propiedad de aditividad (extensividad) establece que la entropía S del sistema compuesto es la suma de las entropías $S^{(\alpha)}$ de los subsistemas constituyentes, siendo $S^{(\alpha)}$ una función únicamente de los parámetros extensivos del subsistema:

$$S = \sum_{\alpha} S^{(\alpha)} \quad S^{(\alpha)} = S^{(\alpha)}(U^{(\alpha)}, V^{(\alpha)}, N_1^{(\alpha)}, \dots, N_r^{(\alpha)})$$

La aditividad de la entropía aplicada a subsistemas espacialmente separados requiere de la siguiente propiedad: la entropía de un sistema simple es una función homogénea de primer orden de los parámetros extensivos,

$$S(\lambda U, \lambda V, \lambda N_1, \dots, \lambda N_r) = \lambda \cdot S(U, V, N_1, \dots, N_r)$$

La propiedad de ser monótona creciente de la energía implica que: $\left(\frac{\partial S}{\partial U}\right)_{V, N_i} > 0$

A medida que la teoría se desarrolle en secciones posteriores, veremos que la recíproca de esta derivada parcial se toma como definición de la temperatura. Así, Callen postula que la temperatura no es negativa. **Ciertos sistemas con energía acotada superiormente pueden tener temperaturas negativas, por lo que la propiedad de monotonía creciente de la entropía debe eliminarse del postulado** (Hevia, 2014; Hevia, 2023). Dichos sistemas con temperaturas negativas suelen llamarse sistemas “anormales” (Biel, 1998).

■ **Postulado IV.** “La entropía de cualquier sistema se anula en el estado para el que

$$\left(\frac{\partial U}{\partial S}\right)_{V, N_i} = 0”.$$

El *Postulado IV* indica que la temperatura cero implica entropía nula. Este postulado es una extensión, debida a Planck, del Postulado de Nernst en el marco del Tercer Principio de la Termodinámica. Se sabe que no es de validez general (Biel, 1998), pero es compatible con la Mecánica Cuántica si el estado cuántico fundamental del sistema es no degenerado.

Al eliminar en el *Postulado III* la exigencia de que la entropía sea monótona creciente de la energía, no es posible invertir de forma general la interdependencia de S y U , y la forma correcta de enunciar el *Postulado IV* es: “La entropía de cualquier sistema se anula en el estado para el que $(\partial S/\partial U)_{V, N_i}$ diverge”.

4.4 La representación entrópica

La formulación más general de la Termodinámica y de la Física Estadística utiliza la **representación entrópica**, o formulación entrópica de la Termodinámica, en la que se utiliza la entropía y sus transformadas de Legendre como potenciales termodinámicos para estudiar los sistemas (Baixauli y Cobos, 1996; Balian, 2017; Hevia, 2014; Hevia, 2023).

Cada potencial termodinámico está caracterizado por tres ecuaciones (Gibbs, Euler, y Gibbs-Duhem). A partir de ahora trataremos con **sistemas monofásicos, monocomponentes y que sólo pueden intercambiar trabajo hidrostático asociado al cambio de volumen.**

■ **Ecuación de Gibbs.** Expresa la diferencial exacta del potencial termodinámico. Los potenciales termodinámicos son funciones de estado, lo que implica que admiten una diferencial total o exacta. La entropía es una función de estado $S(U, V, N)$ cuyas variables naturales (que son extensivas) son la energía interna U , el volumen V y el número de partículas N . Su diferencial exacta es:

$$dS = \left(\frac{\partial S}{\partial U} \right)_{V,N} dU + \left(\frac{\partial S}{\partial V} \right)_{U,N} dV + \left(\frac{\partial S}{\partial N} \right)_{U,V} dN \quad (4.3)$$

Se definen la **temperatura termodinámica** T del sistema y el **potencial químico** μ de su único componente por las siguientes relaciones:

$$\left(\frac{\partial S}{\partial U} \right)_{V,N} = \frac{1}{T} \quad \left(\frac{\partial S}{\partial N} \right)_{U,V} = -\frac{\mu}{T} \quad (4.4)$$

Y la **presión** P (fuerza generalizada) puede calcularse como:

$$\left(\frac{\partial S}{\partial V} \right)_{U,N} = \frac{P}{T} \quad (4.5)$$

De esta forma, la ecuación de Gibbs en la representación entrópica es la siguiente:

$$\boxed{dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN} \quad (4.6)$$

■ **Ecuación de Euler.** Es consecuencia de la homogeneidad de grado 1 de la entropía respecto de las variables extensivas (Tejerina, 1976; Callen, 1985; Biel, 1998). Para el caso de la entropía:

$$\boxed{S = \frac{U}{T} + \frac{PV}{T} - \frac{\mu N}{T}} \quad (4.7)$$

■ **Ecuación de Gibbs-Duhem.** Es la relación de ligadura entre las variables intensivas de la representación. Se obtiene de combinar la ecuación de Gibbs con la ecuación de Euler.

$$\boxed{Ud\left(\frac{1}{T}\right) + Vd\left(\frac{P}{T}\right) - Nd\left(\frac{\mu}{T}\right) = 0} \quad (4.8)$$

Se comprueba que **la ecuación de Gibbs-Duhem es independiente del potencial entrópico utilizado: es única para toda la representación entrópica. Indica que, en el equilibrio, las variables intensivas no son independientes.** Las variables naturales de la entropía, (U, V, N) , son variables extensivas, y las variables a las que acompañan en las ecuaciones anteriores, $(1/T, P/T, -\mu/T)$, son las variables intensivas conjugadas de las extensivas. La importancia que tiene esta distinción en el formalismo entrópico es que las variables intensivas solo están bien definidas a nivel macroscópico. Las variables extensivas, sin embargo, tienen sentido a nivel microscópico y macroscópico. Esto tendrá una importancia fundamental a la hora de describir los distintos microestados, o configuraciones microscópicas, de un sistema en Física Estadística.

4.5 Transformadas de Legendre de la entropía

La transformada de Legendre (Casanova, 2019) es una operación matemática que transforma una cierta función $\varphi(x_1, x_2, \dots, x_n)$ en otra, $\Phi(y_1, x_2, \dots, x_n)$, de forma que ambas contienen la misma información, pero la dependencia respecto a la variable x_1 ha sido sustituida por otra: y_1 . La transformada de Legendre $\Phi(y_1, x_2, \dots, x_n)$ de $\varphi(x_1, x_2, \dots, x_n)$ respecto de x_1 es:

$$\Phi(y_1, x_2, \dots, x_n) = \varphi(x_1, x_2, \dots, x_n) - x_1 \left(\frac{\partial \varphi}{\partial x_1} \right)_{x_2, \dots} \quad \text{con} \quad y_1 = \left(\frac{\partial \varphi}{\partial x_1} \right)_{x_2, \dots, x_n} \quad (4.9)$$

De forma equivalente, se puede definir la transformada de Legendre respecto de varias variables:

$$\Phi(y_1, \dots, y_k, x_{k+1}, \dots, x_n) = \varphi(x_1, \dots, x_n) - \sum_{i=1}^k x_i \left(\frac{\partial \varphi}{\partial x_i} \right)_{x_j \neq x_i} = \varphi(x_1, \dots, x_n) - \sum_{i=1}^k x_i y_i$$

En Termodinámica, las transformadas de Legendre sustituyen una variable extensiva por su canónica conjugada intensiva. Esta operación nos va a permitir deducir nuevos potenciales termodinámicos a partir de los que ya conocemos. El objetivo es obtener las transformadas de Legendre de la entropía $S(U, V, N)$, así que partiremos de la entropía expresada en términos de sus variables naturales U , V y N .

4.5.1 La función de Massieu

La función de Massieu $\Psi = \Psi(1/T, V, N)$ se define como la transformada de Legendre de la entropía respecto de la energía interna:

$$\Psi = S - U \left(\frac{\partial S}{\partial U} \right)_{V, N} = S - \frac{U}{T} = \left(\frac{U}{T} + \frac{PV}{T} - \frac{\mu N}{T} \right) - \frac{U}{T} = \frac{PV}{T} - \frac{\mu N}{T} \quad (4.10)$$

La variable extensiva que se ha eliminado es la energía interna U , y se ha sustituido por su variable conjugada intensiva, que es $1/T$. La ecuaciones de Euler y de Gibbs para Ψ son:

$$\text{Euler: } \boxed{\Psi = \frac{PV}{T} - \frac{\mu N}{T}} \quad \text{Gibbs: } \boxed{d\Psi = -Ud\left(\frac{1}{T}\right) + \frac{P}{T}dV - \frac{\mu}{T}dN} \quad (4.11)$$

4.5.2 La función de Kramers-Landau (potencial gran canónico)

La función de Kramers-Landau $\eta = \eta(1/T, V, \mu/T)$, también conocida como el potencial gran canónico, se define como la transformada de Legendre de Ψ respecto del número de partículas:

$$\eta = \Psi - N \left(\frac{\partial \Psi}{\partial N} \right)_{1/T, V} = \Psi - N \left(-\frac{\mu}{T} \right) = \Psi + \frac{\mu N}{T} = \left(\frac{PV}{T} - \frac{\mu N}{T} \right) + \frac{\mu N}{T} = \frac{PV}{T} \quad (4.12)$$

Las ecuaciones de Euler y de Gibbs para la función de Kramers-Landau son:

$$\text{Euler: } \boxed{\eta = \frac{PV}{T}} \quad \text{Gibbs: } \boxed{d\eta = -Ud\left(\frac{1}{T}\right) + \frac{P}{T}dV + Nd\left(\frac{\mu}{T}\right)} \quad (4.13)$$

4.5.3 La función de Planck

La función de Planck $\phi = \phi(1/T, P/T, N)$ se define como la transformada de Legendre de la función de Massieu respecto del volumen:

$$\phi = \Psi - V \left(\frac{\partial \Psi}{\partial V} \right)_{1/T, N} = \Psi - V \left(\frac{P}{T} \right) = -\frac{\mu N}{T} \quad (4.14)$$

Sus ecuaciones de Euler y de Gibbs son:

$$\text{Euler: } \boxed{\phi = -\frac{\mu N}{T}} \quad \text{Gibbs: } \boxed{d\phi = -Ud\left(\frac{1}{T}\right) - Vd\left(\frac{P}{T}\right) - \frac{\mu}{T}dN} \quad (4.15)$$

Capítulo 5: Sistemas con muchas entidades elementales independientes: el límite termodinámico

La definición de sistema termodinámico con la que estamos trabajando, para la que la energía interna y la entropía son extensivas en estados de equilibrio, permite entender dichos sistemas en Física Estadística como conjuntos de **entidades elementales independientes**, entendidas éstas como los entes conceptuales más simples que, de acuerdo con un modelo de realidad, conservan su independencia estadística, la aditividad de la energía y de la entropía (Cobos, J.C., comunicación personal, 2023; Hevia, 2023). La naturaleza de estas entidades elementales independientes puede ser muy variada. En ocasiones, ni siquiera son entes materiales, como, por ejemplo, en el caso de los modos normales de vibración de una red cristalina en la aproximación armónica. Estos aspectos fueron analizados en profundidad por Cobos, J.C. (comunicación personal, 2023), y recogidos en la Tesis Doctoral de Mozo (2010, sección 5.2.).

Este tipo de sistemas, calificados de “**ideales**”, pues para ellos “el todo es la suma de las partes” (Anderson, 1972), muestran propiedades macroscópicas muy bien definidas cuando aumenta mucho el número de entidades elementales (lo que se denomina **límite termodinámico**), que son consecuencia de su independencia estadística. En particular, el valor de su entropía estadística no depende del conjunto utilizado para su descripción (capítulo 6).

Si el sistema no es “ideal” en el sentido mencionado, debido a que se viola alguna de las condiciones impuestas en el capítulo 4 a la definición de sistema termodinámico, estas propiedades, entre ellas la equivalencia de conjuntos, pueden no satisfacerse, lo que da lugar a una compleja y variada rama de la Física Estadística en la que, a día de hoy, se sigue investigando aún (Campa, Dauxois y Ruffo, 2009; Campa et al., 2018; Costeniuc, Ellis y Touchette, 2006; Hill, 2002; Latella et al., 2015; Loscar y Horowitz, 2018; Touchette, Ellis y Turkington, 2004; Tsallis, 2009).

En este capítulo tratamos de explicar, evitando en lo posible complicaciones matemáticas, en qué consisten estas deseables propiedades de los sistemas “ideales”.

5.1 Gas ideal en un recipiente aislado, parte I

Para ilustrar uno de los comportamientos típicos de un sistema “ideal” en el límite termodinámico, analizaremos un ejemplo propuesto por Baierlein (1978), convenientemente adaptado para ceñirnos a los propósitos de este capítulo.

Supongamos que tenemos N moléculas **independientes** (un típico gas ideal clásico) contenidas en una caja **aislada** (cuyas paredes impiden el intercambio de toda clase de materia y energía) de longitud L (Figura 5.1). ¿Cómo de probable es que todas las N moléculas se encuentren cerca de uno de los extremos de la caja? ¿O de que se encuentren sólo en una mitad? ¿O de que estén distribuidas uniformemente en todo el volumen? ¿Cómo puede fluctuar la distribución? Se

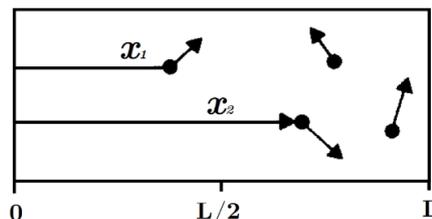


Figura 5.1. Un gas encerrado en una caja

puede hacer una buena estimación a las respuestas considerando la media muestral de la posición de las N moléculas en un instante de tiempo determinado:

$$x_{\text{media muestral}} = \frac{1}{N}(x_1 + x_2 + \dots + x_N)$$

La variable $x_{\text{media muestral}}$ es una forma de describir cómo están situadas las moléculas del gas. Si la mayoría de las moléculas están cerca de $x = 0$, el valor de $x_{\text{media muestral}}$ será pequeño y cercano a 0. De la misma forma, $x_{\text{media muestral}}$ será aproximadamente L si la mayoría de las partículas están cerca de $x = L$.

Ahora, consideramos el valor medio probabilístico de la media muestral de las posiciones de las moléculas, $\langle x_{\text{media muestral}} \rangle$. Puede entenderse en sentido frecuentista como la media que se obtiene observando el sistema muchísimas veces consecutivas. En cada instante, se toma una imagen de las partículas, se calcula su posición media $x_{\text{media muestral}}$, y finalmente se hace el valor medio de $x_{\text{media muestral}}$ en las muchas observaciones realizadas sobre el sistema. Debido a la independencia estadística de las moléculas, y en consecuencias de las distintas posiciones x_i de las distintas moléculas:

$$\langle x_{\text{media muestral}} \rangle = \frac{\langle x_1 + x_2 + \dots + x_N \rangle}{N} = \frac{\langle x_1 \rangle + \langle x_2 \rangle + \dots + \langle x_N \rangle}{N}$$

Para calcular $\langle x_{\text{media muestral}} \rangle$, vamos a razonar de la siguiente forma. Para la molécula 1, las posiciones opuestas centradas en $x_1 = L/2$ son igualmente probables a largo plazo. Entonces, tenemos que $\langle x_1 \rangle = L/2$, y esto mismo es cierto para cualquier otra molécula del gas. Entonces:

$$\langle x_{\text{media muestral}} \rangle = \frac{\langle x_1 \rangle + \langle x_2 \rangle + \dots + \langle x_N \rangle}{N} = \frac{L/2 + L/2 + \dots + L/2}{N} = \frac{L}{2}$$

que es el resultado que podríamos esperar.

Imaginemos ahora cómo es la gráfica de las probabilidades para un valor específico de $x_{\text{media muestral}}$. La curva debería ser suave, simétrica en torno a $L/2$ y con un máximo en $x_{\text{media muestral}} = L/2$. En la Figura 5.2 hemos representado la forma de dicha curva. Ahora, ¿cómo de ancha es la curva? ¿Es razonable que el valor de $x_{\text{media muestral}}$ fluctúe mucho alrededor de $\langle x_{\text{media muestral}} \rangle = L/2$?

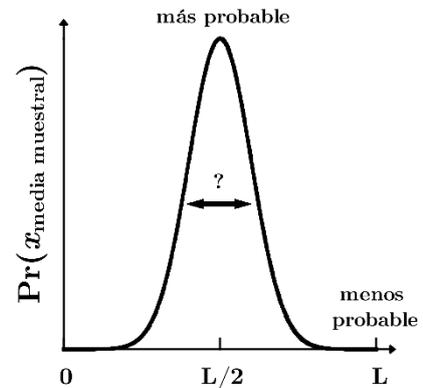


Figura 5.2. Forma de la distribución

Para tener una idea de la anchura de la distribución, podemos calcular su desviación típica, que cuantifica la desviación de un conjunto de datos respecto de su valor medio. La desviación típica, o desviación estándar, de una variable aleatoria X se define a partir de:

$$(\Delta^* X)^2 = \langle (X - \langle X \rangle)^2 \rangle \quad (5.1)$$

Vamos a estimar su valor $\Delta^* x_{\text{media muestral}}$ teniendo en cuenta algunos detalles:

$$(\Delta^* x_{\text{media muestral}})^2 = \left\langle \left[\frac{1}{N}(x_1 + \dots + x_N) - \frac{N}{N} \frac{L}{2} \right]^2 \right\rangle = \frac{1}{N^2} \left\langle \left[\left(x_1 - \frac{L}{2} \right) + \dots + \left(x_N - \frac{L}{2} \right) \right]^2 \right\rangle$$

El primer paso ha sido escribir $\langle x_{\text{media muestral}} \rangle = (N/N)(L/2)$, y repartir los $L/2$, uno para cada x_i . Después, tras aplicar la operación de elevar al cuadrado, aparecen productos que contienen dos moléculas diferentes, del tipo $\left\langle \left(x_1 - \frac{L}{2}\right) \left(x_2 - \frac{L}{2}\right) \right\rangle$, que promedian a cero debido a la independencia estadística de las moléculas.

$$\left\langle \left(x_1 - \frac{L}{2}\right) \left(x_2 - \frac{L}{2}\right) \right\rangle = \left\langle \left(x_1 - \frac{L}{2}\right) \right\rangle \left\langle \left(x_2 - \frac{L}{2}\right) \right\rangle = \left(\langle x_1 \rangle - \frac{L}{2} \right) \left(\langle x_2 \rangle - \frac{L}{2} \right) = \left(\frac{L}{2} - \frac{L}{2} \right) \left(\frac{L}{2} - \frac{L}{2} \right) = 0$$

Por ejemplo, para el término cruzado que hemos escrito, lo podemos ver así: para un valor fijo de x_1 , los valores positivos y negativos de $(x_2 - L/2)$ son igualmente probables, y la media del producto debe ser cero. En consecuencia:

$$\left(\Delta^* x_{\text{media muestral}} \right)^2 = \frac{1}{N^2} \left\{ \left\langle \left(x_1 - \frac{L}{2}\right)^2 \right\rangle + \dots + \left\langle \left(x_N - \frac{L}{2}\right)^2 \right\rangle \right\}$$

El valor máximo que puede tomar uno de los sumandos $\left\langle \left(x_i - L/2\right)^2 \right\rangle$ se da en el caso de que $x_i = 0$ o bien $x_i = L$, es decir, $\left\langle \left(x_i - L/2\right)^2 \right\rangle$ valdrá como mucho $(L/2)^2$ y será, hablando toscamente, del mismo orden de magnitud.

$$\left(\Delta^* x_{\text{media muestral}} \right)^2 \sim \frac{1}{N^2} \cdot N \cdot \left(\frac{L}{2} \right)^2 = \frac{1}{N} \left(\frac{L}{2} \right)^2 = \frac{\langle x_{\text{media muestral}} \rangle^2}{N} \Rightarrow \boxed{\Delta^* x_{\text{media muestral}} \sim \frac{1}{\sqrt{N}} \langle x_{\text{media muestral}} \rangle}$$

Tomando la raíz cuadrada, y suponiendo un número de moléculas razonable para un gas en un recipiente ($N = 10^{20}$), se obtiene la desviación típica:

$$\Delta^* x_{\text{media muestral}} \sim \frac{1}{10^{10}} \langle x_{\text{media muestral}} \rangle$$

La distribución es extremadamente estrecha. Teniendo en cuenta que la altura h de la distribución es inversamente proporcional a la desviación típica ($h \propto 1/\Delta^* x_{\text{media muestral}}$), podemos ilustrar cómo de estrecha es la curva con los siguientes ejemplos:

- Si quisiéramos dibujar en una pizarra la distribución de probabilidad, de forma que su anchura midiese 1 cm, su altura debería ser de unos 10^{10} cm. Esto es **la cuarta parte de la distancia de la Tierra a la Luna**, una distancia enorme comparada con su anchura.
- Si tomamos un tamaño del recipiente contenedor de unos 60 cm, la anchura de la distribución sería $\Delta^* x_{\text{media muestral}} \simeq (1/10^{10}) \cdot (60/2) = 3 \cdot 10^{-9}$ cm = 0.03 nm. Sin embargo, no es posible dibujar la distribución con ninguna tiza en la pizarra, pues los trazos realizados son mucho más anchos que la propia distribución. Y, en el caso de que pudiésemos dibujarla, no podríamos apreciar su anchura, y nos parecería una “delta de Dirac”.

