

LUIS FERNANDO HEVIA DE LOS MOZOS

FUNDAMENTOS DE FÍSICA ESTADÍSTICA

Versión 2.2

10 de julio de 2023



Prólogo y agradecimientos

Este texto ha nacido como material para las clases de la disciplina “Fundamentos de Física Estadística” impartidas en el primer curso del Grado en Física y el segundo curso del Programa de Estudios Conjunto de Grado en Física y Grado en Matemáticas en la Universidad de Valladolid.

Resulta un reto muy interesante elaborar un curso de una disciplina tan compleja como es la Física Estadística que, por un lado, sea suficientemente sencillo a nivel formal y, por otro, sea conceptualmente riguroso, útil para el alumnado y coherente con las asignaturas más avanzadas del plan de estudios (en particular, con las propias de la Termología, a saber, “Termodinámica” y “Física Estadística”).

Siendo esta una tarea difícil, no es imposible, gracias a que la formulación de la Física Estadística basada en el **Principio de Máxima Entropía** de E.T. Jaynes (1957) no requiere de grandes conocimientos de Mecánica Clásica o Cuántica para una primera aproximación. Desde el punto de vista de Jaynes, el problema que se resuelve en Física Estadística es la inferencia de las probabilidades de los microestados (configuraciones microscópicas) de un sistema a partir de las leyes de la Teoría de la Probabilidad y de unas pocas restricciones macroscópicas conocidas. Todo este proceso de inferencia estadística puede comprenderse sin necesidad de conocer en detalle los modelos de realidad mecánicos (clásicos o cuánticos), pues el único papel que juega la Mecánica en el proceso es la enumeración de los microestados compatibles con los valores de unas ciertas variables macroscópicas. La filosofía del curso es que, **comprendiendo las razones conceptuales por las que este proceso de inferencia estadística funciona, los estudiantes puedan hallar toda la información termodinámica de un sistema a partir de un modelo de realidad dado**. De este modo, el curso permite a los estudiantes desarrollar un esquema mental de cómo funcionan la Termodinámica y la Física Estadística y su relación con el resto de la Física, aprender el mecanismo de cálculo de los **potenciales termodinámicos entrópicos** y aplicarlo a modelos sencillos cuyos detalles, a nivel mecánico, no tienen por qué comprender a fondo.

Además, también se centra en combatir preconceptos erróneos (o de aplicabilidad limitada), o de prevenir su aparición. Por ejemplo, **se evita la identificación del calor con una variación de temperatura únicamente** y en su lugar se define como un intercambio de energía de un sistema cerrado que no se puede describir como trabajo mecánico macroscópico (lo que constituye la formulación de M. Born del Primer Principio de la Termodinámica de 1921). También **se evita la identificación de la temperatura con la energía cinética media** de un sistema, pues esta no es más que una interpretación de aplicabilidad limitada en el marco de los modelos basados en la Mecánica Clásica. **Se señalan las limitaciones de la Teoría Cinética de los gases ideales, que, por sí sola, no puede obtener su Ecuación Térmica de Estado**; en cambio, utiliza dicha ecuación, justificada desde la Termodinámica experimental, para inferir la relación entre la temperatura y la velocidad cuadrática media de las moléculas del gas ideal. También **se evita la identificación de entropía con desorden**, pues, como indica P.T. Landsberg, “desorden” es una palabra del lenguaje común, con un significado impreciso, mientras que “entropía” es un concepto científico, cuantitativo y muy preciso; la identificación de ambos conceptos genera paradojas de interpretación, como la de los sistemas biológicos, que, en su crecimiento, aumentan su entropía a la vez que se organizan (se “ordenan”), emergiendo estructuras más complejas.

El origen de este curso, sin el cual este enfoque nunca habría sido posible, son las clases de Física Estadística de cursos superiores del profesor José Carlos Cobos, quien me introdujo en el mundo de la Física y fue mi profesor de Física Estadística, tutor de mi Trabajo Fin de Grado y director de mi Tesis Doctoral. J.C. Cobos siempre ha tratado de encontrar enfoques didácticos para enseñar los métodos de la Física Estadística (o, como él lo llama, resolver “el problema”), y los ha plasmado en sus clases y exámenes: monedas, dados, imágenes que valen más que mil palabras, carniceros malvados, Reyes Magos, preparación de bocadillos... Es obligado agradecer su trabajo de varias décadas y las numerosas horas que ha ejercido de mentor, que han permitido la elaboración de este curso.

Quiero también expresar mi agradecimiento al profesor Daniel Lozano por una concienzuda revisión del texto y por sus oportunas y acertadas sugerencias. También al resto de mis compañeros del Grupo Especializado en Termodinámica de los Equilibrios entre Fases (**GETEF**) y el Departamento de Física Aplicada de la Universidad de Valladolid, Juan Antonio González, Isafías García, Ana Cobos y Víctor Alonso. Por último, a mi amigo y profesor del Departamento de Electricidad y Electrónica, Luis Sánchez-Tejerina, por las interesantes conversaciones sobre temas en la intersección entre su área de especialización y la Física Estadística.

Índice

1. Termodinámica.....	9
1.1. Sistemas macroscópicos. El lugar de la Física Estadística.....	9
1.1.1. Estado, sistema, ley. Sistema macroscópico.....	9
1.1.2. Sistemas macroscópicos y modelos de realidad. Descripción microscópica.....	10
1.1.3. Descripción macroscópica.....	11
1.1.4. Descripción estadística o probabilística: Física Estadística.....	11
1.2. Fases y tipos de variables.....	11
1.2.1. Sistemas homogéneos, heterogéneos. Fases.....	11
1.2.2. Variables intensivas, extensivas u otras.....	12
1.3. Procesos adiabáticos y Primer Principio.....	12
1.3.1. Primer Principio de la Termodinámica: argumento intuitivo.....	12
1.3.2. Primer Principio de la Termodinámica: enunciado.....	13
1.3.3. Trabajo de configuración y trabajo disipativo.....	13
1.3.4. Definición de calor.....	14
1.4. Características de la energía de un sistema.....	15
1.4.1. Tipos de fuerzas macroscópicas.....	15
1.4.2. Posibilidad de definir un sistema según las características de su energía.....	15
1.4.3. Extensividad (o no) de la energía interna.....	16
1.5. Componentes y entidades elementales independientes.....	16
1.6. Paredes.....	17
1.7. Estado de equilibrio.....	18
1.7.1. Definición de equilibrio.....	18
1.7.2. Definición de equilibrio mutuo.....	18
1.7.3. Principio General de la Termodinámica.....	19
1.7.4. Macroestado en un estado de equilibrio.....	19
1.7.5. Procesos termodinámicos.....	19
1.8. Hacia el postulado fundamental: Boltzmann.....	20
1.8.1. Direccionalidad de los procesos: la idea de Boltzmann.....	20
1.8.2. Un ejemplo cuantitativo.....	21
1.8.3. La entropía de un sistema aislado.....	23
1.9. Postulados de Callen.....	25
1.9.1. El postulado de estado de Callen.....	25
1.9.2. El postulado de máxima entropía de Callen.....	25
1.10. Consecuencias del postulado de máxima entropía.....	26
1.10.1. Principio de Aumento de Entropía.....	26
1.10.2. Variables intensivas.....	27
1.10.3. Relación de la entropía con la presión.....	28

1.10.4.	Condiciones de equilibrio mecánico mutuo	29
1.10.5.	Condiciones de equilibrio térmico mutuo.....	30
1.10.6.	Condiciones de equilibrio material mutuo	30
1.11.	Ecuación fundamental y ecuaciones de estado.....	31
1.12.	Sistema cerrado, isócoro e isoterma: función de Massieu	32
	Ejercicios – Termodinámica.....	33
2.	Teoría Cinética de los gases ideales	35
2.1.	Cálculo en el espacio tridimensional.....	35
2.1.1.	Integración en varias variables.....	35
2.1.2.	Coordenadas esféricas en el espacio tridimensional.....	36
2.2.	Conceptos básicos de Probabilidad	37
2.2.1.	Variables aleatorias	39
2.3.	Los gases ideales	42
2.3.1.	Los gases ideales desde el punto de vista termodinámico.....	42
2.3.2.	El modelo cinético del gas ideal.....	42
2.4.	El espacio de las velocidades del gas ideal	43
2.4.1.	Forma funcional de la FDP y FDPM radial y angular	43
2.4.2.	Cálculos estadísticos en el espacio de velocidades.....	44
2.5.	Densidad de flujo molecular	45
2.5.1.	Problema mecánico	45
2.5.2.	Problema estadístico.....	46
2.6.	Presión.....	46
2.6.1.	Problema mecánico	46
2.6.2.	Problema estadístico.....	47
2.7.	Relación entre Teoría Cinética de los gases y Termodinámica.....	47
2.8.	Recorrido libre medio y ecuación de supervivencia	47
	Ejercicios – Teoría Cinética de los gases ideales	50
3.	Principio de Máxima Entropía.....	53
3.1.	Combinatoria	53
3.1.1.	Variaciones sin repetición	54
3.1.2.	Variaciones con repetición	54
3.1.3.	Permutaciones sin repetición.....	54
3.1.4.	Permutaciones con repetición	54
3.1.5.	Combinaciones sin repetición.....	55
3.1.6.	Combinaciones con repetición.....	55
3.2.	Un juego de reparto.....	56
3.2.1.	10 fichas, 10 €.....	56
3.2.2.	20 fichas, 20 €.....	58
3.3.	El Principio de Máxima Entropía (PME o MaxEnt)	59
3.3.1.	La función ignorancia de una distribución de probabilidad.....	59
3.3.2.	Puntos de vista en Teoría de la Probabilidad	60
3.3.3.	Inferencia de probabilidades: PME.....	60

3.3.4.	El ejemplo del dado de Brandeis.....	62
3.4.	Solución del juego de reparto con el PME.....	63
3.5.	Complemento (para valientes): Argumento de Wallis y conexión entropía-frecuencia	66
3.5.1.	Fórmula asintótica para las permutaciones con repetición	66
3.5.2.	El argumento de Wallis	67
3.5.3.	Conexión entropía-frecuencia.....	69
3.6.	Complemento: caracterización de la entropía	70
	Ejercicios – Principio de Máxima Entropía	72
4.	Formalismo de la Física Estadística	73
4.1.	Descripción estadística de un estado de equilibrio	73
4.2.	Conjunto microcanónico: macroestado (U, V, N)	75
4.2.1.	Inferencia de probabilidades.....	75
4.2.2.	Consecuencias.....	75
4.3.	Conjunto canónico: macroestado $(1/T, V, N)$	76
4.3.1.	Inferencia de probabilidades.....	76
4.3.2.	Entropía en el conjunto canónico. Equivalencia de conjuntos	77
4.3.3.	Relación de β con la temperatura y de Z con la función de Massieu	77
4.3.4.	Obtención de Z como una transformada de Laplace.....	79
4.3.5.	Factorización de la función de partición.....	79
4.3.6.	Fluctuaciones de energía en el conjunto canónico	80
4.4.	Entropía de una distribución continua	81
4.5.	Física Estadística semiclásica.....	83
4.5.1.	Discretización del espacio fásico.....	84
4.5.2.	Partículas idénticas distinguibles e indistinguibles.....	84
4.5.3.	Influencia de la indistinguibilidad en la posibilidad de definir Z_1	85
4.5.4.	Validez del límite semiclásico.....	86
4.6.	Ejemplo de cálculo de la ecuación fundamental.....	86
4.6.1.	Análisis en el conjunto microcanónico	86
4.6.2.	Análisis en el conjunto canónico	87
4.6.3.	Obtención de Z a partir de Ω	87
4.7.	Resumen: procedimiento de trabajo en Física Estadística	88
	Ejercicios – Formalismo de la Física Estadística.....	92
5.	Gas ideal monoatómico. Distribución de Maxwell	93
5.1.	Modelo microscópico del gas ideal monoatómico semiclásico	93
5.2.	Solución en el conjunto canónico – macroestado (β, V, N)	93
5.3.	Probabilidades de los estados de una molécula	95
5.3.1.	Densidad de probabilidad de momentos y velocidades de una molécula	95
5.3.2.	Densidad de probabilidad del módulo del momento y de la velocidad	96
5.3.3.	Celeridad más probable	97
5.3.4.	Celeridad media y celeridad cuadrática media	99
5.3.5.	Celeridad reducida.....	99
5.4.	Complemento: integrales y funciones especiales	100

5.4.1.	Integral de Euler-Poisson.....	100
5.4.2.	Integrales de la forma $\int_0^{\infty} x^n e^{-bx^2} dx$	101
5.4.3.	Integrales de la forma $\int_{-\infty}^{\infty} x^n e^{-bx^2} dx$	101
5.4.4.	La función error de Gauss	102
5.4.5.	La función gamma de Euler.....	102
	Ejercicios – Gas ideal monoatómico y distribución de Maxwell	104
6.	Teorema de equipartición de la energía. Gas ideal poliatómico	107
6.1.	Teorema de Equipartición de la Energía (conj. canónico clásico).....	107
6.2.	Grados de libertad	108
6.2.1.	Definición.....	108
6.2.2.	Grados de libertad de sólidos rígidos.....	108
6.2.3.	Grados de libertad moleculares.....	110
6.2.4.	Gas ideal poliatómico semiclásico. Aplicación del Teorema de Equipartición	110
6.2.5.	Gas ideal poliatómico semiclásico: función de partición y ecuaciones de estado	111
6.3.	Predicciones de la teoría semiclásica del gas ideal poliatómico: energía interna y capacidades caloríficas.....	112
6.4.	Comparación de las capacidades caloríficas predichas por el modelo semiclásico con los resultados experimentales	114
6.4.1.	Movimiento de traslación	115
6.4.2.	Movimiento de rotación.....	116
6.4.3.	Movimiento de vibración	117
6.4.4.	Temperaturas características.....	119
6.5.	Complemento: grados de libertad de un sólido rígido	119
	Ejercicios - Teorema de equipartición de la energía y gas ideal poliatómico.....	121
	Ejercicios extra para entrenarse	123
	Bibliografía.....	129

1. Termodinámica

En este tema trataremos de alcanzar dos objetivos básicos:

- **Definir la jerga que usaremos al tratar la Termodinámica y la Física Estadística.** Antes de comenzar a estudiar cualquier ciencia, es necesario conocer su jerga característica, esto es, las palabras que designan los **conceptos abstractos que se tienen que corresponder con fenómenos reales**. Es importante tener en mente que, en muchas ocasiones, diferentes autores utilizan una misma palabra con significados ligeramente diferentes o utilizan diferentes palabras para un mismo concepto.
- **Estudiar los principios y postulados de la Termodinámica y sus consecuencias.** Como veremos, la Física Estadística es un formalismo que permite inferir las probabilidades de las configuraciones microscópicas de un sistema de acuerdo con un modelo de realidad y obtener sus potenciales termodinámicos, como la entropía, de modo que el conocimiento de las propiedades básicas de estos potenciales termodinámicos es clave para la comprensión de la Física Estadística.

1.1. Sistemas macroscópicos. El lugar de la Física Estadística

1.1.1. Estado, sistema, ley. Sistema macroscópico

Llamaremos **sistema** a una parte de la realidad que puede ser estudiada por separado del resto de la realidad, a la que llamamos **entorno**. La posibilidad de seleccionar algo como un sistema implica:

- La selección de ciertas propiedades de éste que sean relevantes para los fenómenos que se quieran estudiar.
- La existencia de magnitudes medibles, llamadas **variables de estado**, que caractericen adecuadamente dichas propiedades.
- La existencia de unas **leyes** que relacionen las variables de estado entre sí y con variables del entorno.

La palabra **estado** designa, de forma algo imprecisa, la “forma de presentarse el sistema”. A veces, el estado de un sistema se puede caracterizar de diferentes maneras (utilizando diferentes tipos de variables), dando lugar a lo que se llaman **distintas descripciones del estado**. Este es el caso de los sistemas macroscópicos, que abordaremos en breve.

El concepto de estado y su descripción siempre se encuentran ligados a las leyes que permiten definirlo; en otras palabras, existe una **interrelación entre los conceptos leyes-sistema-estado**. Por ejemplo, en la Mecánica Clásica el *estado mecánico* de una partícula en un *campo de fuerzas externo independiente* está definido por su posición y su velocidad (o su posición y su momento), pero esta definición no sería útil si no existieran unas *leyes del movimiento* que permitieran, por un lado, separar la partícula de su entorno (que queda representado por el campo de fuerzas externo) y calcular, a partir del estado mecánico y las fuerzas en un instante, el estado mecánico en el instante siguiente. Así pues, estas leyes del movimiento permiten definir el sistema (la partícula), su entorno (campo de fuerzas externo) y su estado (posición y velocidad).

La palabra “independiente” que acompaña al campo de fuerzas externo en el párrafo anterior quiere decir que el movimiento de la partícula no afecta al valor del campo de fuerzas externo, pues de lo contrario deben existir unas leyes que relacionen no solo el movimiento de la partícula en función del valor del campo, sino también valor del campo en función del movimiento de la partícula. En otras palabras, estamos prescindiendo de la “reacción” de la partícula sobre el campo. Si no es posible prescindir de esta “reacción”, la partícula no podrá considerarse como un sistema independiente del campo. Sí podrá, sin embargo, considerarse como un sistema independiente el conjunto formado por la partícula y la fuente de campo. Un **ejemplo**:

- Sería posible considerar como un sistema una piedra pequeña en caída libre en la superficie de la Tierra, ya que el campo gravitatorio *debido a la piedra* no afecta apreciablemente a las propiedades de la Tierra y, por tanto, se puede describir el estado de la piedra en términos de una energía potencial externa gravitatoria que es independiente de los efectos del campo gravitatorio *creado por ella misma*.
- Por el contrario, al estudiar el movimiento de dos cuerpos celestes A y B que se ven atraídos por sus campos gravitatorios, el campo gravitatorio que siente B (debido a A) depende *apreciablemente* de la posición en la que se encuentre A, y al revés, el campo gravitatorio que siente A depende *apreciablemente* de la posición en la que se encuentre B. No es posible, por tanto, definir los estados de A y B independientemente uno de otro. Este es un ejemplo del famoso **problema de los dos cuerpos**, emblemático de la Mecánica Clásica.

Más adelante, al hablar de la energía de un sistema macroscópico, precisaremos mejor cuándo es posible la separación entre sistema y entorno.

Llamaremos **sistema macroscópico** a aquel de escala suficientemente similar a la humana como para poderles asignar magnitudes físicas que tengan sentido a escala humana. A partir de ahora, salvo que se diga explícitamente lo contrario, con la palabra “sistema” nos referiremos a un sistema macroscópico.

1.1.2. Sistemas macroscópicos y modelos de realidad. Descripción microscópica

Hoy en día se acepta que los sistemas macroscópicos tienen una **microestructura** subyacente, esto es, que están formados por diferentes **unidades estructurales** (como átomos o moléculas) que interaccionan entre sí a través de las **fuerzas microscópicas** electromagnética, gravitatoria, nuclear fuerte y nuclear débil (hasta donde la experimentación ha podido comprobar).

En Mecánica Macroscópica prescindimos del detalle de las fuerzas microscópicas entre unidades estructurales y las reemplazamos por unas **fuerzas macroscópicas equivalentes a escala humana** que producen los mismos efectos “globales” sobre porciones macroscópicas de realidad. Por ejemplo, la fuerza macroscópica de contacto entre dos sólidos se puede entender como resultante a partir de las fuerzas electromagnéticas entre los átomos de los sólidos.

La hipótesis de la existencia de microestructura lleva a la elaboración de **modelos de realidad** de los sistemas macroscópicos, que permiten describir el estado del sistema aplicando la Mecánica a sus unidades estructurales. Por ejemplo, podemos elaborar un modelo de un gas en el que consideramos que está formado por moléculas entre las que existen unas ciertas fuerzas y que están contenidas en un recipiente. Dicha forma de describir su estado se denomina **descripción microscópica** (de acuerdo con un modelo de realidad elegido) y cada uno de los posibles estados descritos de este modo se le denomina **MICROESTADO**.

Aquí, con “Mecánica” nos estamos refiriendo a todo el conjunto de teorías físicas al margen de la Termodinámica, o sea que bien podríamos llamarla la “no-Termodinámica”. Puede distinguirse entre **modelos clásicos** o **modelos cuánticos**, en función de si la mecánica utilizada es clásica o cuántica.

En principio, podría parecer que con las leyes mecánicas es suficiente para predecir el comportamiento de cualquier sistema. No obstante, un objeto de nuestra escala tiene del orden de 10^{23} unidades estructurales, un número tan enormemente grande que hace que **sea imposible la aplicación de las leyes mecánicas a su estudio**¹. Por ejemplo, **si tenemos un modelo clásico de un sistema, para caracterizar completamente un microestado del sistema deberíamos especificar los valores de las posiciones y los momentos de cada una de sus unidades estructurales.**

1.1.3. Descripción macroscópica

Se puede intentar describir las propiedades del sistema aplicándole las leyes de la Mecánica como un todo, ignorando su posible microestructura, a lo que llamaremos **Mecánica Macroscópica**. Sin embargo, este intento fracasa: **existen multitud de fenómenos cuyos detalles no pueden caracterizarse utilizando la Mecánica Macroscópica** como, por ejemplo, los fenómenos de calentamiento y enfriamiento, cambios de estado, y otros.

Aquí es donde entra en juego la **Termodinámica**: el formalismo teórico que se encarga de estudiar los fenómenos que ocurren en sistemas macroscópicos. En particular, estamos interesados en la **Termodinámica del Equilibrio**: aquella que se ocupa de unos estados particularmente simples que se llaman de **equilibrio termodinámico**, que se pueden describir con un **número reducido de variables observables a nuestra escala** (como presión, deformación, volumen, etc.) que no son propiedades de sus constituyentes, sino que son propiedades del conjunto del sistema, resultado del comportamiento colectivo del ingente número de unidades estructurales que lo componen. A la descripción del estado con estas variables se la denomina **descripción macroscópica**, y a cada una de las formas de describir el sistema de este modo, **MACROESTADO**.

1.1.4. Descripción estadística o probabilística: Física Estadística

Ante la imposibilidad de la descripción completamente microscópica, para relacionar las variables de estado termodinámicas con la “configuración microscópica” del sistema de acuerdo con un modelo, se opta por una descripción “híbrida” que involucra aspectos de la descripción microscópica y de la macroscópica, relacionadas a través de la Teoría de la Probabilidad.

La descripción estadística es la que se utiliza en Física Estadística. En ella, se especifica un macroestado y se enumeran los microestados que son compatibles con el macroestado elegido. A cada uno de los microestados se le asigna una probabilidad. **El problema que se resuelve en Física Estadística es la inferencia de las probabilidades de los microestados a partir de las leyes de la Teoría de la Probabilidad y de unas pocas restricciones macroscópicas conocidas.** Ello permite obtener, como veremos, **toda la información termodinámica del sistema para el modelo de realidad elegido.**

1.2. Fases y tipos de variables

1.2.1. Sistemas homogéneos, heterogéneos. Fases

Se dice que un sistema macroscópico es **homogéneo** cuando dos partes cualesquiera de igual tamaño y forma son indistinguibles entre sí desde el punto de vista macroscópico. En otras palabras, cuando todas las propiedades observables a escala macroscópica son iguales en ambas partes.

Un sistema macroscópico se llama **heterogéneo** cuando está compuesto de distintas partes macroscópicas homogéneas, llamadas **fases**.

¹ En relación con esto, se propone al lector buscar información sobre el **demonio de Laplace**.

1.2.2. Variables intensivas, extensivas u otras

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.

Nótese que pueden existir variables que no sean ni intensivas ni extensivas. Tal es el caso, por ejemplo, del cuadrado del volumen o el cubo de la masa.

1.3. Procesos adiabáticos y Primer Principio

1.3.1. Primer Principio de la Termodinámica: argumento intuitivo

A continuación, daremos un argumento intuitivo, en el marco de la Mecánica Clásica newtoniana, para entender el significado físico de la magnitud que pasaremos a estudiar: la energía interna.

Un sistema sin microestructura se puede estudiar aplicando las leyes de la Mecánica Macroscópica. De modo que pueden utilizarse las leyes de Newton y sus consecuencias. Supongamos un sistema de este tipo, descrito en un sistema de referencia inercial, que consta de un conjunto de partículas que interactúan entre sí únicamente por medio de fuerzas conservativas (derivadas de una energía potencial interna $E_{p,int}$), y sobre el que además actúan unas fuerzas externas (en general no conservativas). El teorema trabajo-energía aplicado a dicho sistema de partículas indica que:

$$W_{ext} = \Delta E_{propia} \quad (1.1)$$

donde W_{ext} es el trabajo realizado por las fuerzas externas y la **energía propia** del sistema (E_{propia}) es igual a la suma de la energía cinética total y la energía potencial interna ($E_{p,int}$):

$$E_{propia} = \sum_{\substack{a \\ \text{todas las} \\ \text{partículas}}} \left(\frac{1}{2} m_a v_a^2 \right) + E_{p,int} \quad (1.2)$$

siendo m_a la masa de cada partícula y v_a el módulo de su velocidad. A partir de ahora, salvo que se indique lo contrario, nos referiremos siempre al trabajo de las fuerzas externas y omitiremos el subíndice “ext”.

Nos podemos hacer la siguiente pregunta: **para un sistema macroscópico real (con microestructura), ¿se cumple la ecuación (1.1)?** La respuesta es negativa, como veremos a continuación.

Supongamos que para un sistema con microestructura sigue teniendo sentido hablar de “energía propia” (cosa que postularemos a continuación en el Primer Principio). Por un lado, el sistema podría **intercambiar materia** (posibilidad no incluida en la ecuación (1.1)), lo que necesariamente variaría esta energía propia. Por otro lado, dado que el sistema macroscópico está hecho de una microestructura, podría ser que pudiera intercambiar energía de un modo que no está contemplado en la Mecánica Macroscópica y que no se debiera a intercambio de materia. Esta posibilidad es un hecho experimental ampliamente comprobado. A dicho modo de intercambio de energía, que englobaría a todo aquel distinto del trabajo de fuerzas macroscópicas y del intercambio de materia, lo llamaremos **calor** (y lo definiremos cuantitativamente más adelante). Los procesos por los que ocurre la transferencia de calor se clasifican tradicionalmente en tres:

- **Conducción:** interacciones (como “choques”) entre las unidades estructurales, de un modo “caótico” que no puede ser expresado como una fuerza equivalente macroscópica.
- **Radiación:** emisión y absorción de radiación electromagnética.

- **Convección:** originado por movimientos de un fluido debidos a “gradientes” de ciertas propiedades dentro de él o forzados desde el exterior, involucrando a su vez a los otros dos modos de transferencia de calor.

Nótese que un sistema sin microestructura no podría experimentar ni intercambio de calor ni de materia. Y si un sistema estuviera rodeado de un entorno sin microestructura tampoco podría experimentarlos, dado que el entorno no podría dar o recibir ni materia ni calor.

Así, surge la noción de **proceso adiabático**: es un cambio de estado de un sistema en el que **la microestructura del entorno es irrelevante**². Es decir, un proceso es adiabático si, para describir dicho cambio de estado, es posible sustituir el entorno por un *sistema sin microestructura* que en cada instante de tiempo ejerza sobre el sistema las mismas fuerzas macroscópicas externas.

Un sistema que sólo puede realizar procesos adiabáticos se denomina **adiabáticamente aislado** o **térmicamente aislado**.

1.3.2. Primer Principio de la Termodinámica: enunciado

Primer Principio de la Termodinámica. Para todo sistema macroscópico existe una magnitud E_{propia} que llamamos energía propia, tal que su variación durante un proceso adiabático es igual al trabajo de las fuerzas macroscópicas externas.

El Primer Principio es un enunciado universal de la ley de existencia y conservación de la energía. La energía se puede “intercambiar” entre unos sistemas y otros, pero no puede “crearse” ni “destruirse”. Es imposible construir *móviles perpetuos de primera especie*, esto es, dispositivos que proporcionen más energía de la que consumen.

Normalmente se usa una versión ligeramente más restringida de esta ley, que se aplica a sistemas que se encuentran en **reposo macroscópico**, en los que **cada una de sus partes tiene momento lineal y momento angular nulo** (respecto a un sistema de referencia inercial). En tal caso, no existe energía cinética de movimiento de sus partes. Desde el punto de vista microscópico, la energía cinética se debe al movimiento de las unidades estructurales respecto al centro de masas del sistema. En tales circunstancias, llamaremos **energía interna** a la energía propia³. A partir de ahora, salvo que se indique lo contrario, supondremos que el sistema se encuentra en reposo macroscópico y hablaremos de energía interna. El Primer Principio se escribe, en tal caso:

Primer Principio de la Termodinámica (versión restringida). Para todo sistema macroscópico que esté en reposo macroscópico existe una magnitud U , que llamamos energía interna, tal que su variación durante un proceso adiabático es igual al trabajo de las fuerzas macroscópicas externas.

$$\boxed{W^{ad} = \Delta U} \quad (1.3)$$

1.3.3. Trabajo de configuración y trabajo disipativo⁴

En general, el trabajo realizado sobre un sistema en reposo macroscópico puede dividirse en dos clases:

² Una vez se definen con precisión los conceptos de energía propia *del sistema macroscópico* y calor (cosa que aún no hemos hecho), se ve inmediatamente que **un proceso adiabático es aquel en el que la transferencia de energía se debe únicamente a trabajo de fuerzas macroscópicas externas, sin transferencias de calor ni materia.**

³ Para definir la energía interna cuando hay movimiento macroscópico de las partes del sistema hay que hacer consideraciones adicionales que no mencionaremos aquí. El significado de la expresión “energía interna” que estamos usando puede diferir ligeramente del significado con el que la usan algunos autores.

⁴ Para una explicación más detallada del cálculo del trabajo realizado sobre un sistema macroscópico, véase *B. Diu, C. Guthmann, D. Lederer, B. Roulet, “Thermodynamique”, Hermann, 2007.*

- **Trabajo de configuración:** aquel que se puede escribir en función del cambio de una cierta variable del sistema distinta de la energía interna, como su volumen, longitud, carga eléctrica, masa, etc.
- **Trabajo disipativo:** el caso contrario. Por ejemplo, el trabajo realizado por fuerzas de rozamiento, de viscosidad, disipación de energía eléctrica por efecto Joule en una resistencia, etc.

Más precisamente, el trabajo de configuración es aquel que entre dos instantes t_1 y t_2 se puede escribir:

$$W = \sum_{i=1}^k \int_{t_1}^{t_2} A_i \frac{da_i}{dt} dt \quad (1.4)$$

O, dicho de otro modo, tal que el “trabajo elemental” se pueda poner en la forma:

$$\delta W = \sum_{i=1}^k A_i da_i \quad (1.5)$$

En las ecuaciones (1.4) y (1.5):

- 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...).

Pongamos un par de ejemplos sencillos de trabajo de configuración:

- Sistema de volumen V sometido a una presión externa p : $\delta W = -pdV$.
- Hilo elástico de longitud L sometido a una tensión externa τ : $\delta W = \tau dL$.

El trabajo de configuración es importante en el formalismo termodinámico porque sirve para distinguir variables del sistema que son relevantes para definir un macroestado (los desplazamientos generalizados). El trabajo disipativo, sin embargo, no permite identificar ninguna variable del sistema relevante para el macroestado. Para más detalles, véase la sección 1.9.1.

Nótese que estamos considerando positivo el trabajo realizado sobre el sistema y negativo el que el sistema cede al entorno.

1.3.4. Definición de calor

Se define el **calor intercambiado por un sistema macroscópico en un proceso** como la **energía intercambiada que no puede ser calculada ni como trabajo mecánico ni como intercambio de materia**. En el caso sencillo de **un sistema que no intercambia materia**, el calor Q se define simplemente por la relación:

$$Q = \Delta U - W \quad (1.6)$$

Si el sistema puede intercambiar materia, a veces no está claro el “límite” entre las distintas formas de intercambio de energía (en particular, entre la energía intercambiada por calor y por intercambio de materia), y la definición de calor no es evidente.

De igual modo que con el trabajo, consideramos positivo el calor absorbido por el sistema y negativo el cedido al entorno.

1.4. Características de la energía de un sistema

1.4.1. Tipos de fuerzas macroscópicas

Atendiendo a su **alcance**, distinguimos **dos clases de fuerzas macroscópicas**:

- Las **fuerzas macroscópicas de corto alcance** son aquellas que solo son apreciables cuando existe contacto directo entre cuerpos. Son debidas a fuerzas microscópicas cuya acción se aprecia solamente a distancias del orden de la distancia interatómica, por lo que, cuando se ponen dos cuerpos en contacto, solo las unidades estructurales situadas en las superficies en contacto sienten dichas fuerzas. Cuando sólo existen estas fuerzas se dice que hay **acoplamiento débil** entre los cuerpos.
- Las **fuerzas macroscópicas de largo alcance** existen entre cuerpos que no están en contacto directo, conllevando la presencia de **campos macroscópicos gravitatorios o electromagnéticos**. Cuando existen, se dice que hay **acoplamiento fuerte** entre los cuerpos.

Hemos de señalar que las fuerzas microscópicas responsables de las fuerzas macroscópicas a corta distancia pueden dar lugar a lo que se conoce como **efectos de superficie**. Éstos consisten en el comportamiento particular de la superficie de contacto entre dos cuerpos de naturaleza diferente (por ejemplo, límites de grano en un sólido, interfaz entre un líquido y un vapor, etc.), que es distinto del comportamiento de cada uno de los cuerpos. Cuando no se puede despreciar la influencia de estos efectos en relación con las propiedades del sistema que desean ser estudiadas, el modelo macroscópico debe incluirlos. Las propiedades dependerán, en particular, de la forma de los cuerpos. Por **ejemplo**, en una gota de agua suficientemente pequeña los efectos debidos a la tensión superficial hacen que su superficie se comporte de forma diferente que el interior y son suficientemente importantes como para influir en la energía de la gota.

1.4.2. Posibilidad de definir un sistema según las características de su energía

Es condición necesaria para decir que un cuerpo macroscópico en reposo es un verdadero sistema (cuyo estado se puede describir separadamente del de su entorno) que sea posible escribir su energía total como la suma de dos términos independientes:

- **Una energía potencial externa** que se deba a campos macroscópicos eléctricos o gravitatorios **no debidos al propio cuerpo ni influidos por él**.
- **Una energía interna.**

Para ello, **el campo creado por el propio cuerpo debe tener una influencia despreciable en los cuerpos que son fuentes del campo externo** (pues, de lo contrario, dicho campo no sería verdaderamente “externo”, sino que tendría carácter también “interno” por el efecto del propio cuerpo en él).

Al ejemplo de las fuerzas gravitatorias dado en la sección 1.1 podemos añadir otro inspirado en la electrostática:

- No sería posible considerar como sistemas independientes a dos partes de un líquido polar (como el agua) bajo la influencia del campo eléctrico creado por un conductor. El campo eléctrico externo (el creado por el conductor) hace aparecer momentos dipolares macroscópicos en el líquido, que a su vez crean su propio campo. Cuando se alcanza el equilibrio electrostático, el valor del momento dipolar y del campo eléctrico total (suma del externo más el de polarización) en cada punto del líquido depende de cuánto se han polarizado el resto de los puntos, y por ello la energía de un punto del líquido **no** se puede escribir como la suma de una parte estrictamente interna y otra parte estrictamente externa.

- Sí que sería posible considerar como sistemas independientes a dos partes de un gas a baja presión en la situación del punto anterior. Esto se debe a que el campo eléctrico debido a un gas polarizado poco denso es despreciable en comparación con el campo que lo polariza, lo que no ocurre en el caso de un líquido polar.

De acuerdo con ello, cuando hay acoplamiento débil es posible hablar de varios sistemas macroscópicos independientes. De hecho, en este caso, cada una de las partes macroscópicas de un sistema puede considerarse, a su vez, como un sistema, al que se llama **subsistema** del primero.

Sin embargo, cuando existe acoplamiento fuerte hay que analizar más en profundidad la posibilidad de describir un cuerpo como un sistema independiente.

1.4.3. Extensividad (o no) de la energía interna

En acoplamiento débil, la energía interna es extensiva. Sin embargo, **en acoplamiento fuerte** en general no, ya que podría no poderse ni siquiera definir adecuadamente. **Por ejemplo, en una estrella** existen fuerzas gravitatorias intensas entre sus diferentes partes, lo que impide definir adecuadamente la energía interna de cada parte y por ello la energía interna no es extensiva.

1.5. Componentes y entidades elementales independientes

En esta sección nos restringimos al caso de un sistema macroscópico en **situación de acoplamiento débil**. Para estos sistemas, la energía propia puede considerarse como la suma de las energías propias de sus partes macroscópicas. Sin embargo, podemos ir más allá y preguntarnos cuáles son los **entes conceptuales idénticos más simples que, de acuerdo con un modelo de realidad, permiten escribir la energía propia del sistema como la suma de las energías propias de dichos entes**. Éstos reciben el nombre de **entidades elementales independientes**.

La descripción de la cantidad de materia de un sistema es complicada, dada la enorme variedad de estructuras y tipos de interacciones que pueden existir entre las unidades estructurales que lo componen. Desde el punto de vista de la Física Estadística que veremos en este curso, es conveniente cuantificar la cantidad de materia en términos de estas entidades elementales independientes.

Las entidades elementales independientes pueden constar de una o varias unidades estructurales (como átomos, moléculas u otros); o incluso pueden no ser objetos materiales, como en el caso del modelo de cristal aislante monoatómico ideal, en el que pueden considerarse como entidades elementales los distintos modos normales de vibración independientes de la red (se ve más adelante en la asignatura).

Así, por ejemplo, en el modelo de gas ideal monoatómico la entidad elemental sería el átomo, mientras que en un gas ideal diatómico la entidad elemental sería la molécula diatómica. Fowler y Guggenheim clasifican los líquidos en normales, en los que la entidad elemental sería la molécula, y asociados, en los que, debido al acoplamiento entre las variables que describen mecánicamente las diferentes moléculas, la entidad elemental sería un agregado supramolecular, cuyo tamaño podría incluso variar dinámicamente.

Desde este punto de vista podemos llamar **número de componentes de un sistema** al **número mínimo de tipos distintos de entidades elementales independientes que, de acuerdo con un modelo de realidad, es necesario especificar para describir cómo está constituido un sistema**.

La anterior definición habla de “número mínimo de tipos [...] que [...] es necesario especificar” porque pueden existir reacciones químicas (donde unas unidades estructurales se transforman en otras) o coexistencia de distintas fases en interacción (donde unas unidades estructurales pasan de una fase a otra). En tales casos, el número de tipos distintos de entidades elementales independientes presentes no coincide, en general, con el número de componentes, ya que las cantidades relativas de distintos tipos de entidades elementales están relacionadas y no se pueden variar independientemente unas de otras.

Se distingue entre:

- Sistemas **monocomponentes**: compuestos por un solo componente.
- Sistemas **multicomponentes**: compuestos por varios componentes.

1.6. Paredes

Habitualmente, a los distintos tipos de intercambios de energía o materia se les llama interacciones, y se suele distinguir entre:

- **Interacción mecánica** (intercambio de trabajo).
- **Interacción material** (intercambio de materia).
- **Interacción térmica** (intercambio de calor).

Podemos definir las **paredes** como entes conceptuales que establecen el límite entre un sistema y su entorno, y que tienen las siguientes propiedades:

- Pueden ser reales o imaginarias.
- **Establecen o prohíben ciertas interacciones (intercambios de materia y energía).**
- “Lo aguantan todo”: permanecen **inalteradas** al establecer o prohibir los intercambios de materia y energía.
- No se consideran como parte del sistema.

Al ser entes conceptuales, no tienen por qué existir en la realidad. Son idealizaciones que nos permiten desarrollar el formalismo termodinámico. En el caso de paredes reales, un ente físico puede ser considerado como pared en la medida de que pueda establecer o prohibir un cierto intercambio *sin sufrir modificaciones apreciables en su propio estado*.

Tipos de **paredes permisivas** (que permiten algún tipo de interacción):

- **Móviles**: permiten interacción mecánica.
- **Permeables**: permiten interacción material.
- **Diatérmanas**: permiten interacción térmica.

Tipos de **ligaduras o paredes prohibitivas** (impiden algún o algunos tipos de interacción):

- **Fijas y rígidas**: prohíben interacción mecánica.
- **Impermeables**: prohíben interacción material.
- **Semipermeables**: prohíben el intercambio de algunos componentes, pero de otros no.
- **Adiabáticas**: impiden interacción material y térmica.
- **Aislantes**: impiden toda interacción.

Las paredes pueden tener varios de los anteriores atributos siempre que no sean mutuamente excluyentes.

Los sistemas macroscópicos también se pueden clasificar en función de sus paredes:

- **Abierto** (paredes permeables, móviles y diatérmanas).
- **Cerrado** (paredes impermeables, móviles y diatérmanas).
- **Térmicamente aislado** o **adiabáticamente aislado** (paredes adiabáticas y móviles).
- **Aislado** (paredes aislantes).

1.7. Estado de equilibrio

Nosotros denominaremos en lo que sigue **sistema termodinámico** a todo sistema que cumpla las siguientes condiciones:

- **Acoplamiento débil.** Los campos gravitatorios o electromagnéticos macroscópicos debidos al propio cuerpo son de intensidad despreciable para los fenómenos que se desean estudiar.
- **Ausencia de inhomogeneidades debidas a campos externos.** Los campos gravitatorios o electromagnéticos macroscópicos externos no deben hacer que las variables intensivas dependan de la posición. Por ejemplo, el sistema debe ser suficientemente pequeño como para que el campo gravitatorio no produzca dependencia de la presión con la altura.
- **Ausencia de histéresis.** Es posible encontrar relaciones entre las variables de dicho sistema que no dependen de su historia pasada, sino solamente de los valores de dichas variables. Por ejemplo, un imán ferromagnético presenta relaciones entre el campo magnético y la magnetización que dependen de su historia pasada, y no sería un sistema termodinámico.

Hay que mencionar que no existe (por el momento) una teoría termodinámica general aplicable cuando existe acoplamiento fuerte o histéresis.

1.7.1. Definición de equilibrio

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:

- 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.

Pueden existir estados que sean estacionarios, pero no de equilibrio. Es el caso, por ejemplo, de una barra conductora del calor con sus extremos mantenidos a temperaturas constantes pero distintas (Figura 1.1): aunque la temperatura de cada punto de la barra llega a ser constante, existe un flujo constante de calor que atraviesa la barra desde el extremo más caliente hasta el extremo más frío. Para mantener este estado estacionario es necesario un dispositivo externo que aporte la energía necesaria para que el extremo caliente no se enfríe y retire la energía necesaria para que el extremo frío no se caliente.

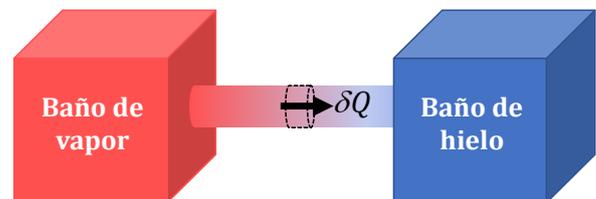


Figura 1.1: Estado estacionario producido tras la unión de un baño frío y un baño caliente por medio de una barra conductora del calor.

1.7.2. Definición de equilibrio mutuo

Muchas veces se dice que “un sistema está en equilibrio con otro”. En este caso, la palabra equilibrio se refiere al concepto de equilibrio mutuo entre dos o más sistemas, que es diferente del de estado de equilibrio. No debe confundirse el equilibrio termodinámico con el equilibrio mutuo.

Se dice que **dos sistemas en equilibrio termodinámico y sin ligaduras internas están en equilibrio mutuo** (o que **uno está en equilibrio con el otro**) cuando si se pusieran en contacto por medio de una pared no se modificaría ninguna de sus propiedades macroscópicas.

- Si la pared solo permite intercambio de calor, se dice que están en **equilibrio térmico mutuo**.
- Si la pared solo permite intercambio de trabajo, se dice que están en **equilibrio mecánico mutuo**.
- Si la pared solo permite intercambio de materia, se dice que están en **equilibrio material mutuo**.

Un sistema puede presentar uno o varios de estos tipos de equilibrio mutuo, en función de las interacciones que permita la pared. Si la pared permite cualquier tipo de intercambio, se dice que están en **equilibrio mutuo total**.

1.7.3. Principio General de la Termodinámica

El Principio General de la Termodinámica se basa en la generalización de hechos experimentales y afirma que:

Principio General de la Termodinámica. Todo sistema termodinámico aislado que no se encuentra en equilibrio termodinámico alcanza tras un intervalo de tiempo finito y no nulo un estado de equilibrio termodinámico que no puede abandonar a no ser que deje de estar aislado.

Es evidente que el estado de equilibrio es una idealización. Surge entonces la siguiente pregunta: en la realidad física, *¿cuándo será válida una descripción basada en estados de equilibrio?*

El estudio de los sistemas físicos es esencialmente experimental y por ello la descripción que se haga de ellos se ha de adaptar a las propiedades que se quieran estudiar y a los tiempos característicos de los procesos físicos que tengan lugar durante el experimento. Así pues, se pueden diferenciar tres situaciones:

1. Cuando *los tiempos característicos de los procesos son muy cortos en comparación con los tiempos entre medidas* (como, por ejemplo, al variar el volumen de un recipiente en el que se encuentra un gas, ya que éste alcanza el equilibrio de nuevo enseguida), el sistema alcanza el equilibrio rápidamente y es posible considerar que se miden estados de equilibrio.
2. Cuando, en el otro extremo, *los tiempos característicos de los procesos son muy largos en comparación con los tiempos durante los que se mide* (como, por ejemplo, en un proceso de oxidación lenta), se puede considerar que las interacciones que sufre el sistema son despreciables y que las propiedades permanecen constantes. En otras palabras, se puede suponer que el sistema no abandona el equilibrio.
3. Entre estos dos casos está la situación intermedia en la que *ambos tiempos son del mismo orden*. En este caso, el sistema está esencialmente fuera del equilibrio y la Termodinámica del Equilibrio no es aplicable.

1.7.4. Macroestado en un estado de equilibrio

El número de variables de estado macroscópicas necesarias para determinar un macroestado de equilibrio se denomina **número de grados de libertad termodinámicos** del sistema.

La expresión "**variables de estado**" se reserva para las variables independientes escogidas para representar el macroestado. Las demás se pueden poner en función de las anteriores y se denominan **funciones de estado**. Si un sistema tiene n grados de libertad, un estado de equilibrio de dicho sistema queda caracterizado por los valores x_1, \dots, x_n de sus n variables de estado en dicho estado. La n -upla (x_1, \dots, x_n) se identifica con un punto de \mathbb{R}^n , espacio al que se le llama **espacio termodinámico** asociado a dicho sistema. Obsérvese que a un mismo sistema se le pueden asociar distintos espacios termodinámicos, en función de las variables que se elijan para caracterizar los estados de equilibrio. Es importante señalar que **en un espacio termodinámico solo se pueden representar estados de equilibrio**.

1.7.5. Procesos termodinámicos

Se denomina **proceso termodinámico** al paso de un sistema de un estado inicial de equilibrio a otro final de equilibrio.

De entre los procesos que sirvieron para desarrollar la Termodinámica, juega un papel especialmente importante el **ciclo** o **proceso cíclico**, que es aquel en el que el estado inicial y el final coinciden. Por otro lado, un **proceso espontáneo** es aquel que se produce al eliminar una ligadura (esto es, al eliminar alguna pared que impida algún tipo de interacción). El tiempo desde que se inicia el proceso espontáneo hasta que se alcanza el equilibrio termodinámico de nuevo se llama **tiempo de relajación**.

El Principio General de la Termodinámica limita el modo en que un proceso puede tener lugar:

- **Un sistema en equilibrio termodinámico no puede salir de él si no interacciona con el entorno.** Esto es debido a que, si pudiera salir sin interacción con el entorno, entonces un sistema aislado podría salir del estado de equilibrio sin dejar de estar aislado, lo que va contra el Principio General de la Termodinámica.
- **No son posibles procesos reales en los que todos los estados intermedios son de equilibrio termodinámico.** En efecto, supóngase que se tiene un sistema en equilibrio termodinámico. Para que pase a un nuevo estado es necesario que interaccione con el entorno de algún modo (como acabamos de deducir del Principio General de la Termodinámica), saliendo así del equilibrio termodinámico. Cuando cesa la interacción, ha de pasar un cierto intervalo de tiempo hasta que alcance de nuevo el equilibrio termodinámico, suponiendo que en ese tiempo no vuelve a interaccionar con el entorno.

No obstante, es útil trabajar con unos procesos idealizados que son casos límite de los procesos reales, llamados **procesos cuasiestáticos**, en los que se supone que todos los estados intermedios son de equilibrio termodinámico. Se pueden representar como curvas continuas en el espacio termodinámico. Se trata de buenas aproximaciones cuando las interacciones del sistema “apenas perturban el equilibrio”, realizándose de “modo lento”. Más concretamente, un proceso puede aproximarse satisfactoriamente por uno cuasiestático cuando la velocidad media de variación de las magnitudes termodinámicas es mucho menor que la velocidad de variación de éstas correspondiente al proceso espontáneo de relajación.

