

UNIVERSIDAD DE LOS ANDES



Revisión de los fundamentos de la mecánica estadística desde la perspectiva de la teoría cuántica de la información

TESIS

QUE PARA OBTENER EL TÍTULO DE

FÍSICO

PRESENTA

GERMÁN EDUARDO OSORIO LEIVA

ASESOR: ALONSO BOTERO MEJÍA

BOGOTÁ, D.C.

2017

AUTOR

FECHA

FIRMA

Índice general

1. Introducción	1
2. Perspectivas	3
2.1. Recuento Histórico	3
2.1.1. Termodinámica	4
2.1.2. Teoría Cinética	7
2.2. Boltzmann	9
2.3. Gibbs	11
2.3.1. Ensemble Canónico	14
2.4. Teorma de Liouville	16
2.4.1. Consecuencias del Teorema de Liouville	17
2.5. Ergodicidad	19
2.6. Jaynes	23
3. El Entrelazamiento y La Mecánica Estadística	29
3.1. Herramientas de la Teoría Cuántica	29
3.1.1. Operador de Densidad	29
3.1.2. Operador de Densidad Reducido	32
3.1.3. Distancia de Traza	34
3.2. Definiciones	38
3.3. Idea Conceptual	39
3.4. Formulación Matemática	41
3.4.1. Lema de Levy	42
3.4.2. Demonstración del Principio General Canónico	45
3.5. Significado Físico	51

4. Evolución Hacia El Equilibrio	55
4.1. Especificación del Equilibrio	56
4.2. Equilibración	58

Bibliografía	73
---------------------	-----------

Capítulo 1

Introducción

Capítulo 2

Perspectivas

2.1. Recuento Histórico

Para poder explicar las visiones más fundamentales de la mecánica estadística se debe dar un contexto histórico; porque es la forma más clara para ver cómo surgieron estas ideas. Por esto, en esta sección se mostrará las circunstancias y personajes en las que la termodinámica se formuló. Este recuento histórico sigue el libro de I. Muller [22]. Este libro muestra la historia de manera dinámica y accesible, por eso se sigue en este texto. La física actual es iniciada por Newton, luego Hamilton y Lagrange crearon métodos muy poderoso para poder atacar diferentes problemas. La mecánica clásica tiene un amplio alcance dando predicciones acordes con los experimentos. Sin embargo ya se mostraba que esta teoría no podía estudiar todos los aspectos de la naturaleza. Ya, desde siglo 18, se habían dado muchos indicios sobre otra teoría que debía tratar con conceptos diferentes a los tratados comúnmente por la mecánica. Los concepto como el calor y la energía fueron emergiendo, en su momento fueron ideas muy vagas y primitivas; pero la teoría sobre el calor fue evolucionando, siendo guiada por los experimentos de Graf von Rumford, Julius Mayer y James Prescott Joule. Por otro lado, Nicolas Carnot trabajaba en motores de calor y sus consecuencias teóricas, esto dio soporte para que un nuevo concepto naciera, la entropía. Poco a poco la termodinámica fue tomando la cara actual con ayuda de Benoît Clapeyron, Lord Kelvin y Rudolf Clausius. La termodinámica fue basada en los experimentos y organizada bajo unos

postulados [8], sus predicciones eran respaldadas experimentalmente ayudando a la innovación e industria. Esto, le da una esencia diferente de las demás teorías físicas porque no propone nuevas leyes, simplemente exponen qué ocurre con las propiedades de un sistema macroscópico en equilibrio. Por ejemplo, la primera ley de la termodinámica simplemente expone cómo la energía del sistema es constante sin importar que proceso se lleve a cabo. Dado su origen, la termodinámica no se interesaba por encontrar una conexión entre sus resultados y los constituyentes de la materia.

De este problema nace la mecánica estadística, sus esfuerzos iniciales eran considerar la materia hecha de partículas y de allí poder predecir todas las propiedades térmicas macroscópicas ya encontradas por la termodinámica. Las personas que más ayudaron a la creación de esta teoría fueron Gibbs, Boltzmann y Maxwell. Ellos le dieron las ideas principales y moldearon la teoría para poder tener métodos generales a aplicaciones de diferentes tipos, los métodos producidos por Gibbs son los que actualmente se utilizan.

2.1.1. Termodinámica

La mecánica estadística es la base conceptual de la termodinámica; pero históricamente primero apareció la termodinámica. La termodinámica apareció primero ya que fue la agrupación de varias ideas en el siglo diecinueve que encerraban varios tipos de experimentos y nociones que existían desde la antigua Grecia. Aunque las ideas, que dieron paso a la termodinámica, surgieron en la antigüedad se han ido refinando. Por ejemplo Klaudios Galenos (133-200 A.C) suponía que la influencia del clima sobre los fluidos del cuerpo podría determinar el carácter de una persona. Así decía que los habitantes del norte frío y húmedo eran salvajes y violentos, al contrario de los habitantes del sur, caliente y seco, eran flácidos y mansos. Esto muestra que la idea de temperatura ha sido conocida de alguna u otra manera, pero de maneras muy intuitivas.