De la estrechez de la distribución podemos deducir varias cosas. Lo primero, que una abismal mayoría de las configuraciones microscópicas de las moléculas del gas producen una posición media muestral $x_{\text{media muestral}}$ increíblemente cercana a $L/2$. Esto no significa que necesariamente las distribuciones de las posiciones de las moléculas sean (más o menos) uniformes, pero la mayoría de las posibles distribuciones que producen un $x_{\text{media muestral}} \simeq L/2$ son (más o menos) uniformes. Podemos decir que:

- Una distribución (más o menos) uniforme de las posiciones de las moléculas individuales es muchísimo más probable que cualquier otra distribución.
- Hay una situación macroscópica, $\langle x_{\text{media muestral}} \rangle = L/2$, que es abrumadoramente más probable que cualquier otra (y que todas las demás juntas).

5.2 Gas ideal en un recipiente aislado, parte II. Analogía del desierto

El estado de un sistema puede describirse macroscópicamente mediante un conjunto de variables termodinámicas (que constituyen su **macroestado**) y microscópicamente mediante los valores detallados de sus variables mecánicas de acuerdo con un modelo de realidad (que constituyen su **microestado**). Algunos macroestados tienen muchos microestados que les corresponden (porque dichos microestados son compatibles con las ligaduras macroscópicas que imponen las paredes del sistema), mientras que a otros macroestados les corresponden pocos microestados.

Para visualizar el significado de estos conceptos, imaginemos un vasto desierto en el que podemos encontrarnos con algún oasis. Una persona “sin mente” comienza su viaje desde un oasis, y deambula por el desierto con movimientos irregulares y erráticos, describiendo un camino aleatorio y sin sentido. Con una probabilidad abrumadora, se adentrará en el desierto (porque hay mucho a su alrededor) y permanecerá allí (por la misma razón). Si hacemos algunas correspondencias, el ejemplo de la persona en el desierto nos servirá para entender el comportamiento de un sistema termodinámico con muchas partículas (Tabla 5.1).

<u>Correspondencias en la analogía del desierto</u>	
Un punto en el mapa	Un microestado específico
El desierto	El macroestado con muchos microestados compatibles
Un oasis	Un macroestado con pocos microestados compatibles
La persona “sin mente”	Un sistema de muchas partículas
El camino de la persona	La secuencia de microestados en la evolución del sistema

Tabla 5.1. Correspondencias en la analogía del desierto

Con ayuda de esta analogía, podemos comprender fenómenos como la difusión del gas en el ejemplo de la sección anterior. Supongamos que el recipiente **aislado** consta en realidad de dos compartimentos separados por una llave. Inicialmente, el gas ocupa uno de los dos compartimentos, mientras que el otro está vacío (Figura 5.3). Al abrirla, el gas se difunde rápidamente hacia la mitad vacía, llenando ambas mitades por igual al final del proceso. Las moléculas del gas parecen no retroceder todas juntas a la primera mitad de nuevo una vez abierta la llave, cosa que nuestra experiencia cotidiana nos hace pensar que no es sorprendente, pero, ¿podrían hacerlo? Claro que podrían, pues no hay nada en las leyes de Newton que impida que todas las moléculas vuelvan a la primera mitad, al igual que tampoco impedirían que al abrir la llave el gas no se difunda a la segunda mitad. Para que esto ocurriera, las colisiones entre las moléculas tendrían que ser “las adecuadas”, pero esto no es probable. Aquí está la clave, en la probabilidad de que se den las colisiones “adecuadas”. Entonces, ¿cómo es de probable una vuelta de las moléculas a la primera mitad?

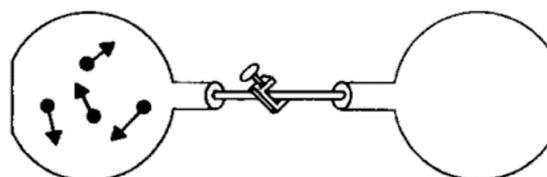


Figura 5.3. Gas encerrado en dos mitades

Dado el análisis de la sección anterior, el sentido común nos dice que al macroestado con $x_{\text{media muestral}} = L/2$ le corresponde, de forma abrumadora, el mayor número de microestados. Cuando se abre la llave que separaba las dos mitades, el gas evoluciona a través de muchos microestados y los correspondientes macroestados. El gas, al final, llegará al macroestado al que le corresponda un mayor número de microestados compatibles con él, sencillamente porque, como hemos visto, el macroestado que cumple esto tiene un número de microestados asociado que es abrumadoramente mayor. Entonces, sin duda habrá cambios en el microestado, pero el razonamiento anterior sugiere que es muy poco probable que se produzcan más cambios en el macroestado. Los continuos cambios de microestado serán casi con toda seguridad cambios de uno a otro de los muchos microestados correspondientes al macroestado con mayor número de microestados. Este macroestado final será el de equilibrio termodinámico.

5.3 Propiedades de un sistema “ideal” en el límite termodinámico

Los razonamientos presentados en las secciones anteriores pueden generalizarse para demostrar, de forma rigurosa, que en un sistema formado por muchísimas entidades elementales independientes se cumplen tres propiedades fundamentales, que enunciamos a continuación.

5.3.1 Estrechez de la distribución de probabilidad de las medias muestrales

A medida que aumenta el número de entidades elementales del sistema, la media muestral de una propiedad de las entidades elementales se aproxima cada vez más a la media probabilística de dicha propiedad.

Por ejemplo, supongamos un gas ideal encerrado en un recipiente aislado con una energía E y formado por N moléculas independientes. La energía de una molécula, ε , sigue una cierta distribución de probabilidad. A medida que $N \rightarrow \infty$, la media muestral $E/N = (\varepsilon_1 + \dots + \varepsilon_N)/N$ se aproxima cada vez más a la media probabilística $\langle \varepsilon \rangle$ calculada a partir de la distribución de probabilidad de la energía de una partícula.

5.3.2 Estrechez de la distribución de probabilidad de las variables extensivas del sistema completo

Supongamos que las ligaduras del sistema no impiden la variación de una propiedad extensiva (del sistema completo, no de una de las entidades elementales) y ésta puede fluctuar en el equilibrio. A medida que aumenta el número de entidades elementales del sistema, la distribución de probabilidad se hace extremadamente estrecha, tomando valores apreciablemente distintos de cero en valores prácticamente iguales al valor medio.

Por ejemplo, si tenemos un sistema que únicamente puede intercambiar energía con un termostato que fija su temperatura, su energía E no es constante (como ocurriría en un sistema aislado), y puede fluctuar. No obstante, a medida que el número N de entidades elementales tiende a infinito, la distribución de probabilidad de la energía (del sistema completo) se estrecha tanto que el valor medio $\langle E \rangle$ y el valor más probable coinciden, siendo las fluctuaciones completamente despreciables (condición que debe cumplirse si el estado es de equilibrio termodinámico, tal y como lo hemos definido).

Este hecho es de vital importancia y es el que permite **identificar el valor macroscópico de las variables extensivas en el equilibrio con su valor medio de acuerdo con la distribución de probabilidad de los microestados del sistema** (lo que será una hipótesis básica de trabajo en Física Estadística, tal y como se verá en el capítulo 6).

5.3.3 Direccionalidad de los procesos en un sistema aislado

Con una probabilidad abrumadora, si se permite efectuar un proceso a un sistema aislado formado por muchas entidades elementales independientes, evolucionará al macroestado con mayor número de microestados y permanecerá en él. En el capítulo 6, postularemos la identificación de la entropía estadística y la entropía termodinámica para un sistema en equilibrio en el límite termodinámico, y veremos que la entropía estadística de un sistema aislado es la entropía de Boltzmann-Planck, proporcional al logaritmo del número de microestados. Todo ello justifica el *Postulado II* de Callen, que explica la direccionalidad natural de los procesos de un sistema aislado hacia el macroestado con mayor entropía.

Debe resaltarse que esta tendencia hacia el macroestado con mayor entropía solo ocurre si el recipiente está aislado: si hay un flujo de materia o energía que puede entrar o salir del sistema, entonces el comportamiento del sistema estará influido, también, por estos intercambios.

A lo largo del trabajo hemos justificado dos principios de máximo relacionados con la entropía, pero hemos trabajado con dos entropías: la entropía termodinámica y la entropía estadística. Aunque ambas deben coincidir en el límite termodinámico, hay sutiles diferencias:

- El Principio de Máxima Entropía (estadística) es un principio que se puede aplicar a todo tipo de sistemas termodinámicos para hallar la distribución de probabilidad de los microestados de un sistema compatibles con un macroestado dado (capítulo 6). Su “dominio de acción” es el conjunto de las distintas distribuciones de probabilidad que podemos asignar a las variables extensivas de un macroestado.
- El *Postulado II* de Callen (de máxima entropía termodinámica) solo se aplica a un sistema termodinámico aislado para hallar su macroestado final de equilibrio de entre todos los compatibles con sus paredes externas aislantes y sus paredes internas.

5.4 Vuelta al juego de reparto. Justificación de la convergencia hacia las probabilidades inferidas mediante el PME

El juego de reparto del capítulo 2 es completamente análogo al ejemplo del gas dado en la sección 5.3.1 si hacemos corresponder:

- La energía del sistema E con el dinero total apostado D .
- Los posibles valores de la energía de una molécula ε con los posibles costes asociados a una ficha d .
- El número total de moléculas N con el número total de fichas N .

En el caso del juego de reparto, cada apuesta de una ficha se corresponde con una variable aleatoria cuyos sucesos elementales consisten en los cuatro posibles colores con sus costes asociados, y la media muestral se corresponde con el cociente D/N . Por lo tanto, a medida que $N \rightarrow \infty$, la media muestral D/N se aproxima cada vez más a la media probabilística $\langle d \rangle$ calculada a partir de la distribución de probabilidad del coste asociado a una ficha. Como consecuencia, la asignación $\langle d \rangle = D/N$ para la inferencia de las probabilidades mediante el PME es cada vez más adecuada a medida que $N \rightarrow \infty$ y el problema es cada vez más parecido al de lanzar N veces una única ficha.

Por otro lado, según el teorema demostrado en el capítulo 3, la distribución de frecuencias relativas obtenida de lanzar N veces 1 ficha, que a su vez se corresponde con la distribución de probabilidad desde el punto de vista frecuentista, se aproxima rápidamente a la distribución de probabilidad inferida mediante el PME a medida que $N \rightarrow \infty$.

Capítulo 6: Los conjuntos estadísticos.

Aplicación y problemas modelo

El **Postulado fundamental de la Física Estadística** consiste en suponer que las propiedades macroscópicas de los sistemas reales son el promedio estadístico de las variables microscópicas que los caracterizan; es decir, las propiedades macroscópicas se identifican con los macroestados.

Para E. T. Jaynes (1957), la Física Estadística es una **teoría de sistemas de muchas partículas** que funciona precisamente gracias a su **carácter estadístico**; y en la que **inferimos**, no postulamos, las probabilidades asociadas a los posibles microestados gracias a la Teoría de la Probabilidad y a unas pocas restricciones macroscópicas conocidas, que son constantes y definen el macroestado en cuestión.

En este capítulo trataremos sobre los así llamados conjuntos estadísticos, también conocidos como *conjuntos de Gibbs* o *colectividades estadísticas*. **Un conjunto estadístico es el espacio muestral de todos los posibles microestados de un sistema sujeto a unas ciertas ligaduras macroscópicas constantes.** Según cómo sean estas ligaduras, tendremos un conjunto u otro. En este capítulo presentaremos las ecuaciones más relevantes de dos conjuntos estadísticos: el conjunto microcanónico y el conjunto canónico. Estas expresiones se obtienen empleando el Principio de Máxima Entropía con las ligaduras macroscópicas que en cada conjunto sean necesarias. El desarrollo formal completo de estos conjuntos se encuentra en el **apéndice B** de este trabajo, junto con otros dos: el gran canónico y el isotermino-isobárico.

En los siguientes apartados consideraremos siempre que el sistema en estudio es un sistema monofásico, monocomponente y que sólo puede intercambiar un tipo de trabajo de configuración, siendo éste el trabajo hidrostático asociado al cambio de volumen: $\delta W = -PdV$.

6.1 El conjunto microcanónico

Se denomina conjunto microcanónico al espacio muestral formado por los microestados de un sistema aislado, para el que los valores de la energía interna U , el volumen V y el número de entidades elementales N son constantes. Los sistemas aislados están rodeados por paredes aislantes, las cuales impiden todo tipo de interacción con el entorno, y por lo tanto, son sistemas que no pueden intercambiar energía (en forma de calor o trabajo hidrostático asociado al cambio de volumen) ni materia.

- **Macroestado:** queda perfectamente determinado por los valores de las variables extensivas (U, V, N) , que son las variables naturales de la entropía $S = S(U, V, N)$, potencial termodinámico natural para describir un sistema aislado.
- **Microestados:** vienen determinados por los $\Omega(U, V, N)$ microestados compatibles con el macroestado (U, V, N) . El microestado i -ésimo tendrá una probabilidad asociada p_i . Estas probabilidades son siempre funciones de las variables del macroestado: $p_i = p_i(U, V, N)$.

6.1.1 Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, \Omega$) asociadas a cada microestado, se emplea el PME considerando únicamente la condición de normalización de las probabilidades:

$$\text{Maximizar } \boxed{S = -k_B \sum_{i=1}^{\Omega} p_i \ln p_i} \text{ sujeta a } \boxed{\sum_{i=1}^{\Omega} p_i = 1} \text{ nos proporciona } \boxed{p_i = \frac{1}{\Omega}} \quad (6.1)$$

Los microestados en el conjunto microcanónico son equiprobables, y la entropía es:

$$S = -k_B \sum_{i=1}^{\Omega} p_i \left(\ln \frac{1}{\Omega} \right) = k_B \ln \Omega \sum_{i=1}^{\Omega} p_i = k_B \ln \Omega \quad \rightarrow \quad \boxed{S = k_B \ln \Omega} \quad (6.2)$$

6.1.2 Ecuaciones de estado y funciones respuesta

Las primeras derivadas del potencial termodinámico natural, **la entropía**, respecto de las variables de las que depende, nos proporcionan las ecuaciones de estado, y las segundas derivadas, las funciones respuesta del sistema.

Conocida $S(U, V, N) = k_B \ln \Omega(U, V, N)$ y considerando $\beta = 1 / k_B T$, las ecuaciones másica de estado (EME), energética de estado (EEE) y térmica de estado (ETE) son:

$$\begin{array}{ccc} \text{(EME)} & \text{(EEE)} & \text{(ETE)} \\ \boxed{\alpha = -\mu\beta = \left(\frac{\partial \ln \Omega}{\partial N} \right)_{U,V}} & \boxed{\beta = \left(\frac{\partial \ln \Omega}{\partial U} \right)_{N,V}} & \boxed{\gamma = P\beta = \left(\frac{\partial \ln \Omega}{\partial V} \right)_{U,N}} \end{array} \quad (6.3)$$

La segunda derivada del potencial termodinámico respecto de U nos proporciona la siguiente función respuesta:

$$\boxed{\left(\frac{\partial^2 \ln \Omega}{\partial U^2} \right)_{V,N} = -\frac{1}{k_B T^2 C_V} = -\frac{k_B \beta^2}{C_V}} \quad (6.4)$$

donde $C_V = \left(\partial U / \partial T \right)_{V,N}$ es la capacidad calorífica a volumen constante.

Cabe destacar el hecho de que, si la entropía S es máxima respecto de los cambios de energía interna U , su segunda derivada debe ser negativa y, por tanto, la capacidad calorífica a volumen constante debe ser positiva: $C_V \geq 0$ (condición de estabilidad energética).

6.2 El conjunto canónico

Se denomina conjunto canónico al espacio muestral formado por los microestados de un sistema cerrado e isócoro en contacto con un termostato que se encarga de fijar su temperatura (sistema isotermo), de forma que la temperatura T , el volumen V y el número de entidades elementales N son estrictamente constantes, pero la energía interna U puede fluctuar, pues habrá distintos microestados con diferentes valores de la energía interna. El sistema puede intercambiar cantidades arbitrariamente grandes de energía con el termostato, de modo que no tenemos ninguna base racional para descartar ningún valor de la energía entre su valor mínimo y máximo. En este caso se fija la variable canónica conjugada de la energía (se fija β y, por tanto, también $1/T$), con lo que queda fijada también la energía media del sistema. De esta forma, el valor macroscópico de la energía interna es igual a su valor medio, es decir, se tiene $U = \langle E \rangle$.

- **Macroestado:** queda perfectamente determinado por las variables $(1/T, V, N)$, que son las variables naturales de la función de Massieu $\Psi = \Psi(1/T, V, N) = \Psi(\beta, V, N)$, potencial termodinámico natural para describir un sistema cerrado isotermo e isócoro.

- **Microestados:** Cada microestado está caracterizado por los valores que toman las variables extensivas del macroestado, en este caso V y N . La temperatura jugará un papel en las probabilidades de que el sistema se encuentre en uno u otro nivel de energía, pero no caracteriza la energía de cada nivel. Caracterizará, sin embargo, que los intercambios de energía entre el sistema y el termostato sean de tal suerte que se fije el valor medio de la energía. Así pues, cada microestado i -ésimo del sistema tendrá una cierta energía $E_i = E_i(V, N)$, y denotaremos al número de microestados como M en este caso. Las probabilidades, como siempre, serán funciones de todas las variables del macroestado: $p_i = p_i(1/T, V, N)$.