1.8. Hacia el postulado fundamental: Boltzmann

1.8.1. Direccionalidad de los procesos: la idea de Boltzmann

El Primer Principio indica que todos los procesos naturales deben conservar la energía. Pero es necesario un Segundo Principio para explicar otros aspectos. Por un lado, está la **direccionalidad e irreversibilidad de los procesos espontáneos**:

- Los cafés se enfrían, pero no se calientan espontáneamente.
- Un gas a presión en una botella tiende a salir, pero no se vuelve a comprimir espontáneamente.
- El café y la leche se mezclan, pero no se separan espontáneamente.

Relacionada con lo anterior, también está la **asimetría en la transformación continua (cíclica) de energía entre trabajo y calor**. En un proceso cíclico:

- Es posible transformar todo el trabajo en calor cedido a un único foco térmico.
- Es imposible transformar todo el calor absorbido de un único foco térmico en trabajo sin que exista una cesión de calor a otro foco térmico más frío (no existen los móviles perpetuos de segunda especie).

El **postulado de máxima entropía** (debido a los físicos László Tisza y Herbert B. Callen) es un “enunciado avanzado” del Segundo Principio de la Termodinámica, que se basa en las ideas estadísticas de Ludwig Boltzmann, y que permite establecer toda la estructura formal de la Termodinámica a partir de un único enunciado adicional a los que ya hemos mencionado. Antes de enunciar dicho postulado, estudiaremos con ayuda de un ejemplo las ideas de Boltzmann (sección 1.8.2).

Partimos de querer analizar la direccionalidad de los procesos espontáneos en la naturaleza. De acuerdo con Callen, esto se reduce a estudiar el **problema básico de la Termodinámica: la determinación del estado de equilibrio al que llegará un sistema aislado al suprimirse alguna de sus ligaduras internas**. En efecto, el conjunto de sistema+entorno se suele considerar un sistema aislado. Cuando hablamos de “direccionalidad”, nos referimos a cómo es la tendencia al equilibrio de dicho conjunto sistema + entorno. El estado final de equilibrio determina el sentido de la evolución.

Al conjunto de sistema + entorno se le suele llamar **universo termodinámico**, que es **un sistema finito y aislado, pero no es el universo en sentido cosmológico**. A escalas suficientemente grandes, las hipótesis iniciales de la Termodinámica, ya mencionadas anteriormente, no se cumplen.

La **idea central de Boltzmann** para explicar la direccionalidad e irreversibilidad de los procesos, que aparentemente contradecía la reversibilidad temporal de las ecuaciones de la Mecánica, es la siguiente:

- Para un macroestado dado, las unidades estructurales del sistema se pueden configurar de muchas maneras: **existen muchos microestados compatibles con un macroestado**.
- Hay **macroestados que pueden ocurrir de MUCHÍSIMAS más formas que otros**.
- Para un **sistema aislado**, cuya evolución *no está gobernada por influencias exteriores*, el **macroestado final de equilibrio** está determinado “probabilísticamente”: es aquel que puede ocurrir de un mayor número de maneras, es decir, **aquel para el que el número de microestados compatible es el máximo (CON MUCHA DIFERENCIA) de entre todos los posibles**.

Insistimos: esta tendencia “probabilística” solo se da en sistemas aislados. **Si el sistema no es aislado**, la influencia del intercambio energético con el exterior compite con la tendencia “probabilística” y **el estado final de equilibrio, en general, no es el de mayor número de microestados**.

1.8.2. Un ejemplo cuantitativo

Para hacernos una idea de cómo de extremas son las expresiones “MUCHÍSIMAS” Y “CON MUCHA DIFERENCIA” de los párrafos anteriores, propongamos un ejemplo sencillo: **un conjunto de moléculas independientes y distinguibles entre sí, encerradas en un recipiente aislado, y que se encuentra en equilibrio termodinámico**.

Sea N el número total de moléculas. Dividamos el recinto en dos mitades 1 y 2, y sean N_1 y N_2 los números de moléculas que hay en sendas mitades, de tal forma que $N = N_1 + N_2$.

¿De cuántas formas, $\Omega(N_1)$, puede ocurrir una distribución de moléculas (N_1, N_2) ? La respuesta se obtiene aplicando la Combinatoria (véase la sección 3.1): es el número de permutaciones con repetición de N elementos (moléculas) en las que hay N_1 iguales (metidas en la mitad 1) y N_2 iguales (en la mitad 2):

$$\Omega(N_1) = \frac{N!}{N_1! N_2!} = \frac{N!}{N_1! (N - N_1)!} \quad (1.7)$$

En nuestro ejemplo, N_1 determinaría el “macroestado”, es decir, cuántas moléculas, globalmente, se encuentran en una de las mitades. Cada una de las posibles distribuciones detalladas de las moléculas que dan lugar a un mismo valor de N_1 sería un “microestado” (compatible con el macroestado N_1). El número Ω sería lo que llamamos “número de microestados”.

El número total de microestados (de formas de repartir las moléculas entre las dos mitades, independientemente del valor de N_1) es 2^N . Este número puede obtenerse pensando que cada molécula puede estar en 2 mitades distintas y que el lugar que ocupe una molécula es independiente del lugar que ocupa cualquier otra, de modo que habrá $2 \cdot 2 \cdot 2 \dots 2 = 2^N$ formas en total. También puede obtenerse este número sumando todos los valores posibles de (1.7), teniendo en cuenta la fórmula del binomio de Newton:

$$(a + b)^N = \sum_{k=0}^N \frac{N!}{k! (N - k)!} a^k b^{N-k} \quad (1.8)$$

en donde $a = b = 1$. Debido a que el recipiente está aislado, el Principio de Indiferencia nos dice que no hay ninguna razón para pensar que una de estas ordenaciones tiene una probabilidad mayor de ocurrir que el resto, de igual modo que no hay razón para pensar que en un dado “perfecto” una de las caras tiene más probabilidad que las otras (para más detalles sobre Probabilidad, véase la sección 2.2). Así que la probabilidad $\Pr(N_1)$ del macroestado con N_1 moléculas en la mitad 1 es:

$$\Pr(N_1) = \frac{\text{(casos favorables)}}{\text{(casos posibles)}} = \frac{\Omega(N_1)}{2^N} = \frac{N!}{2^N \cdot N_1! (N - N_1)!} \quad (1.9)$$

Para $N = 4$ se obtienen los siguientes resultados:

Macroestado, N_1	Número de microestados, $\Omega(N_1)$	Probabilidad del macroestado, $\Pr(N_1)$
0	1	0.0625
1	4	0.25
2	6	0.375
3	4	0.25
4	1	0.0625

Ya en este caso sencillo se aprecia que hay un macroestado más probable y algunos poco probables. Hay menos formas de que todas las moléculas se agrupen en una mitad (solo 2 formas) que de que estén distribuidas de forma homogénea entre ambas mitades (6 formas).

Aumentemos el número de moléculas a $N = 100$. En este caso sigue habiendo 2 formas de que las moléculas se agrupen en una de las dos mitades, pero sin embargo hay $1.01 \cdot 10^{29}$ formas de que se distribuyan homogéneamente entre sus dos mitades (50 moléculas en cada mitad). Para el caso $N = 1000$, **hay $2.7 \cdot 10^{299}$ formas de repartir 500 moléculas en cada mitad**, un número tan fantásticamente grande que se acerca a los límites de la comprensión humana.

La probabilidad de cada macroestado en función de N_1 se representa en la Figura 1.2 ($N = 100$ moléculas) y en la Figura 1.3 ($N = 1000$ moléculas):

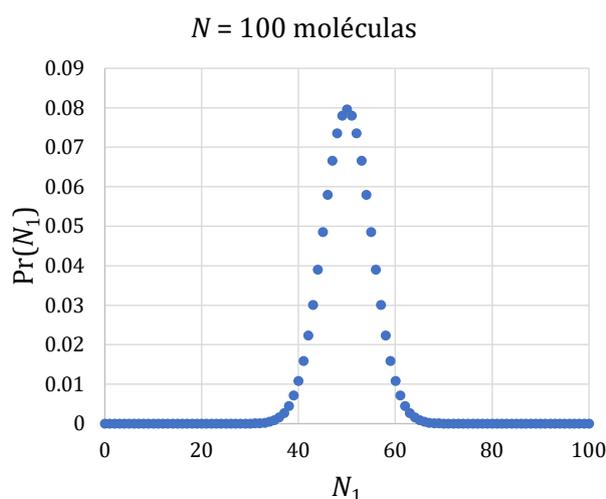


Figura 1.2: Probabilidad de encontrar N_1 moléculas en la primera mitad del recipiente para $N = 100$ moléculas.

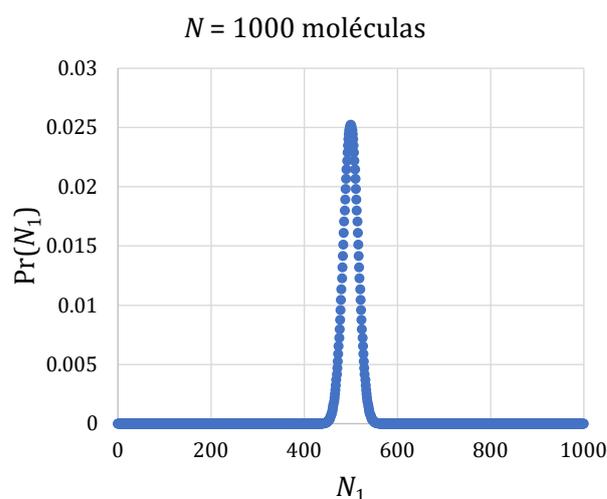


Figura 1.3: Probabilidad de encontrar N_1 moléculas en la primera mitad del recipiente para $N = 1000$ moléculas.

A medida que aumenta el número de moléculas, se observa una tendencia de la distribución de probabilidad a “estrecharse”, es decir, a tomar valores apreciablemente distintos de cero solamente en

la vecindad de un cierto valor, que es el más probable. Observemos los siguientes pasos de dicha tendencia tomando $N = 10\,000$ moléculas (Figura 1.4) y $N = 100\,000$ moléculas (Figura 1.5):

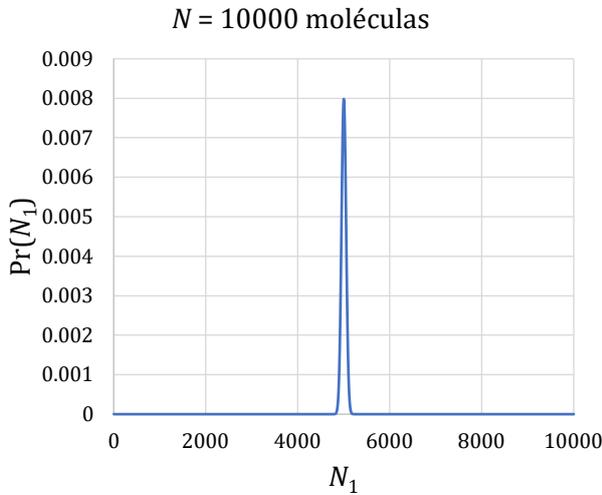


Figura 1.4: Probabilidad de encontrar N_1 moléculas en la primera mitad del recipiente para $N = 10\,000$ moléculas.

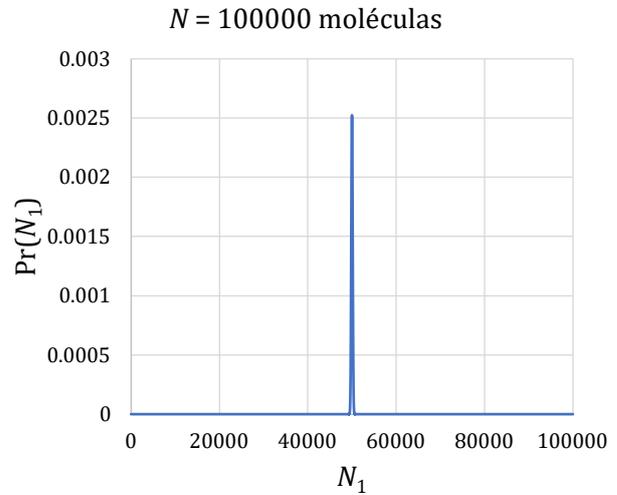


Figura 1.5: Probabilidad de encontrar N_1 moléculas en la primera mitad del recipiente para $N = 100\,000$ moléculas.

Para números de moléculas del orden de los de un sistema macroscópico, $N \sim 10^{23}$, **prácticamente toda la probabilidad recae sobre el macroestado más probable y algunos muy próximos a él**, que difieren en una cantidad relativamente despreciable del valor de $N_1 = N/2$.

Del razonamiento de Boltzmann se extrae la conclusión de que **la direccionalidad de los procesos se debe simplemente a que los procesos inversos son extremadamente improbables** (tan improbables, que, en la práctica, resultan imposibles de observar). Por eso si dejamos 10^{23} moléculas en una mitad del recipiente, tenderán a distribuirse de forma homogénea por todo él, pero nunca observaremos que espontáneamente se agrupen en una de las mitades. El sistema permanecerá fluctuando entre macroestados muy próximos al más probable (tan próximos a él que seremos incapaces de distinguirlos macroscópicamente entre sí).

1.8.3. La entropía de un sistema aislado

La magnitud que hemos llamado “número de microestados” del macroestado de un sistema aislado, Ω , no es una magnitud extensiva. De hecho, es multiplicativa. Supongamos un sistema aislado dividido en 2 partes aisladas entre sí, una con Ω_1 microestados y otra con Ω_2 microestados, el número total de microestados del sistema conjunto, Ω , es el producto:

$$\Omega = \Omega_1 \cdot \Omega_2 \quad (1.10)$$

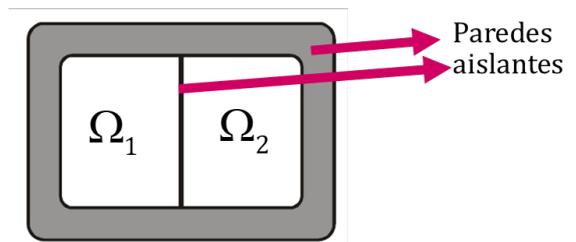
Sin embargo, el logaritmo del número de microestados sí que es una magnitud extensiva:

$$\ln\Omega = \ln\Omega_1 + \ln\Omega_2 \quad (1.11)$$

De aquí que se defina la **entropía del sistema aislado** como una magnitud proporcional al logaritmo del número de microestados:

$$S = k_B \ln\Omega \quad \text{Entropía de un sistema aislado (Ec. de Boltzmann – Planck)} \quad (1.12)$$

donde k_B es una constante denominada **constante de Boltzmann**, cuyo valor según la edición del Sistema Internacional de Unidades del 2019 es:



$$k_B = 1.380\,649 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1} \quad (1.13)$$

El valor de la constante de proporcionalidad se elige de esta forma para que la temperatura termodinámica, que definiremos después, se mida en kelvin (K). La entropía así definida es una **magnitud positiva o nula**, pues el número de microestados es, como mínimo, la unidad.

Lanzamos aquí una advertencia: la ecuación (1.12) que relaciona el número de microestados con la entropía **sólo es estrictamente correcta para el caso de los estados de equilibrio de un sistema aislado**. En el capítulo sobre el formalismo de la Física Estadística explicaremos esta afirmación de forma más precisa.

Dado que el logaritmo es una función creciente, la entropía será máxima cuando el número de microestados sea máximo. Y esto nos permite reformular la condición de equilibrio de un sistema aislado: **el macroestado de equilibrio de un sistema aislado es aquel para el que la entropía toma el valor máximo**. Esta afirmación está basada en que el máximo en el número de microestados es muy agudo y, aunque hemos tratado de ilustrar este hecho con el ejemplo de las moléculas en el recipiente (sección 1.8.2), no es tan simple como parece y, por ello, analizamos sus sutilezas a continuación.

Desde el punto de vista macroscópico, la afirmación de que, en un sistema aislado en equilibrio, la entropía S es máxima es mucho más precisa que la afirmación de que el número de microestados Ω es máximo. Veamos por qué. Como hemos visto en el ejemplo de las moléculas en el recipiente (sección 1.8.2), la probabilidad no nula recae sobre el estado con Ω máximo, pero también sobre los que son relativamente próximos a él, que también pueden ocurrir de un número enorme de formas, por lo que desde el punto de vista microscópico el sistema “fluctuará” entre estos macroestados relativamente muy próximos al máximo de Ω . **No obstante, dado el orden de magnitud de Ω y el carácter logarítmico de S , el valor de S para todos estos macroestados es, CON MUCHA PRECISIÓN, el mismo**. En efecto, supongamos que el macroestado con máximo número de microestados tiene Ω_m microestados y entropía $S_m = k_B \ln \Omega_m$, y supongamos otro macroestado con Ω microestados y entropía S . Entonces:

$$\frac{\Omega_m}{\Omega} = e^{(S_m - S)/k_B} \quad (1.14)$$

Ahora supongamos que $S_m - S$ es del orden de $1 \mu\text{J}/\text{K}$. Este es un valor muy pequeño, aproximadamente igual a la diferencia de entropía entre dos estados de 1 mol de un gas monoatómico a temperatura ambiente y presión cercana a la atmosférica (10^5 Pa) con el mismo volumen, pero cuyas presiones se diferencian en 0.01 Pa . Para dos macroestados que difieran tan poco en la entropía, se tiene:

$$\frac{\Omega_m}{\Omega} = e^{7.243 \cdot 10^{16}} = 10^{(\log_{10} e) 7.243 \cdot 10^{16}} = 10^{3.146 \cdot 10^{16}} \quad (1.15)$$

Es decir, el cociente entre el número de formas en las que puede ocurrir el macroestado más probable y el número de formas en el que puede ocurrir otro macroestado cuya entropía es **INSIGNIFICANTEMENTE MENOR** desde el punto de vista macroscópico es $10^{3.146 \cdot 10^{16}}$ (un “uno” seguido de aproximadamente $3.146 \cdot 10^{16}$ “ceros” en notación decimal). Este ejemplo numérico ilustra que **se requieren variaciones enormes en el número de microestados para obtener variaciones apreciables de entropía**. Por ello, un sistema aislado en equilibrio, aun “fluctuando” entre macroestados muy próximos al de Ω máximo, mantendrá su entropía constante **CON MUCHA PRECISIÓN** y, **desde el punto de vista macroscópico, el macroestado de equilibrio del sistema aislado es el de entropía máxima**.

1.9. Postulados de Callen

1.9.1. El postulado de estado de Callen

¿Qué variables definen el estado de equilibrio de un sistema homogéneo (de una sola fase)?

Notemos que, si el sistema estuviera aislado:

- No podría intercambiar trabajo, luego sus desplazamientos generalizados serían constantes.
- No podría intercambiar materia, luego el número de entidades elementales de sus componentes debería permanecer constante.
- No podría intercambiar tampoco calor, y, en consecuencia, ningún tipo de energía, por lo que su energía interna permanecería constante.

Así pues, podríamos determinar el estado de equilibrio del sistema utilizando su energía interna U , sus desplazamientos generalizados a_1, \dots, a_k y el número de entidades elementales de cada componente N_1, \dots, N_c .

Las variables $(U, a_1, \dots, a_k, N_1, \dots, N_c)$ reciben el nombre de **variables naturales de la entropía del sistema homogéneo**.

Si el sistema es **heterogéneo** (formado por varias fases), entonces las variables naturales de la entropía de tal sistema son el conjunto de **variables naturales de la entropía de cada una de sus fases (subsistemas homogéneos)**.

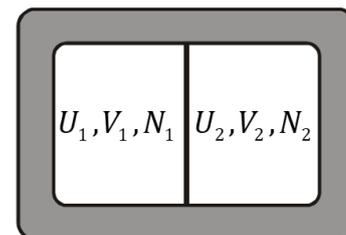
Con esta idea, Callen postula que:

Postulado de estado. El estado de equilibrio de un sistema termodinámico queda definido por los valores que toman las variables naturales de la entropía.

Para un sistema homogéneo, todas estas variables son independientes, pero si el sistema es heterogéneo, podrían no serlo por existir ligaduras entre ellas.

Ejemplo:

Supongamos un sistema aislado, separado en dos partes homogéneas por una pared aislante. Ambas partes solo pueden intercambiar trabajo de expansión ($\delta W = -pdV$) y están compuestas de un único componente, por lo que el estado del sistema aislado total está formado por el conjunto de las energías internas, volúmenes y número de entidades elementales de su único componente en cada uno de los dos subsistemas:



$$(U_1, V_1, N_1, U_2, V_2, N_2)$$

No obstante, estas variables no son independientes, ya que el sistema total está aislado, y la energía total U , el volumen V y el número total de entidades elementales N deben ser constantes:

$$\begin{aligned} U_1 + U_2 = U = \text{cte} &\Rightarrow dU_2 = -dU_1 \\ V_1 + V_2 = V = \text{cte} &\Rightarrow dV_2 = -dV_1 \\ N_1 + N_2 = N = \text{cte} &\Rightarrow dN_2 = -dN_1 \end{aligned} \quad (1.16)$$

de modo que las variables independientes que definen el estado del sistema total pueden tomarse como el conjunto (U_1, V_1, N_1) .

1.9.2. El postulado de máxima entropía de Callen

Aunque ya hemos dicho que la ecuación (1.12) solo es correcta para sistemas aislados, admitamos (sin dar más detalles) que la entropía es una función de estado que se puede definir en cualquier estado de

equilibrio de cualquier sistema aunque no esté aislado, y que su valor es el que tendría si ese sistema estuviera en ese mismo estado de equilibrio pero estuviera aislado.

Inspirado por la entropía de Boltzmann-Planck, Callen propone el siguiente postulado para establecer toda la estructura formal de la Termodinámica:

Postulado de máxima entropía. *Para todo sistema termodinámico, existe una función de estado extensiva, llamada entropía, S , que permite determinar el estado final de equilibrio al que llega el sistema aislado al eliminar una de sus ligaduras internas:*

- *Al eliminar una ligadura interna, existe un conjunto de estados de equilibrio del sistema aislado cuyos valores de las variables naturales de la entropía cumplen las restantes ligaduras.*
- *De entre todos ellos, el estado final de equilibrio es aquel para el que la entropía es máxima como función de sus variables naturales.*

La potencia de este postulado es enorme, ya que permite:

- Encontrar el sentido de los procesos espontáneos (hacia dónde fluye el trabajo, el calor y la materia) y las condiciones que caracterizan el estado final de equilibrio termodinámico.
- Definir la temperatura termodinámica como una magnitud completamente independiente de ningún termómetro o sustancia termométrica.
- Hallar la relación entre la entropía y las fuerzas generalizadas, como la presión.
- Definir el potencial químico como fuerza generalizada asociada al intercambio de materia.
- Calcular ciertas relaciones entre las variables termodinámicas en los estados de equilibrio, denominadas **ecuaciones de estado**, así como la interdependencia que existe entre dichas ecuaciones (**relaciones puente**).
- Encontrar las **condiciones de estabilidad** de los sistemas en equilibrio.

En el ejemplo anterior:

Al reemplazar la pared aislante intermedia por otra diferente que permita algún intercambio de energía o materia (o todo a la vez), se pone en marcha un proceso espontáneo hasta un estado final de equilibrio. Para obtener cuál es el estado final de equilibrio hay que maximizar la entropía del sistema total aislado:

$$S = S_1(U_1, V_1, N_1) + S_2(U_2, V_2, N_2) \quad (1.17)$$

sujeta a las ligaduras (1.16) y a las que imponga la pared intermedia.

1.10. Consecuencias del postulado de máxima entropía

1.10.1. Principio de Aumento de Entropía

Del postulado de máxima entropía se deduce inmediatamente un resultado muy importante que caracteriza la direccionalidad de los procesos: **en un sistema aislado, la entropía no puede disminuir.**

- Si la entropía del sistema aislado aumenta, se dice que ha sufrido un **proceso irreversible**.
- Si la entropía del sistema aislado permanece constante, se dice que ha sufrido un **proceso reversible**. Los procesos reversibles son un caso límite ideal.

Esta ley no es de aplicación si el sistema no es aislado. Pero existe una excepción importantísima, que termina de dar forma matemática al Segundo Principio: el de los sistemas adiabáticamente aislados o, equivalentemente, sistemas que siguen un proceso adiabático.

La idea clave es la siguiente ⁵:

- En un proceso adiabático, la energía solo puede variar por intercambio de trabajo.
- Para describir el trabajo, nos da igual si el entorno está hecho de unidades estructurales o no. A todos los efectos, el sistema “ve” al entorno como una fuerza externa macroscópica. **El proceso se puede describir como si el entorno no tuviera unidades estructurales.**
- Un sistema que no está hecho de unidades estructurales solo puede tener un microestado, y por ello su entropía es igual a cero.
- En conclusión, **el proceso adiabático del sistema puede describirse como si la entropía del entorno fuese cero.**

Esto no significa que la entropía del entorno sea *realmente* cero. Lo que pasa es que *a efectos de calcular cómo varía el estado del sistema durante el proceso adiabático*, podemos reemplazar el entorno por otro hipotético cuya entropía es igual a cero. Por lo tanto, como la entropía de sistema + entorno (sistema total aislado) no puede disminuir y durante el proceso todo ocurre para el sistema como si el entorno tuviera entropía constante e igual a cero, se obtiene la forma general del Principio de Aumento de Entropía:

Principio de Aumento de Entropía. En un proceso adiabático la entropía no puede disminuir:

$$\boxed{\Delta S_{\text{ad}} \geq 0} \quad (1.18)$$

De forma similar a cuando el proceso ocurre en un sistema aislado, podemos decir:

$$\boxed{\begin{array}{l} \Delta S_{\text{ad}} > 0 \Leftrightarrow \text{Proceso adiabático irreversible} \\ \Delta S_{\text{ad}} = 0 \Leftrightarrow \text{Proceso adiabático reversible} \end{array}} \quad (1.19)$$

1.10.2. Variables intensivas

A partir de ahora supondremos por simplicidad que tratamos con **sistemas monocomponentes que solo pueden intercambiar trabajo de expansión**. Las afirmaciones que hagamos se generalizan fácilmente a sistemas con más componentes o que puedan intercambiar otros tipos de trabajo.

Supongamos un sistema homogéneo con esas características. Su entropía vendrá dada por la función $S(U, V, N)$. Su diferencial **en cualquier estado de equilibrio** se escribe:

$$dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \quad (1.20)$$

Se define la **temperatura termodinámica**, T , del sistema por la relación:

$$\boxed{\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N}} \quad (1.21)$$

Por otro lado, se define el **potencial químico**, μ , del componente en el sistema como:

$$\boxed{-\frac{\mu}{T} = \left(\frac{\partial S}{\partial N}\right)_{U,V}} \quad (1.22)$$

Y, además, el postulado de Callen permite demostrar (véase la siguiente sección) que la **presión**, p , esto es, la fuerza generalizada, puede calcularse **en cualquier estado de equilibrio** (pero no fuera del equilibrio) a partir de la ecuación:

⁵ L.D. Landau, E.M. Lifshitz, “Física Estadística”, Volumen 5 del Curso de Física Teórica, Reverté, 1969; y B. Diu, C. Guthmann, D. Lederer, B. Roulet, “Thermodynamique”, Hermann, 2007.

$$\boxed{\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{U,N}} \quad (1.23)$$

De este modo, la ecuación (1.20) se puede escribir como:

$$\boxed{dS = \frac{1}{T}dU + \frac{p}{T}dV - \frac{\mu}{T}dN} \quad (1.24)$$

Normalmente, a las variables naturales de la entropía (U, V, N) se las denomina, abusando un poco del lenguaje, “**las variables extensivas**”. Las variables que acompañan a cada una de ellas en la diferencial de la entropía $\left(\frac{1}{T}, \frac{p}{T}, -\frac{\mu}{T}\right)$ se las denomina “**las variables intensivas**” conjugadas de las extensivas.

Las variables intensivas, a diferencia de las extensivas, solo están bien definidas si el sistema se encuentra en equilibrio termodinámico y no tienen sentido a nivel microscópico (es decir, no pueden caracterizar ningún microestado del sistema).

Ahora puede deducirse la relación que existe entre la entropía y la diferencial del calor cuasiestático para un sistema homogéneo cerrado ($N = \text{cte}$) sin trabajo externo disipativo ($\delta W = -pdV$). Para tal sistema:

$$(dS)_N = \frac{1}{T}(dU)_N + \frac{p}{T}(dV)_N = \frac{(dU)_N + p(dV)_N}{T} = \frac{(dU)_N - \delta W}{T} \quad (1.25)$$

El numerador del último miembro no es más que la diferencial del calor cuasiestático (δQ), por lo que:

$$\boxed{(dS)_N = \frac{\delta Q}{T}} \quad \text{(sistema homogéneo cerrado sin trabajo externo disipativo)} \quad (1.26)$$

De esta relación se deduce que, **para un sistema homogéneo cerrado sin trabajo externo disipativo, un proceso cuasiestático es adiabático si y solo si es isoentrópico.**

1.10.3. Relación de la entropía con la presión

Para poder demostrar la relación (1.23) hay que aplicar el postulado de Callen utilizando las propiedades especiales de los procesos adiabáticos que se comentaron al deducir el Principio de Aumento de Entropía.

Para ello, supongamos un sistema termodinámico homogéneo con variables de estado (U, V, N) y su entorno, ambos rodeados de paredes externas aislantes. El sistema y su entorno están a su vez separados por una pared interna aislante. Si ahora **reemplazamos la pared interna aislante por una adiabática**, el sistema y el entorno abandonan el estado de equilibrio y, tras un tiempo, alcanzan un nuevo estado de equilibrio. Como el proceso del sistema es adiabático, se puede describir como si el entorno tuviera entropía cero. En consecuencia, **la entropía del sistema en estudio, S , en el estado final de equilibrio debe ser la máxima compatible con las ligaduras del problema.**

Hagamos unos comentarios de carácter físico sobre el problema a analizar (que también son aplicables a la sección siguiente). En este problema no podemos considerar que exista rozamiento de la pared interna con la pared externa al moverse, ni de la pared externa con los subsistemas, ya que, en tal caso, parte de la energía disipada sería absorbida por las paredes, transfiriéndose después calor a los subsistemas, y el proceso no sería adiabático. Dicho de otra forma, la presencia de estos rozamientos implicaría que las paredes tienen cierta microestructura, lo que va en contra de la definición de proceso adiabático. Además, por definición, las paredes permanecen inalteradas al establecer o prohibir los intercambios de energía, por lo que si en un caso real el rozamiento produce una cantidad de energía que es relevante para la solución del problema, las paredes no podrían ser consideradas como tales, sino como sistemas termodinámicos, y el problema que habría que resolver sería distinto. Por todo ello, **en este problema los trabajos externos disipativos sobre los dos sistemas deben ser iguales a cero.** No obstante,

dentro de cada subsistema sí pueden existir mecanismos disipativos, que contribuyen a que la pared pueda detenerse finalmente en la posición de equilibrio y están asociados con el aumento de entropía en este proceso espontáneo.

En condiciones adiabáticas y sin trabajo externo disipativo, la energía interna y el volumen no pueden ser independientes, ya que la variación de una magnitud implica la de la otra. Ahora bien, es imposible predecir sus valores finales en los subsistemas individuales, dado que éstos dependen de cómo los mecanismos disipativos dentro de cada uno de ellos transforman el trabajo recibido en energía interna durante el proceso *real* seguido por el sistema y el entorno, lo que a su vez depende de factores que quedan fuera del dominio de la descripción termodinámica (como viscosidades y otros parámetros cinéticos). El Principio de Aumento de Entropía aplicado al proceso adiabático del sistema ($\Delta S \geq 0$) permite afirmar que el estado final de equilibrio deberá estar en una región concreta del espacio termodinámico, pero no permite predecir el valor final de S (más allá de que sea un máximo). En general, existe un conjunto de distintos estados de equilibrio con dicho valor de S y que son compatibles con las ligaduras. Todos estos estados pueden conectarse por procesos adiabáticos sin trabajo externo disipativo (por las ligaduras del problema), cuasiestáticos (pues todos son de equilibrio) e isoentrópicos (pues todos tienen la misma entropía). Los estados iniciales del sistema y del entorno son arbitrarios, por lo que también lo serán los estados finales, y la anterior conclusión sobre los estados con el valor final de S puede generalizarse a un estado de equilibrio arbitrario del sistema: *dado un estado de equilibrio, existe un conjunto de estados que son accesibles desde éste por procesos adiabáticos, sin trabajo externo disipativo, cuasiestáticos e isoentrópicos*. Este es un resultado auxiliar necesario para probar la relación (1.23).

De acuerdo con el Primer Principio, la diferencial de la energía interna a lo largo de uno de estos procesos adiabáticos (para los que $N = \text{cte}$), cuasiestáticos e isoentrópicos es:

$$(dU)_{S,N} = \delta W^{\text{ad}} = -p(dV)_{S,N} \quad (1.27)$$

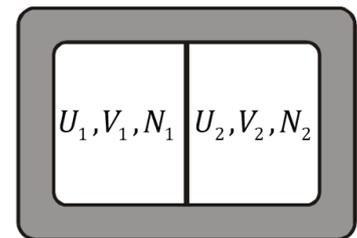
donde el trabajo externo es solo de configuración (no disipativo), y la diferencial de la entropía es:

$$0 = (dS)_{S,N} = \frac{1}{T}(dU)_{S,N} + \left(\frac{\partial S}{\partial V}\right)_{U,N} (dV)_{S,N} = \left[-\frac{p}{T} + \left(\frac{\partial S}{\partial V}\right)_{U,N}\right] (dV)_{S,N} \quad (1.28)$$

de donde se deduce la ecuación (1.23) que relaciona la entropía y la presión.

1.10.4. Condiciones de equilibrio mecánico mutuo

El ejemplo de la sección anterior puede usarse también para determinar las condiciones de equilibrio mecánico mutuo de dos sistemas. Para ello analicemos el sistema aislado de la figura, inicialmente en equilibrio, que está separado en dos partes por una pared aislante. Si **reemplazamos la pared interna aislante por una adiabática**, el sistema abandona el estado de equilibrio y, tras un tiempo, alcanza un nuevo estado de equilibrio. En el fondo, este problema es el mismo que el de la sección anterior y, de acuerdo con lo ya dicho, no es posible determinar completamente el estado final de equilibrio, pero sí que podemos hallar una condición necesaria, aunque no suficiente, para ello.



El estado final de equilibrio debe cumplir las ligaduras:

$$\begin{aligned} U_1 + U_2 = U = \text{cte} &\Rightarrow dU_2 = -dU_1 \quad (\text{pared externa aislante}) \\ V_1 + V_2 = V = \text{cte} &\Rightarrow dV_2 = -dV_1 \quad (\text{pared externa aislante}) \\ N_1 = \text{cte} &\quad (\text{pared interna adiabática y externa aislante}) \\ N_2 = \text{cte} &\quad (\text{pared interna adiabática y externa aislante}) \\ S_1 \text{ máxima, desconocida} &\quad (\text{pared interna adiabática}) \\ S_2 \text{ máxima, desconocida} &\quad (\text{pared interna adiabática}) \end{aligned} \quad (1.29)$$

Para deducir la condición buscada, haremos como en la sección anterior, y analizaremos las condiciones que deben cumplir los estados de equilibrio situados en las regiones isoentrópicas de S_1 máxima y S_2 máxima. La diferencial de las energías internas en estas condiciones cuasiestáticas y adiabáticas es:

$$\begin{aligned} dU_1 &= -p_1 dV_1 \\ dU_2 &= -p_2 dV_2 \end{aligned} \quad (1.30)$$

Combinando las dos primeras ligaduras con estas dos ecuaciones, obtenemos:

$$0 = dU_1 + dU_2 = -p_1 dV_1 - p_2 dV_2 = (p_2 - p_1) dV_1 \Rightarrow \boxed{p_1 = p_2} \quad (1.31)$$

Es decir, como conocíamos ya por la Mecánica Macroscópica, las condiciones de **equilibrio mecánico mutuo** corresponden a la **igualdad de presiones**. Estas condiciones se cumplen en todos los estados de equilibrio situados en las regiones con S_1 máxima y S_2 máxima, por lo que, como ya hemos anunciado antes, esta condición será necesaria, pero no suficiente, para el equilibrio mutuo con una pared interna adiabática.

1.10.5. Condiciones de equilibrio térmico mutuo

En este caso tenemos que analizar qué ocurre al **reemplazar la pared interna aislante por una pared que permita únicamente el intercambio de calor** (diatérmana, fija y rígida, e impermeable). Para ello, consideremos las ligaduras que operan en los hipotéticos estados de equilibrio que puede alcanzar el sistema total:

$$\begin{aligned} U_1 + U_2 = U = \text{cte} &\Rightarrow dU_2 = -dU_1 \quad (\text{pared externa aislante}) \\ V_1 = \text{cte} &\Rightarrow dV_1 = 0 \quad (\text{pared interna fija y rígida}) \\ V_2 = \text{cte} &\Rightarrow dV_2 = 0 \quad (\text{pared interna fija y rígida}) \\ N_1 = \text{cte} &\Rightarrow dN_1 = 0 \quad (\text{pared interna impermeable}) \\ N_2 = \text{cte} &\Rightarrow dN_2 = 0 \quad (\text{pared interna impermeable}) \end{aligned} \quad (1.32)$$

La diferencial de la entropía del sistema total es:

$$dS = dS_1 + dS_2 = \frac{1}{T_1} dU_1 + \frac{p_1}{T_1} dV_1 - \frac{\mu_1}{T_1} dn_1 + \frac{1}{T_2} dU_2 + \frac{p_2}{T_2} dV_2 - \frac{\mu_2}{T_2} dn_2 = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dU_1 \quad (1.33)$$

La última igualdad, en la que solo aparece dU_1 , nos dice que la entropía S del sistema en los hipotéticos estados de equilibrio alcanzables queda determinada únicamente por la variable independiente U_1 . El valor de U_1 en el equilibrio es aquel que maximiza S , para lo cual ha de cumplirse:

$$\frac{dS}{dU_1} = 0 \Leftrightarrow \frac{1}{T_1} = \frac{1}{T_2} \Leftrightarrow \boxed{T_1 = T_2} \quad (1.34)$$

Por tanto, la condición de **equilibrio térmico mutuo** corresponde a la **igualdad de temperaturas**. Por analogía con el caso del equilibrio mecánico mutuo, se dice que **la temperatura es la "fuerza" generalizada asociada al intercambio de calor**.

1.10.6. Condiciones de equilibrio material mutuo

Finalmente, estudiemos lo que acontece al **reemplazar la pared interna aislante por una pared que permita el intercambio de materia** (permeable, fija y rígida). Para ello, consideremos las ligaduras que operan en los hipotéticos estados de equilibrio que puede alcanzar el sistema total:

$$\begin{aligned} U_1 + U_2 = U = \text{cte} &\Rightarrow dU_2 = -dU_1 \quad (\text{pared externa aislante}) \\ V_1 = \text{cte} &\Rightarrow dV_1 = 0 \quad (\text{pared interna fija y rígida}) \\ V_2 = \text{cte} &\Rightarrow dV_2 = 0 \quad (\text{pared interna fija y rígida}) \\ N_1 + N_2 = N = \text{cte} &\Rightarrow dN_1 = -dN_2 \quad (\text{pared externa aislante}) \end{aligned} \quad (1.35)$$

La diferencial de la entropía del sistema total queda:

$$dS = \left(\frac{1}{T_1} - \frac{1}{T_2}\right) dU_1 + \left(-\frac{\mu_1}{T_1} + \frac{\mu_2}{T_2}\right) dN_1 \quad (1.36)$$

La ecuación anterior indica que la entropía es función de dos variables independientes (U_1, N_1). Los valores de éstas en el equilibrio son aquellos que maximizan S , para lo cual han de anularse simultáneamente las derivadas parciales:

$$\begin{aligned} \left(\frac{\partial S}{\partial U_1}\right)_{N_1} &= 0 \Leftrightarrow \frac{1}{T_1} = \frac{1}{T_2} \\ \left(\frac{\partial S}{\partial N_1}\right)_{U_1} &= 0 \Leftrightarrow -\frac{\mu_1}{T_1} = -\frac{\mu_2}{T_2} \end{aligned} \quad (1.37)$$

Las cuales equivalen a:

$$\boxed{\begin{aligned} T_1 &= T_2 \\ \mu_1 &= \mu_2 \end{aligned}} \quad (1.38)$$

Por tanto, en el supuesto de que se ha alcanzado el equilibrio térmico mutuo entre los dos sistemas, la condición de **equilibrio material mutuo** corresponde a la **igualdad de potenciales químicos**. Por eso se dice que **el potencial químico es la "fuerza" generalizada asociada al intercambio de materia**.

1.11. Ecuación fundamental y ecuaciones de estado

La relación $S = S(U, V, N)$ para un sistema concreto recibe el nombre de **ecuación fundamental** del sistema (**primer nivel de información**). Es una ecuación que no puede proporcionar la Termodinámica, sino que puede proceder básicamente de dos fuentes:

- De datos experimentales.
- De un modelo microscópico a través de la aplicación del formalismo de la Física Estadística.

Si se conoce la ecuación fundamental del sistema, se pueden obtener las ecuaciones (1.21), (1.22) y (1.23) por simple derivación de la ecuación fundamental. Estas ecuaciones se conocen con el nombre de **ecuaciones de estado** del sistema (**segundo nivel de información**):

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial U}\right)_{V,N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{T} &= \left(\frac{\partial S}{\partial V}\right)_{U,N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{T} &= \left(\frac{\partial S}{\partial N}\right)_{U,V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (1.39)$$

Si se siguen tomando derivadas a partir de las ecuaciones de estado, se pueden obtener los **coeficientes respuesta (tercer nivel de información)**, tales como los coeficientes de dilatación, las capacidades caloríficas, las compresibilidades, etc.

En definitiva, **a partir de la función $S = S(U, V, N)$ puede obtenerse toda la información termodinámica del sistema**, por lo que recibe el nombre de **potencial termodinámico**. Dado que las variables (U, V, N) están especialmente adaptadas para trabajar con sistemas aislados, se dice que **la entropía es el potencial termodinámico natural de un sistema aislado y que (U, V, N) son sus variables naturales**.

1.12. Sistema cerrado, isócoro e isoterma: función de Massieu

La energía interna es una variable que no se puede medir y resulta en general poco conveniente de usar como variable independiente. Para solventar este problema, se define la **función de Massieu**, ψ , como:

$$\psi = S - U \left(\frac{\partial S}{\partial U} \right)_{V,N} = S - \frac{1}{T} U \quad (1.40)$$

La diferencial de la función de Massieu puede escribirse:

$$d\psi = dS - U d \left(\frac{1}{T} \right) - \frac{1}{T} dU = \left[\frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} dN \right] - U d \left(\frac{1}{T} \right) - \frac{1}{T} dU \quad (1.41)$$

Simplificando, queda:

$$d\psi = -U d \left(\frac{1}{T} \right) + \frac{p}{T} dV - \frac{\mu}{T} dN \quad (1.42)$$

Hemos obtenido así $d\psi$ en función de las variables $\left(\frac{1}{T}, V, N \right)$, que reciben el nombre de **variables naturales de la función de Massieu**. Hemos reemplazado, de esta forma, la energía interna por el inverso de la temperatura como variable independiente.

Puede justificarse matemáticamente que $\psi \left(\frac{1}{T}, V, N \right)$ porta la misma información que $S(U, V, N)$. De hecho, ψ es lo que se conoce como **transformada de Legendre** de S respecto a U . La ecuación $\psi = \psi \left(\frac{1}{T}, V, N \right)$ resulta ser otra forma equivalente de la ecuación fundamental del sistema, a partir de la cual se pueden calcular por derivación las ecuaciones de estado, que se deducen de la ecuación (1.42):

$$\begin{aligned} -U &= \left(\frac{\partial \psi}{\partial [1/T]} \right)_{V,N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{T} &= \left(\frac{\partial \psi}{\partial V} \right)_{\frac{1}{T}, N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{T} &= \left(\frac{\partial \psi}{\partial N} \right)_{\frac{1}{T}, V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (1.43)$$

La función de Massieu se adapta especialmente bien al estudio del equilibrio de **sistemas cerrados isoterms e isócoros**, esto es, aquellos que están en contacto térmico con otro sistema denominado **termostato** con una capacidad calorífica muy elevada (idealmente, infinita). El termostato es capaz de ceder o absorber cantidades arbitrariamente grandes de energía en forma de calor sin variar apreciablemente su temperatura y, por ello, **tiene el papel de mantener la temperatura del sistema constante sin modificar apreciablemente su propio estado**.

Se dice que **la función de Massieu es el potencial termodinámico natural de un sistema cerrado isoterma e isócoro**. Esta función, además, cumple propiedades de máximo similares a las de la entropía en este tipo de sistemas.

Ejercicios - Termodinámica

1.- La ecuación fundamental de un gas ideal monoatómico semiclásico es:

$$S(U, V, N) = k_B \ln \left[\frac{1}{h^{3N}} \left(\frac{V}{N} \right)^N \left(\frac{4\pi m U}{3N} \right)^{3N/2} \right] + \frac{5Nk_B}{2}$$

donde h es la constante de Planck y m es la masa de cada átomo.

- Calcule las ecuaciones de estado del gas y la capacidad calorífica a volumen constante.
- Encuentre la función de Massieu, calculando a partir de ella las ecuaciones de estado y comprobando que coinciden con las obtenidas a partir de la entropía.

2.- Un modelo semiclásico sencillo de un sólido cristalino aislante ideal monoatómico, en el que se supone que los N átomos efectúan oscilaciones armónicas en torno a su posición de equilibrio en la red, tiene por ecuación fundamental:

$$S(U, N) = 3Nk_B \ln \left[\frac{eU}{3Nh f_0} \right]$$

donde h es la constante de Planck, f_0 es una frecuencia característica de vibración de los átomos en la red y no se tiene en cuenta el volumen del sistema (este modelo no puede dar cuenta de la dilatación del sistema). Repita los apartados del ejercicio anterior con este sistema.

2. Teoría Cinética de los gases ideales

La Teoría Cinética de los gases y la Física Estadística que se ven en este curso involucran cálculos probabilísticos con las posiciones, velocidades y/o momentos de las moléculas, que son magnitudes físicas vectoriales y, por tanto, pertenecen a un espacio vectorial tridimensional. Por esta razón, es necesario conocer cálculo integral (básico) de varias variables, el uso del sistema de coordenadas esféricas y conocimientos básicos de Probabilidad y de las variables aleatorias continuas. Para comenzar, comentaremos unas pinceladas sobre estos temas, sin pretender ser rigurosos.

2.1. Cálculo en el espacio tridimensional

2.1.1. Integración en varias variables

La idea de la integral en varias variables es totalmente análoga a la de la integral en una variable. En una variable, se toma una función real f definida en un intervalo $[a, b]$ (que supondremos continua y tan regular como sea necesario) y se procede, a grandes rasgos, de la siguiente forma:

- Se divide el intervalo en trozos (por simplicidad, iguales y de longitud Δx), formando lo que se llama una **partición** del intervalo. Se toma en cada uno de los trozos un punto, x_i , formando un **conjunto de puntos intermedios**. Se calcula la suma $\sum f(x_i)\Delta x$.
- Se repite el procedimiento formando una sucesión de particiones con Δx cada vez más pequeño. El límite de la sucesión de sumas $\sum f(x_i)\Delta x$ a medida que Δx tiende a cero se denomina **integral** de f en $[a, b]$.

En el caso de varias variables se toma una función real f definida en un dominio medible $B \subset \mathbb{R}^n$ y se siguen unos pasos análogos a los anteriores. Por ejemplo, si $B = [x_a, x_b] \times [y_a, y_b]$ es un rectángulo en el plano, se forma una partición de pequeños rectángulos de dimensiones $(\Delta x, \Delta y)$ y un conjunto de puntos intermedios (x_i, y_i) localizados en el interior de cada rectángulo. El límite de la suma $\sum f(x_i, y_i)\Delta x \Delta y$ a medida que las dimensiones de los rectángulos tienden a cero se denomina **integral (doble)** de f en B y se denota:

$$\int_B f(x, y) dx dy \quad (2.1)$$

Es tremendamente importante el hecho de que muchas integrales múltiples se pueden calcular a partir de integrales de una variable, que sabemos resolver gracias al Teorema Fundamental del Cálculo. El resultado que permite hacer esto se denomina **Teorema de Fubini** y para el caso de una función definida en un intervalo n -dimensional $B = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$, nos dice que:

$$\int_B f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_n}^{b_n} dx_n f(x_1, x_2, \dots, x_n) \quad (2.2)$$

Es decir, podemos calcular la integral de forma **iterada**: se integra una variable por separado, manteniendo el resto constantes, el resultado se integra respecto de otra variable y así sucesivamente hasta haber terminado con todas las variables. Además, este proceso puede realizarse con **cualquiera que sea el orden en las variables**, de modo que integrales que resultan difíciles en un orden pueden resultar más fáciles en otro.