Ya en el año 1578 Johannis Hasler presentó una tabla de las temperaturas corporales de las personas en relación a la altura en la que vivían. Pero como se conoce ahora mientras se tenga una buena salud no importa la altura se debe tener la misma temperatura corporal. Este error fue debido

a los instrumentos de medición usados en esa época. Ya en el inicio del siglo 17 ya se tenía el termómetro. Gracias a este instrumento muchas ideas equivocadas dadas por la subjetividad de los sentidos sobre el calor y el frío se fueron eliminando. Ya en los años 1700 empezaron a aparecer las escalas de temperatura, hubo incluso 18 escalas en algún tiempo. En 1848 William Thompson (Lord Kelvin) introdujo la escala absoluta o la escala Kelvin. Esta cuenta desde el cero absoluto en adelante, el punto de ebullición del agua a 1 atm es 373.15 Kelvin.

Otra importante parte de la termodinámica es la energía pero en sus inicios no se habla de eso; era calor o fuerza. Pero no se comprendía qué era el calor. La teoría calórica pretendía dar explicación al calor, entre varias propuesta se pueden ver las de Pierre Gassendi (1592-1655) donde proponía que el calor y el frío eran tipos de materia diferente. Consideraba a los átomos del frío como tetraedros y al entrar a un líquido se solidificaría.

Antoine Laurent Lavoisier (1743-1794) consideró al calor como otro elemento, junto con la luz, y lo consideraba como un fluido que llamaba el calórico. Una de las primeras personas en cuestionar la teoría calórica fue Graf Von Rumford(1753-1814), observando unos cañones y cómo cambiaba el calor liberado dependiendo de si estaban afilados o no. Concluyó que el calor debía ser el mismo que fue administrado debido al movimiento.

Rumford siguió trabajando en su idea del equivalente mecánico del calor, con dos caballos moviendo un cabrestante nota que el calor del barril es igual "al de 9 velas grande", más adelante precisó su hallazgo [22]. Rumford siguió haciendo experimentos sobre el calor, midió de manera meticulosa el peso del agua antes y después de ser congelada. Encontró que el peso no cambiaba pero aún así se dio calor en el proceso. Entonces concluye que si el calórico existía era imponderable. Pese a los experimentos realizados por Rumford la teoría calórica siguió siendo de gran aceptación durante muchos años.

Robert Julius Mayer (1814-1878) estudio medicina pero su entusiasmo hacia la física lo guió para hacer experimentos en este campo. La idea principal que tenía Mayer era que la energía se conservaba de forma general, no solo cuando hay trabajo mecánico. Es decir cualquier fenómeno capaz de aportar energía debía tenerse en cuenta al momento de un experimento. Gracias a los estudios hechos por su cuenta sabía que la energía cinética, fuerza viva en sus palabras, podía convertirse en calor. Experimentalmente

llegó a que la caída de un peso a una altura de 365 metros correspondía a calentar el mismo peso de agua de 0° a 1° . No es muy exacto pero estuvo bastante cerca, incluso llegó a cambiar la altura a 425 metros luego de que Joule hiciera mejores mediciones.

Mayer tenía ideas bastante originales aunque en general no sabía expresarlas, por falta de experiencia matemática y su aislamiento de la comunidad científica. Pero Mayer fue el primero en hablar sobre máquinas térmicas y decir que el calor absorbido por el vapor era siempre mayor que el calor liberado durante la condensación y su diferencia era trabajo útil. Esta idea se expuso antes que Carnot y Clapeyron. Sabía como se manejaba el calor pero no conocía su naturaleza.

James Prescott Joule (1818-1889) proporcionó, en extremo detalle, observaciones sobre el calor y la temperatura. Sus experimentos eran discutidos en sus escritos de forma muy extensa, dando presumibles errores, compensando perdidas, sus experimentos eran tan detallados que algunos de sus artículos se volvieron estándares. Es por eso que Mayer revisó sus experimentos respecto a los de Joule. Uno de los aportes de Joule fue que gracias a sus experimentos se podía aceptar la conservación de la energía.

Pero fue Hermann Ludwig Ferdinand Helmholtz (1821-1894) quien propuso la idea de que lo llamado calor era la energía cinética del movimiento térmico de los átomos. Fue Helmholtz el primero en dejar el término de un fuerza viva y llamarlo energía. El trabajo de Helmholtz fue importante ya que Joule y Mayer no podía dar una visión clara sobre lo que se llama calor, y divagaban entre calor y fuerza; Helmholtz dejó los conceptos más claros en comparación con ellos dos.

Todo este proceso histórico fue el camino para poder dar la primera ley de la termodinámica, se puede ver que la historia de la energía no es un desarrollo tan claro y llegar hasta la idea actual no fue para nada rápido ni sencillo. Puede ponerse en contraposición la mecánica cuántica que aunque fue inspirada también de la radiación del cuerpo negro, un tema que en ese entonces era tratado por la termodinámica, se construyó en menos años; aunque con ideas confusas pero su construcción matemática y teórica no fue demorada por siglos.