6.2.1 Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, M$) asociadas a cada uno de los M microestados, se emplea el Principio de Máxima Entropía considerando la condición de normalización y la ligadura correspondiente a la energía media:

$$\text{Maximizar } \boxed{S = -k_B \sum_{i=1}^M p_i \ln p_i} \text{ sujeta a } \boxed{\sum_{i=1}^M p_i = 1} \text{ y } \boxed{\sum_{i=1}^M p_i E_i = \langle E \rangle} \rightarrow \boxed{p_i = \frac{e^{-\beta E_i}}{Z}} \quad (6.5)$$

La variable Z se denomina “función de partición” (es la **función de estructura** en este conjunto), cuya expresión es:

$$\boxed{Z(\beta, V, N) = \sum_{i=1}^M e^{-\beta E_i(N, V)} = \sum_{j=1}^L \Omega(E_j, V, N) e^{-\beta E_j(N, V)}} \quad (6.6)$$

Como puede que algunos microestados tengan la misma energía E_i (están degenerados), se tendrán $L \leq M$ niveles de energía, cada uno de ellos con un valor diferente de energía E_j y con una degeneración $\Omega_j \equiv \Omega(E_j, V, N)$ tal que $\sum_{j=1}^L \Omega(E_j, V, N) = M$. Así, la función de partición se puede escribir como una suma a microestados, o como una suma a niveles de energía con sus respectivas degeneraciones.

La entropía del conjunto canónico y la función de Massieu $\Psi(\beta, V, N)$ son las siguientes:

$$\boxed{S = k_B (\ln Z + \beta \langle E \rangle)} \quad (6.7)$$

$$\boxed{\Psi = k_B \ln Z} \quad (6.8)$$

6.2.2 Ecuaciones de estado, funciones respuesta y fluctuaciones

Conocida $\Psi(\beta, V, N) = k_B \ln Z(\beta, V, N)$ las ecuaciones de estado son:

$$\begin{array}{ccc} \text{(EME)} & \text{(EEE)} & \text{(ETE)} \\ \boxed{\alpha = -\mu\beta = \left(\frac{\partial \ln Z}{\partial N} \right)_{\beta, V}} & \boxed{\langle E \rangle = U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{N, V}} & \boxed{\gamma = P\beta = \left(\frac{\partial \ln Z}{\partial V} \right)_{\beta, N}} \end{array} \quad (6.9)$$

Las segundas derivadas de la función de Massieu nos proporcionan las funciones respuesta. En particular, la función respuesta respecto de la energía da cuenta de las fluctuaciones de energía del sistema:

$$\boxed{(\Delta^* E)^2 = \langle E^2 \rangle - \langle E \rangle^2 = \left(\frac{\partial^2 \ln Z}{\partial \beta^2} \right)_{V, N} = k_B T^2 C_V = \frac{C_V}{k_B \beta^2} \geq 0} \quad (6.10)$$

Donde, análogamente a lo que pasaba en el conjunto microcanónico (6.11), el que la segunda derivada de la función de Massieu respecto de su variable natural (β) sea definida positiva (coincide con la varianza del sistema), hace que la capacidad calorífica a volumen constante deba de ser positiva $C_V \geq 0$ (condición de estabilidad energética).

Por otro lado, la función respuesta para el volumen, vale:

$$\boxed{\left(\frac{\partial^2 \ln Z}{\partial V^2}\right)_{\beta, N} = -\frac{\beta}{V\kappa_T} = -\frac{1}{k_B T V \kappa_T}} \quad \text{con} \quad \kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_{T, N} \quad (6.12)$$

donde κ_T es el coeficiente de compresibilidad isotérmico. Nuevamente, a destacar el hecho de que, si la función de Massieu (Ψ) es máxima respecto de los cambios de volumen V del sistema, su segunda derivada deber ser negativa y, por tanto, el coeficiente de compresibilidad isoterma debe de ser positiva ($\kappa_T \geq 0$) en los sistemas con temperatura definida positiva (condición de estabilidad térmica).

6.3 Protocolo de trabajo en Física Estadística

Gracias al Principio de Máxima Entropía se obtiene, de manera general, que en todos los conjuntos estadísticos existe una expresión que relaciona el **potencial termodinámico entrópico** natural a las variables macroscópicas que caracterizan el sistema (la función termodinámica correspondiente en la representación entrópica que depende de esas variables), con la llamada **función estructura microscópica**: el número de microestados en el conjunto microcanónico, la función de partición en el conjunto canónico, etc., (función que da cuenta del comportamiento probabilístico de los microestados del sistema), que siempre se escribe de la siguiente forma:

$$\boxed{\text{potencial termodinámico entrópico} = k_B \ln(\text{función estructura})}$$

Es muy importante indicar que, para resolver cualquier problema de Física Estadística, debe seguirse un protocolo de trabajo consistente en 4 etapas fundamentales, en 3 de las cuales (I, III y IV) se necesita un profundo conocimiento del formalismo termodinámico:

- I. Definir un **macroestado** a caracterizar.
- II. Postular un **modelo microscópico** de comportamiento.
- III. Encontrar la **función estructura** que caracteriza al sistema.
- IV. Obtener la **termodinámica del sistema**.

6.4 Problema modelo

Resolvamos un problema modelo que aclare de forma inequívoca cómo se aplica el protocolo de trabajo de la Física Estadística.

Problema de dos niveles resuelto en el conjunto microcanónico

I) Macroestado (U, V, N): Considere un **sistema aislado** (en equilibrio termodinámico) de N partículas, cuya energía interna es U y que se encuentran en un recipiente de volumen V .

II) Modelo microscópico: Cada partícula posee un grado de libertad con dos posibles estados cuánticos no degenerados, de energías $\varepsilon_0 = 0$ y $\varepsilon_1 = \varepsilon > 0$. Las partículas son **independientes** y **distinguibles**. Puede especificarse un (**micro**)estado del sistema $|\varphi\rangle$, mediante la siguiente lista de números:

$$|\varphi\rangle \equiv |n_1, \dots, n_j, \dots, n_N\rangle \quad [n_j = 0, 1] \quad \forall j = 1, \dots, N$$

donde cada uno de los n_j caracteriza el **estado** en el que se encuentra la partícula j -ésima. Así pues, la energía (interna) total de este sistema, que **en este caso no depende de V** , vale:

$$U \equiv E = E_{|\varphi\rangle} = \sum_{j=1}^N n_j \cdot \varepsilon = N_0 \cdot 0 + N_1 \cdot \varepsilon = N_1 \varepsilon$$

donde N_0 es el número de partículas situadas en el nivel inferior ε_0 y N_1 el número de partículas que se encuentran situadas en el nivel superior ε_1 , con $N = N_0 + N_1$.

III) Función estructural: Como el sistema es aislado, se debe trabajar usando el formalismo del conjunto microcanónico. La **función estructural** del sistema es, pues, el número de microestados $\Omega = \Omega(E, N)$ compatibles con el macroestado (E, N) del sistema.

Solución:

El sistema a estudiar está compuesto por N_0 partículas en el nivel inferior, y N_1 partículas en el nivel superior (Figura 6.1). Además, tendremos en cuenta las siguientes relaciones en la resolución:

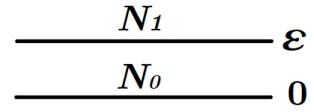


Figura 6.1. Dos niveles de energía

$$\boxed{E = N_1 \varepsilon} \quad \boxed{N = N_0 + N_1} \quad \boxed{N_1 = \frac{E}{\varepsilon}, \quad p_1 = \frac{N_1}{N} = \frac{E}{\varepsilon N}} \quad \boxed{N_0 = N - N_1, \quad p_0 = \frac{N_0}{N} = 1 - \frac{E}{\varepsilon N}}$$

El número de microestados será el número de posibles configuraciones (compatibles con el macroestado) en las que se puede encontrar el sistema. En este problema será el número de formas diferentes en las que se pueden colocar N partículas, de forma que haya N_0 partículas en el nivel inferior, y N_1 partículas en el nivel superior. Esto son las permutaciones con repetición de N elementos, con N_0 elementos iguales, y N_1 elementos iguales:

$$\Omega(E, N) = PR_N^{N_0, N_1} = \frac{N!}{N_0! N_1!} \quad \Rightarrow \quad \ln \Omega = \ln(N!) - \ln(N_0!) - \ln(N_1!)$$

Considerando que el número de partículas es muy grande, podemos aplicar la aproximación de Stirling [ecuación (3.3)] para el logaritmo del factorial de un número muy grande:

$$\ln(N!) \approx N \ln N - N \quad \ln(N_0!) \approx N_0 \ln N_0 - N_0 \quad \ln(N_1!) \approx N_1 \ln N_1 - N_1$$

Llevando estas aproximaciones al número de microestados:

$$\begin{aligned} \ln \Omega &\approx N \ln N - N - N_0 \ln N_0 + N_0 - N_1 \ln N_1 + N_1 = N \ln N - N_0 \ln N_0 - N_1 \ln N_1 \\ &= N_0 \ln N + N_1 \ln N - N_0 \ln N_0 - N_1 \ln N_1 = N_0 \ln \left(\frac{N}{N_0} \right) + N_1 \ln \left(\frac{N}{N_1} \right) \\ &= -N \left[\frac{N_0}{N} \ln \left(\frac{N_0}{N} \right) + \frac{N_1}{N} \ln \left(\frac{N_1}{N} \right) \right] = -N (p_0 \ln p_0 + p_1 \ln p_1) \end{aligned}$$

Es decir, el número de microestados del sistema es: $\boxed{\ln \Omega \approx -N (p_0 \ln p_0 + p_1 \ln p_1)}$

Así, la entropía del sistema completo es: $\boxed{S = k_B \ln \Omega = N [-k_B (p_0 \ln p_0 + p_1 \ln p_1)]}$, de forma que, si ahora definimos la entropía de cada una de las partículas del sistema como:

$$S_1 = -k_B \sum_{i=0}^{i=1} p_i \ln p_i = -k_B (p_0 \ln p_0 + p_1 \ln p_1)$$

tendremos que $\boxed{S = N \cdot S_1}$, demostrando que la entropía del sistema completo es aditiva, siempre que se definan las probabilidades de los estados de las partículas en la forma: $p_1 = N_1/N$ y $p_0 = N_0/N$.

IV) Termodinámica del sistema: Primeras y segundas derivadas del potencial termodinámico natural $S = S(E, N)$ para obtener las ecuaciones de estado y las funciones respuesta.

Solución:

Conviene expresar el número de microestados (función estructura) en función de sus variables naturales:

$$\ln \Omega(E, N) = N \ln N - N_0 \ln N_0 - N_1 \ln N_1 = N \ln N - \left(N - \frac{E}{\varepsilon}\right) \ln \left(N - \frac{E}{\varepsilon}\right) - \left(\frac{E}{\varepsilon}\right) \ln \left(\frac{E}{\varepsilon}\right)$$

▪ **DEPENDENCIA CON E**

Su derivada respecto a la energía nos da la EEE:

$$\beta = \left(\frac{\partial \ln \Omega}{\partial E}\right)_{N,V} = \frac{1}{\varepsilon} \ln \left(\frac{N\varepsilon}{E} - 1\right) \Rightarrow \boxed{\beta\varepsilon = \ln \left(\frac{N\varepsilon}{E} - 1\right)}$$

Invirtiendo esta relación, obtenemos: $\boxed{E(\beta, N) = \frac{N\varepsilon}{1 + e^{\beta\varepsilon}}}$

Antes de estudiar los límites de $E = E(\beta, N)$ en $T \rightarrow 0^+$ y $T \rightarrow +\infty$, escribiremos las expresiones de las probabilidades de los estados de una partícula:

$$\boxed{p_0 = \frac{N_0}{N} = \frac{1}{1 + e^{-\beta\varepsilon}}} \quad \boxed{p_1 = \frac{N_1}{N} = \frac{1}{1 + e^{\beta\varepsilon}}} \quad \boxed{p_0 = e^{\beta\varepsilon} p_1}$$

Ahora, calculemos el límite $E(T \rightarrow 0^+)$:

▪ $\boxed{\lim_{T \rightarrow 0^+} E = \lim_{\beta \rightarrow +\infty} E = \lim_{\beta \rightarrow +\infty} \left(\frac{N\varepsilon}{1 + e^{\beta\varepsilon}}\right) = \frac{N\varepsilon}{\infty} = 0}$

En este límite, las probabilidades de los estados de una partícula son:

$$\boxed{p_0 = 1} \quad \boxed{p_1 = 0}$$

Es decir, todas las partículas del sistema están en su nivel fundamental (el nivel inferior), y además, el sistema completo también está en su nivel fundamental, pues su valor de energía es el mínimo posible de entre todos los valores que puede tomar: $E = E_{\min} = 0$.

Se cumple entonces el Tercer Principio de la Termodinámica: la entropía de un sistema en el cero absoluto alcanza un valor mínimo y constante. **Un sistema a temperatura cero se encuentra en su estado fundamental de energía, y su entropía queda determinada únicamente por la degeneración de dicho estado fundamental.**

Tanto el estado fundamental de una partícula como el estado fundamental del sistema completo son no degenerados. En el caso de una partícula es evidente. En el caso del sistema completo, sólo existe una forma de colocar las N partículas de forma que $E = E_{\min} = 0$, y dicha forma es todas ellas en su nivel fundamental, es decir, $N_0 = N$ y $N_1 = 0$. Si calculamos la entropía, tanto del sistema como de cada una de las partículas:

$$S_1 = -k_B (p_0 \ln p_0 + p_1 \ln p_1) = 0 \quad , \quad S = NS_1 = 0$$

La entropía es nula, resultado que podríamos haber esperado. La entropía de un sistema en el cero absoluto se anula si el estado fundamental es no degenerado.

Ahora, calculemos el límite $E(T \rightarrow +\infty)$

$$\lim_{T \rightarrow +\infty} E = \lim_{\beta \rightarrow 0^+} E = \lim_{\beta \rightarrow 0^+} \left(\frac{N\varepsilon}{1 + e^{\beta\varepsilon}} \right) = \frac{N\varepsilon}{1 + 1} = \frac{N\varepsilon}{2}$$

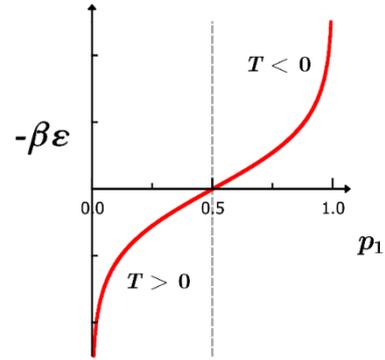
Es el caso equiprobable, pues en este límite las probabilidades de los estados son $p_0 = p_1 = 1/2$.

Finalmente, estudiemos en qué condiciones se podrían tener **temperaturas negativas** en el sistema, es decir, $\beta < 0$. Para ello, tiene que darse la siguiente condición:

$$\ln \left(\frac{N\varepsilon}{E} - 1 \right) < 0 \Rightarrow E > \frac{N\varepsilon}{2}$$

Para verlo mejor, dibujemos ahora la gráfica de $-\beta\varepsilon$ como función de $p_1 = E/(N\varepsilon)$. Para ello, utilizaremos la ecuación de la temperatura del sistema:

$$-\beta\varepsilon = -\ln \left(\frac{N\varepsilon}{E} - 1 \right) \Rightarrow -\beta\varepsilon = -\ln \left(\frac{1}{p_1} - 1 \right) = \ln \left(\frac{p_1}{p_0} \right)$$



Analizando la gráfica (Figura 6.2), observamos que, cuando el nivel superior está más poblado que el inferior ($p_1 > p_0$, $E > N\varepsilon/2$), las temperaturas del sistema son negativas.

Estamos ante un sistema cuya energía está acotada superiormente (esta cota superior corresponde a todas las partículas en el nivel superior, esto es, $E_{\max} = N\varepsilon$). Como ya hemos visto en la sección 4.3, los sistemas con temperaturas negativas suelen llamarse sistemas “anormales” (Biel, 1998).

Calculemos ahora C_V . Para ello utilizaremos la ecuación (6.4):

$$\left(\frac{\partial^2 \ln \Omega}{\partial U^2} \right)_{V,N} = -\frac{k_B \beta^2}{C_V} = -\frac{N}{(N\varepsilon - E)E} \Rightarrow C_V = Nk_B (\beta\varepsilon)^2 \frac{e^{\beta\varepsilon}}{(1 + e^{\beta\varepsilon})^2}$$

DEPENDENCIA CON N

Su derivada respecto a la energía respecto del número de partículas nos da la EME:

$$\alpha = -\mu\beta = \left(\frac{\partial \ln \Omega}{\partial N} \right)_{E,V} = \ln N - \ln \left(N - \frac{E}{\varepsilon} \right) = \ln \left(\frac{N}{N_0} \right) = -\ln p_0 = -\ln(1 - p_1)$$

Utilizando ahora la EEE, obtenemos que:

$$\alpha = -\mu\beta = -\ln p_0 = \ln(1 + e^{-\beta\varepsilon}) \Rightarrow \mu\beta = \ln(1 - p_1)$$

Estudiamos ahora su comportamiento en los límites de alta y baja temperatura:

- $\alpha(T \rightarrow 0^+) = \alpha(\beta \rightarrow +\infty) = \ln(1 + 0) = 0$
- $\alpha(T \rightarrow +\infty) = \alpha(\beta \rightarrow 0^+) = \ln(1 + 1) = \ln 2$

Ahora, dibujamos la gráfica de $\beta\mu$ como función de $p_1 = E/(N \cdot \varepsilon)$ dibujamos también la gráfica de $\alpha = \ln(1 + e^{-\beta\varepsilon})$ como función de $(-\beta)$ (Figura 6.3):

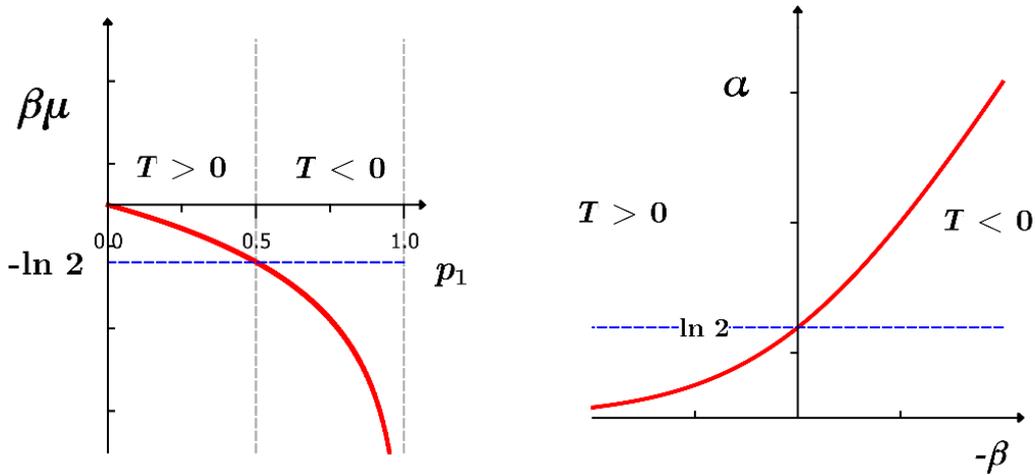


Figura 6.3. Gráficas de la dependencia de la EME con variables de interés.