2.1.2. Coordenadas esféricas en el espacio tridimensional

Un sistema de coordenadas que nos será muy útil es el de coordenadas esféricas (Figura 2.1). En él se sustituyen las coordenadas cartesianas (x, y, z) por una distancia y dos ángulos: el **radio** (r), la **colatitud** (θ) y el **acimut** (ϕ). Estas variables están definidas en los intervalos:

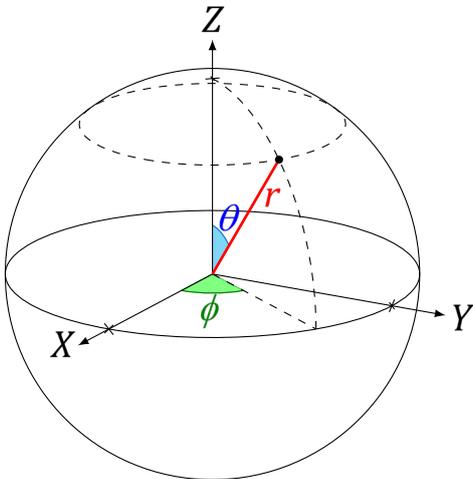


Figura 2.1: Sistema de coordenadas esféricas.

$$\begin{aligned} r &\in [0, \infty) \\ \theta &\in [0, \pi] \\ \phi &\in [0, 2\pi] \end{aligned} \tag{2.3}$$

Las ecuaciones que relacionan el sistema cartesiano con el sistema esférico son:

$$\begin{aligned} x &= r \sin\theta \cos\phi \\ y &= r \sin\theta \sin\phi \\ z &= r \cos\theta \end{aligned} \tag{2.4}$$

Los cálculos en la Teoría Cinética de los Gases se realizan de forma mucho más conveniente en este sistema de coordenadas. Es importante, al hacer un cambio de variable en una integral múltiple, tener en cuenta la aparición de un factor de escala que depende del sistema de coordenadas usado. Así, en el caso del cambio de coordenadas cartesianas a esféricas:

$$\int_B f(x, y, z) dx dy dz = \int_{B'} g(r, \theta, \phi) r^2 \sin\theta dr d\theta d\phi \tag{2.5}$$

donde B es el dominio de integración en cartesianas, B' es el dominio de integración en esféricas y $g(r, \theta, \phi)$ es la función que se obtiene al sustituir en $f(x, y, z)$ el cambio a coordenadas esféricas. Vemos que aparece un factor $r^2 \sin\theta$ adicional. Matemáticamente este factor es el valor absoluto del determinante del jacobiano de la transformación de coordenadas; en otras palabras, es el resultado de una operación con las derivadas parciales de las variables de un sistema de coordenadas respecto de las del otro. Pero este factor también puede obtenerse de forma heurística, como veremos a continuación.

Normalmente la aparición del factor suele expresarse (de modo formal) diciendo que el **elemento de volumen en coordenadas esféricas** toma la forma:

$$d^3x \equiv dx dy dz = r^2 \sin\theta dr d\theta d\phi \tag{2.6}$$

La magnitud d^3x puede entenderse como un elemento infinitesimal de volumen en el espacio (Figura 2.2). Se trata de un prisma con base de área igual al llamado **elemento de superficie** en coordenadas esféricas:

$$dS = (r d\theta)(r \sin\theta d\phi) = r^2 \sin\theta d\theta d\phi \tag{2.7}$$

y altura dr . Su volumen:

$$d^3x = dS dr$$

coincide, por tanto, con el resultado anterior.

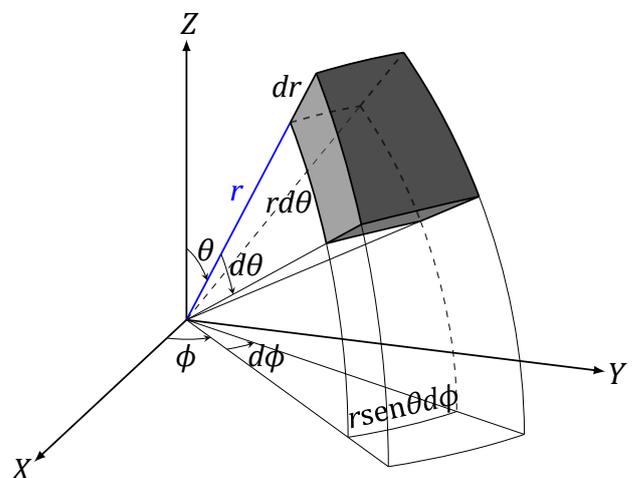


Figura 2.2: Elemento de volumen en coordenadas esféricas.

2.2. Conceptos básicos de Probabilidad

Los experimentos pueden dividirse en dos clases claramente diferenciadas:

- Deterministas: se puede predecir con certeza lo que va a suceder.
- Aleatorios: no se puede predecir el resultado con antelación.

En Física también existen experimentos aleatorios. Por ejemplo, es imposible conocer de antemano las velocidades de todas las moléculas de un gas. Por ello, es conveniente conocer los conceptos siguientes.

Definición. Se llama **espacio muestral** al conjunto formado por todos los posibles resultados de un experimento.

Definición. Se llama **suceso** a cualquier subconjunto del espacio muestral (cualquier conjunto de posibles resultados de un experimento).

Definición. Se llama **suceso elemental** a un subconjunto del espacio muestral formado por un elemento.

Definiciones. Sea E el espacio muestral asociado a un experimento. Sean dos sucesos $A, B \subset E$.

- Llamamos **unión de A y B** , denotado $A \cup B$, al suceso que ocurre si suceden A o B o ambos.
- Llamamos **intersección de A y B** , denotado $A \cap B$, al suceso que ocurre si suceden A y B .
- Llamamos **complementario de A** , denotado \bar{A} , al suceso que ocurre cuando no sucede A .
- Decimos que A y B son **mutuamente excluyentes** si no pueden ocurrir simultáneamente, es decir, si $A \cap B = \emptyset$. Los sucesos elementales son mutuamente excluyentes entre sí.

Si $\{A_i\} \subset E, i = 1, 2, \dots, n$ es una colección de sucesos:

- El suceso $\bigcup_{i=1}^n A_i$ ocurre si sucede algún A_i .
- El suceso $\bigcap_{i=1}^n A_i$ ocurre cuando suceden todos los A_i .

Pero, aunque en los experimentos aleatorios no se sepa lo que va a pasar, no todos los sucesos tienen por qué ser igual de “esperables”. Por ello, los humanos hemos inventado el concepto de probabilidad. La idea de la Teoría de la Probabilidad es asignar ciertos números reales positivos a los sucesos, llamados probabilidades, de forma que:

- La probabilidad del espacio muestral ha de ser igual al máximo valor posible de la probabilidad.
- Cuando se sabe exactamente el suceso que tendrá lugar, entonces ese suceso tendrá el máximo valor posible de la probabilidad y el resto tendrán probabilidad nula.
- Cuando no se tiene ni idea del suceso que tendrá lugar, de forma que todos ellos sean igual de esperables para quien se enfrenta al problema, todos los sucesos tienen la misma probabilidad.

En todas las situaciones intermedias de conocimiento del fenómeno, las probabilidades tomarán otros valores distintos.

Definición. Sea E el espacio muestral asociado a un experimento. Llamamos **probabilidad** a una función que asocia a cada suceso $A \subset E$ un número real $\Pr(A)$ verificando:

1. $0 \leq \Pr(A) \leq 1$.
2. $\Pr(E) = 1$ (**condición de normalización**).
3. Si $A, B \subset E$ son mutuamente excluyentes, $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

Propiedades de la probabilidad (deducibles de su definición):

1. $\Pr(\emptyset) = 0$.
2. $\Pr(\bar{A}) = 1 - \Pr(A)$
3. $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
4. Si $A \subset B$, $\Pr(A) \leq \Pr(B)$.

Ejemplo. Supongamos que disponemos de una caja con N bolas, de las cuales N_R son rojas y N_A son azules. Imaginemos el experimento “sacar una bola”. En este experimento, el espacio muestral está compuesto de N sucesos elementales correspondientes a la extracción de cada una de las bolas. No hay motivos para que unos sucesos elementales sean más esperables que otros, por lo que el **Principio de Indiferencia** indica que la probabilidad de cada suceso elemental “sacar cualquier bola” debe ser igual para todos ellos (esto es, los sucesos elementales deben ser **equiprobables**). Como la suma de las probabilidades de todos estos sucesos elementales, mutuamente excluyentes, debe ser la unidad:

$$\Pr(\text{"sacar cualquier bola"}) = \frac{1}{(\text{número de sucesos elementales})} = \frac{1}{N}$$

Llamemos R al suceso “sacar bola roja” y A al suceso “sacar bola azul”. El suceso R incluye N_R sucesos elementales porque hay N_R bolas rojas, mientras que el suceso A incluye N_A sucesos elementales. Como los sucesos elementales son siempre mutuamente excluyentes, por la propiedad 3 de la definición de probabilidad, tenemos que

$$\Pr(R) = N_R \cdot \Pr(\text{"sacar cualquier bola"}) = \frac{N_R}{N}$$

$$\Pr(A) = N_A \cdot \Pr(\text{"sacar cualquier bola"}) = \frac{N_A}{N}$$

Este ejemplo ilustra la conocida regla que indica que, **cuando los sucesos elementales son equiprobables, la probabilidad de un suceso es igual al cociente entre el número de casos favorables y el número de casos posibles.**

Definición. Sea E el espacio muestral asociado a un experimento y $A, B \subset E$ dos sucesos. Si la ocurrencia del suceso B depende de que haya ocurrido el suceso A , se llama **probabilidad de B condicionada a A** al número real $\Pr(B|A) = \Pr(A \cap B) / \Pr(A)$.

Definición. Sea E el espacio muestral asociado a un experimento y $A, B \subset E$ dos sucesos. Se dice que A y B son **independientes** si la ocurrencia de uno no depende de la ocurrencia del otro:

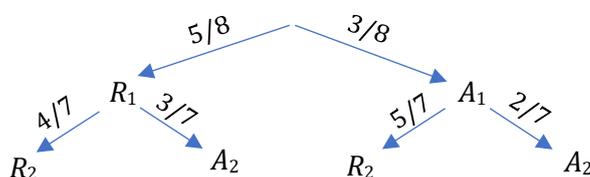
$$\Pr(B|A) = \Pr(B) \Leftrightarrow \Pr(A \cap B) = \Pr(A)\Pr(B)$$

Ejemplo. Para comprender mejor la idea de la probabilidad condicionada, imaginemos un experimento aleatorio que consiste en la extracción consecutiva de dos bolas de una caja, de forma que la segunda bola se extrae sin volver a meter la primera bola.

Supongamos que, inicialmente, hay 8 bolas en la caja, de las cuales 5 son rojas y 3 son azules. Se trata de un experimento compuesto de otros dos:

- **Experimento 1:** “sacar una bola”. Hay 8 bolas en total, de las cuales 5 son rojas y 3 son azules.
- **Experimento 2:** “sacar otra bola”. Hay 7 bolas en total, pero no se sabe con certeza cuántas rojas y azules hay, ya que esto depende del resultado del experimento 1. En otras palabras, **este segundo experimento está condicionado al primero**, es decir, las probabilidades de los posibles resultados son **probabilidades condicionadas** y dependen de los resultados del primer experimento.

Nos podemos preguntar cuál es la probabilidad de “sacar dos rojas”. Una estrategia sencilla para abordar el problema consiste en representarlo mediante un **diagrama de árbol**:



Las dos ramas superiores indican las probabilidades de R y de A en el experimento 1, $\Pr(R_1) = 5/8$ y $\Pr(A_1) = 3/8$. Las cuatro ramas inferiores indican las probabilidades de R y de A en el experimento 2 condicionadas a los resultados del primero, tras el cual hay una bola menos en la caja; de izquierda a derecha, son: $\Pr(R_2|R_1) = 4/7$, $\Pr(A_2|R_1) = 3/7$, $\Pr(R_2|A_1) = 5/7$ y $\Pr(A_2|A_1) = 2/7$.

De acuerdo con la fórmula de la probabilidad condicionada, la probabilidad del suceso $R_1 \cap R_2 =$ “sacar dos rojas” es:

$$\Pr(R_1 \cap R_2) = \Pr(R_1) \Pr(R_2|R_1) = \left(\frac{5}{8}\right) \left(\frac{4}{7}\right) = \frac{20}{56} = \frac{5}{14}$$

Análogamente puede hacerse para calcular la probabilidad de otros sucesos, como $A_1 \cap A_2 =$ “sacar dos azules”. Como puede verse, la probabilidad de los sucesos intersección del experimento compuesto puede obtenerse mediante el producto de las probabilidades de las ramas del diagrama de árbol. Por tanto, se dispone de la misma “información” sobre el experimento compuesto tanto si se conocen las probabilidades de todos los sucesos intersección como si se dispone de las probabilidades del correspondiente diagrama de árbol (es decir, de las probabilidades condicionadas). Esta idea fue utilizada por C.E. Shannon para caracterizar matemáticamente la entropía de información (sección 3.3.1) cuando estableció la Teoría Matemática de la Información.

2.2.1. Variables aleatorias

En Física surge de forma bastante natural el concepto de **variable aleatoria**, que es una variable (posiblemente multidimensional) que asociamos a un espacio muestral. Matemáticamente se representan como funciones del espacio muestral en \mathbb{R}^n . En el ejemplo de las moléculas de un gas en Teoría Cinética, el espacio muestral sería el conjunto de todas las velocidades posibles de las moléculas, a las cuales asociamos tres variables escalares para su descripción y que constituirían dicha variable aleatoria.

Hay que distinguir dos clases de variables aleatorias (VA):

- Discretas (VAD): solo pueden tomar un conjunto numerable de valores.
- Continuas (VAC): toman un conjunto no numerable de valores.

2.2.1.1. Variables aleatorias unidimensionales

Restrinjámosnos por el momento a variables aleatorias unidimensionales. Las **variables aleatorias discretas (VAD)** están caracterizadas por su **función de probabilidad**, que asigna a cada posible valor de la variable, x , su probabilidad $\Pr(x)$. Esta función cumple las siguientes propiedades:

- $0 \leq \Pr(x) \leq 1$.
- Si dos sucesos x_i y x_j son mutuamente excluyentes (es decir, no puede ocurrir uno si ha ocurrido el otro), la probabilidad del suceso “resultado x_i o resultado x_j ” es $\Pr(x_i \text{ o } x_j) = \Pr(x_i) + \Pr(x_j)$.
- Condición de normalización: $\sum \Pr(x_i) = 1$.

Para calcular el valor medio de una función $\phi(x)$ de la VAD x , se utiliza la expresión:

$$\langle \phi(x) \rangle = \sum_i \phi(x_i) \Pr(x_i) \quad (2.8)$$

Por otro lado, en las **variables aleatorias continuas (VAC)** la probabilidad de que la variable tome un valor perfectamente definido es cero, ya que el conjunto de valores es no numerable. Solo tiene sentido hablar de la probabilidad de que la variable x esté entre dos valores dados, digamos a y b . Designemos esta probabilidad por $\Pr(a \leq x \leq b)$. Por lo ya dicho, se sustituye el concepto de función de probabilidad por el de **función de densidad de probabilidad (FDP)**, $f(x)$, que verifica las siguientes propiedades:

- $f(x) \geq 0$.
- $\Pr(a \leq x \leq b) = \int_a^b f(x)dx$.
- Condición de normalización: $\int_{-\infty}^{\infty} f(x)dx = 1$.

Se ha supuesto que x puede tomar cualquier valor real. Esto no resta generalidad, pues si la variable no pudiera tomar todos los valores, a los intervalos imposibles se les podría asignar $f(x) = 0$.

¿Cuál es el significado de la FDP? Intuitivamente, puede entenderse $f(x)dx$ como la “probabilidad infinitesimal” $d\Pr(x)$ de que la variable se encuentre en un intervalo infinitesimal de valores entre x y $x + dx$. Veamos esto de forma algo más rigurosa. Tomemos un intervalo arbitrario de valores $[a, b] \subset \mathbb{R}$ y dividámoslo en pequeños intervalos de extremos $x_0 = a, x_1, \dots, x_M = b$, por simplicidad de igual longitud $\Delta x = x_i - x_{i-1}$. Ahora escojamos el intervalo $[x_{i-1}, x_i]$ y definamos en él la función:

$$F(y) = \int_{x_{i-1}}^y f(x)dx \quad (2.9)$$

Si suponemos que f es continua, por el Teorema Fundamental del Cálculo F es derivable en $[x_{i-1}, x_i]$ y su derivada es f . Además, por el Teorema de los Incrementos Finitos de Lagrange, existe un punto $t_i \in (x_{i-1}, x_i)$ tal que $F(x_i) - F(x_{i-1}) = F'(t_i)(x_i - x_{i-1})$. Notemos que:

$$F(x_i) - F(x_{i-1}) = \int_{x_{i-1}}^{x_i} f(x)dx - \int_{x_{i-1}}^{x_{i-1}} f(x)dx = \int_{x_{i-1}}^{x_i} f(x)dx = \Pr(x_{i-1} \leq x \leq x_i) \quad (2.10)$$

Por tanto:

$$\Pr(x_{i-1} \leq x \leq x_i) = f(t_i)\Delta x \quad (2.11)$$

Para comprender mejor el significado de esta igualdad, construyamos un **histograma**. Para ello, llamemos $\Delta\Pr(t_i) = \Pr(x_{i-1} \leq x \leq x_i)$ y representemos barras de altura $f(t_i) = \Delta\Pr(t_i)/\Delta x$ y anchura Δx para cada intervalo $[x_{i-1}, x_i]$ (Figura 2.3a). El área de esas barras coincidirá con la probabilidad $\Delta\Pr(t_i) = \Pr(x_{i-1} \leq x \leq x_i)$. Al ir haciendo cada vez más pequeña la anchura Δx de las barras del histograma, éste cada vez se aproxima más a la forma de la FDP (Figura 2.3b).

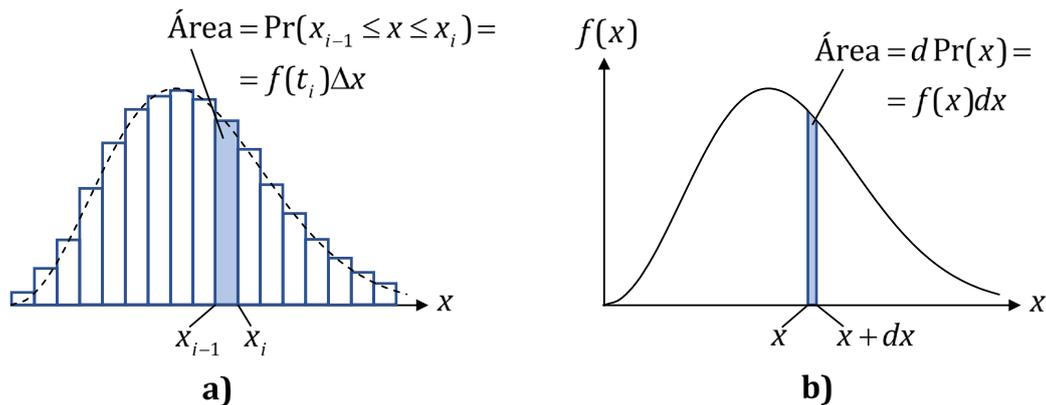


Figura 2.3: Interpretación de la función de densidad de probabilidad de una variable aleatoria continua.

Tomando como base todas las consideraciones anteriores, puede entenderse (formalmente) que las variables aleatorias continuas funcionan de igual modo que las discretas, pero sustituyendo las sumas por integrales. Así, para calcular el valor medio de la función $\phi(x)$ de la VAC x , se utiliza la expresión:

$$\langle \phi(x) \rangle = \int_{-\infty}^{\infty} \phi(x)f(x)dx \quad (2.12)$$

2.2.1.2. Variables aleatorias en más de una dimensión

Para ilustrar el caso de las variables aleatorias multidimensionales, vamos a estudiar el ejemplo concreto que usaremos en Teoría Cinética de los Gases: el espacio de las velocidades de las moléculas. La velocidad es una VAC tridimensional de componentes $\vec{v} = (v_x, v_y, v_z) \in \mathbb{R}^3$. No expondremos un ejemplo del caso discreto, ya que ya conocemos las similitudes con el caso continuo.

La FDP correspondiente la llamaremos $D(v_x, v_y, v_z)$ y verificará las propiedades:

$$D(v_x, v_y, v_z) \geq 0 \quad (2.13)$$

$$\text{Si } B \subset \mathbb{R}^3, \Pr(\vec{v} \in B) = \int_B D(v_x, v_y, v_z) dv_x dv_y dv_z \quad (2.14)$$

$$\int_{\mathbb{R}^3} D(v_x, v_y, v_z) dv_x dv_y dv_z = 1 \quad (2.15)$$

$$\langle \phi(v_x, v_y, v_z) \rangle = \int_{\mathbb{R}^3} \phi(v_x, v_y, v_z) D(v_x, v_y, v_z) dv_x dv_y dv_z \quad (2.16)$$

Uno de los aspectos que diferencian una VA multidimensional de una unidimensional es la posibilidad de definir **FDP marginales (FDPM)**. Estas funciones surgen cuando estamos interesados en la descripción de la distribución de probabilidad de algunas de las variables y el resto no nos interesan (en otras palabras, consideramos que pueden tener cualquier valor), y se construyen integrando la FDP a las variables irrelevantes. Por ejemplo, supongamos que estamos interesados en la distribución de la variable v_x , independientemente del valor que tengan v_y y v_z . La FDPM de v_x se obtiene como:

$$D_x(v_x) = \int_{-\infty}^{\infty} dv_z \int_{-\infty}^{\infty} dv_y D(v_x, v_y, v_z) \quad (2.17)$$

Las FDPM funcionan exactamente igual que las FDP. Por ejemplo, en este caso:

- $D_x(v_x) \geq 0$.
- Si $B \subset \mathbb{R}^3, \Pr(a \leq v_x \leq b) = \int_a^b \phi(v_x) D_x(v_x) dv_x$.
- $\int_{-\infty}^{\infty} D_x(v_x) dv_x = 1$.
- $\langle \phi(v_x) \rangle = \int_{-\infty}^{\infty} \phi(v_x) D_x(v_x) dv_x$.

2.3. Los gases ideales

2.3.1. Los gases ideales desde el punto de vista termodinámico

Desde el punto de vista termodinámico, los sistemas se caracterizan por **ecuaciones de estado**. Para un sistema cerrado y monocomponente que solo puede intercambiar trabajo de expansión, las ecuaciones de estado relevantes son dos: la **ecuación térmica de estado (ETE)** y la **ecuación energética de estado (EEE)**.

En el caso del gas ideal la ETE es la conocida **ecuación de Clapeyron** (1834):

$$pV = nRT \quad (2.18)$$

donde p es la presión, V el volumen, n la cantidad de sustancia, $R = 8.314\,462\,618\,153\,24\text{ J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$ es la **constante universal de los gases** y T la temperatura. Por su parte, la EEE nos indica que la energía interna U es solo función de su temperatura (Primera Ley de Joule):

$$U = U(T) \quad (2.19)$$

En Teoría Cinética no se opera con la cantidad de sustancia, sino con el número de partículas N :

$$N = nN_A \quad (2.20)$$

donde $N_A = 6.022\,140\,76 \cdot 10^{23}\text{ mol}^{-1}$ es la **constante de Avogadro**. Esto permite reescribir la ETE del gas ideal en la forma:

$$pV = Nk_B T \quad (2.21)$$

donde:

$$k_B = R/N_A = 1.380\,649 \cdot 10^{-23}\text{ J} \cdot \text{K}^{-1} \quad (2.22)$$

es la **constante de Boltzmann**.

2.3.2. El modelo cinético del gas ideal

En la segunda mitad del XIX, Clausius, Maxwell, Boltzmann y otros desarrollaron la denominada Teoría Cinética de los Gases, que se basa en la idea de que todos los gases se comportan de la misma manera en lo referente al movimiento molecular. Las hipótesis que definen este modelo son las siguientes:

1. En cualquier porción macroscópica de gas, **el número de moléculas es enorme**.
2. Las moléculas se encuentran **separadas por distancias grandes en comparación con sus dimensiones**.
3. Las moléculas no ejercen fuerzas entre sí, excepto cuando chocan, y no están sujetas a ningún campo externo. Entre choque y choque, las moléculas describen un **movimiento rectilíneo y uniforme**.
4. Los **choques** de moléculas entre sí y con las paredes del recipiente que las contiene son perfectamente **elásticos**. Las paredes del recipiente son perfectamente lisas y no hay cambio en la velocidad tangencial en un choque contra las paredes, por lo que el ángulo de incidencia con la normal a la pared es igual al ángulo con el que rebotan.
5. El gas es **homogéneo** y las moléculas están uniformemente distribuidas por el recipiente. Por tanto, su **densidad volumétrica, N/V , es la misma en cualquier porción macroscópica del gas**.
6. El gas es **isótropo** y todas las posibles orientaciones de las velocidades de las moléculas son igualmente probables.

2.4. El espacio de las velocidades del gas ideal

Para describir el estado de las partículas en Física Estadística clásica es necesario especificar tanto sus posiciones como sus velocidades. Las posiciones caracterizan la **energía potencial**, mientras que las velocidades determinan la **energía cinética**.

Suponiendo que las moléculas puedan ser consideradas como entidades elementales independientes del sistema, su **energía potencial** puede expresarse como la suma de varias contribuciones a la energía potencial:

$$\varepsilon_p = (\varepsilon_p)_{\text{intermolecular}} + (\varepsilon_p)_{\text{intramolecular}} + (\varepsilon_p)_{\text{campo externo}}$$

Las hipótesis del modelo cinético del gas ideal implican:

- $(\varepsilon_p)_{\text{intermolecular}} = 0$, pues no existen fuerzas intermoleculares.
- $(\varepsilon_p)_{\text{campo externo}} = 0$, pues no están sometidos a ningún campo externo.

Además, las paredes del recipiente se encargan de que no escapen moléculas al exterior y las confinan en un recinto de volumen V .

La $(\varepsilon_p)_{\text{intramolecular}}$ determina los movimientos de los átomos que componen la molécula en torno a sus posiciones de equilibrio. Para muchas aplicaciones, se puede suponer que estos movimientos son simplemente rotaciones en torno a un eje que pasa por el centro de masas de la molécula y vibraciones con respecto a dichas posiciones de equilibrio, y que dichos movimientos son independientes unos de otros y pueden estudiarse por separado (lo que se suele expresar diciendo que **los distintos movimientos de traslación, rotación y vibración no están acoplados**). Para estos casos, la **energía cinética** de la molécula puede ponerse muy aproximadamente como la suma de tres contribuciones independientes:

$$\varepsilon_c = (\varepsilon_c)_{\text{traslación}} + (\varepsilon_c)_{\text{rotación}} + (\varepsilon_c)_{\text{vibración}}$$

y la **energía total** de las moléculas del gas ideal dentro del recipiente será:

$$\varepsilon = (\varepsilon_c)_{\text{traslación}} + (\varepsilon_c)_{\text{rotación}} + (\varepsilon_c)_{\text{vibración}} + (\varepsilon_p)_{\text{intramolecular}} \quad (2.23)$$

Nos ceñiremos en este tema a los **gases ideales monoatómicos**, para los que sólo existe **energía cinética de traslación**. Los resultados que se obtengan serán aplicables a los **movimientos de traslación del centro de masas de las moléculas de gases ideales poliatómicos**. El movimiento de traslación está caracterizado por las componentes de las velocidades de los centros de masas de las moléculas y, por ello, el espacio de las velocidades juega un papel esencial.

2.4.1. Forma funcional de la FDP y FDPM radial y angular

Podemos hacer unas cuantas consideraciones sencillas de carácter físico que nos van a conducir a ciertas características de la forma funcional de la FDP de las velocidades, $D(v_x, v_y, v_z)$. Según la hipótesis de **isotropía** del gas, todas las direcciones de la velocidad son igualmente probables. Esto quiere decir que la FDP $D(v_x, v_y, v_z)$ debe ser una función únicamente del módulo v de la velocidad:

$$D(v_x, v_y, v_z) = f(v) \quad (2.24)$$

(al módulo de la velocidad lo llamaremos también **celeridad**). Con esta información, hagamos un cambio a coordenadas esféricas y tratemos de hallar tanto la FDP en esféricas, que llamaremos $\rho(v, \theta, \phi)$, como las FDPM asociadas a v y a la parte angular. Teniendo en cuenta la definición de $\rho(v, \theta, \phi)$ y el factor de escala en el elemento de volumen, la probabilidad de que (v, θ, ϕ) esté en una región B del espacio de velocidades es:

$$\Pr(\vec{v} \in B) = \int_B \rho(v, \theta, \phi) dv d\theta d\phi = \int_B f(v) v^2 \sin\theta dv d\theta d\phi \quad (2.25)$$

de donde:

$$\rho(v, \theta, \phi) = f(v) v^2 \sin\theta \quad (2.26)$$

La función $f(v)$ está restringida por la condición de normalización:

$$1 = \int_0^\infty dv \int_0^{2\pi} d\phi \int_0^\pi d\theta \rho(v, \theta, \phi) = \int_0^\infty dv \cdot v^2 f(v) \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta = 4\pi \int_0^\infty dv \cdot v^2 f(v) \quad (2.27)$$

La **FDPM del módulo de la velocidad**, que denominaremos $F(v)$, se calcula como:

$$F(v) = \int_0^{2\pi} d\phi \int_0^\pi d\theta \rho(v, \theta, \phi) = v^2 f(v) \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta = 4\pi v^2 f(v) \quad (2.28)$$

y la **FDPM de la parte angular**, $G(\theta, \phi)$, es:

$$G(\theta, \phi) = \int_0^\infty dv \rho(v, \theta, \phi) = \sin\theta \int_0^\infty dv \cdot v^2 f(v) = \frac{\sin\theta}{4\pi} \quad (2.29)$$

que no depende de ϕ . Como puede verse, la **FDP en esféricas**:

$$\rho(v, \theta, \phi) = F(v)G(\theta, \phi) \quad (2.30)$$

ha sido factorizada en dos términos, que a su vez son las FDPM del módulo y la dirección (ángulos), respectivamente. Esto será muy útil a la hora de realizar ciertos cálculos. Por ejemplo, si se quiere calcular el valor medio de una función del módulo de la velocidad, $\phi(v)$, simplemente hay que calcular la integral:

$$\langle \phi(v) \rangle = \int_0^\infty dv \phi(v) F(v) \quad (2.31)$$

El caso $\phi(v) = v$ nos permitiría obtener el valor medio del módulo de la velocidad, $\phi(v) = v^2$ nos llevaría al valor medio del cuadrado del módulo de la velocidad, etc.

2.4.2. Cálculos estadísticos en el espacio de velocidades

Nos vamos a encontrar con el problema de hallar cierta magnitud de carácter macroscópico, digamos A , a partir de las velocidades del gas. Naturalmente, para poder resolver este problema, es necesario atribuir a algún mecanismo mecánico microscópico el origen de la magnitud A , lo que no siempre es posible. Estos cálculos pueden realizarse de forma sistemática siguiendo dos pasos:

1. En primer lugar, ha de resolverse el **problema mecánico**: suponiendo que todas las moléculas tienen una velocidad \vec{v} , ¿cuál sería el valor $A_{\vec{v}}$ de la magnitud A ?
2. Una vez conocido el valor de $A_{\vec{v}}$, hay que resolver el **problema estadístico**: suponiendo que la magnitud A es aditiva con respecto a las moléculas del gas (cosa que hay que justificar en cada caso particular), ¿cuál es el valor total de la magnitud A ? Éste puede hallarse como:

$$\begin{aligned} A &= \sum_{\substack{\text{todas las} \\ \vec{v} \text{ posibles}}} (\text{n}^\circ \text{ moléc. con veloc. } \vec{v}) \cdot (\text{contribución a } A \text{ de 1 moléc. con veloc. } \vec{v}) = \\ &= \sum_{\substack{\text{todas las} \\ \vec{v} \text{ posibles}}} (\text{n}^\circ \text{ moléc. con veloc. } \vec{v}) \cdot \frac{(\text{valor de } A \text{ si todas las moléc. tienen veloc. } \vec{v})}{(\text{n}^\circ \text{ total moléc.})} = \\ &= \sum_{\substack{\text{todas las} \\ \vec{v} \text{ posibles}}} \frac{(\text{n}^\circ \text{ moléc. con veloc. } \vec{v})}{(\text{n}^\circ \text{ total moléc.})} \cdot A_{\vec{v}} = \sum_{\substack{\text{todas las} \\ \vec{v} \text{ posibles}}} (\text{fracción de moléc. con veloc. } \vec{v}) \cdot A_{\vec{v}} \end{aligned} \quad (2.32)$$

Es decir, A es la suma de los productos de $A_{\vec{v}}$ por la fracción de moléculas con velocidad \vec{v} . Dada la enormidad del número de moléculas presentes en el gas, esta fracción de moléculas puede identificarse con la probabilidad de que una molécula tenga velocidad \vec{v} . Pero entonces esa suma de productos de $A_{\vec{v}}$ por la probabilidad de que las moléculas tengan velocidad \vec{v} no es más que el valor medio de $A_{\vec{v}}$ y, por tanto:

$$A = \sum_{\substack{\text{todas las} \\ \vec{v} \text{ posibles}}} \text{Pr}(\vec{v}) \cdot A_{\vec{v}} = \langle A_{\vec{v}} \rangle \quad (2.33)$$

Es decir, **el valor macroscópico de una magnitud aditiva se obtiene promediando su valor $A_{\vec{v}}$, solución del problema mecánico asociado, a todas las velocidades posibles de las moléculas del gas.** Para ser precisos, el razonamiento anterior debería matizarse teniendo en cuenta que la velocidad no es una variable aleatoria discreta, sino continua, reemplazando las sumas por integrales donde sea necesario, pero este matiz matemático no afecta a los conceptos subyacentes.

2.5. Densidad de flujo molecular

Se define la densidad de flujo molecular, η , como la cantidad de moléculas que atraviesan una superficie o chocan con ella desde uno de sus lados por unidad de tiempo y por unidad de área.

2.5.1. Problema mecánico

Consideremos una pequeña superficie ΔS macroscópica y calculemos el número $\Delta N_{\vec{v}}$ de moléculas que chocan con ella en un pequeño intervalo de tiempo Δt , en el supuesto de que todas las moléculas del gas tengan una velocidad \vec{v} con cierto módulo y dirección.

Sin pérdida de generalidad, podemos suponer que dicha superficie es perpendicular al eje Z (Figura 2.4). Esta posibilidad se debe a que, debido a la homogeneidad e isotropía del gas, todas las posibles posiciones y orientaciones de la superficie son físicamente equivalentes. Esto también permite que la superficie tenga la forma que más convenga; podemos suponer que se trata, por ejemplo, de un cuadrado. Por otro lado, si la superficie es perpendicular al eje Z, para ángulos $\theta \geq \pi/2$ no choca ninguna molécula, ya que se alejan de la superficie. Es decir:

$$\eta_{\vec{v}} = 0 \quad \text{si } \theta \geq \pi/2 \quad (2.34)$$

Para hallar el número de moléculas que chocan con la pared en el caso $\theta < \pi/2$, comencemos notando que sólo chocarán con ella las que tengan tiempo de llegar a la superficie en un tiempo menor o igual que Δt . Las últimas que llegan son las que tardan un tiempo exactamente igual a Δt , recorriendo una distancia en el eje Z igual a $v_z \Delta t = v \cos \theta \Delta t$. Por lo tanto, este número $\Delta N_{\vec{v}}$ coincide con las que se encuentran en un prisma con base de área ΔS y altura $v \cos \theta \Delta t$. La densidad volumétrica de moléculas en el prisma es igual a su valor macroscópico N/V (donde N es el número total de moléculas consideradas y V el volumen macroscópico que ocupan), pues se trata de una porción macroscópica de gas que contiene un número suficientemente grande de ellas:

$$\frac{N}{V} = \frac{\Delta N_{\vec{v}}}{\Delta S \cdot v \cos \theta \Delta t} \quad \text{si } \theta < \pi/2 \quad (2.35)$$

Por lo tanto, la densidad de flujo molecular será:

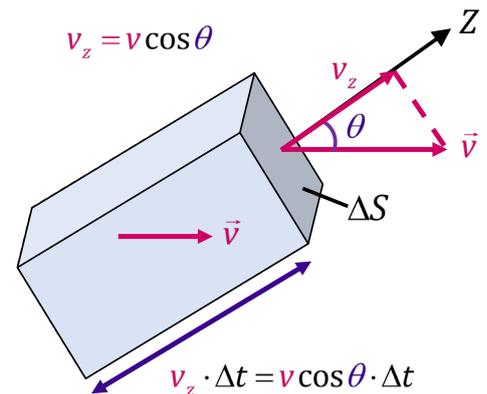


Figura 2.4: Esquema para el problema mecánico de la densidad de flujo molecular.

$$\eta_{\vec{v}} = \frac{\Delta N_{\vec{v}}}{\Delta S \Delta t} = \frac{(N/V) \Delta S \cdot v \cos \theta \Delta t}{\Delta S \Delta t} = \frac{N}{V} v \cos \theta \quad \text{si } \theta < \pi/2 \quad (2.36)$$

Resumiendo los dos casos tratados:

$$\boxed{\begin{array}{ll} \eta_{\vec{v}} = 0 & \text{si } \theta \geq \pi/2 \\ \eta_{\vec{v}} = \frac{N}{V} v \cos \theta & \text{si } \theta < \pi/2 \end{array}} \quad (2.37)$$

2.5.2. Problema estadístico

El valor de la densidad de flujo molecular que buscamos será el promedio de $\eta_{\vec{v}}$ a todas las velocidades de las moléculas. Como para $\theta \geq \pi/2$ el valor de $\eta_{\vec{v}}$ es cero, la variable θ solo hay que integrarla entre 0 y $\pi/2$:

$$\begin{aligned} \eta &= \langle \eta_{\vec{v}} \rangle = \int_0^\infty dv \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta \eta_{\vec{v}} F(v) G(\theta, \phi) \\ &= \frac{N}{V} \frac{1}{4\pi} \left(\int_0^\infty dv \cdot v F(v) \right) \left(\int_0^{2\pi} d\phi \right) \left(\int_0^{\pi/2} d\theta \sin \theta \cos \theta \right) \\ &= \frac{N}{V} \frac{1}{4\pi} (\langle v \rangle) (2\pi) \left(\frac{1}{2} \int_0^{\pi/2} d\theta \sin 2\theta \right) \end{aligned} \quad (2.38)$$

La última integral vale $[-\cos 2\theta/2]_0^{\pi/2} = (-\cos \pi + \cos 0)/2 = 1$ y, por tanto:

$$\boxed{\eta = \frac{1}{4} \frac{N}{V} \langle v \rangle} \quad (2.39)$$

2.6. Presión

Como es bien sabido, la **presión** sobre una superficie se define como la fuerza ejercida perpendicularmente sobre ella por unidad de área. Según su propia definición, la mencionada fuerza es igual a la cantidad de momento lineal transmitida a la superficie por unidad de tiempo.

2.6.1. Problema mecánico

Consideremos una pequeña superficie ΔS macroscópica y calculemos el momento lineal transmitido por $\Delta N_{\vec{v}}$ moléculas que chocan con ella en un pequeño intervalo de tiempo Δt , en el supuesto de que todas las moléculas del gas tengan una velocidad \vec{v} . Procediendo de forma análoga al cálculo de la densidad de flujo molecular, la superficie puede ser considerada perpendicular al eje Z.

Esta cantidad de momento será igual a $\Delta N_{\vec{v}}$ veces la cantidad de momento $\Delta K_{1,\vec{v}}$ transmitida por una molécula. En consecuencia, la presión $p_{\vec{v}}$ es:

$$p_{\vec{v}} = \frac{\Delta N_{\vec{v}} \Delta K_{1,\vec{v}}}{\Delta S \Delta t} = \eta_{\vec{v}} \Delta K_{1,\vec{v}} \quad (2.40)$$

Debido a la hipótesis de choque elástico y de paredes perfectamente lisas, la componente tangencial del momento no se modifica cuando una molécula choca con la superficie. Por lo tanto, las moléculas simplemente invierten la componente normal (componente Z) y, según el principio de conservación del momento lineal:

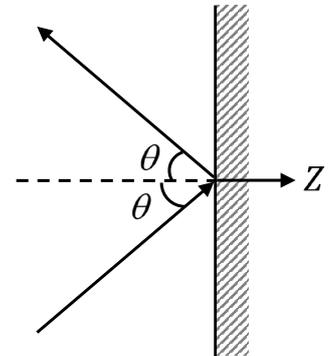


Figura 2.5: Esquema para el problema mecánico de la presión.

$$\Delta K_{1,\vec{v}} = mv\cos\theta - (-mv\cos\theta) = 2mv\cos\theta \quad (2.41)$$

donde m es la masa de una molécula. Finalmente, obtenemos:

$$\boxed{\begin{aligned} p_{\vec{v}} &= 0 && \text{si } \theta \geq \pi/2 \\ p_{\vec{v}} &= 2 \frac{N}{V} mv^2 \cos^2\theta && \text{si } \theta < \pi/2 \end{aligned}} \quad (2.42)$$

2.6.2. Problema estadístico

Calculando el promedio de $p_{\vec{v}}$ a todas las velocidades:

$$\begin{aligned} p &= \langle p_{\vec{v}} \rangle = \int_0^\infty dv \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta p_{\vec{v}} F(v) G(\theta, \phi) \\ &= \frac{N}{V} \frac{2m}{4\pi} \left(\int_0^\infty dv \cdot v^2 F(v) \right) \left(\int_0^{2\pi} d\phi \right) \left(\int_0^{\pi/2} d\theta \sin\theta \cos^2\theta \right) \\ &= \frac{N}{V} \frac{2m}{4\pi} (\langle v^2 \rangle) (2\pi) \left[-\frac{\cos^3\theta}{3} \right]_0^{\pi/2} \end{aligned} \quad (2.43)$$

es decir:

$$\boxed{p = \frac{1}{3} \frac{N}{V} m \langle v^2 \rangle} \quad (2.44)$$

2.7. Relación entre Teoría Cinética de los gases y Termodinámica

Comparando la ETE de los gases ideales, $p = Nk_B T/V$, y el cálculo cinético de la presión, $p = Nm\langle v^2 \rangle/3V$, deducimos que:

$$\boxed{\langle v^2 \rangle = \frac{3k_B T}{m} = \frac{3RT}{M}} \quad (2.45)$$

donde M es la masa molar del gas. Esta ecuación es el puente entre la Teoría Cinética y la Termodinámica. Es importante notar que **hemos tenido que recurrir directamente a la Termodinámica** para obtenerla, lo cual se debe a que **la temperatura es una propiedad macroscópica, de conjunto, del sistema, y no tiene análogo mecánico** al que podamos recurrir. **No tiene sentido hablar de la temperatura de una molécula.**

Además, debido a la isotropía del gas, todas las direcciones son físicamente equivalentes y $\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle$. Como $v^2 = v_x^2 + v_y^2 + v_z^2$, se tiene:

$$\boxed{\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle = \frac{\langle v^2 \rangle}{3} = \frac{k_B T}{m}} \quad (2.46)$$

2.8. Recorrido libre medio y ecuación de supervivencia

En los movimientos de las moléculas que tienen lugar en el gas, las distancias recorridas por una molécula entre colisiones van variando entre colisión y colisión. Una magnitud útil es el valor medio de la distancia entre colisiones, denominado **recorrido libre medio** λ . Un cálculo medianamente riguroso del recorrido libre medio es tremendamente complicado y queda fuera de los límites de este curso. Es por ello por lo que aquí simplemente nos restringiremos a unos razonamientos muy sencillos que ilustren el significado

físico de esta magnitud. Un tratamiento más detallado de estas cuestiones lleva al estudio de **fenómenos de transporte**: conductividad térmica, conductividad eléctrica, viscosidad, autodifusión...

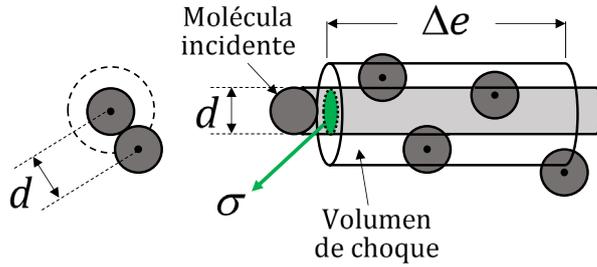


Figura 2.6: Esquema para la estimación del recorrido libre medio.

Para empezar, supondremos que la frecuencia de colisión de una molécula con las demás es independiente del tiempo y que el recorrido libre medio se puede calcular a partir de la expresión:

$$\lambda = \frac{\Delta e}{\Delta N_c} \quad (2.47)$$

donde Δe es el espacio que recorrería una molécula (llamémosla A) en línea recta en un tiempo Δt (suficientemente grande) y ΔN_c es el número (macroscópico) de moléculas con las que puede chocar al recorrer una línea recta en ese tiempo Δt . Podemos obtener una buena estimación del recorrido libre medio utilizando las siguientes hipótesis simplificadoras:

- El espacio recorrido por la molécula (numerador de la ecuación (2.47)) se puede calcular como si la molécula A llevara una celeridad constante e igual a su **celeridad cuadrática media**, también llamada **celeridad rms** (*root-mean-square*) ($v_{\text{rms}} = \sqrt{\langle v^2 \rangle}$). Es decir:

$$\Delta e = v_{\text{rms}} \Delta t \quad (2.48)$$

- Si todas las moléculas fueran estrictamente puntuales, su volumen sería nulo y no chocarían nunca, por lo que consideraremos que son esferas duras con radio r . Si nos fijamos en una cierta molécula A, cuando en su movimiento choca con otra molécula, sus centros quedan separados una distancia $2r$, y todo ocurre como si la molécula A tuviera radio $d = 2r$ y las moléculas con las que choca fueran puntuales. El área efectiva de colisión:

$$\sigma = \pi(2r)^2 = 4\pi r^2 \quad (2.49)$$

se denomina **sección eficaz de choque**.

- El número de choques puede calcularse aproximadamente como si el resto de las moléculas estuvieran paradas y la celeridad de la molécula A fuera constante e igual a la celeridad relativa cuadrática media de la molécula A respecto de otra molécula (que llamaremos $v_{\text{rel,rms}}$). Sea B cualquiera de estas otras moléculas. Entonces:

$$v_{\text{rel,rms}}^2 = \langle (\vec{v}_A - \vec{v}_B) \cdot (\vec{v}_A - \vec{v}_B) \rangle = \langle v_A^2 + v_B^2 - 2 \cdot \vec{v}_A \cdot \vec{v}_B \rangle = \langle v_A^2 \rangle + \langle v_B^2 \rangle - 2 \langle \vec{v}_A \cdot \vec{v}_B \rangle \quad (2.50)$$

Las magnitudes $\langle v_A^2 \rangle$ y $\langle v_B^2 \rangle$ son iguales, ya que las dos moléculas son idénticas, y coinciden con lo que en anteriores apartados hemos llamado simplemente $\langle v^2 \rangle$. Por otro lado, \vec{v}_A y \vec{v}_B son estadísticamente independientes dado que las moléculas no interactúan entre sí entre choque y choque, por lo que $\langle \vec{v}_A \cdot \vec{v}_B \rangle = \langle \vec{v}_A \rangle \cdot \langle \vec{v}_B \rangle$, y su valor promedio debe ser igual a cero debido a la completa isotropía del gas. Esta última afirmación se obtiene como un resultado matemático cuando se estudia la distribución de probabilidad de las velocidades o momentos de las moléculas (sección 5.3, ecuación (5.26)). En consecuencia, se tiene para la $v_{\text{rel,rms}}$:

$$v_{\text{rel,rms}} = \sqrt{2 \langle v^2 \rangle} = \sqrt{2} \cdot v_{\text{rms}} \quad (2.51)$$

El denominador de la ecuación (2.47) puede calcularse a partir de las dos últimas hipótesis. El número de moléculas ΔN_c es igual al de las contenidas en un cilindro cuya base tiene un área σ y cuya longitud es

$v_{\text{rel,rms}}\Delta t$. Como ese cilindro es macroscópico, la densidad volumétrica de moléculas es igual a su valor de equilibrio, N/V . En consecuencia:

$$\frac{N}{V} = \frac{\Delta N_c}{\sigma v_{\text{rel,rms}}\Delta t} \Rightarrow \Delta N_c = \frac{N}{V} \sigma v_{\text{rel,rms}}\Delta t = \frac{N}{V} \sigma \sqrt{2} \cdot v_{\text{rms}}\Delta t \quad (2.52)$$

En consecuencia, el recorrido libre medio es, en esta aproximación:

$$\lambda = \frac{1}{\sqrt{2} \cdot (N/V)\sigma} \quad (2.53)$$

La **frecuencia de colisión**, o número de colisiones por unidad de tiempo, es:

$$f_c = \frac{\Delta N_c}{\Delta t} = \frac{N}{V} \sigma \sqrt{2} \cdot v_{\text{rms}} = \frac{v_{\text{rms}}}{\lambda} \quad (2.54)$$

y el tiempo medio entre colisiones, o **tiempo libre medio** τ , es el recíproco de la frecuencia de choque:

$$\tau = \frac{1}{f_c} = \frac{1}{(N/V)\sigma\sqrt{2} \cdot v_{\text{rms}}} \quad (2.55)$$

El recorrido libre medio de una molécula debe ser el valor medio del recorrido libre de acuerdo con una cierta distribución de probabilidad. Bajo las hipótesis de que los choques se producen siempre entre 2 moléculas máximo y de que la frecuencia de colisión no depende del tiempo, puede obtenerse ⁶ que dicha distribución de probabilidad tiene una dependencia exponencial con la distancia recorrida por la molécula. Estas hipótesis parecen ser razonables cuando se trata de gases diluidos. La **probabilidad de supervivencia** de una molécula tras haber recorrido una distancia x (esto es, la probabilidad de NO haber chocado con ninguna otra molécula en ese recorrido) viene dada por la expresión:

$$P_{\text{sup}}(x) = e^{-x/\lambda} \quad (2.56)$$

Si se considera un conjunto macroscópico de N_0 moléculas, el número de ellas cuyos recorridos libres son mayores que una cantidad x_0 es:

$$N(x \geq x_0) = N_0 P_{\text{sup}}(x_0) = N_0 e^{-x_0/\lambda} \quad (2.57)$$

(relación denominada **ecuación de supervivencia**). Mediante simples razonamientos, que se dejan al lector, puede calcularse la probabilidad de haber chocado tras haber recorrido una distancia x y el número de moléculas de un conjunto dado cuyos recorridos son menores que una cantidad x_0 .

⁶ Véase J. Brey Ábalo, J. de la Rubia Pacheco, J. de la Rubia Sánchez, "Mecánica Estadística", UNED, 2001.

Ejercicios – Teoría Cinética de los gases ideales

1.- La densidad del amoníaco a 1.04 atm y 20 °C es 0.736 kg·m⁻³. Sabiendo que está compuesto de nitrógeno e hidrógeno, cuyas masas atómicas relativas son 14 y 1 respectivamente, deducir su fórmula molecular.

2.- Se sabe que en un cristal de cloruro sódico la malla elemental es un cubo cuya arista mide 5.63 Å (1 Å = 10⁻¹⁰ m) a temperatura 20 °C, y que contiene cuatro átomos de cloro y cuatro de sodio. La densidad del cloruro sódico a esa temperatura es 2.1632 g·cm⁻³. Las masas atómicas relativas del cloro y del sodio son, respectivamente, 35.457 y 22.997. Calcular la constante de Avogadro.

R: 6.06·10²³ mol⁻¹.

3.- La densidad del helio a 0 °C y 760 mmHg es de 0.1785 g·L⁻¹. Estimar la distancia media que separa los centros de las moléculas.

R: 33.4 Å.

4.- Calcular el número de moléculas que hay en un cubo de 1 mm de lado en un gas a 300 K cuando la presión es de 10⁻³ mmHg.

R: 3.21·10¹⁰.

5.- Las bombas de vacío modernas permiten alcanzar una rarificación de hasta 4·10⁻¹⁵ atm a 300 K. Hallar el número de moléculas de un gas en 1 cm³ y estimar la distancia media entre ellas a esa presión.

R: 9.783·10⁴; 2.170·10⁻⁴ m.

6.- Sean las coordenadas esféricas cuya relación con las cartesianas viene dada por las ecuaciones: $x = r \cdot \cos\theta \cdot \cos\varphi$; $y = r \cdot \cos\theta \cdot \sin\varphi$; $z = r \cdot \sin\theta$. La diferencial del ángulo sólido del eje definido por (θ, φ) es: $d\Omega = \sin\theta \cdot d\theta \cdot d\varphi$.

a) ¿Qué fracción de las moléculas de un gas tienen velocidades cuya dirección corta el área limitada por el cono de ángulo sólido $\theta = 45^\circ$, $\Delta\theta = 0.05$ rad, $\varphi = 30^\circ$, $\Delta\varphi = 0.05$ rad?

b) Calcular el número de moléculas en 1 mol de gas cuyas velocidades están en un ángulo sólido definido por $\theta = 30^\circ$; $\Delta\theta = \Delta\varphi = 1^\circ$.

R: a) 1.41·10⁻⁴, b) 5.63·10¹¹.

7.- Se tiene una burbuja de aire esférica, de radio 0.1 mm, a 1 atm y 300 K. Sabiendo que la masa molar media del aire es de 28.97 g·mol⁻¹ y que la celeridad media de las moléculas de un gas de masa molar M a temperatura T es $\langle v \rangle = (8RT/\pi M)^{1/2}$:

a) ¿Cuál es la frecuencia promedio de choque con las paredes por molécula de aire?

b) ¿Cuántas moléculas de aire hay en ella?

R: a) 3.5·10⁶ s⁻¹; b) 1.02·10¹⁴.

8.- A 273.15 K y 1 atm, la densidad del aire es 1.293 kg·m⁻³. Calcular, en esas condiciones, el número de choques de las moléculas de aire durante 1 s sobre 1 mm² de pared, si la celeridad media de las moléculas de un gas de masa molar M a temperatura T es $\langle v \rangle = (8RT/\pi M)^{1/2}$.

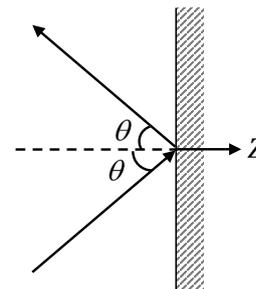
R: 3.00·10²¹.

9.- En un recipiente de capacidad $V = 3.6$ L se hace el vacío hasta $p = 0.001$ mmHg. Este recipiente queda en un ambiente a $T = 300$ K y $p_0 = 1$ atm, y por tener un pequeño orificio la presión aumenta en 10 horas hasta ser de $p' = 1$ mmHg. Sabiendo que la masa molar media del aire es de 28.97 g·mol⁻¹, y que la celeridad media de las moléculas de un gas de masa molar M a temperatura T es $\langle v \rangle = (8RT/\pi M)^{1/2}$, calcular el área del orificio.

R: 1.12·10⁻¹² m².

10.- Un haz paralelo de moléculas de nitrógeno (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$), que se desplaza a una celeridad $v = 400 \text{ m}\cdot\text{s}^{-1}$, incide sobre una pared formando un ángulo $\theta = 30^\circ$ con su normal. La concentración de las moléculas en el haz es $N/V = 0.90 \cdot 10^{19} \text{ cm}^{-3}$. Hallar la presión que el haz ejerce sobre la pared, suponiendo que las moléculas chocan elásticamente con la pared y que el ángulo de incidencia y el de reflexión son iguales.

R: 10^5 Pa .



11.- Calcular la presión de vapor del sodio (de masa molar $22.997 \text{ g}\cdot\text{mol}^{-1}$) a 261°C , sabiendo que a través de un pequeño orificio de 0.25 mm^2 de área, practicado en la pared de un balón conteniendo sodio líquido en equilibrio con su vapor, pasan $6.2 \cdot 10^{-4} \text{ g}$ de sodio en 2 horas. La celeridad media de las moléculas de un gas de masa molar M a temperatura T es $\langle v \rangle = (8RT/\pi M)^{1/2}$.

R: 0.38 Pa .

12.- Se estima que la presión de vapor de la plata (de masa molar $107.87 \text{ g}\cdot\text{mol}^{-1}$) a 2000 K es 170 mmHg . Calcular la masa total de las moléculas de plata que chocan durante 1 s contra una porción de 1 cm^2 de las paredes de un recipiente conteniendo plata en equilibrio con su vapor a dicha temperatura. La celeridad media de las moléculas de un gas de masa molar M a temperatura T es $\langle v \rangle = (8RT/\pi M)^{1/2}$.

R: 2.30 g .

13.- Para un gas en determinadas condiciones, el recorrido libre medio es $\lambda = 0.01 \text{ mm}$. En un mol de este gas, ¿cuántas moléculas existen que posean un recorrido libre superior a 0.02 mm ?

R: $8.15 \cdot 10^{22}$.

14.- El recorrido libre medio de las moléculas de un cierto gas es: $\lambda = 3 \cdot 10^{-2} \text{ cm}$. En un mol de este gas:

a) ¿Cuántas moléculas hay cuyos recorridos libres sean mayores que $x_1 = 4 \cdot 10^{-2} \text{ cm}$?

b) ¿Cuántas moléculas hay cuyos recorridos libres estén comprendidos entre x_1 y $x_2 = 8 \cdot 10^{-2} \text{ cm}$?

R: a) $1.59 \cdot 10^{23}$; b) $1.17 \cdot 10^{23}$.

15.- Un recipiente contiene nitrógeno (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$) a la temperatura de 27°C y 10^{-6} mmHg de presión. El radio medio de la molécula de nitrógeno es 1.88 \AA . Partiendo del número de recorridos libres que realiza una molécula en un minuto (60 s), calcular:

a) Cuántos de estos recorridos libres son mayores que el recorrido libre medio.

b) Cuántos son menores que el recorrido libre medio.

c) Cuántos están comprendidos entre el recorrido libre medio $\pm 1 \text{ m}$.

d) Cuántos son exactamente iguales al recorrido libre medio.

R: a) 230; b) 395; c) 9; d) 0.

3. Principio de Máxima Entropía

La formulación de la Física Estadística que estudiaremos hace uso de un método de inferencia de probabilidades denominado Principio de Máxima Entropía. Para abordarlo con cierto rigor, a los conceptos de Probabilidad ya estudiados hemos de añadir algunos de **Combinatoria**. Por completitud, se incluyen los 6 tipos de agrupaciones típicas de la combinatoria, ya que todas ellas tienen aplicación en algún problema físico. No obstante, **en este curso nos centraremos en aplicaciones que requieren de las permutaciones (especialmente con repetición)**. De hecho, los argumentos que justifican la utilización del Principio de Máxima Entropía para la inferencia de probabilidades se basan precisamente en las propiedades de las permutaciones con repetición.

3.1. Combinatoria

La **Combinatoria** es una rama de las matemáticas que estudia, entre otras cosas, el cálculo del número de formas diferentes en las que se pueden agrupar los elementos de un conjunto dado. Desde sus inicios, ha ido estrechamente ligada al cálculo de probabilidades.

La pregunta que deseamos responder es la siguiente:

Partiendo de un conjunto de N elementos ¿cuántas agrupaciones distintas (de acuerdo con unas reglas o restricciones dadas) de m elementos, seleccionados entre los N originales, podemos construir?

A la hora de decidir qué conjuntos consideramos “distintos”, es necesario tener en cuenta:

- Si dos agrupaciones con distinta ordenación se consideran distintas (**variaciones** y **permutaciones**) o no (**combinaciones**).
- Si cada elemento se puede incluir en cada agrupación sólo una vez (agrupaciones **sin repetición**) o varias veces (agrupaciones **con repetición**); o, visto desde otro punto de vista, si varios elementos de la agrupación se consideran **idénticos** o no.

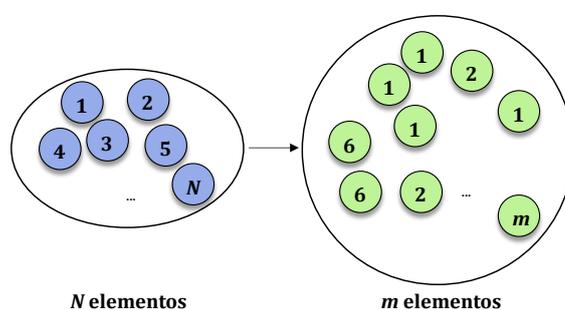


Figura 3.1: Esquemización de la creación de una agrupación de elementos.

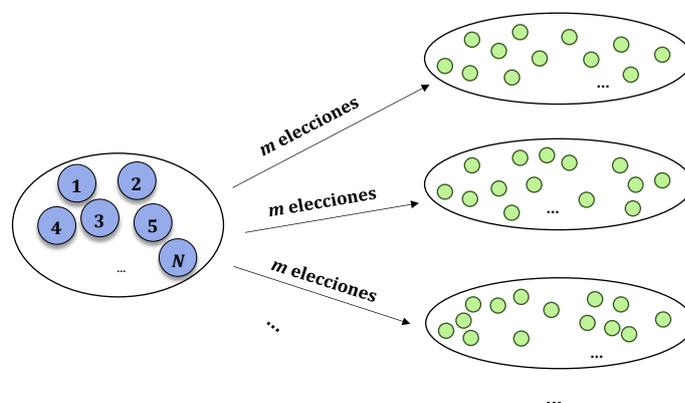


Figura 3.2: Esquemización de la creación de distintas agrupaciones de m elementos a partir de un conjunto de N elementos.

Todos estos problemas se reducen a los 6 cálculos combinatorios básicos:

- Variaciones sin repetición.
- Variaciones con repetición.
- Permutaciones sin repetición.
- Permutaciones con repetición.
- Combinaciones sin repetición.
- Combinaciones con repetición.

3.1.1. Variaciones sin repetición

Se denominan **variaciones sin repetición de N elementos tomados de m en m** ($V_{N,m}$) a las distintas agrupaciones de m elementos que se pueden hacer de un conjunto de N elementos, de tal forma que:

- Dos agrupaciones con distinta ordenación se consideran distintas.
- Cada elemento sólo se puede incluir en cada agrupación una vez.

Cálculo de las variaciones sin repetición. Tomamos el primer elemento: tenemos N posibilidades. Tomamos el segundo: como no podemos volver a coger el primer elemento tenemos $(N - 1)$ posibilidades. Tomamos el tercero: tenemos $(N - 2)$. Y así sucesivamente hasta haber seleccionado m :

$$V_{N,m} = N(N - 1)(N - 2) \dots (N - m + 1) = \frac{N!}{(N - m)!} \quad (3.1)$$

3.1.2. Variaciones con repetición

Se denominan **variaciones con repetición de N elementos tomados de m en m** ($VR_{N,m}$) a las distintas agrupaciones de m elementos que se pueden hacer de un conjunto de N elementos, de tal forma que:

- Dos agrupaciones con distinta ordenación se consideran distintas.
- Se puede incluir varias veces el mismo elemento dentro de cada agrupación.

En este caso, m puede ser mayor que N .

Cálculo de las variaciones con repetición. Tomamos el primer elemento: tenemos N posibilidades. Tomamos el segundo: como sí podemos volver a coger el primer elemento tenemos N posibilidades. Y así sucesivamente hasta haber seleccionado m :

$$VR_{N,m} = N^m \quad (3.2)$$

3.1.3. Permutaciones sin repetición

Se denominan **permutaciones de N elementos** (P_N) a las distintas ordenaciones que se pueden hacer de los N elementos. Se corresponden con las variaciones sin repetición de N elementos tomados de N en N .

$$P_N = V_{N,N} = N! \quad (3.3)$$

Se puede demostrar que, para $N > 0$, el factorial admite el siguiente desarrollo (**desarrollo asintótico del factorial**):

$$\ln(N!) = N \ln N - N + \frac{1}{2} \ln(2\pi N) + \frac{1}{12N} - \frac{1}{360N^3} + \frac{1}{1260N^5} - \frac{1}{1680N^7} \dots \quad (3.4)$$

Cuando N es del orden del número de partículas de un sistema macroscópico, sólo los dos primeros términos son necesarios, pues el tercero es muchos órdenes de magnitud más pequeño. Dicha aproximación recibe el nombre de **aproximación de Stirling**:

$$\ln(N!) \approx N \ln N - N \quad , \quad \text{si } N \gg \ln N \quad (3.5)$$

3.1.4. Permutaciones con repetición

Se denominan **permutaciones con repetición de N elementos, en donde hay n_1 elementos idénticos, n_2 elementos idénticos, ..., n_s elementos idénticos** ($P_N^{n_1, \dots, n_s}$), a las distintas ordenaciones de dicho

conjunto de elementos, considerando iguales todas las ordenaciones que difieran solo en la posición relativa de elementos idénticos entre sí.

Cálculo de las permutaciones con repetición. Hay $P_{n_1} = n_1!$ ordenaciones de los n_1 elementos iguales que dan lugar a la misma ordenación de los N elementos, ..., $P_{n_s} = n_s!$ ordenaciones de los n_s elementos iguales que dan lugar a la misma ordenación de los N elementos. Por tanto:

$$P_N^{n_1, \dots, n_s} = \frac{P_N}{P_{n_1} \dots P_{n_s}} = \frac{N!}{n_1! \dots n_s!} \tag{3.6}$$

3.1.5. Combinaciones sin repetición

Se denominan **combinaciones sin repetición de N elementos tomados de m en m** ($C_{N,m}$) a las distintas agrupaciones de m elementos que se pueden hacer de un conjunto de N elementos, de tal forma que:

- Dos agrupaciones con distinta ordenación se consideran iguales.
- Cada elemento sólo se puede incluir en cada agrupación una vez.

Cálculo de las combinaciones sin repetición. Se pueden calcular como el cociente entre las correspondientes variaciones sin repetición ($V_{N,m}$), para las que dos agrupaciones con distinta ordenación se consideran en la cuenta, entre el número total de ordenaciones de los m elementos (P_m):

$$C_{N,m} = \frac{V_{N,m}}{P_m} = \frac{N(N-1) \dots (N-m+1)}{m!} = \frac{N!}{m! (N-m)!} = \binom{N}{m} \tag{3.7}$$

3.1.6. Combinaciones con repetición

Se denominan **combinaciones con repetición de N elementos tomados de m en m** ($CR_{N,m}$) a las distintas agrupaciones de m elementos que se pueden hacer de un conjunto de N elementos, de tal forma que:

- Dos agrupaciones con distinta ordenación se consideran iguales.
- Se puede incluir varias veces el mismo elemento dentro de cada agrupación.

En este caso, m puede ser mayor que N .

Cálculo de las combinaciones con repetición. La estrategia para este cálculo se basa en la representación del problema en barras y estrellas. Se quieren formar agrupaciones de m elementos sin importar el orden, lo que se representa con una fila de m estrellas:

$$\overbrace{***..**}^{m \text{ estrellas}}$$

Cada uno de estos * puede representar a cualquiera de los N elementos del conjunto. Para distinguirlos, llamémoslos $\{e_1, e_2 \dots e_N\}$. Dado que el orden de los elementos no es relevante, para cada posible agrupación de m elementos coloquemos primero los del tipo e_1 , después los del tipo e_2 , etc. Por ejemplo, si $N = 3$ y $m = 5$, algunas de estas agrupaciones se podrían escribir como:

$$\overbrace{* * * * *}^{m=5 \text{ estrellas}} \quad \overbrace{* * * * *}^{m=5 \text{ estrellas}}$$

$$e_1 \ e_1 \ e_2 \ e_2 \ e_3 \quad , \quad e_1 \ e_1 \ e_1 \ e_3 \ e_3$$

Como se ve en el segundo ejemplo, no tienen por qué estar presentes todos los tipos de elementos en cada agrupación. La representación del problema se simplifica si, habiendo tomado el convenio de escribir los elementos por su "orden de tipo", se representa el cambio de tipo de elemento simplemente

por una barra. En total, para indicar N tipos de elemento, se deben indicar $N - 1$ cambios de tipo de elemento, o sea que debe haber $N - 1$ barras. En los dos ejemplos anteriores, la representación sería:

$$\begin{array}{c} \overbrace{* \quad * \quad / \quad * \quad * \quad / \quad *}^{m=5 \text{ estrellas y } N-1=2 \text{ barras}} \\ e_1 \quad e_1 \quad \quad e_2 \quad e_2 \quad \quad e_3 \end{array}, \quad \begin{array}{c} \overbrace{* \quad * \quad * \quad / \quad / \quad * \quad *}^{m=5 \text{ estrellas y } N-1=2 \text{ barras}} \\ e_1 \quad e_1 \quad e_1 \quad \quad \quad e_3 \quad e_3 \end{array}$$

En el segundo ejemplo, al no haber elementos e_2 , se deben poner 2 barras seguidas. Cada una de estas series de barras y asteriscos representa cada posible agrupación buscada. El número total de estas formas corresponde al número de series de m estrellas y $N - 1$ barras. Esto es precisamente el número de formas de escoger las posiciones de las $N - 1$ barras dentro de una serie de $m + N - 1$ símbolos (barras o estrellas). Y esto, a su vez, es igual al número de agrupaciones de $N - 1$ elementos que se pueden hacer de un conjunto de $m + N - 1$ elementos tales que:

- Dos agrupaciones con distinta ordenación se consideran iguales (pues el orden en el que se eligen las posiciones de las barras no es relevante).
- Cada elemento sólo se puede incluir en cada agrupación una vez (hay siempre un total de $N - 1$ barras).

Se trata, por tanto, del número de combinaciones sin repetición de $m + N - 1$ elementos tomados de $N - 1$ en $N - 1$:

$$CR_{N,m} = C_{m+N-1, N-1} = \frac{(m+N-1)(m+N-2)\dots(N-1)}{(N-1)!} = \frac{(m+N-1)!}{m!(N-1)!} = \binom{m+N-1}{m} = \binom{m+N-1}{N-1} \quad (3.8)$$

3.2. Un juego de reparto

3.2.1. 10 fichas, 10 €

Sea un juego de ruleta muy especial, en el que cada jugador dispone de $N = 10$ fichas idénticas por partida. Hay cuatro colores en la mesa: rojo, negro, amarillo y verde. Para apostar por el rojo hay que pagar $d_4 = 4€$, para apostar por el negro $d_2 = 2€$, para apostar por el amarillo $d_1 = 1€$ y para apostar por el verde $d_0 = 0€$. En cada partida hay que gastar obligatoriamente $D = 10€$ repartiendo las $N = 10$ fichas entre los cuatro colores. *¿Cuál será la probabilidad de que un color determinado reciba una ficha?*

Este problema puede compararse con el caso físico de la adsorción localizada (moléculas idénticas y distinguibles). Supongamos que en una superficie sólida nos encontramos con cuatro emplazamientos adsorbentes de distinta energía, con gastos energéticos $4k_B T$, $2k_B T$, $k_B T$ y $0k_B T$, respectivamente. Se adsorben 10 moléculas con una energía total de $10k_B T$. Nos interesa conocer las probabilidades de ocupación de los distintos emplazamientos. Estos dos problemas son claramente isomorfos.

El primer paso en la resolución del problema consiste en encontrar todas las combinaciones permitidas mediante el método de tabulación directa. Llamemos **número de ocupación** n_i al número de fichas que se han apostado en el color que tiene un coste d_i . Cada combinación está sujeta a las ligaduras:

$$\sum_i n_i = N \quad \Leftrightarrow \quad n_4 + n_2 + n_1 + n_0 = 10 \quad (3.9)$$

$$\sum_i n_i d_i = D \quad \Leftrightarrow \quad 4n_4 + 2n_2 + 1n_1 + 0n_0 = 10 \quad (3.10)$$

De cada una de las posibles combinaciones, hay que tener en cuenta que el experimento aleatorio consiste en apostar las N fichas, por lo que una misma combinación de números de ocupación $\{n_4, n_2, n_1, n_0\}$ puede obtenerse de $g(n_4, n_2, n_1, n_0)$ diferentes formas (en diferente orden). Se trata por tanto del número de

ordenaciones de un conjunto de N elementos en el que hay n_4 , n_2 , n_1 y n_0 elementos iguales respectivamente o, en otras palabras, las correspondientes permutaciones con repetición:

$$g(n_4, n_2, n_1, n_0) = \frac{N!}{n_4! n_2! n_1! n_0!} \quad (3.11)$$

Las 12 posibles agrupaciones se representan en la siguiente tabla:

Número de agrupación (j)	n_{4j} (rojo)	n_{2j} (negro)	n_{1j} (amarillo)	n_{0j} (verde)	g_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

Hemos construido así un espacio muestral para nuestro experimento con un número total de sucesos elementales:

$$\Omega = \sum_{j=1}^{12} g_j = 26113 \quad (3.12)$$

que representan todas las posibles formas en las que se puede realizar la apuesta. La frecuencia absoluta F_i con la que aparece en el espacio muestral la ficha correspondiente a un color se obtiene sumando el número de fichas de ese color en todas las combinaciones posibles:

$$F_i = \sum_{j=1}^{12} n_{ij} g_j \quad (3.13)$$

y el número total de fichas apostadas a todos los colores es $N\Omega = 261130$. Este resultado también se puede obtener sumando las frecuencias absolutas de cada color:

$$\sum_i F_i = \sum_i \sum_{j=1}^{12} n_{ij} g_j = \sum_{j=1}^{12} \left[\left(\sum_i n_{ij} \right) g_j \right] = N \sum_{j=1}^{12} g_j = N\Omega \quad (3.14)$$

La probabilidad p_i de que un color reciba una ficha se obtiene como el cociente entre la frecuencia absoluta de ese color (casos favorables) y el número total de fichas apostadas a todos los colores (casos posibles). Los resultados se muestran en la siguiente tabla:

$N\Omega = 261130$	rojo	negro	amarillo	verde
F_i	18780	53370	79270	109710
p_i	0.071918	0.204381	0.303565	0.420136

3.2.2. 20 fichas, 20 €

Como puede verse en las siguientes tablas, correspondientes a $N = 20$ fichas y $D = 20€$, la resolución del problema por tabulación directa se complica enormemente al aumentar el número de fichas, lo que hace que para sistemas con $N \sim 10^{23}$ no resulte viable usarlo. En su lugar se propondrá otro método.

Número de agrupación (j)	n_{4j} (rojo)	n_{2j} (negro)	n_{1j} (amarillo)	n_{0j} (verde)	g_j
1	5	0	0	15	15504
2	4	2	0	14	581400
3	4	1	2	13	8139600
4	4	0	4	12	8817900
5	3	4	0	13	2713200
6	3	3	2	12	70543200
7	3	2	4	11	211629600
8	3	1	6	10	155195040
9	3	0	8	9	27713400
10	2	6	0	12	3527160
11	2	5	2	11	126977760
12	2	4	4	10	581981400
13	2	3	6	9	775975200
14	2	2	8	8	374130900
15	2	1	10	7	66512160
16	2	0	12	6	3527160
17	1	8	0	11	1511640
18	1	7	2	10	66512160
19	1	6	4	9	387987600
20	1	5	6	8	698377680
21	1	4	8	7	498841200
22	1	3	10	6	155195040
23	1	2	12	5	21162960
24	1	1	14	4	1162800
25	1	0	16	3	19380
26	0	10	0	10	184756
27	0	9	2	9	9237800
28	0	8	4	8	62355150
29	0	7	6	7	133024320
30	0	6	8	6	116396280
31	0	5	10	5	46558512
32	0	4	12	4	8817900
33	0	3	14	3	775200
34	0	2	16	2	29070
35	0	1	18	1	380
36	0	0	20	0	1

$N\Omega = 92522608260$	rojo	negro	amarillo	verde
F_i	7169650380	18222672000	27398662740	39731623140
p_i	0.077491	0.196954	0.296129	0.429426

3.3. El Principio de Máxima Entropía (PME o MaxEnt)

3.3.1. La función ignorancia de una distribución de probabilidad

Sea dada una *distribución discreta y finita de probabilidades*:

$$A = \begin{pmatrix} A_1 & A_2 & \dots & A_i & \dots & A_\Omega \\ p_1 & p_2 & \dots & p_i & \dots & p_\Omega \end{pmatrix} \quad (3.15)$$

donde $\{A_i\} = \{A_1, A_2, \dots, A_i, \dots, A_\Omega\}$ es un conjunto *completo* (es decir, $\sum_{i=1}^{\Omega} p_i = 1$) de sucesos *elementales* (por tanto, *mutuamente excluyentes*) y p_i es la probabilidad del suceso A_i . Llamaremos **ignorancia de la distribución** A a la siguiente función de las probabilidades:

$$I(p_1, \dots, p_\Omega) = -k \sum_{\substack{i=0 \\ p_i \neq 0}}^{\Omega} p_i \ln p_i \quad (3.16)$$

donde k es una constante que determina las unidades en las que se mide la ignorancia. Escoger un valor para k es equivalente a cambiar la base de los logaritmos en la expresión de la ignorancia. Cuando $k = 1$, se mide en nats; cuando $k = 1/\ln(2)$, se mide en bits. En Física Estadística, se toma k igual a la constante de Boltzmann, con lo que se mide en J/K.

Normalmente, la especificación $p_i = 0$ que aparece en las condiciones del sumatorio se omite, entendiendo que cuando $p_i = 0$ el término $p_i \ln p_i$ vale 0. Esto, en el fondo, es admitir la extensión por continuidad de la función $x \mapsto x \ln x$ en $x = 0$, cuyo límite cuando x tiende a 0 por la derecha es precisamente 0.

Cuando todos los sucesos son equiprobables, se tiene que:

$$p_i = \frac{1}{\Omega} \quad \forall i \quad \Rightarrow \quad I(p_1, \dots, p_\Omega) = k \ln \Omega \quad (3.17)$$

Es posible (véase la sección 3.6) dar una caracterización de la función ignorancia como la única función posible $I(A) \equiv I(p_1, \dots, p_\Omega)$ de las probabilidades que verifica las siguientes propiedades:

1. Es continua.
2. Es máxima si todos los sucesos son equiprobables (y sus probabilidades son iguales a $1/\Omega$).
3. La ignorancia de la distribución es igual a la de otra distribución igual a la que se haya añadido un suceso imposible (con probabilidad nula).
4. Si B es otra distribución finita e $I(B|A_i)$ es la ignorancia de B cuando se da el suceso A_i posible (tal que $p_i \neq 0$) de la distribución A , la ignorancia de la distribución AB (cuyos sucesos son las intersecciones $A_i \cap B_i$) es:

$$I(AB) = I(A) + \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} p_i I(B|A_i) \quad (3.18)$$

de modo que, si las distribuciones A y B son independientes, en cuyo caso $I(B|A_i) = I(B)$ (pues la ignorancia de B no dependerá del conocimiento de A), será:

$$I(AB) = I(A) + I(B) \quad (3.19)$$

5. La ignorancia de A es nula si $\Omega = 1$ (si hay un único suceso).

Es importante notar que la propiedad de aditividad expresada por la ecuación (3.19) no es suficiente para determinar la forma funcional (3.16). Existen infinitas funciones que pueden cumplir simultáneamente la ecuación (3.19) y las propiedades 1, 2, 3 y 5 (por ejemplo, las llamadas "entropías de Rényi"). La

propiedad expresada por la ecuación (3.18) es, por tanto, muy importante y representa, dicho de forma coloquial, la **combinación de aditividad y coherencia con las reglas de la probabilidad condicionada**. Esta coherencia fue empleada como propiedad caracterizadora de la función ignorancia por C.E. Shannon cuando estableció la Teoría Matemática de la Información, y viene a indicar que la ignorancia tiene que valer lo mismo independientemente de si se dan como datos directamente las probabilidades de los sucesos A_i o si se dan indirectamente a través de otros sucesos como probabilidades condicionadas (dado que, en ambos casos, el estado de conocimiento que se tiene es el mismo).

Las propiedades mencionadas justifican bien el nombre “ignorancia”, pero esta función también recibe otros nombres, como **entropía de Shannon, entropía de información, entropía de Gibbs...** En adelante, se hará referencia a ella indistintamente como entropía o ignorancia.

En Física Estadística, los sucesos A_i de la distribución de probabilidad (3.15) se corresponden con los posibles microestados correspondientes a un macroestado (estado termodinámico) de un sistema macroscópico de acuerdo con un modelo de realidad.

3.3.2. 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.

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.

3.3.3. Inferencia de probabilidades: PME

El **Principio de Indiferencia** aplicado a las probabilidades indica que, **en ausencia de información que indique lo contrario, los sucesos elementales de un experimento aleatorio deben ser equiprobables**.

Este principio representa un criterio adoptado para resolver el problema de asignación de probabilidades en los casos en los que no existe información. El problema que nos proponemos resolver es una generalización de éste:

Si se tiene una información determinada sobre un experimento aleatorio, ¿cuál es la distribución de probabilidad más adecuada que se debe asignar al conjunto de posibles sucesos elementales, de tal forma que sea compatible con la información disponible?

En la formulación de esta pregunta, se está adoptando necesariamente el *punto de vista bayesiano* de la Teoría de la Probabilidad. A partir de la **interpretación de la función ignorancia como una medida de la falta de información**, E.T. Jaynes propuso que la solución se podía encontrar con ayuda del **Principio de Máxima Ignorancia o de Máxima Entropía (PME)**:

Las probabilidades que han de asignarse a los sucesos elementales de una distribución de probabilidad son aquellas que hacen la ignorancia máxima compatible con la información disponible.

Dicho de forma coloquial, este principio es el equivalente probabilístico de “no sacar conclusiones precipitadas”, asignando una probabilidad a todos los sucesos que no sean a priori imposibles y buscando la distribución de probabilidad menos sesgada (más “ancha”) compatible con la información de partida.

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**. 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.**

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, pero:

- Si cambia la información de partida, las predicciones se modifican.
- Si, al realizar un experimento, los resultados que se obtienen son incompatibles con las predicciones realizadas, esto no significa que el PME sea incorrecto, sino que **la información de partida era incompleta o errónea.**

En la línea de esta última posibilidad, Jaynes 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 termodinámico puede hacerse sobre la base de que esos estados obedecen 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).

Se puede *intentar* poner una objeción al argumento intuitivo que ha llevado al PME. El enunciado del PME se ha fundamentado en que la función ignorancia es la magnitud *correcta* para cuantificar “información” de una distribución de probabilidad, pero, en el fondo, la ignorancia *define* la falta de información desde un punto de vista científico. ¿No es esto, de algún modo, un argumento circular? El hecho de que la falta de información, tal y como la hemos definido, permita obtener una inferencia “correcta y completa” de la distribución de probabilidad se basa en algunas **sólidas razones que provienen de la Teoría de la Probabilidad, que garantizan que el PME es el método adecuado para inferir probabilidades a partir de una información de partida**. Sin entrar en detalles aún, mencionemos que se puede justificar que, **SILA INFORMACIÓN DE PARTIDA ES CORRECTA Y SUFICIENTE:**

- El PME es una consecuencia del Principio de Indiferencia y de las reglas de la Teoría de la Probabilidad (**argumento de Wallis**; véase sección 3.5.2).
- Las frecuencias relativas de los posibles sucesos elementales obtenidas de repetir en condiciones idénticas un experimento (probabilidades frecuentistas) se aproximan “rápidamente” a las que predice el PME (probabilidades bayesianas) a medida que crece el número de experimentos

realizados (**conexión entropía-frecuencia**; véase sección 3.5.3). Dicho de forma más precisa: si W es el número de formas en que pueden ocurrir experimentalmente las frecuencias relativas coincidentes con las probabilidades que predice el PME y W' es el número de formas en que puede ocurrir experimentalmente cualquier otra distribución de frecuencias relativas, entonces **el cociente W/W' tiende exponencialmente a infinito cuando el número de experimentos tiende a infinito**.

3.3.4. El ejemplo del dado de Brandeis

Como ejemplo ilustrativo del significado del PME y de la conexión entre el PME y las frecuencias relativas, mencionemos (sin desarrollar los cálculos) el ejemplo del *problema del dado de Brandeis*, que Jaynes comenta con detalle en respuesta a la solución que J.S. Rowlinson proporciona sin usar el PME. El experimento consiste en lanzar N veces un dado y, si el valor medio obtenido de la puntuación del dado es 4.5, la pregunta es: ¿cuáles son las probabilidades más adecuadas que debemos asignar a cada uno de los posibles resultados de lanzar el dado $\{1,2,3,4,5,6\}$?

- La solución de Jaynes, aplicando el PME, supone únicamente el valor medio de la puntuación del dado y nada más.
- La solución de Rowlinson consiste en razonar como si el experimento fuera a realizarse de un modo muy concreto, imaginando un “modelo de urnas” (cuyos detalles no son relevantes y omitimos), lo que de alguna manera implica suponer más información que simplemente el valor medio de la puntuación del dado.

Supongamos que se lanza el dado $N = 1000$ veces. Llamemos W al número de maneras en que puede realizarse la distribución de frecuencias relativas coincidente con las probabilidades obtenidas del PME, y W' al número de maneras en que puede realizarse la distribución de frecuencias relativas obtenida del razonamiento de Rowlinson. Si se calcula el cociente W/W' , éste arroja un valor de $1.52 \cdot 10^{86}$. Es decir, **para 1000 lanzamientos, por cada forma en que la distribución de Rowlinson puede ocurrir experimentalmente, la distribución del PME puede ocurrir de más de 10^{86} formas** (recordemos que la edad del universo es del orden de 10^{18} s). Y este cociente aumenta exponencialmente a medida que aumenta el número de lanzamientos. Así pues, el “modelo de urnas” de Rowlinson introduce en el problema, sin justificación, información que no estaba en el problema original (es decir, estaría resolviendo *otro problema distinto*), lo que se traduce en una distribución de probabilidad más sesgada que la de máxima ignorancia. Si el experimento no se realiza de acuerdo con su “modelo de urnas”, no hay por qué esperar que la distribución experimental de frecuencias se corresponda con las probabilidades predichas por Rowlinson, ya que el número de formas en que se puede realizar la distribución del PME es muchísimo mayor.

De hecho, se puede ir más allá en el cálculo. La estimación del PME para el valor de que el dado arroje el resultado “1” es $p_1 = 0.0543$. Pues bien, de todas las posibles formas en las que puede ocurrir que el valor medio de la puntuación del dado sea 4.5, en el 99% de ellas la frecuencia relativa del resultado “1” está en el intervalo (0.05435 ± 0.0153) . El valor de p_1 en la distribución de Rowlinson es 0.00243, muy alejado de ese intervalo.

Debemos, sin embargo, insistir en que estas conexiones entre las frecuencias 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.**

3.4. Solución del juego de reparto con el PME

A diferencia del método de tabulación directa, en la que trabajábamos con el espacio muestral correspondiente a todas las posibles formas de distribuir unas fichas con las condiciones de ligadura (3.9) y (3.10), en la aplicación del PME el espacio muestral que hay que construir es aquel relativo a los posibles colores (con su coste d_i asociado) que se pueden asignar a una ficha.

Así, tendremos 4 posibles sucesos correspondientes a cada color, caracterizados por su coste d_i y su probabilidad p_i . La distribución de probabilidad la podemos representar esquemáticamente como:

$$\begin{pmatrix} \text{Rojo} & \text{Negro} & \text{Amarillo} & \text{Verde} \\ d_4 = 4 \text{ €} & d_2 = 2 \text{ €} & d_1 = 1 \text{ €} & d_0 = 0 \text{ €} \\ p_4 & p_2 & p_1 & p_0 \end{pmatrix} \quad (3.20)$$

Hemos de transformar la información que se da en el problema en una relación entre probabilidades que podamos implementar en el proceso de maximización de la ignorancia. Para ello, interpretamos el cociente $D/N = 1\text{€}/\text{ficha}$ como el valor medio $\langle d \rangle$ del dinero que le corresponde a una ficha. Este valor medio es el que se obtiene de los dos casos estudiados por tabulación directa, ($N = 10$ fichas, $D = 10\text{€}$) y ($N = 20$ fichas, $D = 20\text{€}$). Por tanto, hemos de maximizar la función ignorancia condicionada a las ligaduras:

$$\sum_i p_i = 1 \quad (3.21)$$

$$\sum_i p_i d_i = \langle d \rangle \quad (3.22)$$

Es evidente que el valor de la constante k que determina las unidades de la función ignorancia no puede afectar a las probabilidades inferidas con el PME. Trabajaremos por comodidad con lo que llamaremos la **función ignorancia adimensional** $\sigma = I/k$, que numéricamente coincidirá con la ignorancia medida en nats. El objetivo será por tanto hallar el máximo de la función:

$$\boxed{\sigma = \frac{I}{k} = - \sum_i p_i \ln p_i} \quad \text{(Función ignorancia adimensional)} \quad (3.23)$$

sujeta a las 2 condiciones anteriores. Para ello, utilizaremos el método de los multiplicadores de Lagrange, según el cual la maximización propuesta es equivalente a la **maximización de otra función, L , llamada función de Lagrange, sin ninguna ligadura**. La función de Lagrange se construye incluyendo las ecuaciones de ligadura en la forma siguiente:

$$L(\{p_i\}, \lambda, \beta) = - \sum_i p_i \ln p_i - \lambda \left[\sum_i p_i - 1 \right] - \beta \left[\sum_i p_i d_i - \langle d \rangle \right] \quad (3.24)$$

donde λ y β son dos parámetros desconocidos (los multiplicadores de Lagrange) que hay que determinar durante el proceso de optimización. Para maximizar esta función de varias variables (las probabilidades y los multiplicadores), se procede a derivar parcialmente respecto de todas las variables e igualar el resultado a cero. Para el caso de las variables λ y β se vuelven a obtener las condiciones de ligadura, por lo que nos resta derivar respecto de cada una de las probabilidades:

$$\frac{\partial L}{\partial p_k} = -\ln p_k - 1 - \lambda - \beta d_k = 0 \quad (3.25)$$

Despejando la probabilidad:

$$p_k = \frac{1}{Z} e^{-\beta d_k} \quad (3.26)$$

donde se ha definido una nueva constante $Z = e^{1+\lambda}$, que se denomina **función de partición**. Insertando (3.26) en la condición de normalización (3.21) se obtiene:

$$Z = \sum_i e^{-\beta d_i} \quad (3.27)$$

Sustituyendo ahora la ecuación (3.27) en la condición del valor medio (3.22) y reordenando, se llega a:

$$\sum_i (d_i - \langle d \rangle) e^{-\beta d_i} = 0 \quad (3.28)$$

Esta ecuación se puede transformar en una polinómica haciendo el cambio de variable:

$$x = e^{-\beta} \quad (3.29)$$

con lo que la nueva variable x debe ser un número real positivo que cumple:

$$\sum_i (d_i - \langle d \rangle) x^{d_i} = 0 \quad (3.30)$$

Para los datos de nuestro problema, esta ecuación toma la forma:

$$3x^4 + x^2 - 1 = 0 \quad (3.31)$$

La única solución real positiva de esta ecuación bicuadrada es:

$$x = \sqrt{\frac{-1 + \sqrt{13}}{6}} = 0.65898296 \quad (3.32)$$

La función de partición puede calcularse sustituyendo el valor de x en la ecuación (3.27):

$$Z = \sum_i x^{d_i} = x^4 + x^2 + x + 1 = 2.28182199 \quad (3.33)$$

Finalmente, las probabilidades se calculan a partir de los valores obtenidos de x y Z sustituyendo en la ecuación (3.26):

$$p_k = \frac{1}{Z} x^{d_k} \quad (3.34)$$

Las probabilidades calculadas se muestran en la siguiente tabla en comparación con los resultados obtenidos por tabulación directa (por completitud, se incluyen los resultados de dos casos más que los que se han resuelto explícitamente en la sección 3.2):

	rojo	negro	amarillo	verde
$p_i (N = 10 \text{ fichas}, D = 10\text{€})$	0.071918	0.204381	0.303565	0.420136
$p_i (N = 20 \text{ fichas}, D = 20\text{€})$	0.077491	0.196954	0.296129	0.429426
$p_i (N = 30 \text{ fichas}, D = 30\text{€})$	0.079254	0.194661	0.293662	0.432423
$p_i (N = 40 \text{ fichas}, D = 40\text{€})$	0.080118	0.193545	0.292436	0.433900
$p_i (\text{PME}, \langle d \rangle = 1\text{€}/\text{ficha})$	0.082645	0.190312	0.288797	0.438246

Puede apreciarse que al aumentar el número de fichas los resultados se aproximan más a los que predice el PME. Teniendo en cuenta algunas leyes estadísticas de “grandes números” (en las cuales no entraremos por estar fuera de los objetivos del curso) y por la conexión entropía-frecuencia que se ha comentado en la sección 3.3.3 (y que se justifica en la sección 3.5.3), esto es de esperar, pues aumentar el número de fichas equivale a aumentar el número de experimentos “asignar un color a una ficha” realizados sujetos a la ligadura $\langle d \rangle = 1\text{€}/\text{ficha}$, que es la que se ha utilizado en la maximización de la ignorancia. Esta tendencia puede observarse en la Figura 3.3.

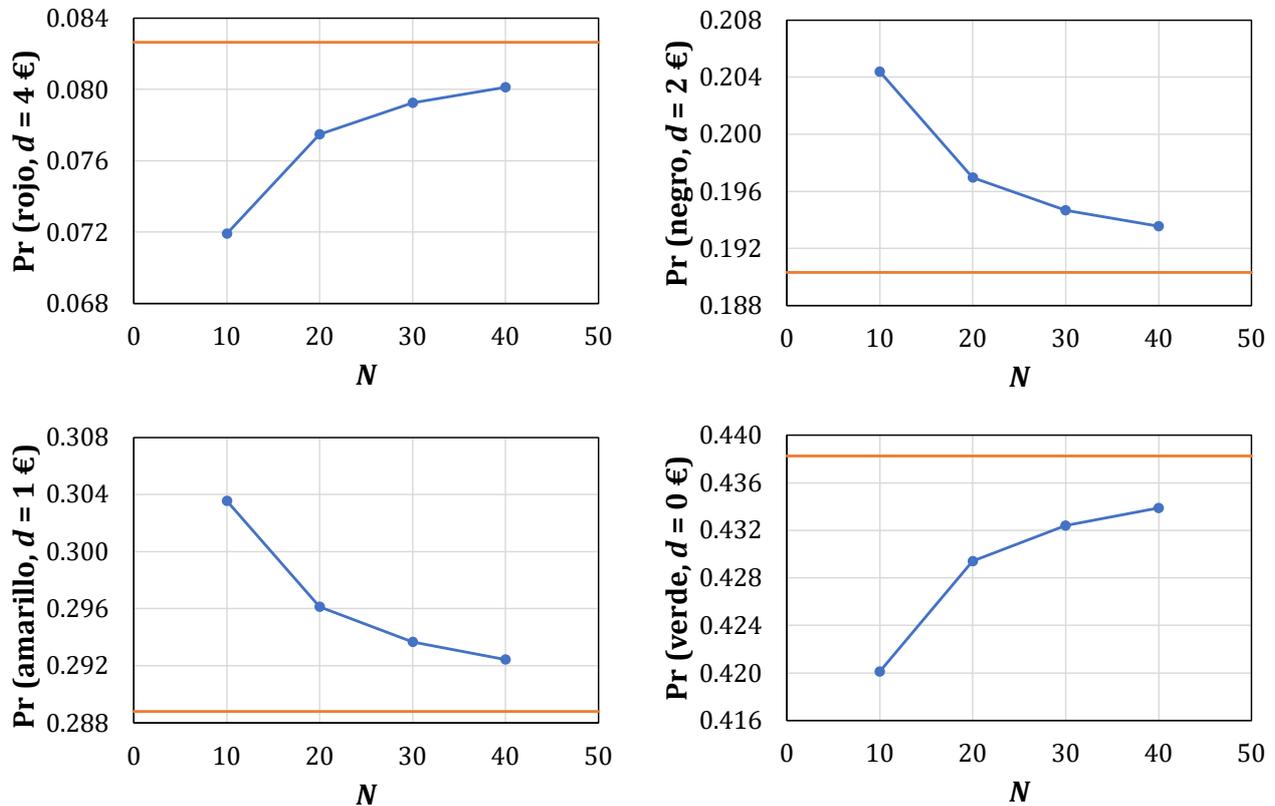


Figura 3.3: Probabilidad calculada de que los 4 colores de la ruleta reciban una ficha para $\langle d \rangle = 1\text{€}/\text{ficha}$. Los puntos representan los resultados del método de tabulación directa para distintos valores del número de fichas (N). Las líneas horizontales continuas son los resultados obtenidos a partir del PME. A medida que N aumenta, los resultados obtenidos por tabulación directa se aproximan cada vez más a los calculados a partir del PME.

El valor de la ignorancia adimensional para esta distribución de probabilidad se obtiene sustituyendo (3.26) en (3.23). Tras operar, se llega a:

$$\sigma(p_1, \dots, p_i, \dots) = \beta \langle d \rangle + \ln Z \quad (3.35)$$

Hasta ahora, lo único que se ha demostrado es que la distribución de probabilidad hallada es un punto crítico, **pero nada se ha dicho de que sea un máximo**. Para demostrarlo, será de ayuda la desigualdad:

$$\ln x \geq \left(1 - \frac{1}{x}\right) \quad (3.36)$$

donde la igualdad se da si y solo si $x = 1$. Considérese otra distribución, arbitraria, de probabilidades q_i de los posibles microestados que sea compatible con las ligaduras, es decir, tal que verifique:

$$\sum_i q_i = 1 \quad (3.37)$$

$$\sum_i q_i d_i = \langle d \rangle \quad (3.38)$$

Se trata de ver que tal distribución conduce a una entropía que es menor o igual que la dada por la distribución calculada anteriormente. La desigualdad (3.36) lleva fácilmente a la relación:

$$\sum_i q_i \ln \frac{q_i}{p_i} \geq \sum_i q_i \left(1 - \frac{p_i}{q_i}\right) = 0 \Rightarrow \sum_i q_i \ln q_i \geq \sum_i q_i \ln p_i \quad (3.39)$$

con igualdad si y solo si $q_i = p_i$. Sustituyendo p_i por su valor:

$$\sum_i q_i \ln q_i \geq \sum_i q_i [-\beta d_i - \ln Z] = -\beta \langle d \rangle - \ln Z \quad (3.40)$$

o lo que es lo mismo:

$$-\sum_i q_i \ln q_i = \sigma(q_1, \dots, q_i, \dots) \leq \sigma(p_1, \dots, p_i, \dots) = \beta \langle d \rangle + \ln Z \quad (3.41)$$

Es decir, cualquier otra distribución de probabilidad compatible con las ligaduras lleva a una entropía que es menor o igual que la calculada con la distribución dada por la ecuación (3.26), por lo que dicha distribución de probabilidad representa un máximo.

3.5. Complemento (para valientes): Argumento de Wallis y conexión entropía-frecuencia

En esta sección nos proponemos mostrar argumentos sólidos que justifiquen el PME más allá de la interpretación de la función ignorancia como medida correcta de la falta de información contenida en una distribución de probabilidad. Si bien los argumentos que vamos a mostrar son cualitativamente simples, matemáticamente no lo son tanto y hacen uso de una fórmula asintótica de las permutaciones con repetición (esto es, una fórmula aproximada que es útil cuando el número de elementos a permutar es muy grande), y por ello exponemos esta fórmula primero, para evitar interrumpir la exposición con muchas matemáticas entre pasos intermedios.