Calculamos la actividad absoluta:

$$\lambda = e^{\mu\beta} = \frac{1}{1 + e^{-\beta\varepsilon}} = 1 - \frac{E}{N\varepsilon} = 1 - p_1 = p_0$$

Por último, representamos las siguientes gráficas (Figura 6.4):

- $\lambda = \lambda(-\beta) = \frac{1}{1 + e^{-\beta\varepsilon}}$
- $\lambda = \lambda(p_1) = 1 - p_1$

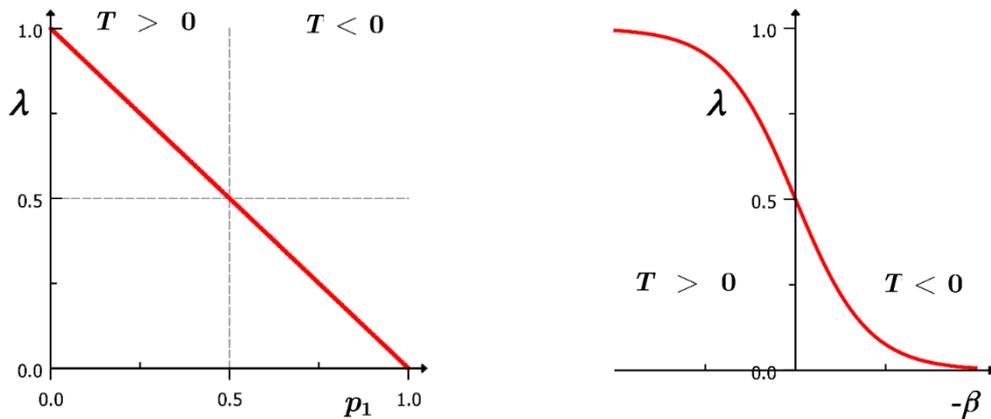


Figura 6.4. Gráficas de la dependencia de la actividad absoluta con dos variables de interés.

Finalmente, calculamos entropía del sistema como función de la temperatura $S = S(\beta, N)$.

$$S(\beta, N) = N \cdot S_1 = Nk_B \left[\frac{\ln(1 + e^{-\beta\varepsilon})}{1 + e^{-\beta\varepsilon}} + \frac{\ln(1 + e^{\beta\varepsilon})}{1 + e^{\beta\varepsilon}} \right]$$

O bien, operando:

$$S(\beta, N) = k_B \left[\beta E + N \ln(1 + e^{-\beta \varepsilon}) \right] = Nk_B \left[\frac{\beta \varepsilon}{1 + e^{\beta \varepsilon}} + \ln(1 + e^{-\beta \varepsilon}) \right]$$

Calculamos ahora los límites de baja y alta temperatura:

- $S(T \rightarrow 0^+) = S(\beta \rightarrow +\infty) = 0$
- $S(T \rightarrow +\infty) = S(\beta \rightarrow 0^+) = Nk_B \ln 2$

Conclusiones

En el trabajo se ha concluido que el Principio de Máxima Entropía (PME) es un método de razonamiento que permite obtener las “mejores” predicciones que pueden hacerse con la información que se tiene sobre un problema a resolver. De esta forma, el PME permite inferir probabilidades desconocidas a partir de la información disponible, maximizando la entropía estadística sujeta a las ligaduras macroscópicas que impone dicha información.

Además, mediante ejemplos didácticos hemos ilustrado diferentes comportamientos:

- **La tendencia de las probabilidades frecuentistas a las que predice el PME cuando el número de experimentos realizados aumenta, justificando así la bondad del PME como método de razonamiento.**
- La estrechez de la distribución de probabilidad de las variables extensivas de un sistema en equilibrio termodinámico cuando su número de entidades elementales es muy elevado. Esto nos permite afirmar que las fluctuaciones de dichas variables en el equilibrio son despreciables, y así se puede **identificar el valor macroscópico de las variables extensivas en el equilibrio con su valor medio** de acuerdo con la distribución de probabilidad de los microestados del sistema, lo que constituye una hipótesis básica de trabajo en Física Estadística.

Con el PME se han podido estudiar también diferentes conjuntos estadísticos, que se han definido como los espacios muestrales de los posibles microestados de un determinado sistema a estudiar. En todos los conjuntos estudiados en el capítulo 6 y en el apéndice B, se llega a una relación del siguiente tipo:

$$\boxed{\text{potencial termodinámico entrópico} = k_B \ln(\text{función estructura})}$$

que establece la conexión entre la descripción microscópica (la de un modelo de realidad) y macroscópica (la descripción termodinámica) de los sistemas reales.

Como último punto, es importante notar que durante todo el trabajo se ha empleado una descripción entrópica, tanto en la parte de Termodinámica, como en la parte de Estadística. **Una descripción estadística de los sistemas físicos formados por muchas entidades elementales (que son los que estudia la Física Estadística) requiere necesariamente una descripción en términos de la entropía y de sus transformadas de Legendre (función de Massieu, función de Kramers-Landau, función de Planck...).**

La descripción en términos de la entropía (representación entrópica de la Termodinámica) es mucho más adecuada para tratar este tipo de sistemas, ya que nos permite obtener conclusiones e ideas de una forma mucho más inmediata y autoconsistente, cosa que no ocurre si se intentase utilizar la representación energética de la Termodinámica.

Bibliografía

- Anderson, P. W. (1972) *More Is Different*. Science, New Series, 177, 393-396.
- Baixauli, L. G. y Cobos, J. C. (1996). *Conjuntos de Gibbs. El Método de Entropía Máxima en Mecánica Estadística*. Grado en Física. Universidad de Valladolid.
- Balian, R. (2017). *François Massieu and the thermodynamic potentials*. C. R. Physique 18, 526-530.
- Biel, J. (1998). *Formalismo y métodos de la Termodinámica* (Vol. 1). Editorial Reverté, S.A.
- Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics* (Second Edition). John Wiley and Sons.
- Campa, A., Dauxois, T. y Ruffo, S. (2009). *Statistical mechanics and dynamics of solvable models with long-range interactions*. Physics Reports, 480, 57-159.
- Campa, A., Casetti, L., Latella, I., Pérez-Madrid, A. y Ruffo, S. (2018). *Concavity, Response Functions and Replica Energy*. Entropy, 20, 907.
- Casanova, C. (2019). *Apuntes de Termodinámica*. Grado en Física. Universidad de Valladolid.
- Costeniuc, M., Ellis, R.S. y Touchette, H. (2006). *Nonconcave entropies from generalized canonical ensembles*. Physical Review E, 74, 010105(R).
- Gibbs, J. W. (1902). *Elementary Principles of Statistical Mechanics*. Yale University Press, New Haven.
- Hevia, L. F. (2014). *Entropía y desorden*. Trabajo Fin de Grado. Facultad de Ciencias. Universidad de Valladolid. <http://uvadoc.uva.es/handle/10324/6094>
- Hevia, L. F. (2023). *Fundamentos de Física Estadística*. Universidad de Valladolid. <https://uvadoc.uva.es/handle/10324/60205>
- Hill, T. L. (1970). *Introducción a la Termodinámica Estadística*. Paraninfo.
- Hill, T.L. (2002). *Thermodynamics of small systems*. Dover Publications, New York.
- Jaynes, E. T. (1957). *Information Theory and Statistical Mechanics*. Physical Review, 106(4), 620-630. <https://doi.org/10.1103/physrev.106.620>
- Jaynes, E. T. (1978). *Where do we stand on Maximum Entropy?* Levine, R. y Tribus, M. (Eds.), The Maximum Entropy Formalism Conference, MIT Press, Cambridge, Massachusetts. <https://bayes.wustl.edu/etj/articles/stand.on.entropy.pdf>
- Jaynes, E. T. (1989). *E.T. Jaynes: Papers on Probability, Statistics, and Statistical Physics*. Springer Science & Business Media.
- Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.
- Latella, I., Pérez-Madrid, A., Campa, A., Casetti, L. y Ruffo, S. (2015). *Thermodynamics of Nonadditive Systems*. Physical Review Letters, 114, 230601.
- Loscar, E.S. y Horowitz, C.M. (2018). *Size effects in finite systems with long-range interactions*. Physical Review E, 97, 032103.
- Lee, L. L. (1988). *Molecular Thermodynamics of Nonideal Fluids*. Butterworth-Heinemann.
- Maxwell, J. C. (1860). V. *Illustrations of the dynamical theory of gases*. Part I. *On the motions and collisions of perfectly elastic spheres*. The London, Edinburgh and Dublin philosophical magazine and journal of science, 19(124), 19-32. <https://doi.org/10.1080/14786446008642818>

- Mozo, I. (2010) *Estudio experimental y teórico de mezclas binarias de n-butanol, di-n-butiléter y celosolvas*. Tesis Doctoral. Departamento de Física Aplicada. Facultad de Ciencias. Universidad de Valladolid.
- O'Connell, J. P., y Haile, J. M. (2005). *Thermodynamics. Fundamentals for Applications* Cambridge University Press.
- Planck, M. (1914). *The Theory of Heat Radiation* (M. Masius, Trad.). P. Blakiston, Son & Co.
- Shannon, C. E. (1948). *A Mathematical Theory of Communication*. Reprinted with corrections from The Bell System Technical Journal, 27, 379-423, 623-659.
- Tejerina, A. F. (1976). *Termodinámica* (Vol. 1). Paraninfo.
- Touchette, H., Ellis, R.S. y Turkington, B. (2004). *An introduction to the thermodynamic and microstate levels of nonequivalent ensembles*. Physica A, 340, 138-146.
- Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics*. Springer.
- von Neumann, J. (1932). *Mathematical Foundations of Quantum Mechanics* (N. A. Wheeler, Ed.; R. T. Beyer, Trad.; New Edition). Amsterdam University Press.

Apéndice A: Complementos al Capítulo 2

A.1. Tablas de tabulación directa del juego de reparto

A.1.1. $N = 5$

Número de apuesta (j)	Rojo: n_{1j} $d_1 = 4\text{€}$	Negro: n_{2j} $d_2 = 2\text{€}$	Amarillo: n_{3j} $d_3 = 1\text{€}$	Verde: n_{4j} $d_4 = 0\text{€}$	Degeneración (Ω_j)
1	0	0	5	0	1
2	0	1	3	1	20
3	0	2	1	2	30
4	1	0	1	3	20
Total fichas a un color: F_i	20	80	115	140	$\Omega = \sum_{j=1}^4 \Omega_j = 71$
Probabilidad	0.056338	0.225352	0.323944	0.394366	$N\Omega = 355$

A.1.2. $N = 15$

Número de apuesta (j)	Rojo: n_{1j} $d_1 = 4\text{€}$	Negro: n_{2j} $d_2 = 2\text{€}$	Amarillo: n_{3j} $d_3 = 1\text{€}$	Verde: n_{4j} $d_4 = 0\text{€}$	Degeneración (Ω_j)
1	0	0	15	0	1
2	0	1	13	1	210
3	0	2	11	2	8190
4	1	0	11	3	5460
5	0	3	9	3	100100
6	1	1	9	4	150150
7	0	4	7	4	450450
8	1	2	7	5	1081080
9	2	0	7	6	180180
10	0	5	5	5	756756
11	1	3	5	6	2522520
12	2	1	5	7	1081080
13	0	6	3	6	420420
14	1	4	3	7	1801800
15	2	2	3	8	1351350
16	3	0	3	9	100100
17	0	7	1	7	51480
18	1	5	1	8	270270
19	2	3	1	9	300300
20	3	1	1	10	60060
Total fichas a un color: F_i	12137580	31968510	47892015	68381250	$\Omega = \sum_{j=1}^{20} \Omega_j = 10691957$
Probabilidad	0.075680	0.199331	0.298617	0.426372	$N\Omega = 160379355$

A.1.3. $N = 20$

Número de apuesta (j)	Rojo: n_{1j} $d_1 = 4\text{€}$	Negro: n_{2j} $d_2 = 2\text{€}$	Amarillo: n_{3j} $d_3 = 1\text{€}$	Verde: n_{4j} $d_4 = 0\text{€}$	Degeneración (Ω_j)
1	0	0	20	0	1
2	0	1	18	1	380
3	0	2	16	2	29070
4	0	3	14	3	775200
5	0	4	12	4	8817900
6	0	5	10	5	46558512
7	0	6	8	6	116396280
8	0	7	6	7	133024320
9	0	8	4	8	62355150
10	0	9	2	9	9237800
11	0	10	0	10	184756
12	1	0	16	3	19380
13	1	1	14	4	1162800
14	1	2	12	5	21162960
15	1	3	10	6	155195040
16	1	4	8	7	498841200
17	1	5	6	8	698377680
18	1	6	4	9	387987600
19	1	7	2	10	66512160
20	1	8	0	11	1511640
21	2	0	12	6	3527160
22	2	1	10	7	66512160
23	2	2	8	8	374130900
24	2	3	6	9	775975200
25	2	4	4	10	581981400
26	2	5	2	11	126977760
27	2	6	0	12	3527160
28	3	0	8	9	27713400
29	3	1	6	10	155195040
30	3	2	4	11	211629600
31	3	3	2	12	70543200
32	3	4	0	13	2713200
33	4	0	4	12	8817900
34	4	1	2	13	8139600
35	4	2	0	14	581400
36	5	0	0	15	15504
Total fichas a un color: F_i	7169650380	18222672000	27398662740	39731623140	$\Omega = \sum_{j=1}^{36} \Omega_j$ $= 4626130413$
Probabilidad	0.077491	0.196954	0.296129	0.429426	$N\Omega =$ 92522608260

Vemos que la resolución por tabulación directa (Hevia, 2023) se complica enormemente al aumentar el número de fichas, lo que hace que para sistemas con $N \sim 10^{23}$ no resulte viable.

A.2. Comentario sobre la maximización de la entropía

Al aplicar el método de los multiplicadores de Lagrange en la resolución del juego de reparto mediante el Principio de Máxima Entropía, se ha encontrado una distribución de probabilidad que es un punto crítico de la entropía (2.2) sujeta a las ligaduras (2.3) y (2.4), pero no se ha dicho nada de que ese punto crítico sea realmente un máximo condicionado de la entropía. Para demostrar que la distribución encontrada $\{p_1, \dots, p_n\}$ es un máximo (Jaynes, 2003; Hevia, 2023), será de ayuda la siguiente desigualdad:

$$\ln x \geq \left(1 - \frac{1}{x}\right) \quad (\text{A.1})$$

La igualdad se da si y solo si $x = 1$. Considérese ahora otra distribución de probabilidad cualquiera $\{q_1, \dots, q_n\}$ de los posibles microestados que sea compatible con las mismas ligaduras que las $\{p_1, \dots, p_n\}$:

$$\sum_{i=1}^n q_i = 1 \quad \sum_{i=1}^n q_i d_i = \langle d_i \rangle \quad (\text{A.2})$$

El objetivo es comprobar que tal distribución $\{q_1, \dots, q_n\}$ conduce a una entropía que es menor o igual que la dada por la distribución $\{p_1, \dots, p_n\}$. Para ello se empleará la desigualdad (A.1), con $x = q_i / p_i$:

$$\sum_i q_i \ln \left(\frac{q_i}{p_i} \right) \geq \sum_i q_i \left(1 - \frac{p_i}{q_i} \right) = \sum_i q_i - \sum_i p_i = 0 \quad (\text{A.3})$$

Esto nos lleva a la siguiente relación:

$$\sum_i q_i \ln \left(\frac{q_i}{p_i} \right) = \sum_i q_i \ln q_i - \sum_i q_i \ln p_i \geq 0 \quad \Rightarrow \quad \sum_i q_i \ln q_i \geq \sum_i q_i \ln p_i \quad (\text{A.4})$$

La igualdad se cumple si y solo si $p_i = q_i$. Sustituyendo p_i por su valor:

$$\sum_i q_i \ln q_i \geq \sum_i q_i \ln p_i = \sum_i q_i (-\beta d_i - \ln Z) = -\beta \langle d \rangle - \ln Z \quad (\text{A.5})$$

o lo que es lo mismo:

$$-k \sum_i q_i \ln q_i = S(q_1, \dots, q_n) \leq S(p_1, \dots, p_n) = k(\beta \langle d \rangle + \ln Z) \quad (\text{A.6})$$

Es decir, cualquier otra distribución de probabilidad $\{q_1, \dots, q_n\}$ compatible con las ligaduras lleva a una entropía que es menor o igual que la calculada con la distribución $\{p_1, \dots, p_n\}$.

La distribución $\{p_1, \dots, p_n\}$ representa por tanto un máximo.

Apéndice B: Desarrollo de los conjuntos estadísticos

Este apéndice consiste en una revisión de los apuntes de la asignatura de Física Estadística (Baixauli y Cobos, 1996), que se redactaron originalmente como una traducción del libro de Lee (1988). Citando dichos apuntes (Baixauli y Cobos, 1996): “*La ecuación de la entropía como función de las probabilidades* [la entropía estadística]

$$S(p_1, \dots, p_n) = -k_B \sum_{i=1}^n p_i \ln p_i$$

*debe considerarse como un postulado físico fundamental, equivalente a la segunda ley de Newton o a la ecuación de Schrödinger, y es necesario introducirla para poder **modelizar los sistemas formados por muchas partículas**. Inferimos las probabilidades asociadas a los microestados con la única ayuda de la información macroscópica conocida; teniendo perfectamente claro que si conociéramos más ligaduras, el tipo de microestados accesibles compatibles con un macroestado, y sus probabilidades asociadas, serán en general distintos. De esta manera, si algún día cambiase la mecánica microscópica subyacente, el procedimiento de trabajo no cambiaría, y únicamente variará la forma de asignar las ligaduras.”*

En este apéndice desarrollaremos las expresiones de cuatro conjuntos estadísticos utilizando el Principio de Máxima Entropía: microcanónico, canónico, gran canónico e isotermo-isobárico. Recordamos del capítulo 6 que un **conjunto estadístico es el espacio muestral de todos los posibles microestados de un sistema sujeto a unas ciertas ligaduras macroscópicas constantes**. Según cómo sean estas ligaduras, tendremos un conjunto u otro.

La definición de los conjuntos microcanónico y canónico ya ha sido revisada por Hevia (2023) en los apuntes de *Fundamentos de Física Estadística*. En este apéndice completamos esa revisión incluyendo los conjuntos gran canónico e isotermo-isobárico. En lo que respecta a los siguientes apartados consideraremos siempre un sistema monofásico, monocomponente y que sólo puede intercambiar trabajo hidrostático asociado al cambio de volumen ($\delta W = -PdV$).

B.1. El conjunto microcanónico

Se denomina **conjunto microcanónico**, o conjunto (U, V, N) , al espacio muestral formado por los microestados de un **sistema aislado**, para el que los valores de la energía interna U , el volumen V y el número de entidades elementales N son estrictamente constantes. Los sistemas aislados están rodeados por paredes aislantes, las cuales impiden todo tipo de interacción con el entorno, y por lo tanto, son sistemas que no pueden intercambiar energía (en forma de calor o trabajo hidrostático asociado al cambio de volumen) ni materia.

- **Macroestado:** queda perfectamente determinado por los valores de las variables extensivas (U, V, N) , que son las variables naturales de la entropía $S = S(U, V, N)$, potencial termodinámico natural para describir un sistema aislado.
- **Microestados:** vienen determinados por los $\Omega(U, V, N)$ microestados compatibles con el macroestado (U, V, N) . El microestado i -ésimo tendrá una probabilidad asociada p_i . Estas probabilidades son siempre funciones de las variables del macroestado, $p_i = p_i(U, V, N)$.

B.1.1. Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, \Omega$) asociadas a cada microestado diferente, se emplea el Principio de Máxima Entropía sin considerar ninguna ligadura macroscópica adicional más allá de la condición de normalización de las probabilidades.

Se parte de la expresión de la entropía estadística:

$$S = -k_B \sum_{i=1}^{\Omega} p_i \ln p_i \quad (\text{B.1.1})$$

con la condición de normalización de las probabilidades:

$$\sum_{i=1}^{\Omega} p_i = 1 \quad (\text{B.1.2})$$

Aplicando el método de los multiplicadores de Lagrange, se construye una función de Lagrange con un parámetro λ_1 acompañando a la ligadura:

$$L(p_1, \dots, p_{\Omega}, \lambda_1) = S(p_1, \dots, p_{\Omega}) + \lambda_1 \left(\sum_{i=1}^{\Omega} p_i - 1 \right) = -k_B \sum_{i=1}^{\Omega} p_i \ln p_i + \lambda_1 \left(\sum_{i=1}^{\Omega} p_i - 1 \right) \quad (\text{B.1.3})$$

Se calcula la derivada parcial de la función L con respecto a p_i y se iguala el resultado a cero:

$$\left(\frac{\partial L}{\partial p_i} \right)_{p_j \neq p_i, \lambda_1} = 0 = -k_B (\ln p_i + 1) + \lambda_1 \quad (\text{B.1.4})$$

Despejando de la ecuación (B.1.4), se obtiene la siguiente expresión para las probabilidades:

$$p_i = e^{\left(\frac{\lambda_1 - 1}{k_B} \right)} = \text{constante} \quad (\text{B.1.5})$$

Como el valor de p_i es una constante, todos los microestados serán equiprobables. Llevando este resultado a la condición (B.1.2):

$$\sum_{i=1}^{\Omega} p_i = \sum_{i=1}^{\Omega} e^{\left(\frac{\lambda_1 - 1}{k_B} \right)} = \Omega \cdot e^{\left(\frac{\lambda_1 - 1}{k_B} \right)} = 1 \quad (\text{B.1.6})$$

Despejando el valor del multiplicador de Lagrange λ_1 de la ecuación anterior se obtiene:

$$\lambda_1 = k_B (1 - \ln \Omega) \quad (\text{B.1.7})$$

Finalmente, de las expresiones (B.1.5) y (B.1.7) se tiene:

$$p_i = e^{-\ln \Omega} = (e^{\ln \Omega})^{-1} = \frac{1}{\Omega} \quad \rightarrow \quad \boxed{p_i = \frac{1}{\Omega}} \quad (\text{B.1.8})$$

De esta forma, se ha deducido que todos los microestados son igualmente probables en el conjunto microcanónico, que es el conjunto natural que se utiliza para describir un sistema aislado.

Ahora, es posible calcular la entropía en el conjunto microcanónico a partir de la expresión (B.1.8) de las probabilidades:

$$S = -k_B \sum_{i=1}^{\Omega} p_i \left(\ln \frac{1}{\Omega} \right) = k_B \ln \Omega \sum_{i=1}^{\Omega} p_i = k_B \ln \Omega \quad \rightarrow \quad \boxed{\boxed{S = k_B \ln \Omega}} \quad (\text{B.1.9})$$

B.1.2. Ecuaciones de estado

Vamos a obtener ahora la descripción microscópica del sistema a partir de la ecuación (B.1.9). Del estudio de los potenciales termodinámicos conocemos que para un sistema aislado monofásico, monocomponente, y que intercambia únicamente trabajo hidrostático asociado al cambio de volumen ($\delta W = -PdV$) su potencial termodinámico natural es la entropía $S = S(U, N, V)$.

Como a cualquier potencial termodinámico, a la entropía se le pueden asociar tres ecuaciones: la ecuación de Euler, que nos da el potencial en función de sus variables naturales; la ecuación de Gibbs, que es la forma fundamental de la representación; y la ecuación de Gibbs-Duhem, que es la relación de ligadura entre los parámetros intensivos de la representación entrópica:

$$S = S(U, V, N) = \frac{U}{T} + \frac{PV}{T} - \frac{\mu N}{T} \quad \text{Ecuación de Euler} \quad (\text{B.1.10})$$

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN \quad \text{Ecuación de Gibbs} \quad (\text{B.1.11})$$

$$Ud\left(\frac{1}{T}\right) + Vd\left(\frac{P}{T}\right) - Nd\left(\frac{\mu}{T}\right) = 0 \quad \text{Ecuación de Gibbs-Duhem} \quad (\text{B.1.12})$$

Como la entropía es una función de estado, admite una diferencial exacta, es decir:

$$dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \quad (\text{B.1.13})$$

También es posible diferenciar la ecuación (B.1.9):

$$dS = k_B \cdot d \ln \Omega = k_B \left[\left(\frac{\partial \ln \Omega}{\partial U}\right)_{V,N} dU + \left(\frac{\partial \ln \Omega}{\partial V}\right)_{U,N} dV + \left(\frac{\partial \ln \Omega}{\partial N}\right)_{U,V} dN \right] \quad (\text{B.1.14})$$

Conocido $\ln \Omega(U, V, N)$, sus primeras derivadas nos proporcionan las ecuaciones de estado, y sus segundas derivadas, las funciones respuesta. Si se identifican términos entre la ecuación de Gibbs (B.1.11) y (B.1.14) considerando $\beta = 1/k_B T$, se llega a las ecuaciones de estado:

$$\text{Ecuación energética de estado (EEE)} \quad \rightarrow \quad \boxed{\frac{1}{k_B T} = \beta = \left(\frac{\partial \ln \Omega}{\partial U}\right)_{N,V}} \quad (\text{B.1.15})$$

$$\text{Ecuación térmica de estado (ETE)} \quad \rightarrow \quad \boxed{\frac{P}{k_B T} = P\beta = \left(\frac{\partial \ln \Omega}{\partial V}\right)_{U,N}} \quad (\text{B.1.16})$$

$$\text{Ecuación másica de estado (EME)} \quad \rightarrow \quad \boxed{-\frac{\mu}{k_B T} = -\mu\beta = \alpha = \left(\frac{\partial \ln \Omega}{\partial N}\right)_{U,V}} \quad (\text{B.1.17})$$

Podemos comprobar ahora cómo, en Física Estadística, β es la variable conjugada de U , $P\beta$ es la variable conjugada de V , y $\alpha = -\mu\beta$ es la variable conjugada de N (es decir, los productos $U\beta$, $P\beta V$ y αN son adimensionales). Recordemos también que, dados dos subsistemas que forman en conjunto un sistema total aislado, las condiciones que caracterizan el equilibrio macroscópico global alcanzado al cabo de cierta evolución, se escriben en la representación entrópica como:

$\beta_1 = \beta_2$	Equilibrio térmico ($T_1 = T_2$)
$P_1\beta_1 = P_2\beta_2$	Equilibrio mecánico ($P_1 = P_2$)
$-\mu_1\beta_1 = -\mu_2\beta_2$	Equilibrio másico ($\mu_1 = \mu_2$)

B.1.3. Funciones respuesta

Se estudian a continuación las segundas derivadas de la función estructura $\ln \Omega(U, V, N)$, que son las funciones respuesta:

$$\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} = k_B \left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right)_{V,N} = \left[\frac{\partial}{\partial U} \left(\frac{1}{T}\right)\right]_{V,N} = -\frac{1}{T^2} \left(\frac{\partial T}{\partial U}\right)_{V,N} = -\frac{1}{T^2 C_V} \quad (\text{B.1.18})$$

donde $C_V = (\partial U / \partial T)_{V,N}$ es la capacidad calorífica a volumen constante. De esta forma, se tiene

$$\boxed{\left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right)_{V,N} = -\frac{1}{k_B T^2 C_V} = -\frac{k_B \beta^2}{C_V}} \quad (\text{B.1.19})$$

También es posible obtener otras funciones respuesta calculando el resto de las segundas derivadas de $\ln \Omega(U, V, N)$:

$$\left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N} = k_B \left(\frac{\partial^2 \ln \Omega}{\partial V^2}\right)_{U,N} = \left[\frac{\partial}{\partial V} \left(\frac{P}{T}\right)\right]_{U,N} = \frac{1}{T^2} \left[T \left(\frac{\partial P}{\partial V}\right)_{U,N} - P \left(\frac{\partial T}{\partial V}\right)_{U,N} \right] \quad (\text{B.1.20})$$

$$\left(\frac{\partial^2 S}{\partial N^2}\right)_{U,V} = k_B \left(\frac{\partial^2 \ln \Omega}{\partial N^2}\right)_{U,V} = \left[\frac{\partial}{\partial N} \left(-\frac{\mu}{T}\right)\right]_{U,V} = -\frac{1}{T^2} \left[T \left(\frac{\partial \mu}{\partial N}\right)_{U,V} - \mu \left(\frac{\partial T}{\partial N}\right)_{U,V} \right] \quad (\text{B.1.21})$$

$$\left(\frac{\partial^2 S}{\partial V \partial U}\right)_N = k_B \left(\frac{\partial^2 \ln \Omega}{\partial V \partial U}\right)_N = \left[\frac{\partial}{\partial V} \left(\frac{1}{T}\right)\right]_{U,N} = \left[\frac{\partial}{\partial U} \left(\frac{P}{T}\right)\right]_{V,N} = \dots \quad (\text{B.1.22})$$

...

B.2. El conjunto canónico

Se denomina **conjunto canónico**, o conjunto $(1/T, V, N)$, al espacio muestral formado por los microestados de un **sistema cerrado isotermo e isócoro**, en contacto con un termostato que se encarga de **fijar su temperatura**. En un sistema cerrado isotermo e isócoro, la temperatura T , el volumen V y el número de entidades elementales N son estrictamente constantes, pero la energía interna U puede fluctuar, pues habrá distintos microestados con diferentes valores de la energía interna.

Dado que el sistema está en contacto con un termostato, puede intercambiar cantidades arbitrariamente grandes de energía con él, de modo que no tenemos ninguna base racional para descartar ningún valor de la energía entre su valor mínimo y máximo. En este caso se fija la temperatura, con lo que queda fijada también la energía media del sistema. De esta forma, el valor macroscópico de la energía interna es igual a su valor medio, es decir, se tiene $U = \langle E \rangle$.

- **Macroestado**: queda perfectamente determinado por las variables $(1/T, V, N)$, que son las variables naturales de la **función de Massieu** $\Psi = \Psi(1/T, V, N) = \Psi(\beta, V, N)$, potencial termodinámico natural para describir un sistema cerrado isotermo e isócoro.

- **Microestados:** Cada microestado está caracterizado por los valores que toman las variables extensivas del macroestado, en este caso V y N . Por ejemplo, los niveles de energía permitidos para un gas ideal formado por moléculas podrán depender del volumen del recipiente y del número de moléculas que lo componen, ya que son variables que tienen sentido a nivel microscópico, pero no tiene sentido decir que dependen de la temperatura, ya que dicha variable no tiene sentido a ese nivel. La temperatura jugará un papel en las probabilidades de que el sistema se encuentre en uno u otro nivel de energía, pero no caracteriza la energía de cada nivel. Así pues, cada microestado i -ésimo del sistema tendrá una cierta energía $E_i = E_i(V, N)$, y denotaremos al número de microestados como M en este caso. Las probabilidades, como siempre, serán funciones de todas las variables del macroestado: $p_i = p_i(1/T, V, N)$.

B.2.1. Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, M$) asociadas a cada uno de los M microestados, se emplea el Principio de Máxima Entropía considerando la condición de normalización y la ligadura correspondiente a la energía media. Para ello, se utiliza la definición de la entropía

$$S = -k_B \sum_{i=1}^M p_i \ln p_i \quad (\text{B.2.1})$$

sujeta a la condición de normalización de las probabilidades

$$\sum_{i=1}^M p_i = 1 \quad (\text{B.2.2})$$

y a la definición estadística de energía media:

$$\sum_{i=1}^M p_i E_i = \langle E \rangle \quad (\text{B.2.3})$$

En este caso se define una función de Lagrange incluyendo las ligaduras (B.2.2) y (B.2.3):

$$L(p_1, \dots, p_M, \lambda_1, \lambda_2) = S + \lambda_1 \left(\sum_{i=1}^M p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^M p_i E_i - \langle E \rangle \right) \quad (\text{B.2.4})$$

Se calcula la derivada parcial de la función L con respecto a p_i y se iguala el resultado a cero:

$$\left(\frac{\partial L}{\partial p_i} \right)_{p_j \neq p_i, \lambda_1, \lambda_2} = 0 = -k_B (\ln p_i + 1) + \lambda_1 + \lambda_2 E_i \quad (\text{B.2.5})$$

Despejando en la ecuación anterior, se llega a la siguiente expresión para las probabilidades:

$$p_i = \exp \left(\frac{\lambda_1}{k_B} + \frac{\lambda_2 E_i}{k_B} - 1 \right) \quad (\text{B.2.6})$$

Definiendo las siguientes constantes:

$$Z^{-1} = \exp \left(\frac{\lambda_1}{k_B} - 1 \right) \quad \beta = -\frac{\lambda_2}{k_B} \quad (\text{B.2.7})$$

la expresión de las probabilidades queda de la siguiente forma:

$$p_i = \frac{1}{Z} e^{-\beta E_i} \quad (\text{B.2.8})$$

Esta distribución de probabilidad se conoce como distribución de Boltzmann, y es la que maximiza la entropía del conjunto canónico sujeta a las ligaduras macroscópicas (B.2.2) y (B.2.3). En Termodinámica, se sabe que cuando la entropía de un sistema aislado es máxima, el sistema está en un estado de equilibrio. Si el sistema ya no es aislado, como es este caso, entonces la entropía asociada al estado de equilibrio ya no tiene por qué ser máxima. Lo que sí será máximo es el potencial entrópico que corresponda a las variables macroscópicas utilizadas para describir el sistema, que en este caso será la función de Massieu: el potencial termodinámico natural para las variables $(1/T, V, N)$.

La función de Massieu $\Psi = \Psi(1/T, V, N)$ es la transformada de Legendre de la entropía $S = S(U, V, N)$ con respecto a la energía interna.

Ahora bien, maximizar la función de Massieu sin ninguna ligadura adicional, operando de la misma forma que se hizo en el conjunto microcanónico, es matemáticamente equivalente a maximizar la entropía estadística sujeta a las ligaduras macroscópicas (B.2.2) y (B.2.3), que son las correspondientes a un sistema cerrado isotermo e isócoro.

La función estructura en el conjunto canónico es la variable que hemos definido anteriormente como Z , y se denomina función de partición. Para calcular su valor, se lleva el resultado (B.2.8) a (B.2.2):

$$\sum_{i=1}^M p_i = 1 = \frac{1}{Z} \sum_{i=1}^M e^{-\beta E_i} \quad \rightarrow \quad Z = \sum_{i=1}^M e^{-\beta E_i} \quad (\text{B.2.9})$$

La ecuación (B.2.9) expresa la función de partición Z como una suma a los M diferentes microestados, cada uno de ellos con una energía E_i . Pero es posible que varios microestados tengan el mismo valor de energía E_i , y por lo tanto, estarán degenerados. De esta forma, se tendrán L niveles de energía, cada uno de ellos con un valor diferente y único de energía E_j y con una degeneración $\Omega_j \equiv \Omega(E_j, V, N)$, que es el número de microestados diferentes para una cierta energía E_j . Es evidente que la suma de todos los $\Omega(E_j, V, N)$ microestados diferentes para cada energía E_j es el número total de microestados M :

$$\sum_{j=1}^L \Omega(E_j, V, N) = M$$

Así, los valores de las energías E_i ($i = 1, \dots, M$) de los M microestados se pueden repetir, mientras que las energías E_j ($j = 1, \dots, L$) de los L niveles son todas diferentes entre sí, siendo $\Omega(E_j, V, N)$ la degeneración de cada nivel.

Es posible expresar la función de partición como una suma en los distintos niveles de energía, teniendo en cuenta la degeneración de cada nivel. De esta forma, se reúnen las sumas de aquellos microestados con la misma energía, reduciendo el número de términos y simplificando el cálculo. Así, la función de partición es:

$$Z(\beta, V, N) = \sum_{\substack{i=1 \\ \text{Microestados}}}^M e^{-\beta E_i(N, V)} = \sum_{\substack{j=1 \\ \text{Niveles}}}^L \Omega(E_j, V, N) e^{-\beta E_j(N, V)} \quad (\text{B.2.10})$$

Para que la idea quede más clara podemos pensar en el siguiente ejemplo: un cierto sistema cerrado isotermo e isócoro tiene tres posibles microestados ($M = 3$), cuyas energías son ε , 2ε y 2ε . Observamos que hay dos microestados que tienen el mismo valor de la energía, es decir, están degenerados. Se tienen por tanto dos niveles ($L = 2$) de energías ε y 2ε . La degeneración

para el nivel inferior (ε) es la unidad, pues sólo hay un microestado posible con esa energía, mientras que la degeneración del nivel superior (2ε) es 2, pues se tienen dos microestados con esa energía. Calculamos la función de partición de las dos formas:

$$Z = \sum_{\substack{i=1 \\ \text{Microestados}}}^{M=3} e^{-\beta E_i} = e^{-\beta \cdot \varepsilon} + e^{-\beta \cdot 2\varepsilon} + e^{-\beta \cdot 2\varepsilon}$$

$$Z = \sum_{\substack{j=1 \\ \text{Niveles}}}^{L=2} \Omega(E_j, V, N) e^{-\beta E_j} = e^{-\beta \cdot \varepsilon} + 2 \cdot e^{-\beta \cdot 2\varepsilon}$$

Con este sencillo ejemplo hemos visto que la suma a microestados y la suma a niveles de energía coinciden. Además, en la ecuación (B.2.10) vemos que **la función estructura del conjunto canónico** $Z(\beta, V, N)$ se puede obtener como una **transformada discreta de Laplace de la función estructura en el microcanónico** $\Omega(U, V, N)$ respecto de la energía:

$$Z(\beta, V, N) = \sum_{\substack{j \\ \text{Niveles}}} \Omega(E_j, V, N) e^{-\beta E_j(N, V)}$$

Una vez conocidas las probabilidades (B.2.8) y la función de partición (B.2.10), vamos a ver cómo es la entropía en el conjunto canónico. Para ello, se toma el logaritmo natural en la expresión de las probabilidades (B.2.8):

$$\ln p_i = \ln \left(\frac{e^{-\beta E_i}}{Z} \right) = -\beta E_i - \ln Z$$

y sustituimos en la ecuación de la entropía (B.2.1):

$$S = -k_B \sum_{i=1}^M p_i (-\beta E_i - \ln Z) = k_B \left(\ln Z \sum_{i=1}^M p_i + \beta \sum_{i=1}^M p_i E_i \right) = k_B \ln Z + k_B \beta \langle E \rangle \quad (\text{B.2.11})$$

Por otro lado, del análisis de los potenciales termodinámicos en la representación entrópica, sabemos que **la función de Massieu es la transformada de Legendre de la entropía respecto de la energía interna**, es decir:

$$\Psi(1/T, V, N) = S(U, V, N) - U \left(\frac{\partial S}{\partial U} \right)_{V, N} = S - \frac{U}{T} \quad (\text{B.2.12})$$

La entropía se puede escribir en términos de la función de Massieu despejando en la ecuación (B.2.12), de forma que:

$$S = \Psi + \frac{U}{T} = \Psi + \frac{\langle E \rangle}{T} \quad (\text{B.2.13})$$

Comparando los sumandos de la ecuación (B.2.13) con $S = k_B \ln Z + k_B \beta \langle E \rangle$ se tiene que:

$$\beta = \frac{1}{k_B T} \quad (\text{B.2.14})$$

$$\boxed{\Psi = k_B \ln Z} \quad (\text{B.2.15})$$

El resultado (B.2.15) es similar a la expresión $S = k_B \ln \Omega$ obtenida en el conjunto microcanónico. Volvemos a encontrarnos en el desarrollo una expresión que relaciona una función de estado macroscópica (potencial termodinámico entrópico), en este caso Ψ , con una microscópica (función estructura), en este caso Z (Tabla B.1).