3.5.1. Fórmula asintótica para las permutaciones con repetición

En la fórmula (3.6) de las permutaciones con repetición, sea $g_i = n_i/N$, con lo que podemos escribir:

$$P_N^{n_1, \dots, n_s} = \frac{N!}{n_1! \dots n_s!} = \frac{N!}{(Ng_1)! \dots (Ng_s)!} \quad (3.42)$$

Utilizando el desarrollo asintótico para el factorial (ecuación (3.4)) y teniendo en cuenta que los términos $0! = 1$ tendrán logaritmo nulo y no contribuirán a las sumas, podemos escribir:

$$\begin{aligned} \frac{1}{N} \ln(P_N^{n_1, \dots, n_s}) &= \ln N - 1 + \frac{1}{2N} \ln(2\pi N) + \frac{1}{12N^2} - \frac{1}{360N^4} + \dots \\ &- \sum_{\substack{i=1 \\ g_i \neq 0}}^s \left(g_i \ln(Ng_i) - g_i + \frac{1}{2N} \ln(2\pi Ng_i) + \frac{1}{12N^2 g_i} - \frac{1}{360N^4 g_i^3} + \dots \right) \end{aligned} \quad (3.43)$$

El sumatorio se puede simplificar de la siguiente manera:

$$\begin{aligned} \sum_{\substack{i=1 \\ g_i \neq 0}}^s \left(g_i \ln(Ng_i) - g_i + \frac{1}{2N} \ln(2\pi Ng_i) \right) &= -\ln N \cdot \sum_{\substack{i=1 \\ g_i \neq 0}}^s g_i - \sum_{\substack{i=1 \\ g_i \neq 0}}^s g_i \ln g_i + \sum_{\substack{i=1 \\ g_i \neq 0}}^s g_i - \frac{1}{2N} \sum_{\substack{i=1 \\ g_i \neq 0}}^s \ln(2\pi N) - \frac{1}{2N} \sum_{\substack{i=1 \\ g_i \neq 0}}^s \ln g_i \\ &= -\ln N + I(g_1, \dots, g_s) + 1 - \frac{s}{2N} \ln(2\pi N) - \frac{1}{2N} \sum_{\substack{i=1 \\ g_i \neq 0}}^s \ln g_i \end{aligned} \quad (3.44)$$

Con lo que la ecuación (3.43) queda:

$$\begin{aligned} \frac{1}{N} \ln(P_N^{n_1, \dots, n_s}) &= \left(\ln N - \chi + \frac{1}{2N} \ln(2\pi N) + \frac{1}{12N^2} - \frac{1}{360N^4} + \dots \right) \\ &\Rightarrow \ln N + I(g_1, \dots, g_s) + \chi - \frac{1}{2} \ln(2\pi N) - \frac{1}{2N} \sum_{\substack{i=1 \\ g_i \neq 0}}^s \ln g_i - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \left(\frac{1}{12N^2 g_i} - \frac{1}{360N^4 g_i^3} + \dots \right) \end{aligned} \quad (3.45)$$

Reagrupando:

$$\frac{1}{N} \ln(P_N^{n_1, \dots, n_s}) = \sigma(g_1, \dots, g_s) + \frac{1}{2N} \left[(s-1) \ln(2\pi N) - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \ln g_i \right] + \frac{1}{12N^2} \left(1 - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \frac{1}{g_i} \right) - \frac{1}{360N^4} \left(1 - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \frac{1}{g_i^3} \right) + \dots \quad (3.46)$$

donde $\sigma = I/k$ es la **ignorancia adimensional**. O bien, multiplicando por N y tomando la exponencial:

$$P_N^{n_1, \dots, n_s} = \exp\{N \cdot \sigma(g_1, \dots, g_s)\} \cdot \sqrt{(2\pi N)^{s-1} \prod_{\substack{i=1 \\ g_i \neq 0}}^s \frac{1}{g_i}} \cdot \exp \left\{ \frac{1}{12N} \left(1 - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \frac{1}{g_i} \right) - \frac{1}{360N^3} \left(1 - \sum_{\substack{i=1 \\ g_i \neq 0}}^s \frac{1}{g_i^3} \right) + \dots \right\} \quad (3.47)$$

3.5.2. El argumento de Wallis

Jaynes da este nombre a una justificación teórica del PME elaborada a partir de una sugerencia que le hizo Graham Wallis en 1962, en la que se usan únicamente las reglas de la Teoría de la Probabilidad y el Principio de Indiferencia (asignación de probabilidades iguales a los sucesos elementales en ausencia de información). En esta justificación juegan un papel importante las fórmulas asintóticas (3.46) y (3.47).

Supongamos que se nos da una cierta información (que podría consistir, como en el ejemplo del juego de la ruleta, en los valores medios de unas ciertas variables aleatorias), que debemos utilizar para inferir las probabilidades de la distribución (3.15). Tenemos una cantidad total de probabilidad:

$$\sum_{i=1}^{\Omega} p_i = 1 \quad (3.48)$$

que distribuir entre los Ω sucesos elementales.

Tomemos un número $n \gg \Omega$ e imaginemos que tenemos n pequeños cuantos (trozos discretos) de probabilidad, cada uno de magnitud $\delta = 1/n$, que tenemos que distribuir de cualquier forma posible entre los sucesos elementales. Este sería un problema análogo a la distribución de n bolas entre Ω cajas. Para efectuar la distribución de forma "justa", sin favorecer innecesariamente a ninguno de los sucesos, podemos proceder así:

- Repartimos los cuantos entre los sucesos totalmente al azar, de tal forma que la probabilidad de que uno de los sucesos reciba un cuanto es la misma ($1/\Omega$) (**Principio de Indiferencia**).
- Si el resultado se adecúa a la información de partida (ciertas relaciones conocidas previamente entre las probabilidades de los sucesos), aceptamos dicho resultado; de lo contrario, lo rechazamos y repetimos el proceso hasta obtener un resultado compatible con esa información.

Si hemos obtenido una distribución válida de los cuantos en la que el suceso A_i ha recibido n_i cuantos, diremos que este experimento aleatorio ha dado como resultado la asignación de probabilidades $p_i = \delta \cdot n_i = n_i/n$.

¿De cuántas formas puede ocurrir una asignación válida de probabilidades? Para contabilizarlas, tenemos que tener en cuenta que, aunque hay $n!$ posibles ordenaciones de los cuantos, son iguales todas las ordenaciones en las que se permutan entre sí solamente los n_i cuantos de cada suceso A_i . En otras

palabras, el número de formas, $W(p_1, \dots, p_\Omega)$, en que puede ocurrir dicha asignación es igual a las permutaciones con repetición de n elementos, en donde hay n_1 elementos iguales, ..., n_Ω elementos iguales:

$$W(p_1, \dots, p_\Omega) = \frac{n!}{n_1! \dots n_\Omega!} = \frac{n!}{(np_1)! \dots (np_\Omega)!} \quad (3.49)$$

Si queremos que en nuestro juego de asignación de probabilidades los posibles valores de éstas sean un continuo entre 0 y 1, tenemos que hacer tender el tamaño de los cuantos de probabilidad a cero. En otras palabras, debemos hacer tender n a infinito. A continuación, demostraremos que, **cuando la información de partida dé lugar a un único máximo condicionado de la función $W(p_1, \dots, p_\Omega)$** , y para n suficientemente grande, una ENORME mayoría de las distribuciones aceptables de los cuantos (es decir, de las asignaciones aceptables de probabilidad a los sucesos) están concentradas en un entorno muy pequeño del máximo de $W(p_1, \dots, p_\Omega)$. Para ello, procederemos en dos etapas:

- Estudiaremos la función $(\ln W)/n$, que es una función monótona creciente de W (y, por tanto, su máximo condicionado debe ocurrir para los mismos valores de las probabilidades) y veremos que para n arbitrariamente grande el valor de esta función es igual a la ignorancia adimensional de la distribución de probabilidad obtenida, por lo que maximizar $(\ln W)/n$ es equivalente a maximizar la ignorancia.
- Compararemos, para n grande, el valor del cociente entre el número de formas en que se puede realizar la distribución del máximo condicionado y el número de formas en que se puede realizar cualquier otra distribución aceptable, y veremos que este cociente tiende exponencialmente a infinito con n tendiendo a infinito.

Procedamos con la primera parte. El desarrollo asintótico de (3.49) viene dado por la ecuación (3.46):

$$\frac{1}{n} \ln W = \sigma(p_1, \dots, p_\Omega) + \frac{1}{2n} \left[(\Omega - 1) \ln(2\pi n) + \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} \ln p_i \right] + \frac{1}{12n^2} \left(1 - \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} \frac{1}{p_i} \right) - \frac{1}{360n^4} \left(1 - \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} \frac{1}{p_i^3} \right) + \dots \quad (3.50)$$

Pero, de acuerdo con (3.51):

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \ln W \right) = \sigma(p_1, \dots, p_\Omega) \quad (3.51)$$

Por ello, **para cuantos de probabilidad suficientemente pequeños, maximizar $W(p_1, \dots, p_\Omega)$ de acuerdo con la información disponible es equivalente a maximizar la función ignorancia de acuerdo con esa misma información.**

Vamos ahora con la segunda parte del argumento. Supongamos que hemos encontrado la distribución de probabilidad $\{p_{1m}, \dots, p_{\Omega m}\}$ que da lugar al único máximo condicionado de $W(p_1, \dots, p_\Omega)$, al que denotaremos por W_m . Sea $\{p'_1, \dots, p'_\Omega\}$ otra distribución **arbitraria** de probabilidad, con la única condición de que sea compatible con la información disponible; denotemos $W' = W(p'_1, \dots, p'_\Omega)$ el número de formas en que puede ocurrir esta asignación de probabilidades. De acuerdo con la fórmula asintótica (3.47), el cociente entre W_m/W' se puede escribir:

$$\frac{W_m}{W'} = \exp\{n \cdot [\sigma(p_{1m}, \dots, p_{\Omega m}) - \sigma(p'_1, \dots, p'_\Omega)]\} \cdot \sqrt{\prod_{\substack{i=1 \\ p_{im}, p'_{i \neq 0}}}^s \frac{p'_i}{p_{im}}} \cdot \exp \left\{ \frac{1}{12n} \left(\sum_{\substack{i=1 \\ p'_{i \neq 0}}}^{\Omega} \frac{1}{p'_i} - \sum_{\substack{i=1 \\ p_{im} \neq 0}}^{\Omega} \frac{1}{p_{im}} \right) - \dots \right\} \quad (3.52)$$

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, que crece enormemente rápido al aumentar n . Así pues, **el número de formas en las que se puede obtener una asignación de probabilidad**

diferente de la del máximo de $W(p_1, \dots, p_\Omega)$ disminuye exponencialmente en comparación con dicho máximo a medida que disminuye el tamaño $\delta = 1/n$ del cuanto de probabilidad.

Combinando ambas partes del argumento, llegamos a la conclusión de que el único criterio lógico –y práctico– para la selección de una distribución de probabilidad compatible con la información disponible es, precisamente, el de maximizar $W(p_1, \dots, p_\Omega)$, que es equivalente al de maximizar la ignorancia, de acuerdo con la información disponible. Así pues, si aceptamos este “juego” como forma “justa” de asignación de probabilidades, llegamos inexorablemente al PME.

3.5.3. Conexión entropía-frecuencia

Desde el punto de vista frecuentista, una distribución de probabilidades necesita ser verificada experimentalmente en términos de frecuencias relativas. Si bien es cierto que en la aplicación del PME a la asignación de probabilidades se adopta el punto de vista bayesiano, eso no implica que el PME no tenga una conexión directa con la obtención de dichas probabilidades en un experimento. Eso es lo que nos proponemos demostrar.

Supongamos, por tanto, que disponemos de una cierta información acerca de un experimento, cuyos sucesos elementales son $\{A_1, \dots, A_\Omega\}$. Con el objeto de obtener experimentalmente las frecuencias relativas de los sucesos, se repite dicho experimento n veces en condiciones idénticas. Como resultado, se obtiene n_i veces el suceso elemental A_i , siendo $n_1 + \dots + n_\Omega = n$. Las frecuencias relativas serán, por tanto, $f_i = n_i/n$. Suponemos que las frecuencias 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.

¿De cuántas formas puede ocurrir una cierta distribución de frecuencias relativas? La respuesta es:

$$W(f_1, \dots, f_\Omega) = \frac{n!}{n_1! \dots n_\Omega!} = \frac{n!}{(nf_1)! \dots (nf_\Omega)!} \quad (3.53)$$

A partir de aquí, el argumento es matemáticamente idéntico al argumento de Wallis, pero reemplazando las probabilidades asignadas por las frecuencias relativas obtenidas del experimento. De esa forma, se demuestra que, cuando la información de partida dé lugar a un único máximo condicionado de la función $W(f_1, \dots, f_\Omega)$:

- Para n suficientemente grande, maximizar $W(f_1, \dots, f_\Omega)$ de acuerdo con la información disponible es equivalente a maximizar la función $\sigma(f_1, \dots, f_\Omega)$ de acuerdo con esa misma información.
- El número de formas en las que se puede obtener una distribución de frecuencias *diferente* de la del máximo de $W(f_1, \dots, f_\Omega)$ disminuye exponencialmente en comparación con dicho máximo al aumentar el número n de veces que se repite el experimento.

Combinando estas dos proposiciones, llegamos a la conclusión de que, para un n suficientemente grande, el número de formas W_m en que puede ocurrir la distribución de frecuencias que maximiza $\sigma(f_1, \dots, f_\Omega)$ condicionada a la información disponible es mucho mayor que el número de formas W' en las que puede ocurrir cualquier otra distribución de frecuencias compatible con la información, y que, además, el cociente W_m/W' tiende **exponencialmente** a infinito cuando n tiende a infinito. Por tanto, para n suficientemente grande es de esperar, y MUCHO, que ocurra la distribución de frecuencias que maximiza $\sigma(f_1, \dots, f_\Omega)$, simplemente porque puede ocurrir de un número ENORMEMENTE MAYOR de veces.

Esta distribución de frecuencias 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 PME.** Y, es más: si experimentalmente se encuentran desviaciones estadísticamente significativas de las predicciones del PME, ello **constituye una evidencia importante (y, si persiste, concluyente) de que hay información que no se ha tenido en cuenta adecuadamente en el proceso de maximización de la ignorancia.**

3.6. Complemento: caracterización de la entropía

En este apéndice demostramos que las cinco propiedades que se indicaron en la sección 3.3.1 como caracterizadoras de la función ignorancia efectivamente la determinan de forma única salvo el valor de la constante k .

a) Sea $L(\Omega)$ la ignorancia equiprobable de la distribución A :

$$L(\Omega) \equiv I\left(\underbrace{\frac{1}{\Omega}, \dots, \frac{1}{\Omega}}_{\Omega}\right) \quad (3.54)$$

Por la propiedad 3 de la función ignorancia:

$$I\left(\underbrace{\frac{1}{\Omega}, \dots, \frac{1}{\Omega}}_{\Omega}\right) = I\left(\underbrace{\frac{1}{\Omega}, \dots, \frac{1}{\Omega}}_{\Omega}, 0\right) \quad (3.55)$$

y por la propiedad 2:

$$I\left(\underbrace{\frac{1}{\Omega}, \dots, \frac{1}{\Omega}}_{\Omega}, 0\right) < I\left(\underbrace{\frac{1}{\Omega+1}, \dots, \frac{1}{\Omega+1}}_{\Omega+1}\right) \quad (3.56)$$

Por tanto, $L(\Omega) < L(\Omega + 1)$, es decir, L es una función creciente.

b) Considérense ahora r distribuciones independientes $A^{(j)}$, cada una con Ω sucesos equiprobables. Por la propiedad 4:

$$I(A^{(1)}A^{(2)} \dots A^{(r)}) = \sum_{j=1}^r I(A^{(j)}) = \sum_{j=1}^r L(\Omega) = rL(\Omega) \quad (3.57)$$

Pero la distribución $A^{(1)}A^{(2)} \dots A^{(r)}$ está constituida por Ω^r sucesos igualmente probables, de modo que:

$$I(A^{(1)}A^{(2)} \dots A^{(r)}) = L(\Omega^r) \quad (3.58)$$

Por tanto:

$$L(\Omega^r) = rL(\Omega) \quad (3.59)$$

c) Sean ahora s , m , n números enteros positivos, con $m, n > 1$ y sea r tal que:

$$n^r \leq m^s \leq n^{r+1} \quad (3.60)$$

Por un lado, se tiene:

$$\ln n^r \leq \ln m^s \leq \ln n^{r+1} \Rightarrow r \ln n \leq s \ln m \leq (r+1) \ln n \Rightarrow \frac{r}{s} \leq \frac{\ln m}{\ln n} \leq \frac{r}{s} + \frac{1}{s} \quad (3.61)$$

y por otro, usando las propiedades encontradas para $L(n)$:

$$L(n^r) \leq L(m^s) \leq L(n^{r+1}) \Rightarrow rL(n) \leq sL(m) \leq (r+1)L(n) \Rightarrow \frac{r}{s} \leq \frac{L(m)}{L(n)} \leq \frac{r}{s} + \frac{1}{s} \quad (3.62)$$

Tanto $\frac{\ln m}{\ln n}$ como $\frac{L(m)}{L(n)}$ están comprendidos entre $\frac{r}{s}$ y $\frac{r}{s} + \frac{1}{s}$, por lo que:

$$\left| \frac{L(m)}{L(n)} - \frac{\ln m}{\ln n} \right| \leq \frac{1}{s} \quad (3.63)$$

Como s es arbitrario:

$$\frac{L(m)}{L(n)} = \frac{\ln m}{\ln n} \Rightarrow \frac{L(m)}{\ln m} = \frac{L(n)}{\ln n} \equiv k \quad (3.64)$$

Es decir, en el caso equiprobable con $\Omega > 1$, la función ignorancia tiene la forma:

$$L(\Omega) = k \ln \Omega \quad (3.65)$$

Esta fórmula cuadra con la propiedad 5, de modo que se puede considerar válida para cualquier valor entero positivo de Ω .

d) Si las probabilidades p_i son números racionales, es decir, pueden escribirse como cociente de dos números enteros $p_i = \frac{q_i}{q}$ (donde q es el mínimo común múltiplo de los denominadores de todas las p_i), entonces, como su suma es la unidad, ha de ser:

$$\sum_{i=1}^{\Omega} q_i = q \sum_{i=1}^{\Omega} p_i = q \quad (3.66)$$

Sea ahora B una distribución finita dependiente de la distribución A y constituida por Ω conjuntos de sucesos tales que, si se da el suceso A_i , en el grupo i hay q_i sucesos equiprobables y todos los demás sucesos de B tienen probabilidad nula. La ignorancia de B cuando se da el suceso posible A_i será:

$$I(B|A_i) = I\left(0, \dots, 0, \underbrace{\frac{1}{q_i}, \dots, \frac{1}{q_i}}_{q_i}, 0, \dots, 0\right) = I\left(\frac{1}{q_i}, \dots, \frac{1}{q_i}\right) = L(q_i) = k \ln q_i = k \ln(q p_i) = k \ln q + k \ln p_i \quad (3.67)$$

Por tanto, en virtud de la propiedad 4:

$$I(AB) = I(A) + k \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} p_i \ln q + k \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} p_i \ln p_i = I(A) + k \ln q + k \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} p_i \ln p_i \quad (3.68)$$

Por otro lado, la probabilidad del suceso $A_i \cap B_l$ de la distribución AB será $p(A_i \cap B_l) = p_i \cdot p(B_l|A_i)$, donde:

$$p(B_l|A_i) = \begin{cases} 0 & \text{si } l \text{ no es del conjunto } i \Rightarrow p(A_i \cap B_l) = 0 \\ \frac{1}{q_i} & \text{si } l \text{ es del conjunto } i \Rightarrow p(A_i \cap B_l) = p_i \frac{1}{q_i} = \frac{1}{q} \end{cases} \quad (3.69)$$

lo que lleva a:

$$I(AB) = L(q) = k \ln q \quad (3.70)$$

Igualando las dos expresiones obtenidas para $I(AB)$, se encuentra:

$$I(A) = -k \sum_{\substack{i=1 \\ p_i \neq 0}}^{\Omega} p_i \ln p_i \quad (3.71)$$

Esta fórmula, que ha sido obtenida para cuando las probabilidades son números racionales, es válida para cualesquiera probabilidades reales, debido a la propiedad 1 de continuidad de la función ignorancia.

Ocurre a menudo que las distribuciones de probabilidad que hay que considerar en Física Estadística no son finitas, sino infinitas numerables. Para estas distribuciones, se define, por extensión, la función ignorancia por la misma expresión que para las distribuciones finitas (se supondrá que la distribución es tal que la suma, que ahora es una serie, converge). La ignorancia verifica todas las propiedades del caso finito, salvo, obviamente, la propiedad 2, que carece de sentido. Se escribirá, por tanto, para una distribución discreta finita o numerable:

$$\boxed{I(A) = -k \sum_i p_i \ln p_i} \quad (3.72)$$

Ejercicios - Principio de Máxima Entropía

- 1.- Se lanza muchas veces un dado y el valor medio obtenido de la puntuación del dado es 3.5. De acuerdo con el PME, ¿cuáles son las probabilidades más adecuadas que debemos asignar a cada uno de los posibles resultados de lanzar el dado $\{1,2,3,4,5,6\}$? ¿Cuál es la ignorancia asociada a este experimento?
- 2.- Resuelva el “problema del dado de Brandeis” utilizando el PME. El experimento consiste en lanzar muchas veces un dado y, si el valor medio obtenido de la puntuación del dado es 4.5, la pregunta es ¿cuáles son las probabilidades más adecuadas que debemos asignar a cada uno de los posibles resultados de lanzar el dado $\{1,2,3,4,5,6\}$? Calcule también la ignorancia asociada a este problema.
- 3.- Utilizando el PME, infiera las probabilidades asociadas a cada uno de los posibles resultados $\{1,2,3,4,5,6\}$ del lanzamiento de un dado trucado, sabiendo que, tras lanzarlo muchas veces, el “6” ha salido el doble de veces que el “1”. Calcule también la ignorancia asociada a esta situación.

4. Formalismo de la Física Estadística

4.1. Descripción estadística de un estado de equilibrio

Dado un sistema termodinámico, cualquiera de sus estados de equilibrio cumple las siguientes propiedades:

- Está completamente determinado especificando los valores de las variables extensivas (las variables naturales de la entropía), que permanecen constantes en el tiempo. También podemos cambiar alguna de las variables extensivas por una de sus intensivas conjugadas (por ejemplo, cambiar U por $1/T$, o cambiar V por p/T).
- Las fluctuaciones de estas variables son despreciables con relación al valor de equilibrio.

Estas dos propiedades permiten construir la descripción estadística de un estado de equilibrio de forma consistente. A lo largo de este tema supondremos que tratamos con sistemas homogéneos sin ligaduras internas, que solo pueden intercambiar trabajo de expansión y con un solo componente. Sus variables características serán:

- Extensivas: energía interna U , volumen V , y número de entidades elementales N .
- Intensivas conjugadas de las extensivas: $1/T$, p/T y $-\mu/T$, donde T es la temperatura, p es la presión y μ es el potencial químico.

En la descripción estadística, se especifica un **macroestado** (unas ciertas variables termodinámicas), normalmente escogido en consonancia con las paredes y el entorno del sistema. Por ejemplo:

- Para un **sistema aislado**, se escogen las variables naturales de la entropía: (U, V, N) .
- Para un **sistema cerrado isotermo e isócoro**, en contacto con un termostato que fija su temperatura, se escogen las variables naturales de la función de Massieu: $(1/T, V, N)$.
- Etc.

Después, se enumeran los microestados A_i que son compatibles con el macroestado elegido y a cada uno se le asigna una probabilidad p_i .

Los microestados están caracterizados por los valores de las variables extensivas (energía, volumen, número de entidades elementales, etc), porque son las que tienen sentido a nivel microscópico. Las variables intensivas ($1/T$, p/T , etc.) no son variables “mecánicas” y no tienen sentido a nivel microscópico, dentro del modelo de realidad.

Las probabilidades dependen de las variables del macroestado elegido y se acepta que, de acuerdo con el Principio de Máxima Entropía, sus valores **maximizan la entropía de información de acuerdo con la información disponible**. En Física Estadística, la constante que aparece en la definición de la entropía se toma igual a la constante de Boltzmann ($k_B = 1.380\ 649 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$):

$$S = -k_B \sum_i p_i \ln p_i \quad (4.1)$$

Esta identificación entre la entropía termodinámica y la entropía de información la podemos considerar como nuestro postulado de partida para el desarrollo de la Física Estadística.

¿Qué podemos considerar como información disponible?

Por un lado, **las paredes y el entorno limitan el número de microestados compatibles con el macroestado**. Por ejemplo:

- Un **sistema aislado** puede tener un **único valor de la energía, del volumen y del número de entidades elementales**. Las paredes aislantes impiden su variación.
- En un **sistema cerrado isotermo e isócoro**, el **volumen y el número de entidades elementales son estrictamente constantes, pero la energía interna puede fluctuar**. 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**.

Por otro lado, necesitamos saber cuál es la relación entre el valor macroscópico de las variables extensivas y sus valores en sus posibles microestados. Para ello, notemos que, al ser las fluctuaciones de las variables extensivas (energía, volumen, número de partículas) despreciables con relación a sus valores de equilibrio, las distribuciones de probabilidad que las describen en los estados de equilibrio son muy “estrechas”, teniendo un **máximo muy acusado que concentra prácticamente toda la probabilidad y que es el que se corresponde con el valor macroscópico**. Siendo la distribución de probabilidad tan estrecha, es equivalente considerar que el valor macroscópico es el valor más probable o el valor medio, ya que ambos coinciden prácticamente. De hecho, para los órdenes de magnitud que se manejan en Física Estadística, esta coincidencia puede calificarse de exacta. Por lo tanto, consideraremos que **los valores macroscópicos de las variables extensivas son iguales a sus valores medios de acuerdo con la distribución de probabilidad de sus microestados**. Esta información debe implementarse en el proceso de maximización de la entropía introduciendo ligaduras del tipo:

$$\sum_i p_i X_i = \langle X \rangle \quad (4.2)$$

siendo X una variable extensiva.

La “estrechez” de las distribuciones de probabilidad también permite deducir que el valor de la entropía es prácticamente el mismo independientemente de las paredes y el entorno del sistema. **Esto es esencial para garantizar la consistencia de nuestro postulado (4.1)**. En el tema sobre Termodinámica, esto lo hemos venido suponiendo implícitamente, pues hemos considerado que la entropía es una función de estado, solo dependiente de las variables de estado del sistema. Veremos con más detalle esta cuestión con el ejemplo concreto del conjunto canónico (sección 4.3.2).

Por último, señalemos que **esta “estrechez” se refiere solamente a las distribuciones de probabilidad del sistema total, pero no a las distribuciones de probabilidad de una de las entidades elementales**, que no comparten esta propiedad. Por ejemplo, tendremos la ocasión de comprobarlo más adelante al ver que la energía de una sola molécula de un gas ideal monoatómico semiclásico ε (cinética de traslación) sigue la distribución de probabilidad descrita por la función de densidad de probabilidad:

$$g(\varepsilon) = \frac{2}{\sqrt{\pi}(k_B T)^{3/2}} \varepsilon^{1/2} e^{-\frac{\varepsilon}{k_B T}} \quad (4.3)$$

que no presenta esas propiedades de “estrechez”. Cada una se mueve con una velocidad y energía cinética, de tal manera que unas velocidades son más probables que otras, con una distribución “ancha” (véase la Figura 5.3 de la sección 5.3). No entraremos en los mecanismos físicos y estadísticos que relacionan esta distribución “ancha” con la distribución “estrecha” que sigue la energía del sistema total, por quedar fuera de los objetivos del curso.

4.2. Conjunto microcanónico: macroestado (U, V, N)

Se denomina **conjunto microcanónico** al espacio muestral formado por los microestados de un **sistema aislado**, para el que los valores de las variables extensivas (energía interna, volumen y número de entidades elementales) son estrictamente constantes.

El macroestado queda perfectamente determinado por los valores de estas variables: **(U, V, N)**, que son las **variables naturales de la entropía**, potencial termodinámico natural utilizado para describir un sistema aislado.

4.2.1. Inferencia de probabilidades

De acuerdo con el modelo de realidad elegido, habrá un número determinado de microestados, $\Omega(U, V, N)$, compatibles con el macroestado. Cada microestado i -ésimo tendrá una probabilidad p_i . Las probabilidades siempre son funciones de las variables del macroestado, $p_i(U, V, N)$.

Se trata, por tanto, de una distribución finita de probabilidad en la que no se dispone de ninguna información adicional más allá de la condición de normalización:

$$\sum_i p_i = 1 \quad (4.4)$$

Como consecuencia, **todos los microestados serán equiprobables**, con una probabilidad:

$$p_i = \frac{1}{\Omega} \quad (4.5)$$

A esta distribución de probabilidad le corresponde una entropía:

$$S = k_B \ln \Omega \quad (4.6)$$

4.2.2. Consecuencias

La ecuación (4.6) es la **ecuación de Boltzmann-Planck**. La hemos obtenido únicamente a partir de la hipótesis de que el estado es de equilibrio y del Principio de Máxima Entropía (de información o ignorancia). Si es posible considerar que el valor de la entropía no depende ni de las paredes ni del entorno, de tal forma que es una función de estado que solo depende de los valores de las variables de estado del sistema (ver sección 4.3.2), esta ecuación permite deducir el postulado de máxima entropía (termodinámica) de Callen, del cual se deriva el Principio de Aumento de Entropía y todo el formalismo termodinámico.

Las ecuaciones (4.5) y (4.6) son muy sencillas y de una gran generalidad. Son válidas, aunque el sistema pueda intercambiar otros tipos de trabajo, tenga más componentes, etc.

La ecuación de Boltzmann-Planck es la **ecuación fundamental del sistema aislado** de acuerdo con el modelo de realidad elegido. A partir de ella podemos obtener las ecuaciones de estado haciendo las derivadas parciales correspondientes. Poniendo $\ln \Omega = S/k_B$ y usando la ecuación (1.39):

$$\begin{aligned} \frac{1}{k_B T} &= \left(\frac{\partial \ln \Omega}{\partial U} \right)_{V, N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{k_B T} &= \left(\frac{\partial \ln \Omega}{\partial V} \right)_{U, N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{k_B T} &= \left(\frac{\partial \ln \Omega}{\partial N} \right)_{U, V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (4.7)$$

4.3. Conjunto canónico: macroestado $(1/T, V, N)$

En muchos casos resulta difícil o inviable obtener la ecuación fundamental de un sistema empleando el conjunto microcanónico. Esa es una de las razones por las que en Física Estadística se utilizan diferentes conjuntos o espacios muestrales. En el conjunto canónico, que pasamos a definir a continuación, en lugar del número de microestados hay que calcular una magnitud denominada función de partición, que tiene ciertas propiedades que simplifican su cálculo en muchos casos prácticos.

Se denomina **conjunto canónico** 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. Como ya hemos dicho al hablar de la descripción estadística, la energía puede tomar cualquier valor entre su valor mínimo y máximo. En este caso, el valor macroscópico de la energía interna es igual a su valor medio, $\langle E \rangle$.

El macroestado adecuado en este caso es aquel que utiliza las variables naturales de la función de Massieu, que es el potencial termodinámico natural que se usa para describir un sistema cerrado isotermo e isócoro: $(1/T, V, N)$.

4.3.1. Inferencia de probabilidades

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 $E_i(V, N)$** . Las probabilidades, como siempre, serán funciones de *todas* las variables del macroestado: $p_i(1/T, V, N)$.

El problema de inferencia consiste en maximizar la entropía sujeta a las ligaduras siguientes:

$$\begin{aligned} \text{Condición de normalización:} & \quad \sum_i p_i = 1 \\ \text{Valor medio de la energía interna:} & \quad \sum_i p_i E_i = \langle E \rangle \end{aligned} \quad (4.8)$$

Estas son las mismas condiciones de ligadura que aparecían en el juego de reparto que se vio en el tema del Principio de Máxima Entropía. Por lo tanto, ya conocemos la solución al problema, tras formar la función de Lagrange para maximizarla sin ninguna ligadura. Las probabilidades son de la forma:

$$p_i = \frac{1}{Z} e^{-\beta E_i} \quad (4.9)$$

donde Z es la llamada **función de partición**:

$$Z = \sum_i e^{-\beta E_i} \quad (4.10)$$

Los multiplicadores de Lagrange y, por tanto, β y Z deben ser funciones de las variables del macroestado, $(1/T, V, N)$.

Existe, no obstante, una diferencia fundamental entre este problema y el del juego de reparto: en este caso no es posible obtener una solución "numérica" para la ecuación que determina el parámetro β (véase la solución del juego de reparto), que es:

$$\sum_i (E_i - \langle E \rangle) e^{-\beta E_i} = 0 \quad (4.11)$$

Sin embargo, vamos a demostrar que β está directamente relacionado con la temperatura del sistema.

4.3.2. Entropía en el conjunto canónico. Equivalencia de conjuntos

La entropía del sistema en el conjunto canónico se puede calcular como:

$$S = -k_B \sum_i p_i \ln p_i = -k_B \langle \ln p_i \rangle = k_B \langle \beta E_i + \ln Z \rangle = k_B \beta \langle E \rangle + k_B \ln Z \quad (4.12)$$

La entropía en el conjunto canónico, como vemos, es diferente de la entropía en el conjunto microcanónico. Esto puede resultar chocante, pero tiene sentido, pues en ambos casos la información sobre el sistema es distinta y al maximizarla obtendremos resultados distintos. Pero fijémonos en que hasta ahora hemos supuesto implícitamente que la entropía en Termodinámica era una función cuyo valor en un macroestado dado no dependía de las paredes que lo delimitan y de su entorno, es decir, que debe tener el mismo valor para sistemas aislados que para sistemas cerrados isotermos e isócoros si ambos se encuentran en macroestados equivalentes. ¿Es esto una contradicción?

La respuesta es que no hay contradicción. Vamos a demostrar que **el hecho de que el valor de la entropía no dependa del conjunto utilizado (canónico o microcanónico) es una consecuencia de la “estrechez” de la distribución de probabilidad de la energía interna del sistema.**

En el contexto del conjunto canónico, la “estrechez” de dicha distribución quiere decir que la probabilidad de que la energía interna sea igual a su valor medio es prácticamente la unidad. De acuerdo con la ecuación (4.9), todos estos microestados cuya energía es igual al valor medio son igualmente probables porque el valor de su energía es el mismo y su probabilidad será $1/\Omega(\langle E \rangle, V, N)$, donde $\Omega(\langle E \rangle, V, N)$ es el número de microestados con esta energía. Por lo tanto, la distribución de probabilidad de los microestados en el conjunto canónico será muy aproximadamente (en la práctica, con los órdenes de magnitud que manejamos en Física Estadística, es exactamente):

$$p_i = \begin{cases} \frac{1}{\Omega(\langle E \rangle, V, N)} & , \text{ si } E_i = \langle E \rangle \\ 0 & , \text{ si } E_i \neq \langle E \rangle \end{cases} \quad (4.13)$$

Y llevando esto a la entropía, obtenemos:

$$S_{\text{canónica}} = k_B \ln \Omega(\langle E \rangle, V, N) = S_{\text{microcanónica}} \quad (4.14)$$

Cuando esto ocurre, se dice que hay **equivalencia de conjuntos**. Sin pretender ser precisos, podemos afirmar que la equivalencia de conjuntos se cumple si hay acoplamiento débil y el número de entidades elementales es suficientemente alto (“límite termodinámico”).

Pueden obtenerse ecuaciones como la (4.14) para otros conjuntos distintos del canónico, que no estudiaremos en este curso. Y son en realidad estas ecuaciones las que garantizan la consistencia lógica de la identificación de la entropía de información con una función de estado termodinámica, pues una función de estado no debe depender de nada más que de sus variables de estado, en este caso, de la energía, del volumen y del número de partículas, y no del carácter de las paredes o del entorno.

4.3.3. Relación de β con la temperatura y de Z con la función de Massieu

A partir de la ecuación (1.40) y la ecuación (4.12) podemos calcular la función de Massieu:

$$\psi\left(\frac{1}{T}, V, N\right) = S - \frac{1}{T} \langle E \rangle = \left(k_B \beta - \frac{1}{T}\right) \langle E \rangle + k_B \ln Z \quad (4.15)$$

donde $\langle E \rangle = U$. La Ecuación Energética de Estado es (ecuación (1.43)):

$$-\langle E \rangle = \left[\frac{\partial \psi}{\partial (1/T)} \right]_{V,N} \quad (4.16)$$

En lo que sigue omitiremos, por simplificar la notación, los subíndices V, N indicando que las derivadas se toman a V y N constantes. Sustituyendo (4.15) en (4.16) y simplificando el término $-\langle E \rangle$ en ambos miembros:

$$0 = k_B \frac{\partial \beta}{\partial (1/T)} \langle E \rangle + \left(k_B \beta - \frac{1}{T} \right) \frac{\partial \langle E \rangle}{\partial (1/T)} + k_B \frac{\partial (\ln Z)}{\partial (1/T)} \quad (4.17)$$

Para simplificar esta ecuación conviene evaluar primero la última derivada:

$$\frac{\partial (\ln Z)}{\partial (1/T)} = \frac{1}{Z} \frac{\partial Z}{\partial (1/T)} = \frac{1}{Z} \left\{ \frac{\partial}{\partial (1/T)} \left(\sum_i e^{-\beta E_i} \right) \right\} \quad (4.18)$$

Para hacer la derivada del corchete, es clave recordar que $E_i(V, N)$ **no depende de $1/T$** (véase la sección 4.1). Entonces:

$$\frac{\partial (\ln Z)}{\partial (1/T)} = \frac{1}{Z} \left(\sum_i e^{-\beta E_i} \right) (-E_i) \frac{\partial \beta}{\partial (1/T)} = - \left(\sum_i \frac{1}{Z} e^{-\beta E_i} E_i \right) \frac{\partial \beta}{\partial (1/T)} \quad (4.19)$$

Teniendo en cuenta la ecuación (4.9):

$$\frac{\partial (\ln Z)}{\partial (1/T)} = - \left(\sum_i p_i E_i \right) \frac{\partial \beta}{\partial (1/T)} = - \langle E \rangle \frac{\partial \beta}{\partial (1/T)} \quad (4.20)$$

Al sustituir (4.20) en (4.17) se obtiene:

$$0 = \left(k_B \beta - \frac{1}{T} \right) \frac{\partial \langle E \rangle}{\partial (1/T)} \quad (4.21)$$

Como la última derivada es distinta de cero, llegamos a la conclusión:

$$\boxed{\beta = \frac{1}{k_B T}} \quad (4.22)$$

Y, sustituyendo en (4.15):

$$\boxed{\psi = k_B \ln Z} \quad (4.23)$$

Esto quiere decir que el cálculo de la función de partición nos permite obtener directamente la función de Massieu y, a partir de ella, se siguen las ecuaciones de estado del sistema por simple derivación parcial. La ecuación (4.23) es la **ecuación fundamental** del sistema cerrado isotermo e isócoro.

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

$$\begin{aligned} -\langle E \rangle &= \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V,N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{k_B T} = p\beta &= \left(\frac{\partial \ln Z}{\partial V} \right)_{\beta,N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{k_B T} = -\mu\beta &= \left(\frac{\partial \ln Z}{\partial N} \right)_{\beta,V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (4.24)$$

4.3.4. Obtención de Z como una transformada de Laplace

Si se ha calculado el número de microestados $\Omega(E, V, N)$ para una cierta energía E (es decir, si se ha logrado obtener la ecuación fundamental del sistema en el conjunto microcanónico), es posible obtener la función de partición a partir de Ω . En la suma (4.10) que define la función de partición, el índice i etiqueta cada posible microestado del sistema. Pero existen $\Omega(E, V, N)$ microestados diferentes para una cierta energía concreta E , de modo que habrá $\Omega(E, V, N)$ sumandos iguales para cada valor E permitido de la energía y, en consecuencia, podemos escribir la suma como:

$$Z(\beta, V, N) = \sum_{\substack{i \\ \text{microestados}}} e^{-\beta E_i} = \sum_{\substack{E \\ \text{niveles}}} \Omega(E, V, N) e^{-\beta E} \quad (4.25)$$

Es decir, la segunda suma de la ecuación (4.25) está escrita en términos de los distintos **niveles de energía** (E) permitidos al sistema, en lugar de a los distintos microestados. En dicha suma, cada término tiene un valor diferente de E .

4.3.5. Factorización de la función de partición

La función de partición tiene una propiedad muy importante que simplifica su cálculo: **siempre que la energía se pueda expresar como una suma de términos independientes, la función de partición podrá ponerse como un producto de las funciones de partición asociadas a cada uno de los sumandos de la energía.**

Ejemplos de lo que pueden representar los distintos sumandos serían:

- Energías asociadas a distintos tipos de movimiento del sistema que se puedan considerar independientes, como traslación y rotación de las moléculas en un gas ideal poliatómico.
- Energías asociadas a distintas entidades elementales independientes de un sistema, como por ejemplo a moléculas de un gas ideal.
- Energías asociadas a distintas porciones independientes de un sistema.

Para demostrar esta propiedad, supongamos que la energía E de un sistema puede expresarse como la suma de dos términos independientes entre sí E_1 y E_2 (si hubiera más términos, la deducción sería análoga):

$$E = E_1 + E_2 \quad (4.26)$$

Al término E_1 se le podrán asignar distintos microestados E_{1i} y de forma análoga al término E_2 se le podrán asignar distintos microestados E_{2j} . Al calcular la función de partición, podremos escribir:

$$Z = \sum_i \left\{ \sum_j e^{-\beta(E_{1i} + E_{2j})} \right\} = \sum_i \left\{ \sum_j e^{-\beta E_{1i}} e^{-\beta E_{2j}} \right\} \quad (4.27)$$

Ahora bien, como E_1 y E_2 son independientes, los índices i y j también lo serán. Podemos extraer como factor común del sumatorio a j todo lo que dependa solo del índice i , con lo que:

$$Z = \sum_i \left\{ e^{-\beta E_{1i}} \sum_j e^{-\beta E_{2j}} \right\} = \left(\sum_i e^{-\beta E_{1i}} \right) \left(\sum_j e^{-\beta E_{2j}} \right) = Z_1 Z_2 \quad (4.28)$$

donde se han definido las funciones de partición de cada sumando independiente de forma evidente. Esta es la relación que queríamos demostrar.

Si un sistema tiene N entidades elementales idénticas e independientes, la energía E del sistema se puede poner como la suma de las energías ε de cada entidad elemental independiente:

$$E = \varepsilon_1 + \dots + \varepsilon_N \quad (4.29)$$

donde, por ser idénticas las entidades elementales, sus posibles microestados (que llamaremos **monoestados**, para diferenciarlos de los microestados del sistema total) son idénticos, con lo que la función de partición de todas ellas vale lo mismo. Designaremos la función de partición de una entidad elemental por Z_1 :

$$Z_1 = \sum_{\text{monoestados}} e^{-\beta\varepsilon_i} \quad (4.30)$$

En el marco de la Mecánica Cuántica, el hecho de poder definir Z_1 está sujeto a que las entidades elementales sean distinguibles (veremos esto en la sección 4.5.3). De acuerdo con la propiedad de factorización:

$$Z = Z_1^N \quad (4.31)$$

Aplicaremos también la propiedad de factorización al cálculo de la función de partición de un gas ideal poliatómico (capítulo 6), descomponiendo su energía en la suma de las energías asociadas a los distintos grados de libertad de traslación, rotación y vibración de sus moléculas.

4.3.6. Fluctuaciones de energía en el conjunto canónico

La **varianza** de una variable aleatoria se define como el **valor medio del cuadrado de la desviación de la variable respecto a su valor medio**. En el caso de la energía en el conjunto canónico, la designaremos ⁷ por $(\Delta^*E)^2$ y será:

$$(\Delta^*E)^2 = \langle (E - \langle E \rangle)^2 \rangle \quad (4.32)$$

Es posible expresar la varianza de otro modo:

$$(\Delta^*E)^2 = \langle E^2 - 2E\langle E \rangle + \langle E \rangle^2 \rangle = \langle E^2 \rangle - 2\langle E \rangle\langle E \rangle + \langle E \rangle^2 = \langle E^2 \rangle - \langle E \rangle^2 \quad (4.33)$$

La ecuación anterior pone de relieve que $\langle E^2 \rangle \neq \langle E \rangle^2$, y que además $\langle E^2 \rangle \geq \langle E \rangle^2$, pues por definición la varianza es una magnitud positiva.

En Física Estadística se suele tomar como medida de las fluctuaciones de una variable su **desviación típica o desviación estándar**, definida como la **raíz cuadrada de su varianza**. Veamos cómo están relacionadas estas fluctuaciones con la función de partición. Calculemos la segunda derivada de $\ln Z$ respecto de β , haciendo uso de la EEE y de la ecuación (4.8), y teniendo en cuenta que las energías de los microestados no dependen de β :

$$\left(\frac{\partial^2 \ln Z}{\partial \beta^2} \right)_{V,N} = - \left(\frac{\partial \langle E \rangle}{\partial \beta} \right)_{V,N} = - \sum_i E_i \left(\frac{\partial p_i}{\partial \beta} \right)_{V,N} \quad (4.34)$$

La derivada de la probabilidad se obtiene más fácilmente tomando el logaritmo de la ecuación (4.9):

$$\ln p_i = -\ln Z - \beta E_i \quad (4.35)$$

y derivando:

$$\frac{1}{p_i} \left(\frac{\partial p_i}{\partial \beta} \right)_{V,N} = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V,N} - E_i = \langle E \rangle - E_i \Rightarrow \left(\frac{\partial p_i}{\partial \beta} \right)_{V,N} = p_i (\langle E \rangle - E_i) \quad (4.36)$$

⁷ Notación del libro *J. Brey Ábalo, J. de la Rubia Pacheco, J. de la Rubia Sánchez, "Mecánica Estadística", UNED, 2001.*

Sustituyendo en (4.34):

$$\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = - \sum_i E_i [p_i (\langle E \rangle - E_i)] = \sum_i p_i E_i^2 - \langle E \rangle \sum_i p_i E_i = \langle E^2 \rangle - \langle E \rangle^2 \quad (4.37)$$

o, según la ecuación (4.33):

$$\boxed{\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = (\Delta^* E)^2} \quad (4.38)$$

Este resultado es general: **las derivadas segundas del potencial termodinámico están relacionadas con las fluctuaciones de las variables extensivas que no estén fijadas por las paredes del sistema.**

Analicemos ahora las **fluctuaciones relativas**, definidas por:

$$(\Delta^* E)_{\text{rel}} = \frac{\Delta^* E}{\langle E \rangle} = \sqrt{\frac{(\Delta^* E)^2}{\langle E \rangle^2}} = \sqrt{-\left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_{V,N} \frac{1}{\langle E \rangle^2}} \quad (4.39)$$

Si el sistema está formado por N entidades elementales idénticas e independientes, entonces se cumple que:

$$\langle E \rangle = N \langle \varepsilon \rangle \quad (4.40)$$

donde $\langle \varepsilon \rangle$ es el valor medio de la energía de una entidad elemental. Sustituyendo en (4.39):

$$\boxed{(\Delta^* E)_{\text{rel}} = \frac{\Delta^* E}{\langle E \rangle} = \sqrt{-\frac{1}{\langle \varepsilon \rangle^2} \left(\frac{\partial \langle \varepsilon \rangle}{\partial \beta}\right)_{V,N} \cdot \frac{1}{N}} \propto \frac{1}{\sqrt{N}}} \quad (4.41)$$

Este resultado también es general: **las fluctuaciones relativas de una variable extensiva son inversamente proporcionales a la raíz del número de entidades elementales.** Por tanto, a medida que aumenta el número de entidades elementales las fluctuaciones en las variables extensivas son cada vez más pequeñas. Este resultado es consistente con nuestra hipótesis de partida: que las fluctuaciones en un sistema en equilibrio son despreciables.

Por último, veamos la relación que existe entre las fluctuaciones de energía y la capacidad calorífica a volumen constante:

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T}\right)_{V,N} = \left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_{V,N} \frac{d\beta}{dT} = -\frac{1}{k_B T^2} \left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_{V,N} = \frac{1}{k_B T^2} \left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = k_B \beta^2 (\Delta^* E)^2 \quad (4.42)$$

Es decir, C_V es proporcional al cuadrado de las fluctuaciones de energía. Debe ser siempre positiva.

4.4. Entropía de una distribución continua

Con el objetivo de abordar el límite semiclásico de la Física Estadística, presentamos brevemente las particularidades que entraña la definición de la entropía para una distribución continua de probabilidad. Por simplicidad, supondremos una variable aleatoria continua unidimensional x que puede tomar cualquier valor en el intervalo $[a,b]$.