<u>Conjunto</u>	<u>Función estructura</u>	<u>Potencial entrópico</u>
Microcanónico	Número de microestados: $\Omega(U, V, N)$	Entropía: $S = k_B \ln \Omega$
Canónico	Función de partición: $Z(\beta, V, N)$	Función de Massieu: $\Psi = k_B \ln Z$

Tabla B.1. Comparación entre los conjuntos microcanónico y canónico.

B.2.2. Ecuaciones de estado

Escribimos a continuación las ecuaciones de Euler (potencial en función de sus variables naturales), Gibbs (forma fundamental) y Gibbs-Duhem (relación de ligadura entre los parámetros intensivos) correspondientes a la función de Massieu:

$$\Psi = \Psi\left(\frac{1}{T}, V, N\right) = -\frac{F}{T} = \frac{PV}{T} - \frac{\mu N}{T} \quad \text{\underline{Ecuación de Euler}} \quad (\text{B.2.16})$$

$$d\Psi = d\left(-\frac{F}{T}\right) = -Ud\left(\frac{1}{T}\right) + \frac{P}{T}dV - \frac{\mu}{T}dN \quad \text{\underline{Ecuación de Gibbs}} \quad (\text{B.2.17})$$

$$Ud\left(\frac{1}{T}\right) + Vd\left(\frac{P}{T}\right) - Nd\left(\frac{\mu}{T}\right) = 0 \quad \text{\underline{Ecuación de Gibbs-Duhem}} \quad (\text{B.2.18})$$

Como la función de Massieu es una función de estado, admite una diferencial exacta, es decir:

$$d\Psi = \left(\frac{\partial\Psi}{\partial(1/T)}\right)_{V,N} d\left(\frac{1}{T}\right) + \left(\frac{\partial\Psi}{\partial V}\right)_{1/T,N} dV + \left(\frac{\partial\Psi}{\partial N}\right)_{1/T,V} dN \quad (\text{B.2.19})$$

También es posible diferenciar la ecuación (B.2.15):

$$d\Psi = k_B \cdot d \ln Z = k_B \left[\left(\frac{\partial \ln Z}{\partial(1/T)}\right)_{V,N} d\left(\frac{1}{T}\right) + \left(\frac{\partial \ln Z}{\partial V}\right)_{1/T,N} dV + \left(\frac{\partial \ln Z}{\partial N}\right)_{1/T,V} dN \right] \quad (\text{B.2.20})$$

Dado que la variable β es proporcional a $1/T$, podemos utilizar β como variable independiente. De esta forma, las ecuaciones de estado y las funciones respuesta se pueden escribir de forma más cómoda en términos de β :

$$\Psi(1/T, V, N) \equiv \Psi(\beta, V, N)$$

La diferencial exacta de la función de Massieu en términos de las variables (β, V, N) es:

$$d\Psi = \left(\frac{\partial\Psi}{\partial\beta}\right)_{V,N} d\beta + \left(\frac{\partial\Psi}{\partial V}\right)_{\beta,N} dV + \left(\frac{\partial\Psi}{\partial N}\right)_{\beta,V} dN \quad (\text{B.2.21})$$

$$d\Psi = k_B \cdot d \ln Z = k_B \left[\left(\frac{\partial \ln Z}{\partial\beta}\right)_{V,N} d\beta + \left(\frac{\partial \ln Z}{\partial V}\right)_{\beta,N} dV + \left(\frac{\partial \ln Z}{\partial N}\right)_{\beta,V} dN \right] \quad (\text{B.2.22})$$

Reescribimos ahora la ecuación de Gibbs (B.2.17) en función de β sabiendo que $d(1/T) = k_B d\beta$ e igualamos el resultado a la ecuación (B.2.22):

$$\begin{aligned}
d\Psi &= -Ud\left(\frac{1}{T}\right) + \frac{P}{T}dV - \frac{\mu}{T}dN \\
&= k_B[-U d\beta + P\beta dV - \mu\beta dN] \\
&= k_B\left[\left(\frac{\partial \ln Z}{\partial \beta}\right)_{V,N} d\beta + \left(\frac{\partial \ln Z}{\partial V}\right)_{\beta,N} dV + \left(\frac{\partial \ln Z}{\partial N}\right)_{\beta,V} dN\right]
\end{aligned} \tag{B.2.23}$$

Conocido $\ln Z(\beta, V, N)$, sus primeras derivadas respecto a sus variables naturales nos proporcionan las ecuaciones de estado. Comparando miembro a miembro las dos últimas líneas de la ecuación (B.2.23) se obtienen las ecuaciones de estado:

$$\text{Ecuación energética de estado (EEE)} \rightarrow \boxed{\langle E \rangle = U = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_{N,V}} \tag{B.2.24}$$

$$\text{Ecuación térmica de estado (ETE)} \rightarrow \boxed{\frac{P}{k_B T} = P\beta = \left(\frac{\partial \ln Z}{\partial V}\right)_{\beta,N}} \tag{B.2.25}$$

$$\text{Ecuación másica de estado (EME)} \rightarrow \boxed{-\frac{\mu}{k_B T} = -\mu\beta = \left(\frac{\partial \ln Z}{\partial N}\right)_{\beta,V}} \tag{B.2.26}$$

B.2.3. Funciones respuesta y fluctuaciones de energía

A continuación se analizarán las segundas derivadas de la función de Massieu Ψ respecto a sus variables naturales. Antes de proceder al cálculo de las funciones respuesta, recordamos la expresión de la capacidad calorífica a volumen constante:

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,N} = \left(\frac{\partial \langle E \rangle}{\partial T}\right)_{V,N} \tag{B.2.27}$$

■ **Fluctuaciones de energía.** Calculamos la segunda derivada parcial de Ψ respecto a β :

$$\begin{aligned}
\left(\frac{\partial^2 \Psi}{\partial \beta^2}\right)_{V,N} &= k_B \left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = k_B \left(\frac{\partial(-\langle E \rangle)}{\partial \beta}\right)_{V,N} \\
&= -k_B \left(\frac{\partial \langle E \rangle}{\partial T}\right) \left(\frac{\partial T}{\partial \beta}\right) = -k_B C_V \left(-\frac{1}{k_B \beta^2}\right) = \frac{C_V}{\beta^2}
\end{aligned} \tag{B.2.28}$$

De esta forma, se tiene que

$$\boxed{\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = \frac{C_V}{k_B \beta^2} = k_B T^2 C_V} \tag{B.2.29}$$

También es posible calcular esta segunda derivada parcial si recurrimos a la definición de la función de partición en la expresión (B.2.9):

$$\begin{aligned}
\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} &= \frac{\partial}{\partial \beta} \left(\frac{\partial \ln Z}{\partial \beta}\right) = \frac{\partial}{\partial \beta} \left(-\frac{1}{Z} \sum_i E_i e^{-\beta E_i}\right) \\
&= \frac{1}{Z} \sum_i E_i^2 e^{-\beta E_i} + \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta}\right) \sum_i E_i e^{-\beta E_i} \\
&= \sum_i p_i E_i^2 - \frac{1}{Z^2} \left(\sum_i E_i e^{-\beta E_i}\right)^2 \\
&= \sum_i p_i E_i^2 - \left(\sum_i p_i E_i\right)^2 = \langle E^2 \rangle - \langle E \rangle^2
\end{aligned} \tag{B.2.30}$$

Ahora bien, es sencillo demostrar que $\langle E^2 \rangle - \langle E \rangle^2 = \langle (E_i - \langle E \rangle)^2 \rangle$. Para ello, es suficiente con operar en la definición habitual del valor medio:

$$\begin{aligned}
\langle (E_i - \langle E \rangle)^2 \rangle &= \sum_i p_i (E_i - \langle E \rangle)^2 = \sum_i p_i (E_i^2 + \langle E \rangle^2 - 2E_i \langle E \rangle) \\
&= \sum_i p_i E_i^2 + \sum_i p_i \langle E \rangle^2 - 2 \sum_i p_i E_i \langle E \rangle \\
&= \sum_i p_i E_i^2 + \langle E \rangle^2 \sum_i p_i - 2 \langle E \rangle \sum_i p_i E_i \\
&= \langle E^2 \rangle + \langle E \rangle^2 - 2 \langle E \rangle \langle E \rangle = \langle E^2 \rangle - \langle E \rangle^2
\end{aligned} \tag{B.2.31}$$

El resultado obtenido nos permite relacionar la derivada segunda del logaritmo natural de la función de partición Z con la varianza de las fluctuaciones de la energía interna en torno a su valor medio. La varianza es una magnitud que siempre es positiva por su propia definición: es el valor medio de los cuadrados de las desviaciones de una variable aleatoria respecto a su valor medio. De esta forma, se puede escribir que

$$\boxed{(\Delta^* E)^2 = \left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = \langle E^2 \rangle - \langle E \rangle^2 = k_B T^2 C_V \geq 0} \tag{B.2.32}$$

La capacidad calorífica a volumen constante, C_V , da cuenta de las fluctuaciones estadísticas de la energía interna respecto a su valor medio.

■ Segunda derivada respecto del volumen:

Operando de la misma forma, se pueden obtener otras segundas derivadas. En este caso, calculamos la segunda derivada de Ψ respecto al volumen V :

$$\left(\frac{\partial^2 \Psi}{\partial V^2}\right)_{\beta,N} = k_B \left(\frac{\partial^2 \ln Z}{\partial V^2}\right)_{\beta,N} = k_B \left(\frac{\partial(P\beta)}{\partial V}\right)_{\beta,N} = k_B \beta \left(\frac{\partial P}{\partial V}\right)_{\beta,N} = -\frac{k_B \beta}{V \kappa_T} \tag{B.2.33}$$

donde se ha tenido en cuenta la definición del coeficiente de compresibilidad isotérmico:

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_{T,N} \tag{B.2.34}$$

De esta forma, se tiene que:

$$\boxed{\left(\frac{\partial^2 \ln Z}{\partial V^2}\right)_{\beta,N} = -\frac{\beta}{V \kappa_T} = -\frac{1}{k_B T V \kappa_T}} \tag{B.2.35}$$

B.3. El conjunto gran canónico

Se denomina **conjunto gran canónico**, o conjunto $(1/T, V, \mu/T)$, al espacio muestral formado por los microestados de un **sistema abierto isotermo e isócoro**, en contacto con un termostato que se encarga de **fijar su temperatura**, y que puede intercambiar materia con una reserva infinita de partículas, que se encarga de **fijar el potencial químico**. En un sistema abierto isotermo e isócoro, la temperatura T , el volumen V y el potencial químico μ son estrictamente constantes, pero la energía interna U y el número de partículas N pueden fluctuar, pues existirán distintos microestados con diferentes valores de la energía interna y del número de partículas.

Dado que el sistema está en contacto con un termostato, puede intercambiar cantidades arbitrariamente grandes de energía con él, de modo que no tenemos ninguna base racional para descartar ningún valor de la energía entre su valor mínimo y máximo. De la misma forma, al estar en contacto con una reserva infinita de partículas, el número de partículas variará en general entre cero y un valor máximo N_{\max} , de forma que no se puede descartar ningún valor del número de partículas N .

Al igual que en el conjunto canónico, se fija la temperatura, con lo que **queda fijada también la energía media del sistema**. De esta forma, el valor macroscópico de la energía interna es igual a su valor medio, es decir, se tiene $U = \langle E \rangle$.

En el conjunto gran canónico se fija también el potencial químico, con lo que **queda fijado el número de partículas medio del sistema**. Así, el valor macroscópico del número de partículas, N_{Macro} , es igual a su valor medio: $N_{\text{Macro}} = \langle N \rangle$.

- **Macroestado:** queda perfectamente determinado por las variables $(1/T, V, \mu/T)$, que son las variables naturales de la **función de Kramers-Landau** (también llamada potencial gran canónico) $\eta = \eta(1/T, V, \mu/T) = \eta(\beta, V, \alpha)$, potencial termodinámico natural para describir un sistema abierto isotermo e isócoro.
- **Microestados:** Cada uno de los M microestados del sistema tendrá una cierta energía $E_i = E_i(V, N_i)$ y un número de partículas N_i . Las probabilidades, como siempre, serán funciones de todas las variables del macroestado: $p_i = p_i(1/T, V, \mu/T)$.

B.3.1. Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, M$) asociadas a cada uno de los M microestados, se emplea el Principio de Máxima Entropía considerando la definición de la entropía para una distribución de probabilidad:

$$S = -k_B \sum_{i=1}^M p_i \ln p_i \quad (\text{B.3.1})$$

sujeta a las siguientes ligaduras:

$$\sum_{i=1}^M p_i = 1 \quad (\text{B.3.2})$$

$$\sum_{i=1}^M p_i E_i = \langle E \rangle \quad (\text{B.3.3})$$

$$\sum_{i=1}^M p_i N_i = \langle N \rangle \quad (\text{B.3.4})$$

Se define una función de Lagrange incluyendo las ligaduras anteriores:

$$L(p_1, \dots, \lambda_1, \lambda_2, \lambda_3) = S + \lambda_1 \left(\sum_{i=1}^M p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^M p_i E_i - \langle E \rangle \right) + \lambda_3 \left(\sum_{i=1}^M p_i N_i - \langle N \rangle \right) \quad (\text{B.3.5})$$

Se calcula la derivada parcial de la función L con respecto a p_i y se iguala el resultado a cero:

$$\left(\frac{\partial L}{\partial p_i} \right)_{p_j \neq p_i, \lambda_k} = 0 = -k_B (\ln p_i + 1) + \lambda_1 + \lambda_2 E_i + \lambda_3 N_i \quad (\text{B.3.6})$$

Despejando en la ecuación anterior, se llega a la siguiente expresión para las probabilidades:

$$p_i = \exp \left(\frac{\lambda_1}{k_B} + \frac{\lambda_2 E_i}{k_B} + \frac{\lambda_3 N_i}{k_B} - 1 \right) \quad (\text{B.3.7})$$

Definiendo las siguientes constantes:

$$\Xi^{-1} = \exp \left(\frac{\lambda_1}{k_B} - 1 \right) \quad \beta = -\frac{\lambda_2}{k_B} \quad \alpha = -\frac{\lambda_3}{k_B} \quad (\text{B.3.8})$$

la expresión de las probabilidades queda de la siguiente forma:

$$p_i = \frac{1}{\Xi} e^{-\beta E_i - \alpha N_i} \quad (\text{B.3.9})$$

Sustituyendo la expresión (B.3.9) de las probabilidades en la ligadura (B.3.2) obtenemos el valor de la **gran función de partición**:

$$\Xi = \sum_{i=1}^M e^{-\beta E_i - \alpha N_i} \quad (\text{B.3.10})$$

El resultado (B.4.10) nos indica que **la función estructura del conjunto gran canónico, $\Xi(\beta, V, \alpha)$ se puede obtener como una transformada discreta de Laplace de la función estructura en el conjunto canónico $Z(\beta, V, N)$ respecto del número de partículas**:

$$\Xi(\beta, V, \alpha) = \sum_i e^{-\beta E_i(V, N_i) - \alpha N_i} = \sum_i e^{-\alpha N_i} \left(\sum_j e^{-\beta E_i(V, N_i)} \right)_{\substack{\text{estados de} \\ n^{\circ} \text{ de partículas fijo}}} = \sum_i Z(\beta, V, N_i) e^{-\alpha N_i}$$

Ahora, para calcular cómo es la entropía en el conjunto gran canónico, vamos a tomar el logaritmo natural de las probabilidades (B.3.9)

$$\ln p_i = -\beta E_i - \alpha N_i - \ln \Xi$$

y vamos a sustituir la expresión obtenida en la entropía estadística (B.3.1):

$$S = -k_B \sum_i p_i (-\beta E_i - \alpha N_i - \ln \Xi) = k_B (\ln Y + \beta \langle E \rangle + \alpha \langle N \rangle) \quad (\text{B.3.11})$$

Por otro lado, del análisis de los potenciales termodinámicos en la representación entrópica, sabemos que **la función de Kramers-Landau es la transformada de Legendre de la función de Massieu respecto del número de partículas**, o bien, es **la transformada de Legendre de la entropía respecto de la energía interna y del número de partículas** (hacemos dos transformadas de Legendre a la entropía) es decir:

$$\eta(1/T, V, \mu/T) = S(U, V, N) - U \left(\frac{\partial S}{\partial U} \right)_{V,N} - N \left(\frac{\partial S}{\partial N} \right)_{U,V} = S - \frac{\langle E \rangle}{T} + \frac{\mu \langle N \rangle}{T} \quad (\text{B.3.12})$$

La entropía se puede escribir en términos de la función de Kramers-Landau despejando en la ecuación (B.3.12), de forma que:

$$S = \eta + \frac{\langle E \rangle}{T} - \frac{\mu \langle N \rangle}{T} \quad (\text{B.3.13})$$

Comparando los sumandos de (B.3.13) con $S = k_B \ln Y + k_B \beta \langle E \rangle + k_B \alpha \langle N \rangle$ se tiene que:

$$\beta = \frac{1}{k_B T} \quad (\text{B.3.14})$$

$$\alpha = -\frac{\mu}{k_B T} = -\mu\beta \quad (\text{B.3.15})$$

$$\boxed{\eta = k_B \ln \Xi} \quad (\text{B.3.16})$$

B.3.2. Ecuaciones de estado

Recordemos las ecuaciones que describen la función de Kramers-Landau, potencial termodinámico entrópico natural para un conjunto gran canónico:

$$\eta = \eta\left(\frac{1}{T}, V, \frac{\mu}{T}\right) = \frac{PV}{T} \quad \text{Ecuación de Euler} \quad (\text{B.3.17})$$

$$d\eta = -Ud\left(\frac{1}{T}\right) + \frac{P}{T}dV + Nd\left(\frac{\mu}{T}\right) \quad \text{Ecuación de Gibbs} \quad (\text{B.3.18})$$

$$Ud\left(\frac{1}{T}\right) + Vd\left(\frac{P}{T}\right) - Nd\left(\frac{\mu}{T}\right) = 0 \quad \text{Ecuación de Gibbs-Duhem} \quad (\text{B.3.19})$$

Si diferenciamos la ecuación (B.3.16) llegamos a la siguiente expresión:

$$d\eta = k_B d \ln \Xi = k_B \left[\left(\frac{\partial \ln \Xi}{\partial (1/T)} \right)_{V, \frac{\mu}{T}} d\left(\frac{1}{T}\right) + \left(\frac{\partial \ln \Xi}{\partial V} \right)_{\frac{1}{T}, \frac{\mu}{T}} dV + \left(\frac{\partial \ln \Xi}{\partial (\mu/T)} \right)_{\frac{1}{T}, V} d\left(\frac{\mu}{T}\right) \right] \quad (\text{B.3.20})$$

Comparando ahora la ecuación de Gibbs (B.3.18) de la representación con (B.3.20), obtenemos las correspondientes **ecuaciones de estado**:

$$\text{Ecuación energética de estado (EEE)} \rightarrow \boxed{\langle E \rangle = U = - \left(\frac{\partial \ln \Xi}{\partial \beta} \right)_{V, \alpha}} \quad (\text{B.3.21})$$

$$\text{Ecuación térmica de estado (ETE)} \rightarrow \boxed{\frac{P}{k_B T} = P\beta = \left(\frac{\partial \ln \Xi}{\partial V} \right)_{\beta, \alpha}} \quad (\text{B.3.22})$$

$$\text{Ecuación másica de estado (EME)} \rightarrow \boxed{N = N_{Macro} = \langle N \rangle = - \left(\frac{\partial \ln \Xi}{\partial \alpha} \right)_{\beta, V}} \quad (\text{B.3.23})$$

B.3.3. Funciones respuesta y fluctuaciones

Veremos brevemente un análisis de las segundas derivadas de la función de Planck del conjunto isoterma-isobárico:

- **Fluctuaciones de la energía:** segunda derivada respecto de $\beta = 1/k_B T$.

$$\left(\frac{\partial^2 \ln \Xi}{\partial \beta^2} \right)_{V,\alpha} = - \left(\frac{\partial \langle E \rangle}{\partial \beta} \right)_{V,\alpha} = (\Delta^* E)^2 = \langle (E_i - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \dots \quad (\text{B.3.24})$$

No vamos a calcular la expresión explícita de esta expresión, pues su desarrollo es bastante complejo. En este caso no es posible establecer una conexión inmediata con la capacidad calorífica a volumen constante, como ocurría en el conjunto canónico. Recordemos para ello la definición de C_V :

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{V,N} \neq \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{V, \alpha = -\mu/kT}$$

Ejemplos como este ilustran la importancia que tienen las variables que permanecen constantes en el proceso de derivación.