Vamos a guiarnos por lo que ya sabemos sobre la entropía de una distribución discreta. Para ello, "discretizamos" el intervalo $[a,b]$, dividiéndolo en M segmentos de igual longitud:

$$\delta x = \frac{b-a}{M} \quad (4.43)$$

Los puntos que dividen el intervalo tienen por coordenadas:

$$a_m = a + m \cdot \frac{b-a}{M} = a + m \cdot \delta x \quad , \quad m = 0, 1, 2, \dots, M \quad (4.44)$$

donde $a_0 = a$ y $a_M = b$. La probabilidad de que x esté entre dos puntos consecutivos, $\delta p_m = p(a_m \leq x \leq a_{m+1})$, puede expresarse en términos de la función de densidad de probabilidad $\rho(x)$ como:

$$\delta p_m = p(a_m \leq x \leq a_{m+1}) = \int_{a_m}^{a_{m+1}} \rho(x) dx \quad , \quad m = 0, 1, \dots, M-1 \quad (4.45)$$

De acuerdo con el teorema del valor medio para integrales, existe un $x_m \in [a_m, a_{m+1}]$ tal que:

$$\delta p_m = \rho(x_m) \delta x \quad , \quad m = 0, 1, \dots, M-1 \quad (4.46)$$

La condición de normalización es:

$$\sum_{m=0}^{M-1} \delta p_m = \sum_{m=0}^{M-1} \rho(x_m) \delta x = 1 \quad (4.47)$$

La entropía de esta distribución “discretizada” de probabilidad, en la que los sucesos elementales no son puntos del intervalo $[a, b]$, sino los segmentos $[a_m, a_{m+1}]$, es:

$$S = -k \sum_{m=0}^{M-1} \delta p_m \ln \delta p_m = -k \sum_{m=0}^{M-1} [\rho(x_m) \delta x] \cdot \ln [\rho(x_m) \delta x] \quad (4.48)$$

El paso al “límite continuo” consiste en hacer tender a cero la longitud de los segmentos para que cada suceso elemental sea un punto del intervalo $[a, b]$. No obstante, al hacer tender $\delta x \rightarrow 0$, la entropía diverge. Es fácil verlo haciendo una pequeña transformación en la ecuación (4.48), separando en dos sumandos el logaritmo $\ln[\rho(x_m) \delta x]$ y usando la condición de normalización:

$$S = -k \sum_{m=0}^{M-1} \delta x [\rho(x_m) \cdot \ln[\rho(x_m)] - k \ln \delta x] \quad (4.49)$$

Tomando el límite, el primer término se transforma en una integral (por definición), pero el segundo tiende a infinito:

$$S \rightarrow -k \int_a^b dx \rho(x) \cdot \ln \rho(x) + \infty = +\infty \quad (4.50)$$

Es decir, **la entropía de una distribución continua es infinita**. Esto es lógico si interpretamos la entropía como medida de la falta de información de una distribución de probabilidad, pues necesitaríamos infinita información para localizar un punto matemático en el segmento $[a, b]$.

Existe, sin embargo, una alternativa para tratar el problema, que consiste en **aceptar que un punto matemático (valor de x con una precisión infinita) no existe nunca en Física**. En un determinado problema, habrá un intervalo mínimo $\delta_0 x$ correspondiente a la mejor precisión posible en la medida de la variable x . El paso al límite continuo es, por tanto, “ilusorio” y la “aproximación” discretizada de la ecuación (4.48) es más próxima a la realidad que el paso al límite. Tomando $\delta x = \delta_0 x$, la entropía de la distribución continua se puede definir como:

$$S = -k \sum_{m=0}^{M-1} [\rho(x_m) \delta_0 x] \cdot \ln [\rho(x_m) \delta_0 x] \quad (4.51)$$

Si $\delta_0 x$ corresponde realmente a la mejor precisión posible sobre las medidas de la variable x , las variaciones de $\rho(x)$ en intervalos de longitud $\delta_0 x$ son necesariamente despreciables y, por ello, para un intervalo $[a, b]$ más grande que $\delta_0 x$, la suma anterior será indistinguible de una integral (pero en este caso

no hacemos tender $\delta_0 x$ a cero dentro del logaritmo). Por otro lado, si el intervalo $[a, b]$ es más pequeño que $\delta_0 x$, la entropía debe ser nula porque en el intervalo $[a, b]$ solo existiría un único suceso. Resumiendo:

$$S = \begin{cases} -k \int_a^b dx \rho(x) \cdot \ln[\rho(x)\delta_0 x] & , \text{ si } b - a > \delta_0 x \\ 0 & , \text{ si } b - a \leq \delta_0 x \end{cases} \quad (4.52)$$

Si tratáramos de aplicar la fórmula integral al caso $b - a < \delta_0 x$ obtendríamos una entropía negativa, cosa que no es posible.

En lo que sigue veremos que, **en Física Estadística semiclásica, el parámetro que juega el papel de $\delta_0 x$ en el espacio fásico está relacionado con la constante de Planck de la Mecánica Cuántica.**

Veamos ahora dos ejemplos extremos de la aplicación de la fórmula (4.52).

- Si la información de que disponemos es máxima, entonces conocemos x con una precisión $\delta_0 x$. En tal caso, hay un punto x_0 tal que:

$$\rho(x) = \begin{cases} \frac{1}{\delta_0 x} & \text{si } x \in \left[x_0 - \frac{\delta_0 x}{2}, x_0 + \frac{\delta_0 x}{2} \right] \\ 0 & \text{si } |x - x_0| > \frac{\delta_0 x}{2} \end{cases} \quad (4.53)$$

$$S = -k \int_{x_0 - \frac{\delta_0 x}{2}}^{x_0 + \frac{\delta_0 x}{2}} dx \frac{1}{\delta_0 x} \ln \left[\frac{1}{\delta_0 x} \delta_0 x \right] = 0 \quad (4.54)$$

- Si $b - a > \delta_0 x$ y la información de que disponemos es mínima, entonces, todos los “segmentos” son igualmente probables y la densidad de probabilidad corresponde a una distribución uniforme:

$$\rho(x) = \frac{1}{b - a} \quad (4.55)$$

$$S = -k \int_a^b dx \frac{1}{b - a} \ln \left[\frac{1}{b - a} \delta_0 x \right] = k \ln \left(\frac{b - a}{\delta_0 x} \right) \quad (4.56)$$

La magnitud $(b - a)/\delta_0 x$ es el número de segmentos elementales en los que se ha dividido el intervalo o, dicho de otra forma, el número de valores de x distinguibles en la práctica. Sería el equivalente al “número de microestados” en el caso discreto.

4.5. Física Estadística semiclásica

La aplicación del formalismo de la Física Estadística a modelos de realidad que obedecen puramente a la Mecánica Clásica da resultados inconsistentes. Estas inconsistencias tienen su germen en limitaciones físicas fundamentales que la Mecánica Cuántica impone a la precisión con que pueden medirse simultáneamente ciertos pares de variables, como la posición y el momento lineal, lo que se conoce como **Principio de Incertidumbre de Heisenberg**.

La **Física Estadística semiclásica** es el resultado de aplicar ciertas “correcciones cuánticas” a la Mecánica Clásica. Estas correcciones se obtienen tomando el “límite semiclásico” de la Mecánica Cuántica (es decir, trabajando con la Mecánica Cuántica en los casos en los que es válida la Mecánica Clásica con un alto grado de aproximación). El análisis detallado de estas cuestiones se sale de los límites de este curso, pero expondremos ahora, de forma simplificada, en qué consisten dichas correcciones.

4.5.1. Discretización del espacio fásico

En Mecánica Clásica, cada uno de los posibles microestados de un sistema viene dado por el conjunto de valores posibles de las coordenadas $x = (x_1, \dots, x_s)$ y momentos $p = (p_1, \dots, p_s)$ de todas sus partículas. El espacio formado por las coordenadas y momentos se denomina **espacio fásico** y el número de coordenadas necesarias, s , se llama **número de grados de libertad**. La energía del sistema se pone como función de sus coordenadas y momentos: $E = E(x, p)$.

El conjunto de posibles microestados, por tanto, forma un continuo y ya hemos visto que el paso al límite continuo en la entropía es imposible. Por suerte, el análisis de la función de onda cuántica en el límite semiclásico revela que **a cada estado cuántico le corresponde en el espacio de las fases una determinada "celdilla" de volumen h^s , donde h es la constante de Planck**. Dicho de otro modo, existe una mínima incertidumbre con la que podemos determinar los valores (x, p) del estado de un sistema.

La primera consecuencia de esto es que si queremos calcular el número de microestados Ω_A que hay en una cierta región A del espacio fásico, tenemos que dividir su "volumen" $\Delta\Gamma_A$ (el volumen que ocupan en el espacio de $2s$ dimensiones formado por x y p) entre el volumen que "ocupa" cada microestado, h^s :

$$\Omega_A = \frac{\Delta\Gamma_A}{h^s} = \frac{1}{h^s} \int_A d^s x d^s p \quad (4.57)$$

Esta expresión es análoga a $(b - a)/\delta_0 x$ en la ecuación (4.56). La ecuación (4.57) no es correcta siempre, porque existen casos en que hay que aplicar otra corrección cuántica, que veremos en la siguiente sección. Por otro lado, si A es el conjunto de microestados compatibles con un cierto macroestado, la entropía de ese macroestado vendrá dada por la generalización de la ecuación (4.52):

$$S = \begin{cases} -k_B \int_A d^s x d^s p \cdot \rho(x, p) \cdot \ln[\rho(x, p)h^s] & , \text{ si } \Delta\Gamma_A > h^s \\ 0 & , \text{ si } \Delta\Gamma_A \leq h^s \end{cases} \quad (4.58)$$

Insistimos en que la entropía vale cero si $\Delta\Gamma_A \leq h^s$ porque en tal caso en la región A solo hay un microestado.

4.5.2. Partículas idénticas distinguibles e indistinguibles

Dos partículas se denominan **idénticas** si tienen las mismas propiedades intrínsecas, como la masa, la carga y el espín. En ciertos casos hay que aplicar una corrección más a la ecuación (4.58) debido a una propiedad específicamente cuántica de las partículas idénticas: la **indistinguibilidad**.

Para entender esta propiedad, supongamos un conjunto de dos partículas idénticas que están deslocalizadas, esto es, que no están "sujetas" a ninguna posición concreta del espacio. En Mecánica Clásica, las dos partículas tienen una trayectoria bien definida, con lo que tiene sentido "etiquetar" las partículas como "1" y "2" y seguirles la pista en su movimiento por el espacio. Sin embargo, en Mecánica Cuántica, el Principio de Incertidumbre hace que carezca de sentido hablar de trayectoria de una partícula. Esto quiere decir que, cuando estas partículas se mueven les "perdemos la pista" y, como son idénticas, no podemos estar seguros de cuál es cuál, con lo que no tiene sentido etiquetar las partículas como "1" y "2". Se dice entonces que **las partículas son indistinguibles**.

Cuando se tiene un sistema de partículas indistinguibles, sólo son posibles aquellos microestados (del sistema total) que no se alteran al permutarlas entre sí. Dicho de otro modo, solo importa cuántas partículas hay en cada estado (el llamado **número de ocupación del estado**), pero no qué partícula está en cada uno. Esto implica una **reducción drástica del número de microestados** posibles.

Un sistema de partículas idénticas e indistinguibles puede estar compuesto de dos tipos de partículas:

- Los **bosones** (partículas de espín entero), que pueden ocupar el mismo estado.

- Los **fermiones** (partículas de espín semientero), que cumplen el Principio de Exclusión de Pauli: no puede haber dos fermiones de un sistema en el mismo estado.

No hemos de confundir las expresiones “partículas idénticas” y “partículas indistinguibles”. **Varias partículas idénticas son distinguibles si están localizadas** (como los iones en una red cristalina), **y son indistinguibles si están deslocalizadas** (como las moléculas de un gas).

Por ejemplo, supongamos un sistema de partículas idénticas e independientes que pueden ocupar dos estados distintos: el estado “1” y el estado “2”. Supongamos que hay N_1 partículas en el estado 1 y N_2 partículas en el estado 2. ¿Cuántos microestados corresponden a este “macroestado”? Esta contabilización depende de si las partículas son distinguibles, bosones o fermiones.

- Si las partículas fueran **fermiones**, la única solución posible sería $N_1 = N_2 = 1$ (por el Principio de Exclusión) y solo habría un microestado.
- Si las partículas fueran **bosones**, cualquier valor de N_1 y N_2 estaría permitido y solo habría un microestado para unos valores dados de N_1 y N_2 (pues cualquier permutación de las partículas generaría el mismo microestado).
- Si las partículas fueran **distinguibles**, se considerarían iguales dos microestados que se diferencian en una permutación de dos partículas dentro del estado “1” o dentro del estado “2”, ya que las partículas son idénticas y su estado no cambia con la permutación. Pero si permutamos una partícula del 1 por otra del 2, el microestado del sistema sí que cambia. El número de microestados se obtiene, por tanto, como las permutaciones con repetición $\frac{(N_1+N_2)!}{N_1! N_2!}$.

La Mecánica Cuántica tiene en cuenta en su formalismo esta reducción drástica del número de microestados, pero en Mecánica Clásica no existe tal reducción. No obstante, se puede demostrar que, en el límite semiclásico, la siguiente receta es válida: **se hacen los cálculos como si las partículas fueran distinguibles y después se corrige el resultado dividiendo por el número total de permutaciones sin repetición ($N!$) de las N partículas que forman el sistema**. Así pues, en el límite semiclásico:

$$\Omega_A = \begin{cases} \frac{\Delta\Gamma_A}{h^s} = \frac{1}{h^s} \int_A d^s x d^s p & \text{partículas distinguibles} \\ \frac{\Delta\Gamma_A}{N! h^s} = \frac{1}{N! h^s} \int_A d^s x d^s p & \text{partículas indistinguibles} \end{cases} \quad (4.59)$$

Esta receta se traslada también a la función de partición de la siguiente manera:

$$Z = \begin{cases} \frac{1}{h^s} \int_A d^s x d^s p \cdot e^{-\beta E(x,p)} & \text{partículas distinguibles} \\ \frac{1}{N! h^s} \int_A d^s x d^s p \cdot e^{-\beta E(x,p)} & \text{partículas indistinguibles} \end{cases} \quad (4.60)$$

La densidad de probabilidad de un microestado en el conjunto canónico semiclásico se obtiene dividiendo la probabilidad de un microestado (4.9) entre su “volumen” en el espacio fásico, h^s :

$$\rho(x,p) = \frac{1}{h^s Z} e^{-\beta E(x,p)} \quad (4.61)$$

4.5.3. Influencia de la indistinguibilidad en la posibilidad de definir Z_1

Volvamos a la ecuación (4.30) que define la función de partición de una entidad elemental idéntica e independiente. Desde el punto de vista de la Mecánica Cuántica, dicha definición solo tiene sentido si las entidades elementales son distinguibles. Cuando éstas son indistinguibles (deslocalizadas), carece de

sentido hablar de Z_1 ya que, incluso aunque la energía sea aditiva, no tiene sentido hablar de una entidad elemental concreta porque no se la puede “etiquetar” ni seguir la pista.

Las fórmulas (4.60) permiten escribir una versión de la ecuación (4.31) que es válida para el caso semiclásico, incluso aunque las entidades elementales sean indistinguibles. Z_1 se calcula como si fueran distinguibles:

$$Z_1 = \frac{1}{h^q} \int_{\text{mono-estados}} d^q x d^q p e^{-\beta \varepsilon(x,p)} \quad (4.62)$$

y la Z del sistema total se obtiene como:

$$Z = \begin{cases} Z_1^N & \text{entidades elementales distinguibles} \\ \frac{Z_1^N}{N!} & \text{entidades elementales indistinguibles} \end{cases} \quad (4.63)$$

4.5.4. Validez del límite semiclásico

Una de las razones fundamentales que dieron lugar al descubrimiento de la Mecánica Cuántica es el hecho de que, a temperaturas suficientemente bajas, la Mecánica Clásica deja de ser válida. Y esta deficiencia persiste incluso con las correcciones cuánticas introducidas en el límite semiclásico.

Es una condición de estabilidad termodinámica que **la energía interna debe ser una función creciente de la temperatura** (no lo demostraremos aquí, pero se puede deducir a partir de las propiedades de la entropía). Cuando la temperatura disminuye, la energía también lo hace, por lo que el número de microestados disminuye (ya que, dicho toscamente, hay menos energía para repartir entre sus unidades estructurales y, por tanto, menos formas de repartirla). **En el caso semiclásico, existe una temperatura por debajo de la cual el volumen fásico de los microestados compatibles se vuelve más pequeño que h^s . Por debajo de esta temperatura, solo puede haber un microestado, la entropía es siempre cero y los resultados obtenidos no se corresponden con la realidad** (comprobaremos esto en la sección 5.2 con el ejemplo del gas ideal monoatómico semiclásico).

4.6. Ejemplo de cálculo de la ecuación fundamental

Para entender mejor las ecuaciones que hemos obtenido en los conjuntos microcanónico y canónico, calculemos el número de microestados y la función de partición para un modelo concreto y veamos las relaciones que hay entre estas dos “funciones estructura”.

Modelo: sistema de N entidades elementales **idénticas, independientes y distinguibles** que pueden distribuirse en dos monoestados cuánticos con energías respectivas $+\varepsilon$ y $-\varepsilon$. No se tiene en cuenta el volumen del sistema.

4.6.1. Análisis en el conjunto microcanónico

① **MACROESTADO:** (U, N) .

② **MODELO:** el ya indicado. Si en un macroestado determinado hay N_+ entidades elementales en el estado con energía $+\varepsilon$ y N_- entidades elementales en el estado con energía $-\varepsilon$, se tendrá:

$$\left\{ \begin{array}{l} N = N_+ + N_- \\ U = N_+(\varepsilon) + N_-(-\varepsilon) \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} N_+ = \frac{1}{2} \left(N + \frac{U}{\varepsilon} \right) \\ N_- = \frac{1}{2} \left(N - \frac{U}{\varepsilon} \right) \end{array} \right\} \quad (4.64)$$

③ **FUNCIÓN ESTRUCTURA: número de microestados.** Como las partículas son distinguibles, el número de microestados se calcula como:

$$\Omega(U, N) = \frac{N!}{(N_+)! (N_-)!} \quad (4.65)$$

Para calcular la entropía o, equivalentemente, el logaritmo del número de microestados, conviene hacer uso de la aproximación de Stirling para los factoriales. Tras operar, se obtiene:

$$\ln \Omega = -N_+ \ln \left(\frac{N_+}{N} \right) - N_- \ln \left(\frac{N_-}{N} \right) = -\frac{1}{2} \left(N + \frac{U}{\varepsilon} \right) \ln \left[\frac{1}{2} \left(1 + \frac{U}{N\varepsilon} \right) \right] - \frac{1}{2} \left(N - \frac{U}{\varepsilon} \right) \ln \left[\frac{1}{2} \left(1 - \frac{U}{N\varepsilon} \right) \right] \quad (4.66)$$

④ **CÁLCULOS:** haciendo las derivadas parciales correspondientes, se obtienen las ecuaciones de estado.

4.6.2. Análisis en el conjunto canónico

① **MACROESTADO:** (β, N) .

② **MODELO:** el ya indicado.

③ **FUNCIÓN ESTRUCTURA: función de partición.** Como las entidades elementales son distinguibles, podemos calcular la función de partición de una de ellas:

$$Z_1 = \sum_{\substack{i \\ \text{monoestados}}} e^{-\beta \varepsilon_i} = e^{-\beta \varepsilon} + e^{\beta \varepsilon} = 2 \cosh(\beta \varepsilon) \quad (4.67)$$

y la función de partición del sistema total será:

$$Z = Z_1^N = [2 \cosh(\beta \varepsilon)]^N \Rightarrow \ln Z = N \ln [2 \cosh(\beta \varepsilon)] \quad (4.68)$$

④ **CÁLCULOS:** haciendo las derivadas parciales correspondientes, se obtienen las ecuaciones de estado.

4.6.3. Obtención de Z a partir de Ω

Podemos utilizar la ecuación (4.25) para obtener Z como una transformada de Laplace de Ω :

$$Z(\beta, N) = \sum_{\substack{E \\ \text{niveles}}} \Omega(E, N) e^{-\beta E} \quad (4.69)$$

Los distintos niveles de energía del sistema son:

$$E = N_+ (+\varepsilon) + N_- (-\varepsilon) \quad , \quad \text{con } N_+ = 0, 1, \dots, N \quad (4.70)$$

donde $N_- = N - N_+$. Por tanto:

$$Z(\beta, N) = \sum_{N_+=0}^N \frac{N!}{(N_+)! (N_-)!} e^{-\beta [N_+ (+\varepsilon) + N_- (-\varepsilon)]} = \sum_{N_+=0}^N \frac{N!}{(N_+)! (N - N_+)!} (e^{-\beta \varepsilon})^{N_+} (e^{\beta \varepsilon})^{N - N_+} \quad (4.71)$$

Recordando la fórmula del binomio de Newton:

$$(a + b)^N = \sum_{k=0}^N \frac{N!}{k! (N - k)!} a^k b^{N-k} \quad (4.72)$$

Obtenemos:

$$Z = (e^{-\beta \varepsilon} + e^{\beta \varepsilon})^N = [2 \cosh(\beta \varepsilon)]^N \quad (4.73)$$

que coincide con el resultado hallado al resolver el problema directamente en el conjunto canónico.

4.7. Resumen: procedimiento de trabajo en Física Estadística

Como se ha puesto de manifiesto en el ejemplo de la sección 4.6, en todo problema de Física Estadística han de seguirse 4 etapas. Para el caso de un **sistema homogéneo y monocomponente que sólo intercambia trabajo de expansión**, son las siguientes:

1. Elección del **macroestado**, esto es, de las variables de estado termodinámicas con las que se va a trabajar. En esta asignatura hemos trabajado con 2 macroestados distintos:
 - a. (U, V, N) , correspondiente a las variables naturales de la **entropía S** , que es el potencial termodinámico correspondiente a un **sistema aislado** (con energía interna, volumen y número de partículas constantes).
 - b. $(1/T, V, N)$, o bien (β, V, N) , siendo $\beta = 1/(k_B T)$, las variables naturales de la **función de Massieu $\psi = S - (1/T)U$** , que es el potencial termodinámico correspondiente a un **sistema cerrado isotermo e isócoro** (con volumen y número de partículas constantes, pero cuya energía interna puede fluctuar). En este caso, el valor macroscópico de la energía interna, U , es igual a su valor medio, $\langle E \rangle$.
2. Especificación de un **modelo de realidad**, que permite formar el espacio muestral o conjunto de microestados (configuraciones mecánicas microscópicas) compatibles con el macroestado dado.
 - a. Para un **sistema aislado**, dicho espacio muestral se denomina **conjunto microcanónico** y consta de todos los microestados con una energía, volumen y número de partículas fijos.
 - b. Para un **sistema cerrado isotermo e isócoro**, el espacio muestral se denomina **conjunto canónico** y está formado por todos los microestados con un volumen y número de partículas fijos, pero cuya energía puede tomar valores entre su valor mínimo e infinito.

Por otro lado, hay que hacer distinción entre modelos cuánticos y clásicos (que obedecen, respectivamente, a la Mecánica Cuántica o Clásica). Para los conjuntos estudiados:

- a. En el **caso cuántico**, los estados se especifican mediante una simple enumeración de estados compatibles con el macroestado (microcanónico) o enumerando los valores $E_i(V, N)$ de la energía asociada a cada microestado compatible (canónico).
- b. En el **caso clásico o semiclásico**, deben especificarse todos los valores permitidos de las coordenadas $x = (x_1, \dots, x_s)$ y momentos $p = (p_1, \dots, p_s)$ que sean compatibles con el macroestado. El espacio formado por las x y las p recibe el nombre de espacio fásico.

Una vez especificado el modelo, conviene identificar las **entidades elementales independientes (EEI)** del sistema, que son las que garantizan la aditividad de la energía interna, entropía, etc. Encontrar las EEI consiste en buscar una forma de escribir la energía interna como una **suma de energías desacopladas** (que no dependan unas de otras a través de ningún grado de libertad) y **que puedan tomar idénticos valores**; entonces, se llama EEI a la entidad física cuya energía sea uno de esos sumandos.

Normalmente, en el modelo se especifica que el sistema está formado por N partículas (tales como átomos, moléculas, iones, etc). **Las EEI no tienen por qué coincidir con las partículas cuyo número estamos designando inicialmente por N** . Supongamos que el sistema consta de N_{EEI} de estas EEI. Frecuentemente, N_{EEI} es proporcional a N (ejemplo: cristal aislante monoatómico ideal), y no hay impedimento en seguir usando la variable N para cuantificar la cantidad de sustancia, siempre recordando, en caso de que no coincidan, que la aditividad de la energía se refiere a las EEI, y no necesariamente se tiene que cumplir para las partículas cuyo número es N .

3. Cálculo de la **función estructura asociada al conjunto escogido**.

Conjunto	Función estructura	Potencial termodinámico
Microcanónico	Número de microestados: $\Omega(U, V, N)$	$S = k_B \ln \Omega$
Canónico	Función de partición: $Z(\beta, V, N)$	$\psi = k_B \ln Z$

- **Número de microestados.** En el caso **cuántico**, se suele obtener por métodos combinatorios. En el caso **semiclásico**, debe calcularse una integral en el espacio fásico; no obstante, existen algunas sutilezas algo engorrosas relativas al conjunto microcanónico semiclásico y, por ello, en este curso no haremos cálculos de este estilo.

- **Función de partición.** Dado que el cálculo directo de la función de partición no suele poderse hacer cómodamente, lo normal es realizarlo a partir de la **función de partición de una de las EEI**, $Z_1(\beta, V)$.

- a. En el **caso cuántico**, los *monoestados* (microestados de una EEI) compatibles con el macroestado se especifican enumerando sus energías $\varepsilon_i(V)$. Si las EEI son **distinguidos (localizadas)**:

$$Z_1(\beta, V) = \sum_{\text{monoestados}} e^{-\beta\varepsilon_i(V)} \Rightarrow Z(\beta, V, N) = [Z_1(\beta, V)]^{N_{\text{EEI}}} \quad \text{EEI distinguibles} \quad (4.74)$$

Cuando las EEI son **indistinguibles (deslocalizadas)**, carece de sentido hablar de Z_1 , ya que, incluso aunque la energía sea aditiva, no tiene sentido hablar de una EEI concreta porque no se la puede “etiquetar” ni seguir la pista (ausencia de trayectoria definida en Mecánica Cuántica). Los microestados del sistema total se especifican por los números de ocupación de los monoestados, esto es, por cuántas EEI hay en cada monoestado. El problema hay que resolverlo en el conjunto grancanónico, cuyo macroestado está caracterizado por las variables $(1/T, V, -\mu/T)$.

- b. En el **caso semiclásico**, los monoestados se especifican con todos los valores permitidos de las posiciones $x = (x_1, \dots, x_q)$ y momentos $p = (p_1, \dots, p_q)$ que sean compatibles con el macroestado. Llamemos q al número de grados de libertad de la EEI. Si las EEI son **distinguidos (localizadas)**:

$$Z_1(\beta, V) = \frac{1}{h^q} \int_{\text{mono-estados}} d^q x d^q p e^{-\beta\varepsilon(V)} \Rightarrow Z(\beta, V, N) = [Z_1(\beta, V)]^{N_{\text{EEI}}} \quad \text{EEI distinguibles} \quad (4.75)$$

donde h es la constante de Planck. Si las partículas son **indistinguibles (deslocalizadas)**, estrictamente hablando existe el mismo problema que en el caso cuántico. No obstante, se puede demostrar (tomando la solución cuántica en el conjunto macrocanónico y calculando su límite semiclásico) que el siguiente procedimiento es válido: se calcula $Z_1(\beta, V)$ como si fueran distinguibles y se hace:

$$Z(\beta, V, N) = \frac{[Z_1(\beta, V)]^{N_{\text{EEI}}}}{N_{\text{EEI}}!} \quad \text{EEI indistinguibles} \quad (4.76)$$

4. **Cálculos** (ecuaciones de estado, probabilidades, coeficientes respuesta, fluctuaciones, etc.).

- a. En el **conjunto microcanónico**, las ecuaciones de estado se deducen de la diferencial de la entropía, consecuencia del Primer y Segundo Principios de la Termodinámica:

$$dS = \frac{1}{T}dU + \frac{p}{T}dV - \frac{\mu}{T}dN = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \quad (4.77)$$

que permite identificar:

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial U}\right)_{V,N} \Leftrightarrow \beta = \frac{1}{k_B T} = \left(\frac{\partial \ln \Omega}{\partial U}\right)_{V,N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{T} &= \left(\frac{\partial S}{\partial V}\right)_{U,N} \Leftrightarrow p\beta = \frac{p}{k_B T} = \left(\frac{\partial \ln \Omega}{\partial V}\right)_{U,N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{T} &= \left(\frac{\partial S}{\partial N}\right)_{U,V} \Leftrightarrow -\mu\beta = -\frac{\mu}{k_B T} = \left(\frac{\partial \ln \Omega}{\partial N}\right)_{U,V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (4.78)$$

En el conjunto microcanónico, todos los **microestados del sistema son igualmente probables**, con probabilidad $1/\Omega$.

Sin embargo, ¡los **monoestados de una EEI no son igualmente probables!** Esta aparente contradicción se debe a que la EEI **no se comporta como aislada**, ya que está en interacción con el resto del sistema (que, a estos efectos, hace el papel de "termostato"). En otras palabras, la información de la que disponemos para calcular las probabilidades de los monoestados es la energía promedio por EEI. Si las EEI están **deslocalizadas** y el volumen accesible a cada una es igual al volumen total (V fijo), se obtiene una distribución de probabilidad de tipo canónico:

$$\begin{aligned} p_i &= \frac{e^{-\beta \varepsilon_i}}{Z_1} && \text{Prob. monoestado,} \\ &&& \text{EEI deslocalizada en microcanónico} \\ \rho(x, p) &= \frac{1}{h^q Z_1} e^{-\beta \varepsilon(x, p)} && \text{Densidad de prob. monoestado,} \\ &&& \text{EEI deslocalizada en microcanónico} \end{aligned} \quad (4.79)$$

Si las EEI están **localizadas**, estas distribuciones pueden no ser válidas.

- b. En el **conjunto canónico**, las ecuaciones de estado se deducen de la diferencial de la función de Massieu:

$$d\psi = -Ud\left(\frac{1}{T}\right) + \frac{p}{T}dV - \frac{\mu}{T}dN = \left(\frac{\partial \psi}{\partial [1/T]}\right)_{V,N} dU + \left(\frac{\partial \psi}{\partial V}\right)_{\frac{1}{T},N} dV + \left(\frac{\partial \psi}{\partial N}\right)_{\frac{1}{T},V} dN \quad (4.80)$$

de donde:

$$\begin{aligned} -U &= \left(\frac{\partial \psi}{\partial [1/T]}\right)_{V,N} \Leftrightarrow -U = \left(\frac{\partial \ln Z}{\partial \beta}\right)_{V,N} && \text{Ecuación Energética de Estado (EEE)} \\ \frac{p}{T} &= \left(\frac{\partial \psi}{\partial V}\right)_{\frac{1}{T},N} \Leftrightarrow p\beta = \frac{p}{k_B T} = \left(\frac{\partial \ln Z}{\partial V}\right)_{\beta,N} && \text{Ecuación Térmica de Estado (ETE)} \\ -\frac{\mu}{T} &= \left(\frac{\partial \psi}{\partial N}\right)_{\frac{1}{T},V} \Leftrightarrow -\mu\beta = -\frac{\mu}{k_B T} = \left(\frac{\partial \ln Z}{\partial N}\right)_{\beta,V} && \text{Ecuación Material de Estado (EME)} \end{aligned} \quad (4.81)$$

La entropía del sistema puede obtenerse como:

$$S = \psi + \frac{1}{T}U \Leftrightarrow \frac{S}{k_B} = \ln Z + \beta U \quad (4.82)$$

En el conjunto canónico, la distribución de probabilidad de **los microestados del sistema** está caracterizada por:

$$\begin{aligned}
 p_i &= \frac{e^{-\beta E_i}}{Z} && \text{Probabilidad microestado, conj. canónico} \\
 \rho(x, p) &= \frac{1}{h^s Z} e^{-\beta E(x, p)} && \text{Densidad de prob. microestado, conj. canónico}
 \end{aligned}
 \tag{4.83}$$

De modo análogo a lo que ocurría en el conjunto microcanónico, si las EEI están **deslocalizadas** la **distribución de probabilidad de monoestados de una EEI** es de tipo canónico:

$$\begin{aligned}
 p_i &= \frac{e^{-\beta \varepsilon_i}}{Z_1} && \text{Prob. monoestado,} \\
 &&& \text{EEI deslocalizada en conj. canónico} \\
 \rho(x, p) &= \frac{1}{h^q Z_1} e^{-\beta \varepsilon(x, p)} && \text{Densidad de prob. monoestado,} \\
 &&& \text{EEI deslocalizada en conj. canónico}
 \end{aligned}
 \tag{4.84}$$

Si las EEI están **localizadas**, estas distribuciones pueden no ser válidas.

Una vez obtenida toda esta información, se puede seguir con los cálculos obteniendo, a través de las segundas derivadas del potencial termodinámico, las capacidades caloríficas, coeficientes de dilatación, coeficientes de compresibilidad, calores latentes, fluctuaciones de energía, etc. En este curso no ahondaremos en los cálculos que involucran las segundas derivadas, que serán vistos con más detalle en la asignatura de Termodinámica (2° curso) y Física Estadística (3er curso).

Ejercicios – Formalismo de la Física Estadística

1.- En la sección 4.6 se ha introducido el siguiente modelo (y se han calculado las funciones estructura en el conjunto microcanónico y en el canónico): un sistema de N entidades elementales idénticas, independientes y distinguibles que pueden distribuirse en dos monoestados cuánticos con energías respectivas $+\varepsilon$ y $-\varepsilon$. No se tiene en cuenta el volumen del sistema. Calcule las ecuaciones de estado asociadas a dicho modelo:

- a) En el conjunto microcanónico. Demuestre la existencia de macroestados con **temperaturas termodinámicas negativas**, representando $\beta\varepsilon = \varepsilon/k_B T$ en función de $U/N\varepsilon$.
- b) En el conjunto canónico. Compruebe que tanto las ecuaciones de estado como la entropía coinciden con las calculadas en el conjunto microcanónico.

Nota: Las temperaturas negativas aparecen en ciertos sistemas cuya energía interna está acotada superiormente, como ocurre en este caso. Esto solo es posible en modelos en los que no se considere la energía cinética de las partículas del sistema.

2.- Vamos a estudiar las vibraciones de los átomos de la red cristalina de un sólido aislante monoatómico cuántico según un modelo desarrollado por Einstein (1907), en el **conjunto canónico (macroestado (β, V, N))**. Trabajaremos con la **aproximación armónica**, es decir, supondremos que las vibraciones pueden describirse como un conjunto de N átomos efectuando oscilaciones armónicas en las tres direcciones del espacio, por supuesto, acopladas entre sí: las oscilaciones de un átomo en una dirección dependen de sus oscilaciones en las otras dos direcciones y también de las oscilaciones de los átomos vecinos. Es posible demostrar (cosa que se verá en la asignatura “Mecánica y Ondas” de 2º curso) que se puede hacer un cambio de coordenadas para encontrar un **conjunto equivalente de $3N$ osciladores armónicos unidimensionales desacoplados**, denominados **modos normales de vibración**.

De acuerdo con el modelo de Einstein, se supone que los modos normales vibran con una única frecuencia f y sus energías pueden tomar los valores:

$$\varepsilon_n = \left(n + \frac{1}{2}\right)hf \quad , \quad \text{para } n = 0,1,2,\dots$$

siendo h la constante de Planck. Esta aproximación no da cuenta del volumen del sistema (Ecuación Térmica de Estado), por lo que no aparece como variable en el macroestado. En consecuencia, el macroestado se simplifica en la forma **(β, N)** .

a) Calcule la función de partición. Para ello:

- Encuentre la entidad elemental independiente (EEI) del sistema, justificando la respuesta a partir de la definición de EEI, e indique claramente: su energía, el número de ellas que hay en el sistema (N_{EEI}) y si son distinguibles o indistinguibles.
- Identifique los valores de los monoestados de la EEI que son compatibles con su macroestado (β) . Calcule la función de partición de una EEI, $Z_1(\beta)$ y, finalmente, la función de partición del sistema, $Z(\beta, N)$.

b) Encuentre las ecuaciones de estado y la entropía del sistema.

c) Determine la capacidad calorífica a volumen constante del sistema (que es una magnitud extensiva), demostrando que puede escribirse mediante la función de Einstein, $\epsilon(x)$, como:

$$C_V(\beta, N) = 3Nk_B \cdot \epsilon(\beta hf) \quad , \quad \epsilon(x) = \frac{x^2 e^x}{(e^x - 1)^2}$$

d) Halle la capacidad calorífica a volumen constante molar, C_{V_m} , que es una magnitud intensiva, y verifique que en el límite de altas temperaturas se cumple la **ley de Dulong y Petit**:

$$C_{V_m} = \frac{C_V}{n} = 3R$$

siendo n la cantidad de átomos (en mol) y R la constante de los gases.

5. Gas ideal monoatómico.

Distribución de Maxwell

5.1. Modelo microscópico del gas ideal monoatómico semiclásico

El modelo de gas objeto de estudio obedece a la Mecánica Clásica y tiene las siguientes características:

- Hay N moléculas (entidades elementales independientes) deslocalizadas (indistinguibles), cada una de masa m .
- Ideal: $(E_p)_{\text{intermolecular}} = 0$.
- Monoatómico: $(E_p)_{\text{intramolecular}} = 0$.
- No sometido a campo externo: $(E_p)_{\text{campo externo}} = 0$.
- Además, las paredes del recipiente se encargan de que no escapen moléculas al exterior y las confinan en un recinto de volumen V .

Así, la energía microscópica del gas será:

$$E = \sum_{i=1}^N \varepsilon_i \quad (5.1)$$

donde, debido a las hipótesis del modelo, la energía de una partícula, ε , solo tiene contribución cinética de traslación:

$$\varepsilon = \frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \quad (5.2)$$

5.2. Solución en el conjunto canónico - macroestado (β, V, N)

La función de partición de una molécula es:

$$Z_1 = \frac{1}{h^3} \int_{\substack{\text{estados} \\ \text{compatibles} \\ \text{con } (\beta, V)}} dx dy dz dp_x dp_y dp_z e^{-\frac{\beta}{2m}(p_x^2 + p_y^2 + p_z^2)} \quad (5.3)$$

Aplicando el Teorema de Fubini:

$$Z_1 = \frac{1}{h^3} \left(\int_{\substack{\text{recinto de} \\ \text{volumen } V}} dx dy dz \right) \left(\int_{-\infty}^{\infty} dp_x e^{-\frac{\beta}{2m}p_x^2} \right) \left(\int_{-\infty}^{\infty} dp_y e^{-\frac{\beta}{2m}p_y^2} \right) \left(\int_{-\infty}^{\infty} dp_z e^{-\frac{\beta}{2m}p_z^2} \right) \quad (5.4)$$

La primera integral es igual al volumen del recinto, V . Las tres últimas integrales entre paréntesis son idénticas, por lo que podemos escribir:

$$Z_1 = \frac{1}{h^3} V \left(\int_{-\infty}^{\infty} dp_x e^{-\frac{\beta}{2m}p_x^2} \right)^3 \quad (5.5)$$

Según la ecuación (5.57) (véase el apéndice del tema), esta integral se puede hacer exactamente y el resultado es:

$$Z_1 = \frac{1}{h^3} V \left(\sqrt{\frac{2\pi m}{\beta}} \right)^3 = V \left(\frac{2\pi m}{h^2 \beta} \right)^{3/2} \quad (5.6)$$

Como las moléculas están deslocalizadas y son indistinguibles, la función de partición del gas total será:

$$Z = \frac{Z_1^N}{N!} = \frac{1}{N!} V^N \left(\frac{2\pi m}{h^2 \beta} \right)^{3N/2} \quad (5.7)$$

Y su logaritmo es, aplicando la aproximación de Stirling:

$$\frac{\psi}{k_B} = \ln Z = N \ln Z_1 - \ln(N!) \approx N \ln Z_1 - N \ln N + N = N \ln \left(\frac{Z_1}{N} \right) + N \quad (5.8)$$

donde ψ es la función de Massieu. O bien:

$$\frac{\psi}{k_B} = \ln Z = N \ln \left[\frac{V}{N} \left(\frac{2\pi m}{h^2 \beta} \right)^{3/2} \right] + N = N \ln V - \frac{3N}{2} \ln \beta + \frac{3N}{2} \ln \left(\frac{2\pi m}{h^2} \right) - N \ln N + N \quad (5.9)$$

Ecuación Energética de Estado (EEE):

$$U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V,N} = \frac{3N}{2\beta} \Rightarrow U = \frac{3}{2} N k_B T \quad (5.10)$$

Ecuación Térmica de Estado (ETE):

$$\frac{p}{k_B T} = p\beta = \left(\frac{\partial \ln Z}{\partial V} \right)_{\beta,N} = \frac{N}{V} \Rightarrow pV = N k_B T \quad (5.11)$$

Ecuación Material de Estado (EME):

$$-\frac{\mu}{k_B T} = -\mu\beta = \left(\frac{\partial \ln Z}{\partial N} \right)_{\beta,V} = \ln \left(\frac{Z_1}{N} \right) \Rightarrow -\frac{\mu}{k_B T} = \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] \quad (5.12)$$

Capacidad calorífica a volumen constante:

$$C_V = \left(\frac{\delta Q}{dT} \right)_{V,N} = \left(\frac{\partial U}{\partial T} \right)_{V,N} = \frac{3}{2} N k_B \quad (5.13)$$

La entropía se puede calcular a través de la relación con la función de Massieu:

$$S = \psi + \frac{U}{T} \Leftrightarrow \frac{S}{k_B} = \frac{\psi}{k_B} + \frac{U}{k_B T} \Leftrightarrow \frac{S}{k_B} = \ln Z + \beta U \quad (5.14)$$

Es decir:

$$\frac{S}{k_B} = N \ln \left[\frac{V}{N} \left(\frac{2\pi m}{h^2 \beta} \right)^{3/2} \right] + N + \frac{3N}{2} \Rightarrow S = k_B N \left\{ \frac{5}{2} + \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] \right\} \quad (5.15)$$

La entropía se anula a una temperatura T^* , por debajo de la cual el modelo deja de ser válido (pues nos daría entropías negativas). Se halla igualando la entropía a cero en la ecuación (5.15):

$$0 = \frac{5}{2} + \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T^*}{h^2} \right)^{3/2} \right] \Rightarrow T^* = \left(\frac{N}{V} \right)^{2/3} \frac{h^2 e^{-5/3}}{2\pi m k_B} \quad (5.16)$$

Por debajo de esta temperatura, el modelo semiclásico debe ser sustituido por el cuántico.

5.3. Probabilidades de los estados de una molécula

5.3.1. Densidad de probabilidad de momentos y velocidades de una molécula

De acuerdo con las ecuaciones generales del conjunto canónico, la densidad de probabilidad de una configuración de la molécula dada por (x, y, z, p_x, p_y, p_z) viene dada por:

$$\begin{aligned} D_{x,p}(x, y, z, p_x, p_y, p_z) &= \frac{1}{h^3 Z_1} e^{-\frac{\beta}{2m}(p_x^2 + p_y^2 + p_z^2)} = \frac{1}{V} \left(\frac{\beta}{2\pi m} \right)^{3/2} e^{-\frac{p_x^2 + p_y^2 + p_z^2}{2mk_B T}} \\ &= \frac{1}{V} \left(\frac{1}{2\pi mk_B T} \right)^{3/2} e^{-\frac{p_x^2 + p_y^2 + p_z^2}{2mk_B T}} \end{aligned} \quad (5.17)$$

Para hallar la densidad de probabilidad de una configuración de momentos $D_p(p_x, p_y, p_z)$, independientemente de cuáles sean las coordenadas (x, y, z) , hay que sumar (integrar) a todos los posibles valores de las coordenadas (x, y, z) . Teniendo en cuenta que $D_{x,p}(x, y, z, p_x, p_y, p_z)$ no depende de (x, y, z) :

$$\begin{aligned} D_p(p_x, p_y, p_z) &= \frac{1}{V} \left(\frac{1}{2\pi mk_B T} \right)^{3/2} e^{-\frac{\beta}{2m}(p_x^2 + p_y^2 + p_z^2)} \int_{\text{recinto de volumen } V} dx dy dz \\ &= \left(\frac{1}{2\pi mk_B T} \right)^{3/2} e^{-\frac{p_x^2 + p_y^2 + p_z^2}{2mk_B T}} \end{aligned} \quad (5.18)$$

Esta es la densidad de probabilidad de los momentos (p_x, p_y, p_z) . Realizando una integración doble a dos de las componentes del momento en la ecuación (5.18), por ejemplo, p_y y p_z , se obtiene que la densidad de probabilidad asociada a la otra componente del momento, p_x , es la función:

$$D_{p_x}(p_x) = \left(\frac{1}{2\pi mk_B T} \right)^{1/2} e^{-\frac{p_x^2}{2mk_B T}} \quad (5.19)$$

que es una gaussiana. Esta función es par (simétrica respecto al eje vertical). Sin embargo, p_x es función impar, por lo que el producto $p_x D_{p_x}(p_x)$ es también una función impar. Por tanto, la integral que hay que realizar para calcular el valor medio de p_x es cero:

$$\langle p_x \rangle = \int_{-\infty}^{\infty} p_x D_{p_x}(p_x) dp_x = 0 \quad (5.20)$$

(esto se puede comprobar también haciendo explícitamente el cálculo). Debido a la isotropía del gas, **la densidad de probabilidad dada por la ecuación (5.19) es aplicable a la componente del momento en cualquier otra dirección del espacio. Por ello, el valor medio de la componente del momento en cualquier dirección debe ser también nulo.**

Para calcular la probabilidad de que p_x se encuentre en un intervalo $[p_{1x}, p_{2x}]$ hay que calcular la integral de $D_{p_x}(p_x)$ en ese intervalo:

$$\Pr(p_{1x} \leq p_x \leq p_{2x}) = \int_{p_{1x}}^{p_{2x}} D_{p_x}(p_x) dp_x \quad (5.21)$$

Lo que hay que hacer para hallar la función $D_{v_x}(v_x)$ de densidad de probabilidad de la componente x de la velocidad es hacer un cambio de variable, $p_x = mv_x$:

$$p_x = mv_x \Rightarrow \frac{dp_x}{dv_x} = m \quad (5.22)$$

Con lo que en la ecuación (5.21) se transforma en:

$$\Pr(v_{1x} \leq v \leq v_{2x}) = \int_{v_{1x}}^{v_{2x}} D_{p_x}(mv_x) \frac{dp_x}{dv_x} dv_x = \int_{v_{1x}}^{v_{2x}} D_{p_x}(mv_x) m dv_x \quad (5.23)$$

donde $p_{1x} = mv_{1x}$ y $p_{2x} = mv_{2x}$. Dado que esto se cumple para cualquier intervalo de valores de v_x , podemos identificar el integrando de la ecuación (5.23) como la **FDP de v_x** . La denotaremos por $D_{v_x}(v_x)$:

$$D_{v_x}(v_x) = D_{p_x}(mv_x) \frac{dp_x}{dv_x} = D_{p_x}(mv_x) \cdot m = \left(\frac{1}{2\pi m k_B T} \right)^{1/2} e^{-\frac{mv_x^2}{2k_B T}} \cdot m \quad (5.24)$$

Simplificando, queda finalmente:

$$\boxed{D_{v_x}(v_x) = D_{p_x}(mv_x) \frac{dp_x}{dv_x} = \left(\frac{m}{2\pi k_B T} \right)^{1/2} e^{-\frac{mv_x^2}{2k_B T}}} \quad (5.25)$$

que es otra gaussiana. De la misma manera que ocurre con p_x :

$$\boxed{\langle v_x \rangle = \int_{-\infty}^{\infty} v_x D_{v_x}(v_x) dv_x = 0} \quad (5.26)$$

Además, la FDP dada por la ecuación (5.25) es aplicable a la componente de la velocidad en cualquier otra dirección del espacio. Por ello, el valor medio de la componente de la velocidad en cualquier dirección debe ser también nulo.

5.3.2. Densidad de probabilidad del módulo del momento y de la velocidad

Notemos que $D_p(p_x, p_y, p_z)$ sólo depende del módulo del momento y no de su dirección (como era de esperar, ya que el gas es isótropo). Podemos indicar esta densidad como una función $f_p(p)$:

$$D_p(p_x, p_y, p_z) = f_p(p) = \left(\frac{1}{2\pi m k_B T} \right)^{3/2} e^{-\frac{p^2}{2m k_B T}} \quad (5.27)$$

De igual forma que se hizo con la densidad de probabilidad de las velocidades $D(v_x, v_y, v_z) = f(v)$ para hallar la densidad de probabilidad del **módulo** de la velocidad $F(v)$ (véase la sección 2.4.1):

$$F(v) = 4\pi v^2 f(v) \quad (5.28)$$

simplemente tendremos que multiplicar por $4\pi p^2$ para obtener la correspondiente **función de densidad de probabilidad del módulo del momento**, $F_p(p)$:

$$\boxed{F_p(p) = 4\pi p^2 f_p(p) = 4\pi p^2 \left(\frac{1}{2\pi m k_B T} \right)^{3/2} e^{-\frac{p^2}{2m k_B T}}} \quad (5.29)$$

Con un procedimiento análogo al seguido para obtener $D_{v_x}(v_x)$ a partir de $D_{p_x}(p_x)$, se obtiene que la función $F(v)$ de la densidad de probabilidad del módulo de la velocidad es:

$$F(v) = F_p(mv) \frac{dp}{dv} = F_p(mv) \cdot m = 4\pi (mv)^2 \left(\frac{1}{2\pi m k_B T} \right)^{3/2} e^{-\frac{(mv)^2}{2m k_B T}} \cdot m \quad (5.30)$$

Simplificando, queda finalmente:

$$\boxed{F(v) = \frac{4}{\sqrt{\pi}} \left(\frac{m}{2k_B T} \right)^{3/2} v^2 e^{-\frac{mv^2}{2k_B T}}} \quad (5.31)$$

La distribución de probabilidad representada por la densidad $F(v)$ se denomina **distribución de Maxwell**. Su forma cualitativa puede verse en la Figura 5.1.

Por brevedad, llamaremos **celeridad** al módulo de la velocidad.

5.3.3. Celeridad más probable

La celeridad a la cual $F(v)$ es máxima se denomina **celeridad más probable** y la denotaremos por v_p . Para calcularla, utilizamos las condiciones:

$$F'(v_p) = 0$$

$$F''(v_p) < 0$$

Derivando $F(v)$:

$$F'(v) = \frac{8}{\sqrt{\pi}} \left(\frac{m}{2k_B T} \right)^{3/2} e^{-\frac{mv^2}{2k_B T}} \left(v - \frac{mv^3}{2k_B T} \right) \quad (5.32)$$

Omitiendo factores constantes, la condición $F'(v_p) = 0$ equivale a:

$$v_p e^{-\frac{mv_p^2}{2k_B T}} \left(1 - \frac{mv_p^2}{2k_B T} \right) = 0 \quad (5.33)$$

Dado que $F(0) = 0$ no es un máximo y que la exponencial nunca se anula, la única solución posible para el máximo es que el paréntesis se anule:

$$v_p = \sqrt{\frac{2k_B T}{m}} = \sqrt{\frac{2RT}{M}} \quad (5.34)$$

donde M es la masa molar del gas. Falta comprobar que efectivamente se trata de un máximo con la condición de la derivada segunda. Antes, reescribimos $F(v)$ en una forma más abreviada introduciendo el valor de v_p :

$$F(v) = \frac{4}{v_p^3 \sqrt{\pi}} v^2 e^{-v^2/v_p^2} \quad (5.35)$$

Con esta notación, la primera y segunda derivadas de $F(v)$ son:

$$F'(v) = \frac{8}{v_p^3 \sqrt{\pi}} e^{-v^2/v_p^2} \left(v - \frac{v^3}{v_p^2} \right) \quad (5.36)$$

$$F''(v) = \frac{8}{v_p^3 \sqrt{\pi}} e^{-v^2/v_p^2} \left(1 - 5 \frac{v^2}{v_p^2} + 2 \frac{v^4}{v_p^4} \right) \quad (5.37)$$

Podemos ver que, en efecto, se trata de un máximo:

$$F''(v_p) = \frac{8}{v_p^3 \sqrt{\pi}} e^{-1} (1 - 5 + 2) = -\frac{16}{e v_p^3 \sqrt{\pi}} < 0 \quad (5.38)$$

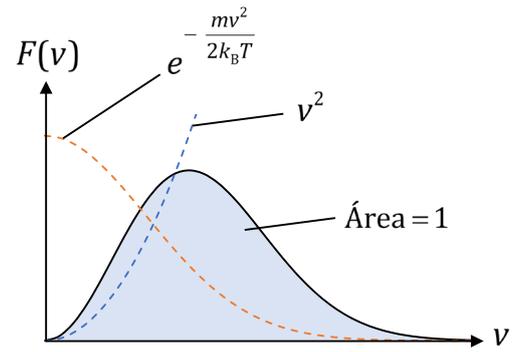


Figura 5.1: Densidad de probabilidad marginal del módulo de la velocidad.

El valor del **máximo de la densidad de probabilidad** es:

$$F(v_p) = \frac{4}{v_p e\sqrt{\pi}} = \frac{4}{e\sqrt{\pi}} \sqrt{\frac{m}{2k_B T}} \tag{5.39}$$

A partir de las ecuaciones (5.29) y (5.34) pueden deducirse cualitativamente ciertas características de la forma de la curva $F(v)$. Las coordenadas del máximo dependen de m y de T en la forma:

$$v_p \propto \sqrt{\frac{T}{m}} \tag{5.40}$$

$$F(v_p) \propto \sqrt{\frac{m}{T}} \tag{5.41}$$

En consecuencia, el efecto en el máximo de aumentar la temperatura es equivalente al de disminuir la masa de las moléculas (Figura 5.2).

- Para un mismo gas (misma m), al aumentar T el máximo disminuye y se desplaza hacia la derecha.
- A una misma T , al aumentar m el máximo aumenta y se desplaza hacia la izquierda.

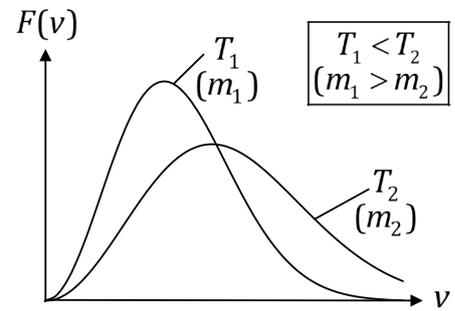


Figura 5.2: Variación de la forma de la curva de $F(v)$ para distintas temperaturas o masas de las moléculas.

En cualquier caso, como ya sabemos, el área total bajo la curva de $F(v)$ tiene que ser igual a la unidad siempre (condición de normalización).

En la Tabla 5.1 se recogen las probabilidades de ciertos intervalos de la celeridad. Como puede verse, una mayoría de las moléculas tienen velocidades en un pequeño entorno de la velocidad más probable (alrededor del máximo de $F(v)$).

Tabla 5.1: Probabilidad de ciertos intervalos de celeridad de una molécula.

a	b	$p(a \leq v/v_p \leq b)$
0	0.5	8.1 %
0.5	1.5	70.7 %
1.5	2	16.6 %
2	3	4.6 %
3	∞	0.04 %
5	∞	$8 \cdot 10^{-9}$ %

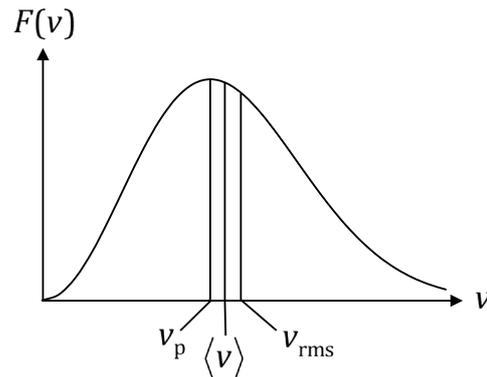


Figura 5.3: Relación entre la celeridad más probable, la celeridad media y la celeridad cuadrática media.

5.3.4. Celeridad media y celeridad cuadrática media

La **celeridad media** se obtiene de:

$$\begin{aligned}\langle v \rangle &= \int_0^{\infty} vF(v)dv = \frac{4}{v_p^3\sqrt{\pi}} \int_0^{\infty} v^3 e^{-v^2/v_p^2} dv = \frac{4}{v_p^3\sqrt{\pi}} \cdot I_3 \left(\frac{1}{v_p^2} \right) = \\ &= \frac{4}{v_p^3\sqrt{\pi}} \frac{1!}{2(1/v_p^2)^{1+1}} = \frac{4}{v_p^3\sqrt{\pi}} \frac{v_p^4}{2} = \frac{2}{\sqrt{\pi}} v_p\end{aligned}\quad (5.42)$$

(donde se ha usado la ecuación (5.59)). Sustituyendo el valor de la celeridad más probable:

$$\boxed{\langle v \rangle = \frac{2}{\sqrt{\pi}} v_p = \sqrt{\frac{8k_B T}{\pi m}} = \sqrt{\frac{8RT}{\pi M}}}\quad (5.43)$$

Realizando un cálculo similar se obtiene el valor medio del cuadrado de la celeridad como:

$$\begin{aligned}\langle v^2 \rangle &= \int_0^{\infty} v^2 F(v) dv = \frac{4}{v_p^3\sqrt{\pi}} \int_0^{\infty} v^4 e^{-v^2/v_p^2} dv = \frac{4}{v_p^3\sqrt{\pi}} \cdot I_4 \left(\frac{1}{v_p^2} \right) = \\ &= \frac{4}{v_p^3\sqrt{\pi}} \frac{1 \cdot 3}{2^{2+1}} \sqrt{\frac{\pi}{(1/v_p^2)^5}} = \frac{3}{2} v_p^2\end{aligned}\quad (5.44)$$

O bien, llamando $v_{\text{rms}} = \sqrt{\langle v^2 \rangle}$ (**celeridad cuadrática media o root-mean-square = rms**)

$$\boxed{v_{\text{rms}} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3}{2}} v_p = \sqrt{\frac{3k_B T}{m}} = \sqrt{\frac{3RT}{M}}}\quad (5.45)$$

La proporción entre la celeridad más probable, la celeridad media y la celeridad cuadrática media es una constante:

$$\begin{aligned}v_{\text{rms}} &= \sqrt{\frac{3\pi}{8}} \langle v \rangle = \sqrt{\frac{3\pi}{8}} \left(\frac{2}{\sqrt{\pi}} v_p \right) \\ v_{\text{rms}} &= \sqrt{\frac{3\pi}{8}} \langle v \rangle = \sqrt{\frac{3}{2}} v_p \\ v_{\text{rms}} &\approx 1.085 \langle v \rangle \approx 1.225 v_p\end{aligned}\quad (5.46)$$

La localización de estas magnitudes en la representación gráfica de $F(v)$ puede verse en la Figura 5.3.

5.3.5. Celeridad reducida

Puede ser conveniente (para efectuar ciertas integrales) trabajar con la **celeridad reducida** v_r :

$$\boxed{v_r = \frac{v}{v_p}}\quad (5.47)$$

Veamos cómo podemos expresar la densidad de probabilidad $F(v)$ en términos de la celeridad reducida. Para ello, tenemos que efectuar un cambio de variable.

Para calcular la probabilidad de que la celeridad se encuentre en un intervalo $[v_1, v_2]$, hay que calcular la integral de $F(v)$ en ese intervalo:

$$\text{Pr}(v_1 \leq v \leq v_2) = \int_{v_1}^{v_2} F(v) dv\quad (5.48)$$

Dado que existe una relación biunívoca entre v_r y v se puede hacer un cambio de variable:

$$v = v_r v_p \Rightarrow \frac{dv}{dv_r} = v_p \quad (5.49)$$

Con lo que la ecuación (5.43) se transforma en:

$$\Pr(v_{r1} \leq v_r \leq v_{r2}) = \int_{v_{r1}}^{v_{r2}} F(v) \frac{dv}{dv_r} dv_r = \int_{v_{r1}}^{v_{r2}} F(v_r v_p) v_p dv_r \quad (5.50)$$

donde $v_1 = v_{r1} v_p$ y $v_2 = v_{r2} v_p$. Dado que esto se cumple para cualquier intervalo de valores de la celeridad, podemos identificar el integrando de la ecuación (5.45) como la **función de densidad de probabilidad de la celeridad reducida**. La designaremos por $F_r(v_r)$:

$$\boxed{F_r(v_r) = F(v) \frac{dv}{dv_r} = F(v_r v_p) v_p = \frac{4}{\sqrt{\pi}} v_r^2 e^{-v_r^2}} \quad (5.51)$$

5.4. Complemento: integrales y funciones especiales

5.4.1. Integral de Euler-Poisson

Se denomina **integral de Euler-Poisson** a la siguiente integral impropia dependiente de un parámetro positivo ($b > 0$):

$$\int_{-\infty}^{\infty} e^{-bx^2} dx \quad \text{con } b > 0 \quad (5.52)$$

Se puede demostrar que esta integral es absolutamente convergente. Aunque el integrando e^{-bx^2} no tiene una primitiva que pueda ser escrita en términos de funciones elementales, existen métodos para evaluar la integral exactamente. Uno muy sencillo consiste en trasladar el cálculo al dominio del análisis de dos variables, aplicando el Teorema de Fubini:

$$\left(\int_{-\infty}^{\infty} e^{-bx^2} dx \right)^2 = \left(\int_{-\infty}^{\infty} e^{-bx^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-by^2} dy \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-b(x^2+y^2)} dx dy \quad (5.53)$$

Ahora podemos hacer un cambio a coordenadas polares en el plano:

$$\begin{aligned} x &= r \cos \phi \\ y &= r \sin \phi \end{aligned} \quad (5.54)$$

con $r \in [0, \infty)$ y $\phi \in [0, 2\pi)$. El "factor de escala" del elemento de superficie en este sistema de coordenadas es igual a r . Es decir:

$$dx dy = r dr d\phi \quad (5.55)$$

Por lo tanto:

$$\left(\int_{-\infty}^{\infty} e^{-bx^2} dx \right)^2 = \int_0^{2\pi} d\phi \int_0^{\infty} dr \cdot r e^{-br^2} = 2\pi \left[-\frac{e^{-r^2}}{2b} \right]_{r=0}^{r \rightarrow \infty} = \frac{\pi}{b} \quad (5.56)$$

de donde deducimos el valor de la integral buscada:

$$\boxed{\int_{-\infty}^{\infty} e^{-bx^2} dx = \sqrt{\frac{\pi}{b}}} \quad (5.57)$$

Como e^{-bx^2} es una función par (es decir, simétrica respecto del eje de ordenadas), las contribuciones a la integral en los intervalos $(-\infty, 0]$ y $[0, \infty)$ son iguales. En consecuencia:

$$\boxed{\int_{-\infty}^0 e^{-bx^2} dx = \int_0^{\infty} e^{-bx^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{b}}} \quad (5.58)$$

5.4.2. Integrales de la forma $\int_0^{\infty} x^n e^{-bx^2} dx$

Definamos como $I_n(b)$ las integrales del tipo:

$$I_n(b) = \int_0^{\infty} x^n e^{-bx^2} dx \quad \text{donde } n \in \mathbb{Z}, n \geq 0, b > 0 \quad (5.59)$$

Podemos ver que el integrando tiene derivada continua respecto del parámetro b , por lo que se puede aplicar la regla de Leibniz de derivación de integrales paramétricas y:

$$\frac{d}{db} \left(\int_0^{\infty} x^n e^{-bx^2} dx \right) = \int_0^{\infty} \frac{d}{db} (x^n e^{-bx^2}) dx = \int_0^{\infty} x^n e^{-bx^2} (-x^2) dx = - \int_0^{\infty} x^{n+2} e^{-bx^2} dx \quad (5.60)$$

Con lo que se tiene la relación de recurrencia:

$$I_{n+2}(b) = - \frac{dI_n(b)}{db} \quad (5.61)$$

En particular, para $n = 0$:

$$I_0(b) = \int_0^{\infty} e^{-bx^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{b}} \quad (5.62)$$

Y para $n = 1$:

$$I_1(b) = \int_0^{\infty} x e^{-bx^2} dx = \left[-\frac{e^{-bx^2}}{2b} \right]_{b=0}^{b \rightarrow \infty} = \frac{1}{2b} \quad (5.63)$$

Aplicando la relación de recurrencia a partir de $I_0(b)$ y $I_1(b)$, se llega a las fórmulas generales para $I_{2n}(b)$ (es decir, para índices pares) y $I_{2n+1}(b)$ (es decir, para índices impares):

$$I_{2n}(b) = \begin{cases} \frac{1}{2} \sqrt{\frac{\pi}{b}} & \text{para } n = 0 \\ \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2^{n+1}} \sqrt{\frac{\pi}{b^{2n+1}}} = \frac{(2n-1)!!}{2^{n+1}} \sqrt{\frac{\pi}{b^{2n+1}}} & \text{para } n \neq 0 \end{cases} \quad (5.64)$$

$$I_{2n+1}(b) = \frac{n!}{2b^{n+1}} \quad (5.65)$$

5.4.3. Integrales de la forma $\int_{-\infty}^{\infty} x^n e^{-bx^2} dx$

El valor de las integrales de la forma:

$$\int_{-\infty}^{\infty} x^n e^{-bx^2} dx \quad \text{donde } n \in \mathbb{Z}, n \geq 0, b > 0 \quad (5.66)$$

puede deducirse de los ya calculados de $I_n(b)$ y del carácter par o impar del integrando. En efecto, si n es par el integrando será una función par, y las contribuciones a la integral en los intervalos $(-\infty, 0]$ y $[0, \infty)$ son iguales. Por el contrario, si n es impar el integrando será una función impar, y las contribuciones a la integral en los intervalos $(-\infty, 0]$ y $[0, \infty)$ serán opuestas, con lo que su suma se anulará. En consecuencia:

$$\int_{-\infty}^{\infty} x^n e^{-bx^2} dx = \begin{cases} 2 \int_0^{\infty} x^n e^{-bx^2} dx = 2I_n(b) & \text{si } n \text{ es par} \\ 0 & \text{si } n \text{ es impar} \end{cases} \quad (5.67)$$

5.4.4. La función error de Gauss

La **función error** (también conocida como función error de Gauss) es una función no elemental que se utiliza en muchos campos de la Física y las Matemáticas. La función queda definida por la expresión:

$$\operatorname{erf}x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad \text{con } x \in \mathbb{R} \quad (5.68)$$

Una representación gráfica de esta función puede verse en la Figura 5.4. Notemos que, debido a la ecuación (5.53):

$$\lim_{x \rightarrow \infty} \operatorname{erf}x = 1 \quad (5.69)$$

Suele definirse también la **función error complementaria** por la expresión:

$$\operatorname{erfc}x = 1 - \operatorname{erf}x = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-t^2} dt - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad (5.70)$$

Los valores de la función error o de su complementaria no pueden hallarse analíticamente y requieren de técnicas numéricas, como la integración numérica o el desarrollo en serie de potencias.

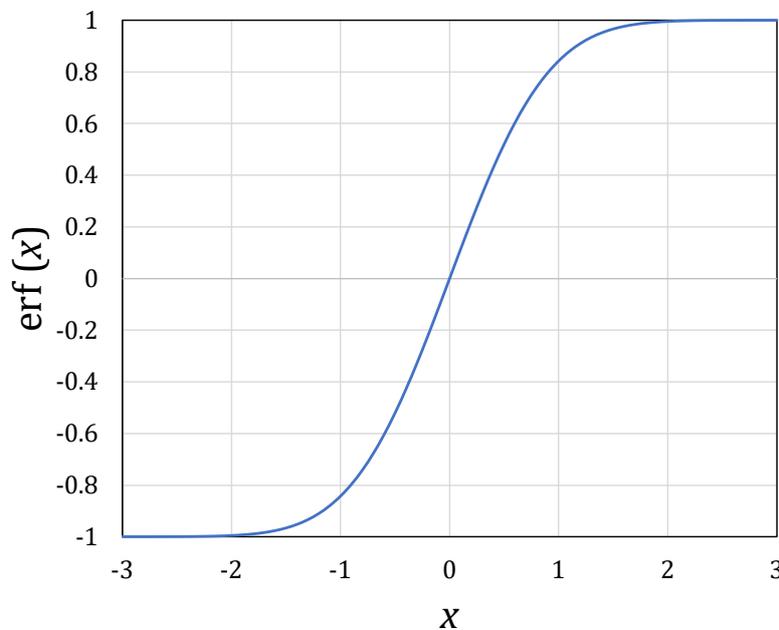


Figura 5.4: Función error de Gauss.

5.4.5. La función gamma de Euler

Adrien-Marie Legendre (1752-1833) propuso, en 1814, llamar **función gamma** y representar con la letra correspondiente, Γ , a una función que había sido introducida por primera vez en una carta que escribió Leonard Euler (1707-1783) a Christian Goldbach (1690-1764) en el año 1729. Esta función está definida para todos los números complejos excepto los enteros negativos y el cero (en donde diverge). Esta función fue escrita inicialmente en forma de límite:

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{(n!)(n^z)}{z(z+1)(z+2)\dots(z+n)} \quad (5.71)$$

El propio Euler obtuvo una expresión integral válida para complejos con **parte real positiva**:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \quad \text{para } \text{Re}(z) > 0 \tag{5.72}$$

Podemos hacer el cambio de variable $t = s^2 \Rightarrow \frac{dt}{ds} = 2s ds$ y obtener:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt = \int_0^\infty s^{2z-2} e^{-s^2} 2s ds = 2 \int_0^\infty s^{2z-1} e^{-s^2} ds \tag{5.73}$$

De esta expresión se obtiene que cuando z es un número n **entero no negativo**:

$$\Gamma(n) = 2 \int_0^\infty s^{2n-1} e^{-s^2} ds = 2 I_{2n-1}(1) = 2 \frac{(n-1)!}{2} = (n-1)! \tag{5.74}$$

Este resultado no es casual, y es que **la función gamma puede ser considerada una extensión del concepto de factorial** a los números complejos. De hecho, es fácil demostrar que se cumple la relación de recurrencia:

$$\Gamma(z + 1) = z\Gamma(z) \tag{5.75}$$

que cuadra a la perfección con esta idea.

Para cuando z es un número $(n + 1/2)$ semientero positivo:

$$\Gamma\left(n + \frac{1}{2}\right) = 2 \int_0^\infty s^{2n} e^{-s^2} ds = 2 I_{2n}(1) = \begin{cases} 2 \frac{\sqrt{\pi}}{2} = \sqrt{\pi} & \text{para } n = 0 \\ 2 \frac{(2n-1)!!}{2^{n+1}} \sqrt{\pi} = \frac{(2n-1)!!}{2^n} \sqrt{\pi} & \text{para } n \neq 0 \end{cases} \tag{5.76}$$

La función gamma cumple otras interesantes propiedades que permiten hallar más valores numéricos de la misma de forma exacta, pero en las cuales no entraremos. Además, también está relacionada con otras funciones matemáticas importantes. En la Figura 5.5 se representa la función gamma para cuando su argumento es real.

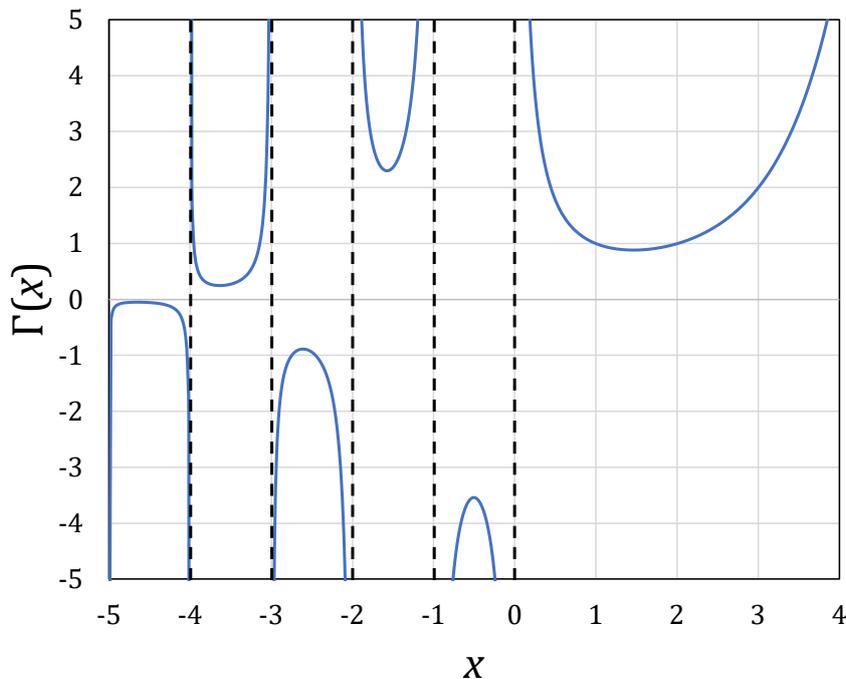


Figura 5.5: Función gamma de Euler en la recta real.

Ejercicios – Gas ideal monoatómico y distribución de Maxwell

1.- Calcular la celeridad más probable, la celeridad media y la celeridad cuadrática media de las moléculas de hidrógeno (masa molar $2.016 \text{ g}\cdot\text{mol}^{-1}$) a $20 \text{ }^\circ\text{C}$ y a $50 \text{ }^\circ\text{C}$ y comparar los resultados obtenidos para cada temperatura.

R: $20 \text{ }^\circ\text{C}$: 1555, 1755 y $1905 \text{ m}\cdot\text{s}^{-1}$; $50 \text{ }^\circ\text{C}$: 1632, 1842 y $1999 \text{ m}\cdot\text{s}^{-1}$.

2.- Por un cierto procedimiento se mide el número de moléculas de oxígeno (gas perfecto diatómico de masa molar $31.999 \text{ g}\cdot\text{mol}^{-1}$) que se mueve con una cierta celeridad. La incertidumbre máxima en la medida de la celeridad es $0.5 \text{ m}\cdot\text{s}^{-1}$, de modo que se admite que las moléculas de $v = 500 \text{ m}\cdot\text{s}^{-1}$ son las de celeridades comprendidas entre 499.5 y $500.5 \text{ m}\cdot\text{s}^{-1}$. La razón entre el número de moléculas que se mueven con 500 y $300 \text{ m}\cdot\text{s}^{-1}$ es 5/3. Calcular la temperatura del gas.

R: 602.72 K.

3.- Hallar en el nitrógeno gaseoso (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$) la temperatura a la cual a las celeridades de las moléculas $v_1 = 300 \text{ m}\cdot\text{s}^{-1}$ y $v_2 = 600 \text{ m}\cdot\text{s}^{-1}$ les corresponden valores idénticos de la densidad de probabilidad.

R: 328.06 K.

4.- Calcular la celeridad de las moléculas de un gas, de masa m , a la que el valor de la densidad de probabilidad, a una temperatura T , será el mismo que para la temperatura n veces mayor.

R: $[(3k_B T/m)(n \cdot \ln n / (n - 1))]^{1/2}$.

5.- Hallar la celeridad de las moléculas con igual densidad de probabilidad en el hidrógeno (masa molar $2.016 \text{ g}\cdot\text{mol}^{-1}$) y en el helio (masa molar $4.003 \text{ g}\cdot\text{mol}^{-1}$) a 300 K .

R: $1607 \text{ m}\cdot\text{s}^{-1}$.

6.- Calcular la fracción de moléculas de nitrógeno (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$) que a $27 \text{ }^\circ\text{C}$ tienen una celeridad superior a $700 \text{ m}\cdot\text{s}^{-1}$.

R: 13.85% (Requiere integrar por métodos numéricos).

7.- Calcular la fracción de moléculas de nitrógeno (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$) que a $27 \text{ }^\circ\text{C}$ tienen una celeridad comprendida entre 200 y $400 \text{ m}\cdot\text{s}^{-1}$.

R: 31.41% (Requiere integrar por métodos numéricos).

8.- Calcular la fracción de moléculas de un gas que se mueven con celeridad no superior a:

a) La celeridad media.

b) La celeridad más probable.

c) La celeridad cuadrática media.

¿Cuántas moléculas corresponderán a cada situación si se trata de un mol?

R: a) 53.31% ($3.21 \cdot 10^{23}$); b) 42.76% ($2.58 \cdot 10^{23}$); c) 60.83% ($3.66 \cdot 10^{23}$) (Requiere integrar por métodos numéricos).

9.- Calcular el tanto por ciento de moléculas de nitrógeno (de masa molar $28.01 \text{ g}\cdot\text{mol}^{-1}$) que a $27 \text{ }^\circ\text{C}$ tienen una celeridad en el intervalo definido por la celeridad más probable $\pm 0.5 \text{ m}\cdot\text{s}^{-1}$.

R: 0.20%.

10.- Deducir la forma de la función de densidad de probabilidad de la energía cinética de traslación de las moléculas de un gas ideal.

$$R: g(\varepsilon) = \frac{2}{\sqrt{\pi}(k_B T)^{3/2}} \varepsilon^{1/2} e^{-\frac{\varepsilon}{k_B T}}.$$

11.- ¿Cuál es la energía cinética de traslación más probable y cuál es la energía cinética de traslación de las moléculas que tienen celeridad igual a la celeridad más probable?

$$R: k_B T/2; k_B T.$$

12.- Deducir, a partir de la densidad de probabilidad de la energía cinética, la energía cinética media de las moléculas de un gas monoatómico que se encuentra en equilibrio.

$$R: (3/2)k_B T \text{ (Requiere usar la función gamma de Euler o integrar por métodos numéricos).}$$

13.- ¿Qué porcentaje de las moléculas de un gas monoatómico, que se encuentra en equilibrio, tiene una energía cinética que difiere del valor medio de la energía en no más de 1%?

$$R: 0.93\%.$$

14.- Calcular el valor medio del cuadrado de la energía cinética de las moléculas de un gas monoatómico.

$$R: (15/4)(k_B T)^2 \text{ (Requiere usar la función gamma de Euler o integrar por métodos numéricos).}$$

6. Teorema de equipartición de la energía. Gas ideal poliatómico

6.1. Teorema de Equipartición de la Energía (conj. canónico clásico)

El Teorema de Equipartición de la Energía del conjunto canónico clásico (o semiclásico) permite obtener la **Ecuación Energética de Estado** de un sistema clásico que cumple unas condiciones ideales, pero que en muchos casos pueden considerarse válidas o aproximadamente válidas.

Este teorema indica que: **si la energía de un sistema que obedece a la Mecánica Clásica es una suma de términos desacoplados y proporcionales al cuadrado de variables dinámicas (coordenadas de posiciones, velocidades, momentos, velocidades angulares, etc.) entonces a cada uno de esos términos le corresponde, en promedio, una energía igual a $k_B T/2$.**

Este teorema se demuestra a partir de la función de partición. Supongamos que la energía del sistema se puede escribir como la suma:

$$E = \sum_{i=1}^n k_i a_i^2 \quad (6.1)$$

donde k_i son constantes y a_i representan variables dinámicas (posiciones, velocidades, momentos, etc.). Sabemos que, cuando la energía es aditiva, la función de partición se puede factorizar. En el caso semiclásico, además, pueden aparecer factores debidos a la indistinguibilidad de las moléculas o la constante de Planck, pero dichos factores no dependen de la temperatura (que es lo importante, lo que nos interesa en esta sección). Por tanto, la función de partición será proporcional al producto de las funciones de partición de cada variable dinámica:

$$Z = C \cdot Z_1 Z_2 \dots Z_n \quad (6.2)$$

donde C representa los factores mencionados, siendo:

$$Z_i = \int_{a_{i,\min}}^{a_{i,\max}} da_i e^{-\beta k_i a_i^2} \quad (6.3)$$

Como C no depende de la temperatura (no depende de β), calculando la Ecuación Energética de Estado tenemos:

$$U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V,N} = - \sum_{i=1}^n \left(\frac{\partial \ln Z_i}{\partial \beta} \right)_{V,N} \quad (6.4)$$

Vemos entonces que podemos calcular la energía interna como la suma de las contribuciones de cada término cuadrático.

Introduciendo el cambio de variable $y_i = \sqrt{\beta k_i} a_i$:

$$Z_i = \frac{1}{\sqrt{\beta k_i}} \int_{y_{i,\min}}^{y_{i,\max}} dy_i e^{-y_i^2} \Rightarrow \ln Z_i = -\frac{1}{2} \ln \beta - \ln \left(\frac{1}{\sqrt{k_i}} \int_{y_{i,\min}}^{y_{i,\max}} dy_i e^{-y_i^2} \right) \quad (6.5)$$

Como el término entre paréntesis no depende de β :

$$\boxed{-\left(\frac{\partial \ln Z_i}{\partial \beta}\right)_{V,N} = \frac{1}{2\beta} = \frac{1}{2}k_B T} \quad (6.6)$$

De las ecuaciones (6.4) y (6.6) concluimos que, efectivamente, cada término cuadrático contribuye a la energía interna, en promedio, con un sumando $k_B T/2$.

6.2. Grados de libertad

6.2.1. Definición

Se denomina **número de grados de libertad** de un sistema mecánico al número mínimo de **coordenadas independientes** necesarias para especificar su configuración mecánica (la posición de todas sus partículas o puntos). En otras palabras, se trata de las coordenadas necesarias para especificar la posición de su centro de masas respecto de un sistema de referencia inercial, pero también la orientación (los ángulos de Euler) de los respectivos ejes principales de inercia para una posición dada del centro de masas, sus posibles vibraciones, etc., ya que el movimiento mecánico más general de un sistema consiste en que realice una traslación como un todo, seguida por una rotación como un todo más una deformación pura final (lo que constituye la llamada Mecánica de Medios Deformables, a partir de la cual pueden obtenerse incluso las ecuaciones de Navier-Stokes de la Física de Fluidos). Así:

- Una **partícula libre** (partícula puntual no sujeta a ninguna interacción) requiere de 3 coordenadas independientes para especificar su posición (por ejemplo, cartesianas o esféricas) y por tanto tiene **3 grados de libertad**.
- Un sistema de **N partículas libres** (esto es, que no interaccionan entre ellas ni con nada más) requiere 3 coordenadas por cada partícula, lo que quiere decir que tiene un total de **$3N$ grados de libertad**.

No obstante, las partículas que constituyen un sistema no tienen por qué ser libres. Por ejemplo, si una partícula está sujeta a moverse por un plano, tendrá 2 grados de libertad en lugar de 3. Matemáticamente hablando, está forzada a cumplir la ecuación de ese plano, de forma que dicho plano transforma las 3 coordenadas en dependientes a través de esa ecuación; conocidas 2 de ellas, dicha ecuación del plano determina la restante.

En general, un sistema de N partículas tendrá un $k \leq 3N$ grados de libertad, debido a la presencia de **ligaduras**, o relaciones que deben cumplir sus coordenadas. Las ligaduras pueden ser de muy diverso tipo (ecuaciones algebraicas, inecuaciones algebraicas, ecuaciones diferenciales integrables o no integrables, etc.) y depender o no del tiempo. En general, las ligaduras son consecuencia de unas fuerzas (conocidas o desconocidas) que reciben el nombre de **fuerzas de ligadura**.

6.2.2. Grados de libertad de sólidos rígidos

Un **sólido rígido** se define como un sistema de partículas tal que tanto las distancias entre ellas como sus orientaciones relativas permanecen constantes en el tiempo (es, pues, un sólido **no deformable**). En otras palabras, los únicos movimientos permitidos a un sólido rígido son traslaciones, rotaciones en torno a un eje y composiciones de traslaciones y rotaciones en torno a un eje.

La constancia de las distancias entre partículas $i, j = 1, \dots, N$, se representa con ligaduras de la forma:

$$|\vec{r}_i - \vec{r}_j| = \text{constante}$$

Para un sólido rígido de N partículas, hay un total de ligaduras igual a:

$$\binom{N}{2} = \frac{N!}{2! \cdot (N-2)!} = \frac{N \cdot (N-1)}{2}$$

Pero no todas son independientes y el cálculo del número de grados de libertad debe hacerse con cuidado.

Ejemplo. En una molécula diatómica rígida (sólido rígido formado por 2 átomos), $N = 2$, de forma que:

$$\binom{2}{2} = \frac{2!}{2! \cdot (2-2)!} = \frac{2 \cdot (2-1)}{2} = 1$$

Así pues, una molécula diatómica rígida sólo tendría 1 ligadura, que consistiría en la obligación de mantener constante la distancia internuclear. Por ello:

$$\left. \begin{array}{l} \text{Coordenadas} = 3N = 3 \cdot 2 = 6 \\ \text{Ligaduras independientes} = 1 \end{array} \right\} \Rightarrow \text{n}^\circ \text{ grados de libertad} = \text{Coordenadas} - \text{Ligaduras independientes} = 5$$

5 grados de libertad $\left\{ \begin{array}{l} 3 \text{ traslación (centro de masas)} \\ 2 \text{ rotación (respecto de los 2 ejes ortogonales al eje internuclear)} \end{array} \right\}$

Si anulamos la exigencia de que la molécula sea rígida, permitiendo su vibración (respecto de su enlace molecular internuclear), la ligadura de distancia internuclear constante desaparece, en cuyo caso tendremos:

$$\left. \begin{array}{l} \text{Coordenadas} = 3N = 3 \cdot 2 = 6 \\ \text{Ligaduras independientes} = 0 \end{array} \right\} \Rightarrow \text{n}^\circ \text{ grados de libertad} = \text{Coordenadas} - \text{Ligaduras independientes} = 6$$

6 grados de libertad $\left\{ \begin{array}{l} 3 \text{ traslación (centro de masas)} \\ 2 \text{ rotación (respecto de los 2 ejes ortogonales al enlace molecular)} \\ 1 \text{ vibración (en la dirección del enlace molecular internuclear)} \end{array} \right\}$

De forma más general, en el cálculo del número de grados de libertad de sólidos rígidos hay dos casos a considerar: el de los **sólidos rígidos lineales** (cuyas partículas están situadas sobre una recta) y el de los **sólidos rígidos no lineales**. Se puede demostrar (véase la sección 6.5) que **el número de grados de libertad de un sólido rígido es 5 para uno lineal y 6 para uno no lineal**.

De los grados de libertad de un sólido rígido, 3 de ellos pueden asociarse a las coordenadas del centro de masas del sistema, caracterizando por tanto el movimiento de traslación del sistema como un todo. Por ello, se dice que hay **3 grados de libertad de traslación**. El **resto de los grados de libertad (2 para lineales o 3 para no lineales)** pueden escogerse como ángulos que determinan la orientación del sistema para una posición dada del centro de masas y, por ello, se denominan **grados de libertad de rotación**. Los designaremos por f_{rot} y se cumple:

$$f_{\text{rot}} = \begin{cases} 2 & \text{si el sistema es lineal} \\ 3 & \text{si el sistema es no lineal} \end{cases} \quad (6.7)$$

El grado de libertad de menos que tienen los sólidos rígidos lineales puede entenderse fácilmente desde el punto de vista físico: la rotación en torno al eje de simetría del sólido lineal no produce ningún cambio apreciable en su configuración, de modo que el ángulo de rotación en torno a ese eje deja de ser relevante para su descripción y ese grado de libertad desaparece.

6.2.3. Grados de libertad moleculares

Hasta ahora, nos hemos restringido únicamente al estudio de los grados de libertad de traslación de las moléculas. Esto es suficiente si las moléculas que estamos considerando son monoatómicas y puntuales (es decir, estrictamente hablando, átomos). Un refinamiento de la teoría requiere tener en cuenta que si el gas está formado por moléculas (poliatómicas), cuyos átomos trataremos a efectos mecánicos como puntos materiales (esta suposición se prueba válida en los experimentos de medida de capacidades caloríficas molares de los gases, como comentaremos más adelante), entonces al movimiento de **traslación** del centro de masas de la molécula hemos de añadir otros dos: las **rotaciones** y las **vibraciones** (para una justificación más detallada de esto, véase la sección 6.2.4).

En primer lugar, prescindamos de la vibración y supongamos que tratamos con una **molécula q -atómica con enlaces rígidos**. Como hemos comentado en el apartado anterior, dicha molécula tendrá un número total de grados de libertad f_{rig} dado por:

$$f_{\text{rig}} = 3 + f_{\text{rot}} \quad (6.8)$$

Por otra parte, **si la molécula no es rígida**, es decir, los átomos pueden vibrar en torno a sus posiciones de equilibrio bajo la acción de las fuerzas intramoleculares (**enlaces elásticos**), entonces el número total de grados de libertad, f_{elas} , deberá ser:

$$f_{\text{elas}} = 3 + f_{\text{rot}} + f_{\text{vib}} = 3q \quad (6.9)$$

donde f_{vib} designa el **número de grados de libertad de vibración** y es igual a:

$$f_{\text{vib}} = 3q - 3 - f_{\text{rot}} = \begin{cases} 3q - 5 & \text{si la molécula es lineal} \\ 3q - 6 & \text{si la molécula es no lineal} \end{cases} \quad (6.10)$$

6.2.4. Gas ideal poliatómico semiclásico. Aplicación del Teorema de Equipartición

Suponiendo que las moléculas puedan ser consideradas como entidades elementales independientes del sistema, su **energía potencial** puede expresarse como la suma de varias contribuciones a la energía potencial:

$$\varepsilon_p = (\varepsilon_p)_{\text{intermolecular}} + (\varepsilon_p)_{\text{intramolecular}} + (\varepsilon_p)_{\text{campo externo}}$$

Las hipótesis del modelo cinético del gas ideal implican:

- $(\varepsilon_p)_{\text{intermolecular}} = 0$, pues no existen fuerzas intermoleculares (gas ideal).
- $(\varepsilon_p)_{\text{campo externo}} = 0$, pues no están sometidos a ningún campo externo.

Además, las paredes del recipiente se encargan de que no escapen moléculas al exterior y las confinan en un recinto de volumen V . La $(\varepsilon_p)_{\text{intramolecular}}$ determina los movimientos de los átomos que componen la molécula en torno a sus posiciones de equilibrio. Para muchas aplicaciones, se puede suponer que estos movimientos son simplemente rotaciones en torno a un eje que pasa por el centro de masas de la molécula y vibraciones con respecto a dichas posiciones de equilibrio, y que dichos movimientos son independientes unos de otros y pueden estudiarse por separado (lo que se suele expresar diciendo que **los distintos movimientos de traslación, rotación y vibración no están acoplados**). Para estos casos, la **energía cinética** de la molécula puede ponerse muy aproximadamente como la suma de tres contribuciones independientes:

$$\varepsilon_c = (\varepsilon_c)_{\text{traslación}} + (\varepsilon_c)_{\text{rotación}} + (\varepsilon_c)_{\text{vibración}}$$

y la **energía total** de cada molécula del gas ideal dentro del recipiente será:

$$\varepsilon = (\varepsilon_c)_{\text{traslación}} + (\varepsilon_c)_{\text{rotación}} + (\varepsilon_c)_{\text{vibración}} + (\varepsilon_p)_{\text{intramolecular}} \quad (6.11)$$

donde la $(\varepsilon_p)_{\text{intramolecular}}$ se supone, en primera aproximación, como una energía potencial compuesta de términos de tipo oscilador armónico, proporcionales al cuadrado del desplazamiento de los átomos respecto de sus posiciones de equilibrio. Así pues, vemos que **todos los términos de la energía son cuadráticos**:

- $(\varepsilon_c)_{\text{traslación}} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$ tiene 3 términos cuadráticos (proporcionales al cuadrado del momento lineal).
- $(\varepsilon_c)_{\text{rotación}} = \sum_{j=1}^{f_{\text{rot}}} \left(\frac{1}{2} I_j \omega_j^2 \right)$ tiene 2 términos cuadráticos si la molécula es lineal y 3 términos cuadráticos si es no lineal, proporcionales al cuadrado de la velocidad angular de rotación ω_j en torno a cada uno de los ejes principales (I_j es el momento de inercia).
- $(\varepsilon_c)_{\text{vibración}}$ tiene tantos términos como grados de libertad de vibración, y dichos términos son proporcionales al cuadrado de la velocidad de vibración relativa a las posiciones de equilibrio de los átomos.
- $(\varepsilon_p)_{\text{intramolecular}}$ tiene tantos términos como grados de libertad de vibración, proporcionales al cuadrado del desplazamiento de los átomos respecto de sus posiciones de equilibrio.

En consecuencia, para **cada molécula** hay una **suma de n_{TC} términos cuadráticos en la energía**, siendo:

$$n_{\text{TC}} = 3 + f_{\text{rot}} + 2f_{\text{vib}} \quad (6.12)$$

Apliquemos ahora el Teorema de Equipartición al caso de las moléculas del gas ideal poliatómico:

- A cada grado de libertad de **traslación** le corresponde una energía cinética de traslación en una cierta dirección (uno de los ejes coordenados escogidos), con lo que le corresponde una energía promedio $k_B T/2$.
- A cada grado de libertad de **rotación** le corresponde una energía cinética de rotación proporcional al cuadrado de la velocidad angular en torno a un eje, de modo que hay que asignarle una energía promedio $k_B T/2$.
- A cada grado de libertad de **vibración**, sin embargo, le corresponden dos tipos de energía: una cinética de vibración, proporcional al cuadrado de la velocidad de vibración en torno a la posición de equilibrio, y una potencial de vibración, proporcional al cuadrado de la distancia a la posición de equilibrio. A cada uno de los tipos de energía le corresponde una energía promedio de $k_B T/2$, de modo que, en total, al grado de vibración le corresponde $k_B T$.

La energía media de una molécula, $\langle \varepsilon \rangle$, que tiene f_{rot} grados de libertad de rotación y f_{vib} grados de libertad de vibración, se obtiene como la suma de las contribuciones promedio de cada grado de libertad:

$$\langle \varepsilon \rangle = (3 + f_{\text{rot}} + 2f_{\text{vib}}) \cdot \frac{1}{2} k_B T = \frac{n_{\text{TC}}}{2} k_B T \quad (6.13)$$

donde se ha usado la ecuación (6.12).

6.2.5. Gas ideal poliatómico semiclásico: función de partición y ecuaciones de estado

De acuerdo con la aditividad de la energía expresada en la ecuación (6.11) y las propiedades de factorización de la función de partición, la función de partición de 1 molécula se podrá expresar como el producto:

$$Z_1 = (Z_1)_{\text{traslación}} (Z_1)_{\text{rotación}} (Z_1)_{\text{vibración}} \quad (6.14)$$

donde cada uno de los factores se calcula usando por separado las energías de traslación (cinética), rotación (cinética) y vibración (cinética + potencial), respectivamente. La función de partición del gas total será:

$$Z = \frac{Z_1^N}{N!} = \left[\frac{(Z_1)_{\text{traslación}}^N}{N!} \right] [(Z_1)_{\text{rotación}}]^N [(Z_1)_{\text{vibración}}]^N = Z_{\text{traslación}} Z_{\text{rotación}} Z_{\text{vibración}} \quad (6.15)$$

Notemos que:

$$Z_{\text{traslación}} = \frac{(Z_1)_{\text{traslación}}^N}{N!} \quad (6.16)$$

se corresponde con la función de partición del gas ideal **monoatómico**, pero tomando como coordenadas y momentos los del centro de masa de la molécula poliatómica.

La función de Massieu del sistema será, por tanto, aditiva:

$$\frac{\psi}{k_B} = \ln Z(\beta, V, N) = \ln Z_{\text{traslación}}(\beta, V, N) + \ln Z_{\text{rotación}}(\beta, N) + \ln Z_{\text{vibración}}(\beta, N) \quad (6.17)$$

y, al calcular las ecuaciones de estado, las derivadas parciales se obtienen como sumas de las derivadas parciales de cada término.

La **Ecuación Energética de Estado (EEE)** ya la hemos estudiado mediante el Teorema de Equipartición de la Energía y no es necesario volver a calcular las derivadas:

$$U = - \left(\frac{\partial \ln Z}{\partial \beta} \right)_{V, N} = N \langle \varepsilon \rangle = \frac{n_{\text{TC}}}{2} N k_B T \quad (6.18)$$

La **Ecuación Térmica de Estado (ETE)** se obtiene derivando parcialmente $\ln Z$ respecto de V pero, como se ha indicado en la ecuación (6.17), **los grados de libertad internos de la molécula** (rotación y vibración) **no dependen del volumen del recipiente**, por lo que el resultado depende solo del término de traslación. Esto es así porque la integración a las coordenadas de los centros de masa, que es la que hace que nos salga en las cuentas el volumen V , se hace en el término de traslación, pero no en los otros dos. En otras palabras, **la ETE del gas ideal poliatómico clásico es idéntica a la del gas ideal monoatómico clásico**:

$$\frac{p}{k_B T} = p\beta = \left(\frac{\partial \ln Z}{\partial V} \right)_{\beta, N} = \frac{N}{V} \Rightarrow \boxed{pV = N k_B T} \quad (6.19)$$

La **Ecuación Material de Estado**, por el contrario, dependerá de los tres tipos de grados de libertad, pero en este curso no ahondaremos en su estudio.

6.3. Predicciones de la teoría semiclásica del gas ideal poliatómico: energía interna y capacidades caloríficas

La energía interna molar del gas ideal poliatómico será:

$$\boxed{U_m = N_A \langle \varepsilon \rangle = \frac{n_{\text{TC}}}{2} N_A k_B T = \frac{n_{\text{TC}}}{2} RT} \quad (6.20)$$

Esta ecuación es la llamada **Ecuación Energética de Estado del gas perfecto**, que solo depende de la temperatura, concordando así con la Primera Ley de Joule de los gases ideales que se estudia en Termodinámica.

La **capacidad calorífica isócara molar**, C_{V_m} , se obtiene derivando la energía interna molar respecto de la temperatura a volumen constante:

$$C_{Vm} = \left(\frac{\partial U_m}{\partial T} \right)_V = \frac{n_{TC}}{2} R \quad (6.21)$$

y la **capacidad calorífica isóbara molar**, C_{pm} , se obtiene a partir de la relación de Mayer (válida para los gases ideales):

$$C_{pm} = C_{Vm} + R = \frac{n_{TC}}{2} R + R = \frac{n_{TC} + 2}{2} R \quad (6.22)$$

A partir de estas dos, es inmediato obtener el coeficiente adiabático γ del gas:

$$\gamma = \frac{C_{pm}}{C_{Vm}} = \frac{n_{TC} + 2}{n_{TC}} \quad (6.23)$$

En la Tabla 6.1 se recogen las predicciones teóricas de las capacidades caloríficas de gases perfectos para ciertas moléculas sencillas.

Tabla 6.1: Valores teóricos de las capacidades caloríficas molares a volumen y presión constante, junto con su correspondiente coeficiente adiabático, para algunas moléculas sencillas.