- **Fluctuaciones en el número de partículas:** segunda derivada respecto de $\alpha = -\mu/k_B T$.

$$\left(\frac{\partial^2 \ln \Xi}{\partial \alpha^2} \right)_{\beta,V} = - \left(\frac{\partial \langle N \rangle}{\partial \alpha} \right)_{\beta,V} = (\Delta^* N)^2 = \langle (N_i - \langle N \rangle)^2 \rangle = \dots = k_B T \langle N \rangle \rho \kappa_T \quad (\text{B.3.25})$$

Las fluctuaciones relativas son:

$$\frac{\Delta^* N}{\langle N \rangle} = \left(\frac{k_B T \rho \kappa_T}{\langle N \rangle} \right)^{1/2} \quad (\text{B.3.26})$$

De forma general, se puede demostrar que las fluctuaciones relativas son inversamente proporcionales a \sqrt{N} :

$$\frac{\Delta^* N}{\langle N \rangle} \propto \frac{1}{\sqrt{N}} \quad (\text{B.3.27})$$

B.4. El conjunto isoterma-isobárico

Se define el **conjunto isoterma-isobárico**, o conjunto $(1/T, P/T, N)$, al espacio muestral formado por los microestados de un **sistema cerrado isoterma e isobárico**, en contacto con un termostato que se encarga de **fijar su temperatura**, y en contacto con un pistón móvil que permite variar el volumen del sistema y que se encarga de **fijar la presión**. En un sistema cerrado isoterma e isobárico, la temperatura T , la presión P y el número de partículas N son estrictamente constantes, pero la energía interna U y el volumen V pueden fluctuar, pues existirán distintos microestados con diferentes valores de la energía interna y del volumen.

Dado que el sistema está en contacto con un termostato, puede intercambiar cantidades arbitrariamente grandes de energía con él, de modo que no tenemos ninguna base racional para descartar ningún valor de la energía entre su valor mínimo y máximo. De la misma forma, al

estar en contacto con un pistón móvil, el volumen variará en general entre un mínimo y un máximo, de forma que no se puede descartar ningún valor del volumen V .

Al igual que en el conjunto canónico, se fija la temperatura, con lo que **queda fijada también la energía media** del sistema. De esta forma, el valor macroscópico de la energía interna es igual a su valor medio, es decir, se tiene $U = \langle E \rangle$.

En el conjunto isoterma-isobárico también se fija la presión, con lo que **queda fijado el volumen medio** del sistema. Así, el valor macroscópico del volumen, V_{Macro} , es igual a su valor medio: $V_{Macro} = \langle V \rangle$.

- **Macroestado:** queda perfectamente determinado por las variables $(1/T, P/T, N)$, que son las variables naturales de la **función de Planck** $\phi = \phi(1/T, P/T, N) = \phi(\beta, \gamma, N)$, potencial termodinámico natural para describir un sistema cerrado isoterma e isobárico.
- **Microestados:** Cada uno de los M microestados del sistema tendrá una cierta energía $E_i = E_i(V_i, N)$ y un volumen V_i . Las probabilidades serán funciones de todas las variables del macroestado: $p_i = p_i(1/T, P/T, N)$.

B.4.1. Inferencia de probabilidades

Para inferir las probabilidades p_i ($i = 1, \dots, M$) asociadas a cada uno de los M microestados, se emplea el Principio de Máxima Entropía considerando la definición de la entropía estadística:

$$S = -k_B \sum_{i=1}^M p_i \ln p_i \quad (\text{B.4.1})$$

sujeta a las siguientes ligaduras:

$$\sum_{i=1}^M p_i = 1 \quad (\text{B.4.2})$$

$$\sum_{i=1}^M p_i E_i = \langle E \rangle \quad (\text{B.4.3})$$

$$\sum_{i=1}^M p_i V_i = \langle V \rangle \quad (\text{B.4.4})$$

Las dos últimas ligaduras son las correspondientes a la energía media y el volumen medio característicos del conjunto. En la ligadura (B.4.4) consideramos que el volumen es discreto.

Definimos una función de Lagrange incluyendo las ligaduras anteriores:

$$L(p_1, \dots, \lambda_1, \lambda_2, \lambda_3) = S + \lambda_1 \left(\sum_{i=1}^M p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^M p_i E_i - \langle E \rangle \right) + \lambda_3 \left(\sum_{i=1}^M p_i V_i - \langle V \rangle \right) \quad (\text{B.4.5})$$

Calculamos la derivada parcial de L con respecto a p_i e igualamos el resultado a cero:

$$\left(\frac{\partial L}{\partial p_i} \right)_{p_j \neq p_i, \lambda_k} = 0 = -k_B (\ln p_i + 1) + \lambda_1 + \lambda_2 E_i + \lambda_3 V_i \quad (\text{B.4.6})$$

Despejando en la ecuación anterior, llegamos a la expresión de las probabilidades:

$$p_i = \exp \left(\frac{\lambda_1}{k_B} + \frac{\lambda_2 E_i}{k_B} + \frac{\lambda_3 V_i}{k_B} - 1 \right) \quad (\text{B.4.7})$$

Definiendo las siguientes constantes:

$$Y^{-1} = \exp\left(\frac{\lambda_1}{k_B} - 1\right) \quad \beta = -\frac{\lambda_2}{k_B} \quad \gamma = -\frac{\lambda_3}{k_B} \quad (\text{B.4.8})$$

la expresión de las probabilidades queda de la siguiente forma:

$$p_i = \frac{1}{Y} e^{-\beta E_i - \gamma V_i} \quad (\text{B.4.9})$$

Sustituyendo la expresión (B.4.9) de las probabilidades en la ligadura (B.4.2) obtenemos el valor de la **función de partición isoterma-isobárica**:

$$Y = \sum_{i=1}^M e^{-\beta E_i - \gamma V_i} \quad (\text{B.4.10})$$

El resultado (B.4.10) nos indica que **la función estructura del conjunto isoterma-isobárico** $Y(\beta, \gamma, N)$ **se puede obtener como una transformada discreta de Laplace de la función estructura en el conjunto canónico** $Z(\beta, V, N)$ **respecto de la energía**:

$$Y(\beta, \gamma, N) = \sum_i e^{-\beta E_i(V_i, N) - \gamma V_i} = \sum_i e^{-\gamma V_i} \left(\sum_j e^{-\beta E_i(V_i, N)} \right)_{\text{estados de volumen fijo}} = \sum_i Z(\beta, V_i, N) e^{-\gamma V_i}$$

Ahora, para calcular cómo es la entropía en el conjunto isoterma-isobárico, vamos a tomar el logaritmo natural de las probabilidades (B.4.9)

$$\ln p_i = -\beta E_i - \gamma V_i - \ln Y$$

y vamos a sustituir la expresión obtenida en la entropía estadística (B.4.1):

$$S = -k_B \sum_i p_i (-\beta E_i - \gamma V_i - \ln Y) = k_B (\ln Y + \beta \langle E \rangle + \gamma \langle V \rangle) \quad (\text{B.4.11})$$

Por otro lado, del análisis de los potenciales termodinámicos en la representación entrópica, sabemos que **la función de Planck es la transformada de Legendre de la función de Massieu respecto del volumen**, o bien, es **la transformada de Legendre de la entropía respecto de la energía interna y del volumen** (hacemos dos transformadas de Legendre a la entropía) es decir:

$$\phi(1/T, P/T, N) = S(U, V, N) - U \left(\frac{\partial S}{\partial U} \right)_{V, N} - V \left(\frac{\partial S}{\partial V} \right)_{U, N} = S - \frac{\langle E \rangle}{T} - \frac{P \langle V \rangle}{T} \quad (\text{B.4.12})$$

La entropía se puede escribir en términos de la función de Planck despejando en la ecuación (B.4.12), de forma que:

$$S = \phi + \frac{\langle E \rangle}{T} + \frac{P \langle V \rangle}{T} \quad (\text{B.4.13})$$

Comparando los sumandos de (B.4.13) con $S = k_B \ln Y + k_B \beta \langle E \rangle + k_B \gamma \langle V \rangle$ se tiene que:

$$\beta = \frac{1}{k_B T} \quad (\text{B.4.14})$$

$$\gamma = \frac{P}{k_B T} \quad (\text{B.4.15})$$

$$\phi = k_B \ln Y \quad (\text{B.4.16})$$

B.4.2. Ecuaciones de estado

Recordemos las ecuaciones que describen la función de Planck, potencial termodinámico entrópico natural para un conjunto isotermo-isobárico:

$$\phi = \phi\left(\frac{1}{T}, \frac{P}{T}, N\right) = -\frac{G}{T} = -\frac{\mu N}{T} \quad \text{\underline{Ecuación de Euler}} \quad (\text{B.4.17})$$

$$d\phi = -Ud\left(\frac{1}{T}\right) - Vd\left(\frac{P}{T}\right) - \frac{\mu}{T}dN \quad \text{\underline{Ecuación de Gibbs}} \quad (\text{B.4.18})$$

$$Ud\left(\frac{1}{T}\right) + Vd\left(\frac{P}{T}\right) - Nd\left(\frac{\mu}{T}\right) = 0 \quad \text{\underline{Ecuación de Gibbs-Duhem}} \quad (\text{B.4.19})$$

Si diferenciamos la ecuación (B.4.16) llegamos a la siguiente expresión:

$$d\phi = k_B d \ln Y = k_B \left[\left(\frac{\partial \ln Y}{\partial (1/T)} \right)_{\frac{P}{T}, N} d\left(\frac{1}{T}\right) + \left(\frac{\partial \ln Y}{\partial (P/T)} \right)_{\frac{1}{T}, N} d\left(\frac{P}{T}\right) + \left(\frac{\partial \ln Y}{\partial N} \right)_{\frac{1}{T}, \frac{P}{T}} dN \right] \quad (\text{B.4.20})$$

Comparando ahora la ecuación de Gibbs (B.4.18) de la representación con (B.4.20), obtenemos las correspondientes **ecuaciones de estado**:

$$\text{Ecuación energética de estado (EEE)} \quad \rightarrow \quad \langle E \rangle = U = - \left(\frac{\partial \ln Y}{\partial \beta} \right)_{\gamma, N} \quad (\text{B.4.21})$$

$$\text{Ecuación térmica de estado (ETE)} \quad \rightarrow \quad V = V_{Macro} = \langle V \rangle = - \left(\frac{\partial \ln Y}{\partial \gamma} \right)_{\beta, N} \quad (\text{B.4.22})$$

$$\text{Ecuación másica de estado (EME)} \quad \rightarrow \quad -\frac{\mu}{k_B T} = -\mu\beta = \left(\frac{\partial \ln Y}{\partial N} \right)_{\beta, \gamma} \quad (\text{B.4.23})$$

B.4.3. Funciones respuesta y fluctuaciones

Veremos brevemente un análisis de las segundas derivadas de la función de Planck:

■ **Fluctuaciones de la energía:** segunda derivada respecto a $\beta = 1/k_B T$:

$$\left(\frac{\partial^2 \ln Y}{\partial \beta^2} \right)_{\gamma, N} = - \left(\frac{\partial \langle E \rangle}{\partial \beta} \right)_{\gamma, N} = (\Delta^* E)^2 = \langle (E_i - \langle E \rangle)^2 \rangle = k_B T^2 C_V + k_B T \frac{(\alpha_T T - \kappa_T P)^2 V}{\kappa_T} \quad (\text{B.4.24})$$

No vamos a detallar la derivación de esta expresión, pues su cálculo resulta bastante complejo. Dicho cálculo es complejo precisamente por las variables que han de ser constantes al derivar el potencial termodinámico. En este caso, al igual que en el conjunto gran canónico, no es posible establecer una conexión inmediata con la capacidad calorífica a volumen constante, como ocurría en el conjunto canónico. Recordemos para ello la definición de C_V :

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{V, N} \neq \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{\gamma = P/kT, N}$$

Ejemplos como este ilustran la importancia que tienen las variables que permanecen constantes en el proceso de derivación.

■ **Fluctuaciones en el volumen:** segunda derivada respecto a $\gamma = P / k_B T$:

$$\left(\frac{\partial^2 \ln Y}{\partial \gamma^2} \right)_{\beta, N} = - \left(\frac{\partial \langle V \rangle}{\partial \gamma} \right)_{\beta, N} = (\Delta^* V)^2 = \langle (V_i - \langle V \rangle)^2 \rangle = k_B T \langle V \rangle \kappa_T = \frac{\langle V \rangle \kappa_T}{\beta} \quad (\text{B.4.25})$$

Las fluctuaciones relativas son:

$$\frac{\Delta^* V}{\langle V \rangle} = \left(\frac{k_B T \kappa_T}{\langle V \rangle} \right)^{1/2} \quad (\text{B.4.26})$$

Para un gas perfecto clásico, se tiene que $P \langle V \rangle = N k_B T$, y el coeficiente de compresibilidad isoterma es:

$$\kappa_T = - \frac{1}{\langle V \rangle} \left(\frac{\partial \langle V \rangle}{\partial P} \right)_{T, N} = \frac{1}{P} \quad (\text{B.4.27})$$

De esta forma, las fluctuaciones relativas son:

$$\frac{\Delta^* V}{\langle V \rangle} = \left(\frac{k_B T}{\langle V \rangle P} \right)^{1/2} = \frac{1}{\sqrt{N}} \quad (\text{B.4.28})$$

B.5. Resumen de los diferentes conjuntos estadísticos

En este apéndice hemos estudiado los conjuntos microcanónico, canónico, gran canónico, e isoterma-isobárico. Pero estos no son los únicos conjuntos estadísticos que existen. Para cada grupo nuevo de ligaduras que tomemos a la hora de aplicar el Principio de Máxima Entropía, obtendremos un nuevo conjunto. En todos los conjuntos estadísticos hemos obtenido un resultado de la forma:

$$\boxed{\text{potencial termodinámico entrópico} = k_B \ln(\text{función estructura})}$$

Las funciones macroscópicas son los potenciales termodinámicos entrópicos, y las funciones microscópicas son las funciones estructura, que dan cuenta del comportamiento microscópico y se escriben como una suma a los microestados del sistema.

Hemos visto también que a cada conjunto se le asocia una función estructura, y un potencial entrópico macroscópico. En este hecho es donde mejor se puede apreciar el papel de la Física Estadística como puente entre la Mecánica y la Termodinámica. En la siguiente tabla se resumen los diferentes conjuntos estudiados, sus funciones estructura, sus potenciales entrópicos y las variables que permanecen constantes en cada conjunto:

<u>Conjunto</u>	<u>Variables</u>	<u>Función estructura</u>	<u>Potencial entrópico</u>
Microcanónico	U, V, N	Número de microestados: $\Omega = \sum_i 1$	Entropía: $S = k_B \ln \Omega$
Canónico	$\frac{1}{T}, V, N$	Función de partición: $Z = \sum_i e^{-\beta E_i}$	Función de Massieu: $\Psi = k_B \ln Z$
Gran canónico	$\frac{1}{T}, V, \frac{\mu}{T}$	Gran función de partición: $\Xi = \sum_i e^{-\beta E_i - \alpha N_i}$	Función de Kramers-Landau: $\eta = k_B \ln \Xi$
Isoterma-isobárico	$\frac{1}{T}, \frac{P}{T}, N$	F. estructura isoterma-isobárica: $Y = \sum_i e^{-\beta E_i - \gamma V_i}$	Función de Planck: $\phi = k_B \ln Y$

Además, hemos visto también las siguientes relaciones entre los conjuntos:

<u>Funciones</u>	<u>Operación matemática para cambiar de conjunto</u>
Funciones estructura	Transformadas discreta de Laplace
Potenciales entrópicos	Transformadas de Legendre

$$\boxed{\begin{array}{l} S(U, V, N) \xrightarrow{T.Legendre(U)} \Psi(1/T, V, N) \xrightarrow{T.Legendre(N)} \eta(1/T, V, \mu/T) \\ S(U, V, N) \xrightarrow{T.Legendre(U)} \Psi(1/T, V, N) \xrightarrow{T.Legendre(V)} \phi(1/T, P/T, N) \end{array}}$$

$$\boxed{\begin{array}{l} \Omega(U, V, N) \xrightarrow{T.Laplace(U)} Z(1/T, V, N) \xrightarrow{T.Laplace(N)} \Xi(1/T, V, \mu/T) \\ \Omega(U, V, N) \xrightarrow{T.Laplace(U)} Z(1/T, V, N) \xrightarrow{T.Laplace(V)} Y(1/T, P/T, N) \end{array}}$$

Apéndice C: El método del término máximo

En Física Estadística se emplea frecuentemente la aproximación de sustituir el logaritmo de una suma por el logaritmo del mayor término de la suma. Esto puede no parecer razonable, pero en la práctica y en condiciones adecuadas, la “aproximación” de hecho no es una aproximación (Hill, 1970), en relación con los términos de un orden de magnitud significativo en Termodinámica. Como ejemplo típico examinemos con detalle una suma que es sencilla, pero que tiene la misma forma y propiedades cualitativas que las sumas que se encuentran en los problemas reales. Consideremos

$$\Sigma = \sum_{N=0}^M t_N \quad (\text{C.1})$$

donde

$$t_N = \frac{M! x^N}{N!(M-N)!} \quad (\text{C.2})$$

En este caso $x = O(1)$ y $M = O(10^{20})$. Σ puede ser, por ejemplo, una función de partición gran canónica (apéndice B). N es el número (fluctuante) de moléculas del sistema, y M el mayor número de moléculas del sistema.