Tipo de molécula	f_{rot}	f_{vib}	n_{TC}	C_{Vm}	C_{pm}	$\gamma = C_{pm}/C_{Vm}$
Monoatómica	0	0	3	$\frac{3}{2}R$	$\frac{5}{2}R$	$\frac{5}{3} \approx 1.67$
Diatómica rígida	2	0	5	$\frac{5}{2}R$	$\frac{7}{2}R$	$\frac{7}{5} = 1.40$
Diatómica elástica	2	1	7	$\frac{7}{2}R$	$\frac{9}{2}R$	$\frac{9}{7} \approx 1.29$
N -atómica no lineal rígida con $N \geq 3$	3	0	6	$\frac{6}{2}R = 3R$	$\frac{8}{2}R = 4R$	$\frac{4}{3} \approx 1.33$

Volviendo de nuevo a la relación de la energía del gas con su Ecuación Térmica de Estado, notemos que:

$$pV = \frac{2}{3} \langle E_{tras} \rangle \quad (6.24)$$

donde $\langle E_{tras} \rangle = N \langle \epsilon_{tras} \rangle = N \cdot (3k_B T/2)$ es la energía cinética de traslación media total del gas. Esta ecuación recibe el nombre de **ecuación de Bernoulli** y tiene la particularidad de que se cumple también para los gases ideales cuánticos.

Mencionemos, a modo de curiosidad, que el (2) que aparece en el numerador de la ecuación (6.24) es el exponente del momento lineal en la función que lo relaciona con la energía cinética molecular, mientras que el (3) que aparece en el denominador es el número de dimensiones geométricas espaciales donde se mueven las partículas. Puede entonces generalizarse fácilmente a otros tipos de partículas (variando entonces el numerador) y a un número diferente de dimensiones (cambiando el denominador). Por ejemplo, un gas ultrarrelativista [donde $E = pv_s$ (en el caso de los fonones, sonido) o $E = pc$ (en el caso de los fotones, luz)] que se mueve en tres dimensiones, tendría una ecuación de Bernoulli dada por: $pV = \frac{1}{3} \langle E_{tras} \rangle$.

6.4. Comparación de las capacidades caloríficas predichas por el modelo semiclásico con los resultados experimentales

En la Tabla 6.2 se muestran los valores experimentales a temperatura ambiente de las capacidades caloríficas molares de algunos gases sencillos, así como su coeficiente adiabático, y se comparan con valores calculados según el modelo cinético de gas ideal.

En primer lugar, notemos que el modelo cinético clásico da muy buenas aproximaciones para las capacidades caloríficas molares de los gases monoatómicos y diatómicos, pero no para el vapor del agua (triatómico). Pero es que, aunque para gases monoatómicos y diatómicos la concordancia es buena, no es perfecta. En particular, supera el valor de la incertidumbre experimental esperable. Teniendo en cuenta que el modelo predice valores que son exactamente múltiplos de $R/2$, cualquier desviación de esos múltiplos es importante.

En segundo lugar, el acuerdo aproximado del modelo y los experimentos a temperatura ambiente para moléculas diatómicas sólo ocurre cuando las consideramos rígidas, es decir, sin tener en cuenta todos los posibles grados de libertad reales. En el caso de la molécula triatómica, las mejores aproximaciones se consiguen también con un número de grados de libertad menor que el total de la molécula.

Tabla 6.2: Valores experimentales y teóricos a temperatura ambiente de las capacidades caloríficas y coeficientes adiabáticos de algunos gases.

Gas	Modelo de molécula	$C_{V,m}/J \cdot mol^{-1} \cdot K^{-1}$		$C_{p,m}/J \cdot mol^{-1} \cdot K^{-1}$		γ	
		Teórico	Exp.	Teórico	Exp.	Teórico	Exp.
He	Monoatómica	12.5	12.5	20.8	20.9	1.67	1.67
O ₂	Diatómica rígida	20.8	20.9	29.1	28.9	1.40	1.4
CO	Diatómica rígida	20.8	21.0	29.1	29.3	1.40	1.4
H ₂ O	Triatómica rígida	25.0	27.8	33.2	36.2	1.33	1.30
	Triatómica con $f_{vib} = 1$	33.2		41.5		1.25	

Las discrepancias entre el modelo y los experimentos se hacen especialmente sorprendentes si se analiza la dependencia de la capacidad calorífica con la temperatura. En la Figura 6.1 está representada cualitativamente la curva experimental de $C_{V,m}(H_2)$ en función de T . La escala de la temperatura en la gráfica es logarítmica.

De acuerdo con el modelo semiclásico, la capacidad calorífica no debería depender de la temperatura. Sin embargo, en la figura puede verse que esto es cierto sólo en ciertos intervalos de temperatura. Y, además, $C_{V,m}$ tiene, en esos intervalos, los valores que corresponden a diferente valor de n_{TC} :

- **Tramo 1-1'**: la molécula se comporta como si $n_{TC} = 3$.
- **Tramo 2-2'**: la molécula aparenta tener $n_{TC} = 5$.
- **Tramo 3-3'**: la molécula muestra el comportamiento correspondiente a $n_{TC} = 7$.

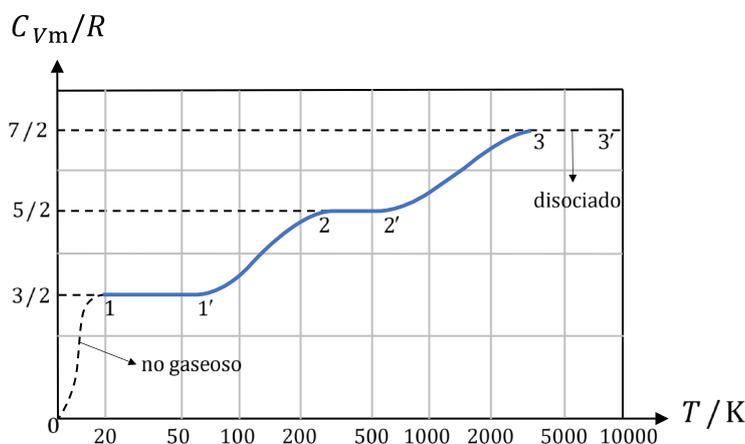


Figura 6.1: Comportamiento cualitativo de la capacidad calorífica isocora molar del H_2 en función de la temperatura a presión atmosférica. El comportamiento entre 3 y 3' es hipotético, pues, a temperaturas tan altas, las moléculas de H_2 se disocian en H atómico (están disociadas en un 95% a 5000 K).

En los tramos 0-1, 1'-2 y 2'-3, C_{vm} crece de forma continua con la temperatura, comportándose como si n_{TC} no fuese un número entero; en otras palabras, **estando en completo desacuerdo con el modelo.**

Llegamos entonces a la conclusión de que existen profundas deficiencias conceptuales en el modelo de gas ideal (en el límite semiclásico) que hemos desarrollado.

La explicación teórica de este comportamiento experimental del gas se encuentra fuera del alcance de la Mecánica Clásica, debiendo recurrir a una nueva mecánica, la Mecánica Cuántica. El tratamiento riguroso de estas cuestiones está fuera de los límites de este curso, de modo que nos limitaremos a comentar los resultados que se obtienen.

De acuerdo con la Mecánica Cuántica, tanto los posibles valores de las energías de traslación de las moléculas confinadas en un recipiente como sus posibles energías de rotación y vibración están **cuantizadas** (adoptan valores discretos de energía). Las condiciones macroscópicas en las que estos movimientos pueden considerarse clásicos marcan las regiones planas en la curva de la capacidad calorífica. Analicemos cada uno de estos tipos de movimiento por separado.

6.4.1. Movimiento de traslación

Se demuestra rigurosamente que el movimiento de traslación puede considerarse clásico cuando:

$$\lambda = e^{\frac{\mu}{k_B T}} = \frac{N}{V} \lambda_t^3 \ll 1 \quad (6.25)$$

donde λ es la llamada actividad absoluta y $\mu = k_B T \ln(\lambda)$ es el **potencial químico del gas** (la fuerza termodinámica generalizada asociada al intercambio de materia).

La magnitud λ_t tiene dimensiones de longitud y se denomina **longitud de onda térmica de "de Broglie"**:

$$\lambda_t = \frac{h}{\sqrt{2\pi m k_B T}} = \left(\frac{h^2}{2\pi m k_B T} \right)^{1/2} \quad (6.26)$$

y h es la **constante de Planck**.

Vemos que **el límite clásico viene determinado por el cociente del potencial químico y la temperatura, y que se alcanza para altas temperaturas y bajas densidades de partículas.** Es decir, cuando $\lambda \rightarrow 0$; $\mu \rightarrow -\infty$.

Pero vamos a analizar con más cuidado qué significa esta condición. De acuerdo con la **hipótesis de "de Broglie"**, a una partícula se le puede asociar una onda de materia cuya longitud de onda es del orden de $\lambda_{DB} = h/p = h/(mv)$. Para una molécula del gas, el módulo de la velocidad será del orden de su valor cuadrático medio, $\sqrt{\langle v^2 \rangle}$, por lo que la **longitud de onda "de Broglie"** será del orden de:

$$\begin{aligned} \lambda_{DB} &= \frac{h}{p} = \frac{h}{(2m\varepsilon_{tras})^{1/2}} \\ \lambda_{DB} &\sim \frac{h}{m \cdot \sqrt{\langle v^2 \rangle}} = \frac{h}{m \cdot (3k_B T/m)^{1/2}} = \left(\frac{h^2}{3mk_B T} \right)^{1/2} \sim \lambda_t \end{aligned} \quad (6.27)$$

Podemos ver entonces que la longitud de onda térmica es del orden de la longitud de onda de las ondas de materia de las moléculas. Por otro lado, la distancia media entre moléculas en el gas, d , puede estimarse en orden de magnitud como:

$$d \sim \left(\frac{V}{N} \right)^{1/3} = \left(\frac{N}{V} \right)^{-1/3} \quad (6.28)$$

En consecuencia, la condición de movimiento de traslación clásico es equivalente a:

$$\left(\frac{N}{V}\right)\lambda_t^3 \ll 1 \Rightarrow d^{-3} \cdot \lambda_t^3 \ll 1 \Rightarrow (\lambda_t/d)^3 \ll 1 \Rightarrow \boxed{\lambda_t \ll d} \quad (6.29)$$

Es decir, para que podamos considerar como clásica la traslación de las moléculas de un gas, **la longitud de onda de “de Broglie” de las moléculas tiene que ser muy pequeña en comparación con las distancias intermoleculares**. En otras palabras, las partículas tienen que estar lo suficientemente separadas para que los fenómenos de interferencia entre **sus ondas de materia no sean relevantes**. Cuando esto no ocurre, se producen fenómenos de coherencia cuántica que hacen que las moléculas estén correlacionadas entre sí, aunque no existan fuerzas intermoleculares.

Las condiciones enunciadas para poder aplicar el límite clásico de la traslación obtenido para el modelo cinético clásico se pueden extender a otros sistemas aparte de dichos gases ideales. En la Tabla 6.3 se recogen los valores aproximados de $\lambda = (N/V)\lambda_t^3$ para algunos sistemas, a las temperaturas indicadas. Se puede apreciar que, para los gases ordinarios (reales), el movimiento de traslación ya se puede considerar clásico a temperaturas muy bajas, próximas al cero absoluto. Cabe destacar el comportamiento cuántico de los electrones de conducción en los sólidos, incluso a temperatura ambiente.

Se puede demostrar que la distancia entre los niveles de energía cercanos es despreciable en las condiciones del límite clásico y que, por tanto, la energía cinética de traslación puede considerarse como un continuo desde temperaturas muy bajas. El caso clásico se corresponde con la zona con temperaturas mayores que la del punto 1 de la Figura 6.1.

Tabla 6.3: Carácter (clásico o cuántico) del movimiento de traslación para distintos sistemas.

Sistema	T/K	$\lambda = (N/V)\lambda_t^3$	Comportamiento
He líquido	4	1.8	Cuántico
He gaseoso	4	0.15	Cuántico
He gaseoso	20	$2.7 \cdot 10^{-3}$	Clásico
He gaseoso	100	$4.8 \cdot 10^{-5}$	Clásico
Ne líquido	27	$1.5 \cdot 10^{-2}$	Clásico
Ne gaseoso	27	$1.1 \cdot 10^{-4}$	Clásico
Ar líquido	86	$7.0 \cdot 10^{-6}$	Clásico
Ar gaseoso	86	$2.2 \cdot 10^{-6}$	Clásico
Electrones de conducción en Ag	300	$6.4 \cdot 10^3$	Cuántico

6.4.2. Movimiento de rotación

Los posibles valores cuánticos de la energía cinética de rotación en torno a un eje de una molécula con momento de inercia I respecto de ese eje son (rotor cuántico):

$$\begin{aligned} \varepsilon_{\text{rot},\ell} &= \ell(\ell + 1) \frac{h^2}{8\pi^2 I} = \ell(\ell + 1) k_B T_{\text{rot}}^* \\ \text{con: } \ell &= 0, 1, 2, \dots \text{ y degeneración: } \Omega_\ell = (2\ell + 1) \\ m_\ell &= -\ell, \dots, 0, \dots, +\ell \end{aligned} \quad (6.30)$$

donde se ha definido una temperatura característica de rotación T_{rot}^* mediante:

$$\frac{h^2}{8\pi^2 I} = k_B T_{\text{rot}}^* \Rightarrow T_{\text{rot}}^* = \frac{h^2}{8\pi^2 I k_B} \quad (6.31)$$

La energía disponible en el gas a temperatura T se reparte por los distintos grados de libertad de todas las moléculas. La mayor parte de las moléculas tiene energías próximas a su valor promedio y hay

relativamente muy pocas moléculas que superen considerablemente este valor. A temperaturas bajas, no hay suficiente energía en el gas para poder excitar los niveles de energía rotacionales de un número suficientemente grande de moléculas. Cuando la temperatura alcanza valores próximos a T_{rot}^* , empieza a aumentar el número de moléculas con energía suficiente para excitar sus niveles rotacionales. El límite clásico, en el que la gran mayoría de moléculas ya pueden rotar, se alcanza cuando $T \gg T_{\text{rot}}^*$. Cuando esto se cumple, la energía promedio por molécula es ya suficientemente grande como para que la distancia entre niveles rotacionales sea despreciable en comparación con ella y el espectro de energías de rotación pueda considerarse como continuo. Esta región clásica se corresponde con la región de la Figura 6.1 para temperaturas superiores a la del punto 2.

En la Tabla 6.4 se recogen las temperaturas características de rotación de algunas moléculas. Como se ve, salvo ciertas moléculas pequeñas que involucran hidrógeno, todas tienen temperaturas características muy próximas al cero absoluto. Eso significa que a temperatura ambiente la inmensa mayoría de las moléculas tienen excitados los niveles rotacionales y su espectro puede ya considerarse continuo, comportándose clásicamente.

6.4.3. Movimiento de vibración

Las oscilaciones acopladas de cualquier sistema mecánico compuesto por varias partículas pueden descomponerse en sumas ponderadas de ciertos modos de oscilación independientes, denominados **modos normales de vibración**, cada uno con una frecuencia bien definida llamada **frecuencia normal**. En el caso de las moléculas de los gases que estamos estudiando, el número de modos normales coincide con el número de grados de libertad de vibración posibles.

La energía de uno cualquiera de estos modos normales de vibración, de frecuencia f , está cuantizada y puede tomar los valores que proporciona el oscilador armónico cuántico (u oscilador de Einstein):

$$\varepsilon_{\text{vib},n} = \left(n + \frac{1}{2}\right) hf = \left(n + \frac{1}{2}\right) k_B T_{\text{vib}}^* \quad (6.32)$$

con: $n = 0, 1, 2, \dots$ y degeneración: $\Omega_n = 1$

Análogamente al caso de la rotación, se ha definido una temperatura característica de vibración del modo normal T_{vib}^* :

$$hf = k_B T_{\text{vib}}^* \Rightarrow T_{\text{vib}}^* = hf/k_B \quad (6.33)$$

Las moléculas empezarán a vibrar según este modo normal en cantidades macroscópicas cuando la temperatura alcance valores del orden de T_{vib}^* , y el límite clásico se alcanzará cuando $T \gg T_{\text{vib}}^*$, en cuyo caso podrá considerarse continuo el espectro de vibración. Esto corresponderá a la región de temperaturas mayores que las del punto 3 de la Figura 6.1.

En la Tabla 6.4 se recogen las temperaturas y energías características de vibración de algunos gases. Como puede verse, todas ellas son **mayores que la temperatura ambiente**, y algunas de ellas de varios miles de Kelvin. En consecuencia, **a temperatura ambiente NO están excitados la mayoría de los modos normales de vibración, y este movimiento no puede considerarse como clásico**. En el caso del O_2 y el CO (moléculas diatómicas que se recogen en Tabla 6.2), las temperaturas características son mayores que 2000 K, lo que justifica que los resultados teóricos más próximos a los experimentales que aparecen en la Tabla 6.2 se obtengan suponiendo que las moléculas son rígidas.

Tabla 6.4: Temperaturas y energías características de rotación y vibración de algunas moléculas. La degeneración de los modos normales de vibración se indica entre paréntesis junto a la temperatura o energía característica del modo normal correspondiente. Las energías características se suelen expresar en electronvoltios (eV). $1 \text{ eV} = 1.602 \ 176 \ 634 \times 10^{-19} \text{ J}$.

Gas	$T_{\text{rot}}^*/\text{K}$	$k_{\text{B}}T_{\text{rot}}^*/\text{eV}$	$T_{\text{vib}}^*/\text{K}$	$k_{\text{B}}T_{\text{vib}}^*/\text{eV}$
Moléculas diatómicas				
H ₂	85.35	$7.36 \cdot 10^{-3}$	5977.7	0.515
D ₂	43.03	$3.71 \cdot 10^{-3}$	4302.4	0.371
O ₂	2.07	$1.78 \cdot 10^{-4}$	2238.9	0.193
N ₂	2.86	$2.47 \cdot 10^{-4}$	3351.6	0.289
F ₂	1.28	$1.10 \cdot 10^{-4}$	1282.0	0.111
Cl ₂	0.345	$2.97 \cdot 10^{-5}$	795.8	0.069
Br ₂	0.116	$1.00 \cdot 10^{-5}$	462.0	0.040
I ₂	0.054	$4.66 \cdot 10^{-6}$	306.9	0.026
HF	29.58	$2.55 \cdot 10^{-3}$	5696.3	0.491
HCl	15.02	$1.29 \cdot 10^{-3}$	4151.6	0.358
HBr	12.01	$1.04 \cdot 10^{-3}$	3681.5	0.317
HI	9.25	$7.97 \cdot 10^{-4}$	3207.9	0.277
HD	64.26	$5.54 \cdot 10^{-3}$	5218.8	0.450
CO	2.76	$2.38 \cdot 10^{-4}$	3082.8	0.266
NO	2.44	$2.10 \cdot 10^{-4}$	2698.7	0.233
Moléculas con más de 2 átomos lineales				
CO ₂	0.563	$4.85 \cdot 10^{-5}$	960(2)	0.083(2)
			1932	0.167
			3380	0.291
N ₂ O	0.606	$5.22 \cdot 10^{-5}$	848	0.073
			1837	0.158
			3199	0.276
HCN	2.137	$1.84 \cdot 10^{-4}$	1024(2)	0.088(2)
			3222	0.278
			4770	0.411
S ₂ C	0.157	$1.35 \cdot 10^{-5}$	571(2)	0.049(2)
			947	0.082
			2205	0.190
OSC	0.292	$2.52 \cdot 10^{-5}$	754(2)	0.065(2)
			1236	0.107
			2970	0.256
C ₂ H ₂	1.693	$1.46 \cdot 10^{-4}$	880(2)	0.076(2)
			1049(2)	0.090(2)
			2840	0.245
			4722	0.407
			4854	0.418
Moléculas con más de 2 átomos no lineales				
H ₂ O	13.4	$1.15 \cdot 10^{-3}$	2290	0.197
	20.9	$1.80 \cdot 10^{-3}$	5160	0.445
	40.1	$3.46 \cdot 10^{-3}$	5360	0.462

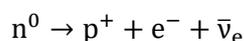
6.4.4. Temperaturas características

Como se puede inferir de todo lo anterior, siempre es posible definir una temperatura característica para cada energía de excitación de un sistema físico, cuya interpretación física es el orden de magnitud de la temperatura a la cual el fenómeno caracterizado comienza a “activarse”. Veamos algunos ejemplos.

Al considerar la estructura interna de los átomos que forman las moléculas aparecen otros grados de libertad además de los de traslación, rotación y vibración vistos hasta ahora. Estos nuevos grados de libertad tienen que ver con la contribución a la entropía del espín nuclear o de los niveles electrónicos moleculares a todas las funciones termodinámicas (se verá en cursos posteriores).

Un orden de magnitud típico de la temperatura característica de excitación de los niveles electrónicos es 10^4 K (bastante grande); la energía de excitación correspondiente sería de 0.862 eV.

Para hacernos una idea del orden de magnitud de la temperatura de excitación de los estados nucleares, consideremos un neutrón libre, que es inestable y sufre un proceso de desintegración β^- emitiendo un protón, un electrón y un antineutrino electrónico:



La energía de este proceso, basándonos en las masas en reposo de neutrón, protón y electrón, es 0.782343 MeV. La diferencia de masas en reposo entre el neutrón y los productos de desintegración se transforma en una energía cinética de 0.782343 MeV. La temperatura característica de este proceso es, por tanto, de $9.075 \cdot 10^9$ K. ¡Enorme! De modo que estos grados de libertad requieren muchísima energía para activarse.

Otro ejemplo en el que se puede definir una temperatura característica es la absorción infrarroja de los sólidos. Supongamos un sólido que absorbe un fotón de frecuencia $f = 10$ THz. La energía característica de esta excitación es de $hf = 0.041$ eV, y su temperatura característica de 479.4 K.

6.5. Complemento: grados de libertad de un sólido rígido

Consideremos en primer lugar los **sólidos rígidos lineales**, esto es, aquellos en los que todas las partículas están localizadas sobre una misma recta. Consideremos primero una sola partícula. Para dar su posición necesitamos 3 variables, por ejemplo, sus coordenadas cartesianas. Situamos ahora una segunda partícula. Su posición tiene 3 coordenadas, pero una de ellas es conocida, ya que sabemos que la distancia a la primera partícula es constante. La posición de la segunda partícula se encuentra sobre una esfera de radio conocido alrededor de la primera y para dar una posición sobre una esfera solo necesitamos 2 variables (latitud y longitud, por ejemplo). Colocamos ahora una tercera partícula colineal con las otras dos. Su posición debe encontrarse en la intersección entre las dos esferas cuyos radios son las distancias a la primera y a la segunda partícula. Como las tres partículas son colineales, esta intersección debe pertenecer a la misma recta que los centros de las otras dos esferas. Para que la intersección de dos esferas esté alineada con sus centros, es necesario que las esferas sean tangentes (pues, de lo contrario, su intersección sería una circunferencia cuyo centro estaría en la línea que une los centros de las esferas). En consecuencia, la intersección de las esferas se reduce a un punto, que es la posición de esa tercera partícula, y por tanto no necesitamos más variables para caracterizarla. Lo mismo ocurre con las siguientes partículas. En consecuencia, **un sólido rígido lineal tiene 5 grados de libertad**.

Antes de proceder con los no lineales, demostremos un resultado auxiliar: *la posición de un punto cualquiera en un plano queda completamente determinada especificando las distancias de dicho punto a otros tres puntos del plano no colineales*. Veamos por qué. Sean A, B y C tres puntos no colineales del plano, y sea D un punto cualquiera de ese plano, situado a distancias r_A, r_B y r_C de los puntos anteriores respectivamente. Las circunferencias de radios r_A, r_B y r_C centradas en A, B y C respectivamente deben

cortarse, al menos, en el punto D . Probemos que D es el único punto en el que pueden cortarse. Hay dos posibilidades:

- Si dos de las distancias son iguales, entonces las circunferencias correspondientes son tangentes, y el punto de tangencia, que es el único punto en el que las tres circunferencias pueden cortarse, debe ser D .
- Si todas las distancias son distintas, entonces dos cualesquiera de las circunferencias, pongamos las centradas en A y B , deben cortarse en dos puntos (Figura 6.2). Para que la circunferencia centrada en C pase por esos dos puntos de corte, la distancia de C a ambos puntos debería ser la misma; en otras palabras, C debería ser parte de la mediatriz de esos dos puntos. Pero tanto A como B , por la misma razón, pertenecen a esa mediatriz, por lo que C debería ser colineal con ellos, lo que es absurdo ya que hemos supuesto que A , B y C no son colineales. En consecuencia, la circunferencia centrada en C solo puede pasar por uno de los dos puntos de corte, que por tanto será el punto D .

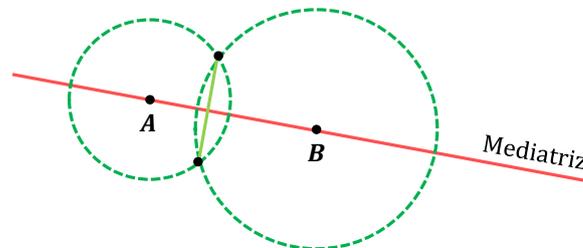


Figura 6.2: La mediatriz de los puntos de corte de dos circunferencias no tangentes pasa por sus centros.

Ahora, calculemos el número de grados de libertad de un **sólido rígido no lineal**. Consideremos primero una sola partícula. Para las dos primeras partículas, el razonamiento es igual que para los lineales: necesitamos 3 variables para la primera y 2 para la segunda. Una tercera partícula no colineal con las anteriores se encuentra a distancias conocidas de la primera y de la segunda y, por tanto, la posición de esta tercera partícula se encuentra sobre la circunferencia intersección de dos esferas de radio conocido alrededor de la primera y la segunda partícula, y para localizar una posición concreta sobre una circunferencia solo necesitamos 1 variable.

En este punto vemos que necesitamos al menos 6 variables. Veamos que no hacen falta más y que, dadas las posiciones de tres partículas no colineales, queda completamente determinada la posición del sólido rígido. Para ello, razonaremos por reducción al absurdo. Supongamos que existen dos posiciones diferentes del sólido para las cuales las tres partículas 1, 2 y 3 (no colineales) ocupan la misma posición. Tomemos otra partícula 4, arbitraria, de la cual conocemos, por supuesto, las distancias a las otras tres partículas. Si la partícula estuviese situada en el plano determinado por las tres anteriores, entonces ya hemos demostrado que su posición está unívocamente determinada dadas las posiciones de 1, 2 y 3 y las distancias de 4 a ellas; supondremos por tanto que la partícula 4 está fuera del plano determinado por éstas. Llamemos A , B y C respectivamente a los puntos donde se encuentran las partículas 1, 2 y 3.

Como las distancias a la partícula 4 son conocidas, ésta solo puede ocupar dos posiciones, D_1 y D_2 , que son los vértices de dos pirámides triangulares cuya base es el triángulo ABC (Figura 6.3). Pero dos posiciones diferentes de un sólido rígido han de poder estar relacionadas por una composición de traslaciones y rotaciones y, como las pirámides son simétricas respecto del plano determinado por ABC , dicha relación no es posible. Llegamos por tanto a un absurdo y hemos de concluir que la posición de la cuarta partícula está unívocamente determinada por las posiciones de las otras tres.

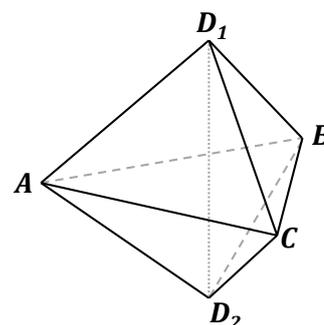


Figura 6.3: Posibles posiciones de la cuarta partícula del sólido rígido.

Por tanto, **un sólido rígido no lineal tiene 6 grados de libertad**.

Ejercicios - Teorema de equipartición de la energía y gas ideal poliatómico

1.- Calcular la energía cinética media de traslación de una molécula de gas helio a 40 °C en electronvoltios.
R: $4.05 \cdot 10^{-2}$ eV.

2.- En un balón de 1 L existen 10^{20} moléculas monoatómicas cuya energía cinética media es de 10^{-20} J por molécula. Calcular la presión del gas.
R: 666.7 Pa.

3.- Un vaso cerrado contiene agua líquida en equilibrio con su vapor a 100 °C y 1 atm. Una masa de 1 g de vapor de agua a esta temperatura y presión ocupa un volumen de 1610 cm³. El calor latente de vaporización específico del agua en estas condiciones es 539.55 cal·g⁻¹.

a) ¿Cuántas moléculas hay en 1 cm³ de vapor?

b) ¿Qué porcentaje representa la energía cinética media de traslación de una molécula de vapor respecto de la energía promedio que necesita para escapar del líquido?

R: a) $1.97 \cdot 10^{19}$; b) 10.84%.

4.- Suponiendo que la temperatura máxima de la superficie de la Luna sea de 150 °C, discutir si es posible que nuestro satélite pueda tener una atmósfera de hidrógeno. Datos: radio de la Luna: $R_L = 0.27 \cdot R_T$; radio de la Tierra: $R_T = 6.37 \cdot 10^6$ m; masa de la Luna: $M_L = 0.012 M_T$; masa de la Tierra: $M_T = 5.98 \cdot 10^{24}$ kg; constante de gravitación universal: $G = 6.67 \cdot 10^{-11}$ kg⁻¹·m³·s⁻²; masa del átomo de H: $m_H = 1.6725 \cdot 10^{-27}$ kg.

5.- Un gas que se compone de moléculas q -atómicas se encuentra a una temperatura T a la que se excitan todos los grados de libertad de las moléculas (de traslación, de rotación y vibracionales). Hallar, para moléculas lineales y no lineales:

a) La energía media de una molécula de este gas.

b) La fracción de esta energía que corresponde al movimiento de traslación.

c) La capacidad calorífica isócara molar.

d) La capacidad calorífica isóbara molar.

e) El coeficiente adiabático.

R: Lineales: a) $(3q - 5/2)k_B T$; b) $1/(2q - 5/3)$; c) $(3q - 5/2)R$; d) $(3q - 3/2)R$; e) $(6q - 3)/(6q - 5)$.

No lineales: a) $(3q - 3)k_B T$; b) $1/(2q - 2)$; c) $(3q - 3)R$; d) $(3q - 2)R$; e) $(3q - 2)/(3q - 3)$.

6.- Determinar el número de grados de libertad totalmente excitados de una molécula de gas si su densidad en condiciones normales es $\rho = 1.3$ kg·m⁻³ y la velocidad de propagación del sonido en él es $c = (\gamma \cdot p / \rho)^{1/2} = 330$ m·s⁻¹ (p es la presión y γ el coeficiente adiabático).

R: 5.

7.- Determinar la relación entre la velocidad del sonido en un gas, $c = (\gamma \cdot p / \rho)^{1/2}$, y la velocidad cuadrática media de las moléculas de éste, si estas son: a) monoatómicas; b) diatómicas rígidas.

R: a) 74.54%; b) 68.31%.

8.- Dados los calores específicos a volumen y presión constante de un gas perfecto ($c_V = 0.65$ J·g⁻¹·K⁻¹ y $c_p = 0.91$ J·g⁻¹·K⁻¹), hallar el n° de grados de libertad totalmente excitados de las moléculas y su masa molar.

R: 5; 32 g·mol⁻¹.

9.- Un gas de moléculas diatómicas rígidas, que se encuentra a una temperatura a la que todos sus posibles grados de libertad están excitados, se expande adiabáticamente hasta que su velocidad cuadrática media disminuye 1.5 veces. ¿En qué factor ha aumentado su volumen?

R: 7.59.

Ejercicios extra para entrenarse

1) El carnicero poco ético

Objetivo: entrenarse en la aplicación del Principio de Máxima Entropía para inferir probabilidades en casos en los que se dispone de información parcial. Esta cuestión fue diseñada por el profesor J.C. Cobos (Departamento de Física Aplicada, Universidad de Valladolid) para el examen final ordinario de Física Estadística del año 1999 (lo que explica por qué los precios del enunciado están en pesetas, y no en euros) a raíz de una crisis de salud pública que tuvo lugar aquel año (la contaminación por dioxinas).

El enunciado original decía así:

“Un carnicero poco ético de nuestra ciudad, alertado por las noticias que han aparecido recientemente en los medios de comunicación, decide eliminar sus existencias de **carne de pollo** importadas desde Bélgica, para lo cual las pone a la venta al público al precio reducido de **100 pts/kg** sin indicar su procedencia.

Los responsables de la salud pública se enteran posteriormente de su dudosa actuación, y procuran conocer el número de personas que pueden caer enfermas como consecuencia de esta. Como el carnicero jura y perjura que ha actuado correctamente, y no proporciona ninguna información adicional que les permita calcular el número de afectados, deciden contratarle a Vd. como científico profesional -¡he aquí uno de *los yacimientos de empleo* para su futuro próximo de los que tanto se habla!- para que les ayude a estimar estos datos (aplicando la metodología *MaxEnt*, evidentemente). La información de partida que le transmiten para su análisis consiste escuetamente en decirle que: en la citada pollería se venden diariamente unos **300 kg** de pollo, siendo habitualmente del tipo "**normal**" (a **200 pts/kg**) o de "**pollo de corral**" (a **600 pts/kg**); y que, en el día que se puso a la venta la partida de pollos belgas, el carnicero hizo una caja de **60000 pts** por este concepto.

Utilizando el **Método de la Entropía Máxima** y la información anterior, *elabore un informe* para las autoridades sanitarias donde aparezca su **estimación** del número de kilogramos de cada tipo de pollo (belga, normal y de corral) que pudo vender el desalmado carnicero.”

Se pide que resuelva esta cuestión interpretando el cociente entre el número de kg de pollo de cada tipo y el número total de kilos como la probabilidad asociada a cada tipo de pollo, que hay que *inferir*. Esta cuestión es formalmente idéntica al juego de la ruleta resuelto en clase, siendo en este caso la **información disponible** el **valor medio del precio del kg de pollo** (que hay que calcular a partir de los datos del enunciado). Se recomienda que busque unas unidades adecuadas para trabajar, de modo que los números en el problema de inferencia estadística sean próximos a la unidad.

Este tipo de ejercicios ilustra la **generalidad del Principio de Máxima Entropía**, cuya aplicación no se limita, ni mucho menos, al campo de la Física Estadística.

2) Distribución de velocidades en la efusión a través de un orificio

Objetivo: familiarizarse con el uso de las funciones de densidad de probabilidad, que cuantifican la probabilidad cuando los valores posibles de una variable forman un continuo.

En un recipiente que contiene un gas ideal monoatómico de moléculas de masa m , en equilibrio termodinámico a una temperatura T , se abre un orificio de área ΔS . Como consecuencia de ello, comienzan a salir partículas del recipiente. Se supone que el orificio es muy pequeño y el tiempo de observación (Δt) es suficientemente corto, de tal modo que se puede considerar que el gas en el interior está constantemente en equilibrio y que el número de moléculas en el interior no ha disminuido apreciablemente. Sin pérdida de generalidad, se supone que la superficie del orificio es perpendicular al eje Z. El objeto de este problema será el **estudio de la distribución de probabilidad de las velocidades de las moléculas que salen por el orificio, que será diferente que la distribución en el interior, debido a que el gas que emerge no se encuentra en equilibrio termodinámico.**

Llamemos: $D_f(\vec{v}) = D_f(v_x, v_y, v_z)$ a la función de densidad de probabilidad (FDP) de las velocidades de las moléculas que salen del recipiente, en coordenadas cartesianas (el subíndice “f” hace referencia a “fuera” del recipiente). A continuación, damos los detalles de cómo podemos obtener esta función haciendo uso de las reglas de la probabilidad condicionada y de lo que hemos estudiado en esta asignatura, aunque esta parte no es imprescindible para las preguntas que se le pide resolver después (lo que interesa de este ejercicio es operar con FDPs). La función $D_f(v_x, v_y, v_z)$ indica la densidad de probabilidad del suceso [tener velocidad \vec{v}] **condicionada** a que haya ocurrido el suceso [salir], lo que indicamos con la notación:

$$D_f(v_x, v_y, v_z) = \text{DPr}(\text{tener } \vec{v} | \text{salir})$$

donde DPr es una forma abreviada de indicar “densidad de probabilidad”. De acuerdo con las reglas de la probabilidad condicionada, tenemos:

$$\text{DPr}(\text{tener } \vec{v} | \text{salir}) = \frac{\text{DPr}(\text{tener } \vec{v} \cap \text{salir})}{\text{Pr}(\text{salir})} = \frac{\text{DPr}(\text{tener } \vec{v}) \cdot \text{Pr}(\text{salir} | \text{tener } \vec{v})}{\text{Pr}(\text{salir})}$$

Esto no es más que la aplicación del célebre **Teorema de Bayes** de la Teoría de la Probabilidad a nuestra situación particular (busque información sobre ello para comprobarlo). Se trata de un razonamiento muy típico que se utiliza para obtener una probabilidad desconocida cuando se conocen otras (en otras palabras, de **inferir probabilidades a partir de información conocida**). Y, efectivamente, todos los factores del último miembro los conocemos del estudio de un gas ideal monoatómico en equilibrio:

- La densidad de probabilidad del suceso [tener velocidad \vec{v}], $\text{DPr}(\text{tener } \vec{v})$, es, simplemente, la densidad de probabilidad de las velocidades de las moléculas del gas *en equilibrio*, $D(v_x, v_y, v_z) = f(v) = F(v)/(4\pi v^2)$ (se ha usado la notación del capítulo de Teoría Cinética de los gases ideales).
- La probabilidad del suceso [salir del recipiente], $\text{Pr}(\text{salir})$, será igual a la fracción de moléculas que atraviesan el orificio. El numerador de esta fracción lo podemos obtener a partir de la **densidad de flujo molecular promedio** η (número de moléculas que, para el gas en equilibrio, atraviesan la unidad de superficie en la unidad de tiempo), estudiada en el capítulo de Teoría Cinética de los gases ideales. El denominador es el número total de moléculas, N .

$$\text{Pr}(\text{salir}) = \frac{\eta \Delta S \Delta t}{N} = \frac{\Delta S \Delta t}{N} \cdot \frac{1}{4} n \langle v \rangle, \quad \text{donde} \quad \begin{cases} n = N/V \\ \langle v \rangle = \text{valor medio del módulo de } \vec{v} \text{ en equilibrio} \end{cases}$$

- La probabilidad del suceso [salir del recipiente] condicionada a que haya ocurrido el suceso [tener velocidad \vec{v}], $\text{Pr}(\text{salir} | \text{tener } \vec{v})$, se obtiene como la fracción de moléculas que salen del recipiente, en el supuesto previo de que todas tienen velocidad \vec{v} . El numerador de esta fracción

lo podemos obtener a partir de la **densidad de flujo molecular supuesto que todas las moléculas tienen velocidad** \vec{v} , $\eta_{\vec{v}}$, también estudiada en el capítulo de Teoría Cinética de los gases ideales. Recuerde que $\eta_{\vec{v}}$ es la **solución del problema mecánico** asociado a la densidad de flujo molecular. El denominador es el número total de moléculas, N .

$$\Pr(\text{salir}|\text{tener } \vec{v}) = \frac{\eta_{\vec{v}} \Delta S \Delta t}{N} = \begin{cases} \frac{\Delta S \Delta t}{N} n v \cos \theta & , \theta \leq \pi/2 \\ 0 & , \theta > \pi/2 \end{cases} , \theta = \text{ángulo entre } \vec{v} \text{ y el eje Z}$$

Sustituyendo, obtenemos:

$$D_f(v_x, v_y, v_z) = \frac{F(v) \eta_{\vec{v}}}{4\pi v^2 \eta} = \begin{cases} \frac{F(v) \cos \theta}{\pi v \langle v \rangle} & , \theta \leq \pi/2 \\ 0 & , \theta > \pi/2 \end{cases}$$

Esta es la función sobre la que vamos a trabajar. No va a ser necesario sustituir el valor de $F(v)$ en los apartados a, b y c, aunque sí en el apartado d. Recuerda que $F(v)$ es la **FDP de la celeridad** (módulo de la velocidad) **de un gas ideal clásico en equilibrio**, que define la llamada "**distribución de velocidades de Maxwell**" (ecuación (5.26))

Se pide resolver las siguientes cuestiones:

a) Haciendo un cambio de variables en tres dimensiones, calcule la correspondiente **FDP en coordenadas esféricas**, que llamaremos $\rho_f(v, \theta, \phi)$.

b) Calcule, mediante una integración apropiada de la función $\rho_f(v, \theta, \phi)$, la FDP del **módulo** de la velocidad de las moléculas que salen, que llamaremos $F_f(v)$, escribiéndola en función de $F(v)$, v y $\langle v \rangle$.

c) A partir de $F_f(v)$ y conociendo los valores medios $\langle v \rangle = \sqrt{\frac{8k_B T}{\pi m}}$ y $\langle v^2 \rangle = \frac{3k_B T}{m}$ en equilibrio, calcule el **valor medio de la celeridad de las moléculas que salen**, que llamaremos $\langle v \rangle_f$. ¿Es mayor o menor que $\langle v \rangle$? ¿Podría dar una explicación física cualitativa de este hecho?

d) Escriba explícitamente la forma de la función $F_f(v)$ sustituyendo el valor de $F(v)$ en la expresión obtenida en el apartado b. Haciendo un cambio de variable $v_r = v/v_p$, donde $v_p = \sqrt{\frac{2k_B T}{m}}$ es la celeridad más probable del gas en equilibrio, calcule la **FDP de la celeridad reducida del gas en efusión**, que llamaremos $F_{f,r}(v_r)$, **demostrando que en ella no aparecen explícitamente m , k_B ni T** . **Represente gráficamente $F_{f,r}(v_r)$ junto con la FDP de la celeridad reducida del gas en equilibrio ($F_r(v_r)$, ecuación (5.51))** y analice las diferencias encontradas.

3) Cristal aislante monoatómico ideal en el límite semiclásico

Objetivo: comprender mejor el formalismo general de la Física Estadística, que permite **obtener las propiedades termodinámicas macroscópicas asociadas a un modelo microscópico (UN OBJETIVO CENTRAL DE LA ASIGNATURA)**, en un caso concreto en el que el modelo obedece a la Mecánica Clásica. Para ello, debe estudiar con detalle el **procedimiento de trabajo en Física Estadística**.

① **MACROESTADO:** (β, V, N) .

② **MODELO:** Vamos a estudiar las vibraciones de los átomos de la red cristalina de un sólido aislante monoatómico en el límite semiclásico, en el **conjunto canónico**.

Trabajaremos con la **aproximación armónica**, es decir, supondremos que las vibraciones pueden describirse como un conjunto de N átomos efectuando oscilaciones armónicas en las tres direcciones del espacio, por supuesto, acopladas entre sí: las oscilaciones de un átomo en una dirección dependen de sus oscilaciones en las otras dos direcciones y también de las oscilaciones de los átomos vecinos. Es posible demostrar (cosa que se verá en la asignatura “Mecánica y Ondas” de 2º curso) que se puede hacer un cambio de coordenadas para encontrar un **conjunto equivalente de $3N$ osciladores armónicos unidimensionales desacoplados**, denominados **modos normales de vibración**.

Supongamos que el conjunto de modos normales tiene coordenadas $x = (x_1, \dots, x_{3N})$ y momentos $p_x = (p_{x,1}, \dots, p_{x,3N})$. La energía del sistema puede escribirse:

$$E(x, p_x) = \sum_{i=1}^{3N} \left(\frac{p_{x,i}^2}{2m} + \frac{1}{2} k_i x_i^2 \right) = \sum_{i=1}^{3N} \left(\frac{p_{x,i}^2}{2m} + \frac{1}{2} m \omega_0^2 x_i^2 \right)$$

donde m es la masa atómica y k_i es la constante elástica de cada oscilador. Se han tomado todas las constantes elásticas iguales a un mismo valor $k = m \omega_0^2$, lo que es equivalente a tomar iguales las frecuencias angulares de oscilación $\omega_0 = 2\pi f_0$ (donde f_0 es la frecuencia lineal de oscilación).

Esta aproximación no da cuenta del volumen del sistema (Ecuación Térmica de Estado), por lo que no aparece como variable en el macroestado. En consecuencia, el macroestado se simplifica en la forma **(β, N)** .

③ **FUNCIÓN ESTRUCTURA:**

3.a) Encuentre la **entidad elemental independiente (EEI)** del sistema, justificando la respuesta a partir de la definición de EEI, e indique claramente:

- Su energía, $\varepsilon(x, p_x)$.
- El número de ellas que hay en el sistema, N_{EEI} .
- Si son distinguibles o indistinguibles.
- El número de grados de libertad q tiene cada EEI, indicando cuántos hay de cada tipo (traslación, rotación, vibración).
- El número de grados de libertad, s , del sistema total y cuántos hay de cada tipo.

3.b) Identifique los valores de los **monoestados** (x, p_x) de la EEI que son compatibles con su macroestado (β) . Calcule la **función de partición de una EEI**, $Z_1(\beta)$ (aparecerán unas integrales que puede resolver recurriendo a la sección 5.4).

3.c) Calcule la **función de partición del sistema**, $Z(\beta, N)$, prestando atención a la distinguibilidad o indistinguibilidad de las EEI.

④ **CÁLCULOS:**

4.a) Calcule la **función de Massieu del sistema**, ψ , y la función de Massieu de una EEI, ψ_1 , verificando que es extensiva: $\psi = N_{EEI} \cdot \psi_1$.

4.b) Obtenga las **ecuaciones de estado** del sistema: la Ecuación Energética de Estado (EEE) y la Ecuación Material de Estado (EME). Razone si se cumple el **Teorema de Equipartición de la Energía**, relacionando las hipótesis del modelo con las hipótesis del teorema.

4.c) Como el modelo no incorpora la posibilidad de intercambio de trabajo de expansión (variación del volumen V), la única capacidad calorífica que se puede calcular es la **capacidad calorífica a volumen constante**, C_V , que es una magnitud extensiva (aditiva). Calcúlela, y halle también la **capacidad calorífica a volumen constante molar**, C_{Vm} , que es una magnitud intensiva, demostrando que se cumple la **ley de Dulong y Petit**,

$$C_{Vm} = \frac{C_V}{n} = 3R,$$

siendo n la cantidad de átomos (en mol) y R la constante de los gases.

4.d) Encuentre la **entropía** S del sistema, y discuta en qué intervalo de temperaturas es válido el modelo, suponiendo que la frecuencia de los osciladores es $f_0 = 5$ THz.

Utilice, para la discusión, los valores de las constantes de acuerdo con la última edición del Sistema Internacional de Unidades (del 20 de mayo de 2019):

- Constante de Planck: $h = 6,626\ 070\ 15 \cdot 10^{-34}$ kg·m²·s⁻¹.
- Constante de Boltzmann: $k_B = 1,380\ 649 \cdot 10^{-23}$ kg·m²·s⁻²·K⁻¹.

4.e) Transforme la anterior expresión de la entropía hasta convertirla en un **potencial termodinámico** $S = S(U, N)$, y obtenga a partir de ella las ecuaciones de estado, comprobando que coinciden con las obtenidas en el apartado 4.b.

Bibliografía

- J. Aguilar Peris.** “Curso de Termodinámica”, Alhambra Longman, 1989.
- M. Alonso, y E.J. Finn.** “Física. Volumen III: Fundamentos cuánticos y estadísticos”. Fondo Educativo Iberoamericano, 1971.
- A. Ben-Naim.** “La entropía desvelada”, Colección Metatemas (nº 118). Tusquets. Barcelona, 2011.
- J. Biel Gayé.** “Formalismo y métodos de la Termodinámica”, Vol. 1 y 2, Reverté, 1998.
- J. Brey Ábalo, J. de la Rubia Pacheco, J. de la Rubia Sánchez,** “Mecánica Estadística”, UNED, 2001.
- H.B. Callen.** “Termodinámica”. Alfa Centauro, 1985.
- M. Díaz Peña.** “Termodinámica estadística”. Alhambra, 1979.
- B. Diu, C. Guthmann, D. Lederer, B. Roulet.** “Physique Statistique”, Hermann, 1989.
- B. Diu, C. Guthmann, D. Lederer, B. Roulet.** “Thermodynamique”, Hermann, 2007.
- L. González Baixauli, J.C. Cobos Hernández.** “Conjuntos de Gibbs. El Método de Entropía Máxima en Mecánica Estadística”. Universidad de Valladolid, 1996.
- E.T. Jaynes.** “Where Do We Stand on Maximum Entropy?” Páginas 15-118 del libro: “The Maximum Entropy Formalism”, editado por D. Levine, M. Tribus. The M.I.T. Press, 1979.
- L.D. Landau, E.M. Lifshitz,** “Física Estadística”, Volumen 5 del Curso de Física Teórica. Reverté, 1969.
- P.T. Landsberg.** “Thermodynamics and Statistical Mechanics”, Dover Publications, Inc., 1990.
- F.W. Sears, G.L. Salinger.** “Termodinámica, Teoría Cinética y Mecánica Estadística” (Segunda Edición), Reverté, 1978.
- C.E. Shannon, W. Weaver.** “The Mathematical Theory of Communication”. Univ. Illinois Press, 1969.
- N.O. Smith.** “Elementary Statistical Thermodynamics: A problems approach”. Plenum Press, New York, 1982.
- P.A. Tipler, G. Mosca.** “Física para la Ciencia y la Tecnología, Vol. 1”, Reverté, 6ª edición, 2010.