En primer lugar, advirtamos que la suma de la ecuación (C.1) se puede realizar exactamente y el resultado es

$$\Sigma = (1+x)^M \quad \text{o bien} \quad \boxed{\ln \Sigma = M \ln(1+x)} \quad (\text{C.3})$$

En Termodinámica es siempre el logaritmo de una suma (logaritmo de una función estructura) la magnitud que está relacionada con una función termodinámica (potencial termodinámico entrópico) (Baixauli y Cobos, 1996), y por tanto, es la función que nos interesa.

A continuación, veamos cuál es el valor de N , N^* , que proporciona el término mayor de la suma. Como el t_N máximo es también el $\ln t_N$ máximo, por conveniencia se puede localizar N^* con referencia a $\ln t_N$, en lugar de a t_N .

N^* y $M - N^*$ serán del mismo orden de magnitud que M , esto es, serán números muy grandes y, por tanto, se puede emplear sin error apreciable la aproximación de Stirling (3.3):

$$\ln t_N = M \ln M - N \ln N - (M-N) \ln(M-N) + N \ln x \quad (\text{C.4})$$

$$\frac{\partial \ln t_N}{\partial N} = 0 = -\ln N + \ln(M-N) + \ln x$$

$$\frac{N^*}{M - N^*} = x \quad N^* = \frac{xM}{1+x} \quad (\text{C.5})$$

Para encontrar el logaritmo del término máximo, sustituyamos la expresión de N^* dada por la ecuación (C.5) en lugar de N en la ecuación (C.4), y se encontrará:

$$\boxed{\ln t_{N^*} = M \ln(1+x)} \quad (\text{C.6})$$

Así, pues, respecto a términos de orden M , y estos son los únicos términos de importancia termodinámica, el logaritmo de la suma completa Σ (C.3) tiene el mismo valor que el logaritmo del término máximo solamente (C.6).

Para entender más claramente qué es lo que detrás de esta concordancia sorprendente, avancemos un poco más en el desarrollo efectuado. Consideremos el desarrollo en serie de Taylor de $\ln t_N$ alrededor de $N = N^*$:

$$\ln t_N = \ln t_{N^*} + \frac{1}{2} \left(\frac{\partial^2 \ln t}{\partial N^2} \right)_{N=N^*} (N - N^*)^2 + \dots$$

El término lineal en $N - N^*$ no aparece porque la primera derivada en $N = N^*$ es nula. De la ecuación (C.6), se tiene

$$\frac{\partial^2 \ln t_N}{\partial N^2} = -\frac{1}{N} - \frac{1}{M - N}$$

Entonces, empleando (C.5),

$$\ln t_N = \ln t_{N^*} - \frac{(1+x)^2 (N - N^*)^2}{2xM} + \dots$$

o

$$t_N = t_{N^*} \exp \left(-\frac{(N - N^*)^2}{2\sigma^2} \right) \quad (\text{C.7})$$

donde la desviación estándar es

$$\sigma = \frac{x^{1/2} M^{1/2}}{1+x} \quad (\text{C.8})$$

Así pues, se obtiene una gaussiana muy aguda, centrada en N^* , con una desviación estándar que es extremadamente pequeña, relativa al valor más probable,

$$\frac{\sigma}{N^*} = \frac{1}{x^{1/2} M^{1/2}} = O(M^{-1/2}) = O(10^{-10}) \quad (\text{C.9})$$

siendo $M = O(10^{20})$. Dentro de la exactitud de los términos que se han retenido anteriormente $\bar{N} = N^*$. Los otros términos introducirían una discrepancia entre \bar{N} y N^* de un orden completamente despreciable.

La función gaussiana t_N disminuye rápidamente a cero cuando se aleja de $N = N^*$ en ambas direcciones una distancia del orden de algunas desviaciones estándar. Por tanto el número de términos que contribuyen significativamente a la Σ es de varias veces σ , o sea, del orden de $M^{1/2}$; todos los términos restantes se pueden ignorar. El orden de magnitud de cada uno de esos términos es el mismo que el de t_{N^*} , o sea, $O(e^M)$.

Por tanto

$$\Sigma = O(e^M) O(M^{1/2})$$

y

$$\ln \Sigma = O(M) + O(\ln M) = \ln t_{N^*} + O(\ln M) \quad (\text{C.10})$$

Pero, debido al orden de magnitud de M , el término $O(\ln M)$ de esta ecuación es totalmente despreciable comparado con el otro término. Así pues, incluso aunque Σ sea 10^{10} veces mayor que t_{N^*} , la diferencia entre $\ln \Sigma$ y $\ln t_{N^*}$ es insignificante (comprobado en las ecuaciones C.3 y C.6).

Como comentario final, subrayemos que todos los términos significativos de la suma tienen valores de N que difieren despreciablemente de N^* . Esto es, N puede diferir de la media en $O(\sigma)$ o $O(M^{1/2})$, pero respecto al valor de la media, esto es solamente del $O(M^{-1/2})$ como en la ecuación (C.9).

Apéndice D: Las probabilidades a priori.

Los antecedentes históricos del PME.

En este apéndice introducimos, como lectura adicional para los alumnos, ciertos antecedentes históricos del Principio de Máxima Entropía como solución al problema de asignar probabilidades a priori en Teoría de la Probabilidad. En su artículo “*Where do we stand on Maximum Entropy?*” (1978), E. T. Jaynes intenta dar una visión general del origen del PME como una extensión natural (generalización) del Principio de Indiferencia de Bernoulli.

Aunque esta lectura no es necesaria para comprender los fundamentos detrás del PME (los cuales ya se han expuesto con total claridad y precisión en el capítulo 3), puede servir como complemento para aquellos interesados en conocer más sobre el origen de este método de razonamiento.

Para profundizar todavía más en este asunto, se recomienda también la lectura de la primera sección (“*A. Historical Background*”) del artículo de Jaynes mencionado.

D.1. Introducción. Las probabilidades a priori

Como ya se ha comentado en este trabajo, el primer paso en un problema que requiera de la Teoría de la Probabilidad es asignar unos valores iniciales a las probabilidades de los diferentes sucesos. En el caso de no tener información sobre el problema a estudiar, se aplica el Principio de Indiferencia: en ausencia de información que indique lo contrario, los sucesos elementales que conforman el espacio muestral de un experimento aleatorio han de ser equiprobables.

Bernoulli, Laplace y Bayes intentaron evitar esta asignación inicial de probabilidades, en lo que se conoce como el “problema de la inversión” (Jaynes, 1978). Los dos últimos lograron resolverlo, pero las “probabilidades a priori” que se pretendían evitar se volvían a introducir inevitablemente en las ecuaciones. No disponían de ninguna herramienta o principio para hallar probabilidades a priori que no fuese el Principio de Indiferencia.

Sin embargo, como se ha visto en el problema del juego de reparto del capítulo 2 de este trabajo, el Principio de Máxima Entropía sí que permite inferir probabilidades utilizando la información disponible. Para entender el origen histórico detrás de este método de razonamiento, se exponen a continuación los argumentos que proporcionaron Maxwell y Boltzmann para tratar problemas que requieren de la aplicación de la Teoría de la Probabilidad.

D.2. El argumento de Maxwell

Maxwell elaboró el primer trabajo sobre la aplicación del análisis probabilístico a la teoría cinética de los gases. Se enfrentó de inmediato con el problema de tener que asignar probabilidades iniciales a las posiciones y velocidades de las moléculas del gas.

Para encontrar la distribución de probabilidad de la dirección de la velocidad de una molécula tras el impacto con otra, Maxwell (1860) enunció la siguiente proposición: “*Para que se lleve a cabo una colisión, la línea del movimiento de una de las bolas debe pasar por el centro de la otra a una distancia menor a la suma de sus radios, es decir, debe atravesar un círculo cuyo centro sea el de la otra bola, y cuyo radio sea la suma de los radios de las bolas. Dentro de este círculo cada posición es igualmente probable (...)*”. Aquí también, como primer paso necesario en un análisis de probabilidades, Maxwell tuvo que aplicar el Principio de Indiferencia. La diferencia en este caso es que las cantidades con las que trabaja tienen un

significado físico: posición dentro de una sección circular de colisión. De esta forma, la probabilidad de incidir en una región determinada debe tomarse proporcional al área de dicha región, y no al cubo del área, o al logaritmo del área. El simple hecho de conocer el significado físico de nuestros parámetros ya constituye una información previa muy relevante.

Aunque no se detallará el desarrollo matemático de Maxwell en esta lectura, con su razonamiento fue capaz de encontrar la ley de distribución de velocidades, los coeficientes de difusión, viscosidad y conductividad térmica del gas, entre otros. Además, los experimentos confirmaban las predicciones realizadas por Maxwell.

D.3. El argumento de la distribución de Boltzmann

El siguiente argumento fue el primer precedente del Principio de Máxima Entropía. Boltzmann dio el primer paso hacia la búsqueda de una solución explícita al problema de las “probabilidades a priori”. Su objetivo era encontrar cómo se distribuyen unas moléculas en un campo de fuerzas conservativo (que puede ser, por ejemplo, un campo gravitatorio). La fuerza que actúa en una molécula que se encuentra en una cierta posición \vec{r} viene dada por:

$$\vec{F}(\vec{r}) = -\nabla\Phi(\vec{r}) \quad (\text{D.1})$$

donde $\Phi(\vec{r})$ es su energía potencial. Teniendo en cuenta también su energía cinética, cada molécula de masa m , posición \vec{r} y velocidad \vec{v} tiene la siguiente energía total:

$$E = \frac{1}{2}mv^2 + \Phi(\vec{r}) \quad (\text{D.2})$$

Boltzmann supone que la interacción entre moléculas es despreciable, y que están encerradas en un recipiente de volumen V . Pero él no ignoraba por completo cómo se distribuyen las moléculas, porque sabía que, se muevan como se muevan, el número total N de moléculas presentes no puede cambiar, y la energía total del conjunto debe permanecer constante:

$$E = \sum_{i=1}^N \left(\frac{1}{2}mv_i^2 + \Phi(\vec{r}_i) \right) \quad (\text{D.3})$$

Debido a la restricción (ligadura) en la energía, no todas las posiciones y velocidades serán igualmente probables.

En este punto del razonamiento, Boltzmann dividió el espacio fásico (espacio de posiciones y momentos) disponible para las moléculas en celdas discretas. Con esta división, la celda k -ésima será una cierta región R_k tan pequeña que la energía E_k de una molécula no varía apreciablemente dentro de la celda, pero lo suficientemente grande para que en cada celda haya un elevado número de moléculas, es decir, $N_k \gg 1$. El conjunto de celdas $\{R_k, 1 \leq k \leq s\}$ debe llenar por completo el espacio fásico accesible (cuyo volumen es finito por la restricción en la energía total) sin solaparse unas celdas con otras, y todas deben tener el mismo tamaño.

El problema a resolver es el siguiente: dado un número de moléculas N , una energía total E y la energía potencial $\Phi(\vec{r})$ debida al campo conservativo, **¿cuál es la mejor predicción que se puede hacer sobre el número de moléculas N_k que hay en la celda R_k ?**

En este punto del razonamiento de Boltzmann es donde se halla el origen del Principio de Máxima Entropía. Primero se preguntó: **¿De cuántas formas diferentes podría realizarse un conjunto dado de números de ocupación $\{N_k, 1 \leq k \leq s\}$?** La respuesta son las permutaciones con repetición:

$$PR_N^{N_1, \dots, N_s} = W(N_1, \dots, N_s) = \frac{N!}{N_1! N_2! \dots N_s!} \quad (\text{D.4})$$

La ligadura correspondiente a la energía total de la distribución es:

$$E = \sum_{k=1}^s N_k E_k \quad (\text{D.5})$$

Además, los N_k también están sujetos a la siguiente ligadura:

$$N = \sum_{k=1}^s N_k \quad (\text{D.6})$$

Cualquier conjunto $\{N_k, 1 \leq k \leq s\}$ de números de ocupación para el que E y N se ajustan a la información dada (ligaduras), representa una distribución posible del sistema. De entre todas las muchas posibles distribuciones que existen, ¿cuál es más probable que ocurra?

La respuesta de Boltzmann fue que la distribución más probable es aquella que puede realizarse de un mayor número de formas, es decir, aquella que maximiza $W(N_1, \dots, N_s)$ sujeta a las ligaduras $E = \sum_{k=1}^s N_k E_k$ y $N = \sum_{k=1}^s N_k$.

Tomando el logaritmo natural en la expresión (D.4):

$$\ln W = \ln \left(\frac{N!}{N_1! N_2! \dots N_s!} \right) = \ln(N!) - \sum_{k=1}^s \ln(N_k!) \quad (\text{D.7})$$

Como los números de ocupación N_k son muy grandes (y en consecuencia N también es muy grande), se puede aplicar la aproximación de Stirling para los factoriales:

$$\ln n! \approx n \ln n - n \quad (\text{D.8})$$

Aplicando dicha aproximación a la ecuación (D.7):

$$\begin{aligned} \ln W &\approx N \ln N - N - \sum_{k=1}^s (N_k \ln N_k - N_k) = N \ln N - \sum_{k=1}^s N_k \ln N_k \\ &= (N_1 + \dots + N_s) \ln N - \sum_{k=1}^s N_k \ln N_k = \sum_{k=1}^s (N_k \ln N - N_k \ln N_k) \\ &= \sum_{k=1}^s N_k \ln \left(\frac{N}{N_k} \right) = - \sum_{k=1}^s N_k \ln \left(\frac{N_k}{N} \right) \end{aligned} \quad (\text{D.9})$$

Con esta aproximación, la expresión (D.4) de $W(N_1, \dots, N_s)$ se puede escribir de la siguiente forma:

$$\ln W = -N \sum_{k=1}^s \left(\frac{N_k}{N} \right) \ln \left(\frac{N_k}{N} \right) \quad (\text{D.10})$$

La expresión (D.10) resulta especialmente familiar si la comparamos con las expresiones de la entropía estadística que conocemos:

$$S(p_1, \dots, p_n) = -k \sum_{i=1}^n p_i \ln p_i \quad \sigma(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \ln p_i$$

La solución matemática al problema utilizando los multiplicadores de Lagrange es sencilla. Se construye en primer lugar la función de Lagrange

$$L(N_1, \dots, N_s, \lambda_1, \lambda_2) = -N \sum_{k=1}^s \frac{N_k}{N} \ln \frac{N_k}{N} + \lambda_1 \left(\sum_{k=1}^s N_k - N \right) + \lambda_2 \left(\sum_{k=1}^s N_k E_k - E \right) \quad (\text{D.11})$$

Derivando respecto de λ_1 y λ_2 se vuelven a obtener las ligaduras (D.5) y (D.6). Si se deriva respecto a N_k , y se definen dos nuevas variables Z y β :

$$Z^{-1} = \exp(\lambda_1 - 1) \qquad \beta = -\lambda_2 \qquad (D.12)$$

$$\frac{\partial L}{\partial N_k} = 0 = -\left(\ln \frac{N_k}{N} + 1\right) + \lambda_1 + \lambda_2 E_k \quad \rightarrow \quad N_k = N \exp(\lambda_1 - 1 + \lambda_2 E_k) = \frac{N}{Z} e^{-\beta E_k} \quad (D.13)$$

se obtiene que el resultado del valor “más probable” de N_k es

$$\hat{N}_k = \frac{N}{Z} \exp(-\beta E_k) = \frac{N}{Z} e^{-E_k/kT} \quad (D.14)$$

donde $\beta = (k_B T)^{-1}$ y $Z(\beta, V, N)$ es

$$Z(\beta, V, N) = \sum_{k=1}^s \exp(-\beta E_k) = \sum_{k=1}^s e^{-E_k/kT} \quad (D.15)$$

Las expresiones (D.14) y (D.15) son las que describen la estadística de Maxwell-Boltzmann. En Física Estadística, la estadística de Maxwell-Boltzmann describe la distribución estadística de partículas materiales clásicas en diferentes estados de energía en equilibrio térmico. Se aplica cuando la temperatura es lo suficientemente alta o la densidad de partículas lo suficientemente baja como para que los efectos cuánticos sean despreciables.

Para llegar a la solución al problema, Boltzmann ha utilizado la información que tenía, y aprovechando dicha información ha conseguido hacer predicciones correctas en unas pocas líneas de cálculo, que serían imposibles de obtener si intentásemos realizar un cálculo detallado de las trayectorias individuales de 10^{23} moléculas diferentes. El argumento de Boltzmann tiene en cuenta la dinámica del sistema que se está estudiando, pero lo hace de una forma realmente eficiente. La información sobre la dinámica entra en sus ecuaciones en dos puntos del razonamiento seguido:

- i. La conservación de la energía total.
- ii. La definición las celdas en términos del volumen fásico, el cual permanece constante durante la evolución temporal del sistema (teorema de Liouville).

El hecho de que esta información haya sido suficiente para predecir la distribución correcta de energías de las moléculas muestra que los complejÍsimos detalles dinámicos que no se tuvieron en cuenta eran irrelevantes para las predicciones, y se habrían anulado de todos modos si Boltzmann se hubiera tomado la gran molestia de calcularlos.

Su razonamiento fue realmente eficiente. Introdujo en sus ecuaciones únicamente la información sobre la dinámica del sistema que resultó ser relevante para resolver el problema. Si somos capaces de reconocer y eliminar aquella información que resulta irrelevante al inicio de un problema, nos ahorraremos tener que realizar cálculos inmensos, sólo para descubrir al final que prácticamente todo lo que calculamos era irrelevante para la resolución del problema.

En el método de Boltzmann sobre la distribución más probable se encuentra la esencia matemática del Principio de Máxima Entropía. Pero, como se ha observado, en las ecuaciones de Boltzmann no aparece ninguna probabilidad, pues su objetivo no era calcular una distribución de probabilidades, sino intentar estimar los números de ocupación por medio de un criterio (el valor de las diferentes formas $W(N_k)$ en las que se puede realizar un set de números de ocupación $\{N_k\}$) que tiene en cuenta el número real de posibilidades físicas, que es un número definido que no tiene nada que ver con el estado de conocimiento de nadie.

La transición de este argumento de Boltzmann al presente y más abstracto Principio de Máxima Entropía, aunque matemáticamente es trivial, requirió un gran esfuerzo conceptual para los físicos de la época.