A diferencia de la energía, la entropía era un concepto más provocador al momento de darle un sentido microscópico. La historia de la entropía empieza por los motores térmicos hay especulaciones que en el primer siglo

antes de cristo ya existían algunas máquinas que funcionaban con vapor de agua, Hero de Alexandría se especula que hizo una de esas máquinas. Pero en el siglo 18 Thomas Newcome y Thomas Savery crearon una máquina de vapor que inicialmente ayudaba a sacar agua de las minas. Fue James Watt(1736-1819) quien mejoró esa máquina de vapor haciendo una que era de 3 a 4 veces más eficiente.

Nicolas Léonard Sadi Carnot (1796-1832) se preguntó sobre que tanto se puede mejorar la eficiencia de una máquina térmica. Carnot publicó un libro donde muestra aseveraciones que intentan resolver la pregunta hecha; Carnot encuentra que: una máquina que trabaja entre dos temperaturas y solo intercambia calor entre ellas, su eficiencia solo depende de la temperatura, $e = F(T)dT$ para un delta de temperatura, incluso aunque las máquinas se manejen con agentes diferentes para generar trabajo.

Carnot seguía creyendo en la teoría calórica para llegar a estas ideas no fue necesario saber que esta teoría estaba incorrecta, aún así sus resultados se mantienen como ciertos. Carnot no pudo encontrar exactamente los valores de la eficiencia. Clapeyron y Kelvin no pudieron tampoco encontrar los valores de esta eficiencia ni por mediciones ni por cálculos.

Ya en 1850 se estaba asegurando que la teoría calórica estaba mal formulada y se debía hacer algo. Clausius retocó algunas ideas de Carnot y pudo encontrar la eficiencia del ciclo de Carnot. Clausius le dio forma a la termodinámica actual, hasta los cursos actuales de termodinámica siguen un artículo de Clausius que habla de gases ideales y vapor húmedo. Clausius amplió sus investigaciones a ciclos no infinitesimales y a procesos no reversibles, él fue quien dio la idea de la entropía y sus propiedades. Con eso propuso la segunda ley de la termodinámica: El calor no puede pasar solo de un cuerpo frío a un cuerpo caliente. Clausius llamó $S = \frac{Q}{T}$ la entropía y vio que esto era lo que se conservaba en un ciclo de Carnot. Pero en términos microscópicos no se conocía el significado de S .

2.1.2. Teoría Cinética

Paralelamente a mitad del siglo 19 se trabajaba en la teoría cinética de los gases. Esta teoría estudiaba las propiedades macroscópicas tales como la presión, temperatura, viscosidad, conductividad térmica etc. de un

número muy grande de partículas que seguían las ecuaciones clásicas del movimiento. Habían dos suposiciones que se hacía en esta teoría. La primera que los gases eran sistemas mecánicos de muchas partículas idénticas. La segunda era que debido a la gran cantidad de partículas se debía introducir probabilidades para así ver algunas regularidades que salían debido a las configuraciones de las moléculas. Personas como Calusius, Maxwell (1831-1879) y Boltzmann (1844-1906) trabajaron en esta área y produjeron resultados importantes pese a que concordaban con los resultados experimentales, hubo muchas discusiones sobre las hipótesis usadas para llegar a estos resultados.

Por ejemplo en los trabajos de Clausius se hace uso de los siguientes supuestos para gases en reposo y en equilibrio térmico: Las moléculas que están dentro de un recipiente se encuentran distribuidas con la misma densidad en todo el recipiente, la distribución de velocidades es igual en todo el recipiente, todas las direcciones de velocidad son igual de probables. Estas tres hipótesis son el inicio para ver los sistemas mecánicos con una perspectiva probabilística.

Maxwell convirtió las ideas de Clausius sobre la poca dispersión de la distribución de velocidad en algo que se podía calcular. Así que Maxwell propuso en 1859 su ley de distribución

$$f(u, v, w) \Delta u \Delta v \Delta w = A e^{-B(u^2 + v^2 + w^2)} \Delta u \Delta v \Delta w. \quad (2.1)$$

Donde $f(u, v, w) \Delta u \Delta v \Delta w$ es el número de moléculas entre esos límites de velocidades, cada límite representa un componente de la velocidad.

Boltzmann en 1868 generalizó el resultado de Maxwell, ahora se tiene la misma situación de un gas en equilibrio y en reposo, pero ahora se tiene un campo de fuerza externo actuando sobre las moléculas. Boltzmann tuvo en cuenta que las moléculas eran compuestas de otras partículas lo cual afecta los diferentes valores de energía potencial. Entonces denotando a $\Delta\tau$ como el rango de variaciones muy pequeñas que puede tener el estado de la molécula, o sea $\Delta\tau = \Delta x \Delta y \Delta z \Delta u \Delta v \Delta w$. La generalización dada por Boltzmann es

$$f(x, y, z, u, v, w) \Delta\tau = \alpha e^{-\beta E} \Delta\tau. \quad (2.2)$$

E es la energía total que tiene la molécula esto es la energía cinética, energía potencial interna y energía potencial externa. Esta generalización es llamada la distribución de Maxwell-Boltzmann. Boltzmann además mostró que

cualquier otra distribución bajo la condición de colisiones entre partículas evoluciona hacia la distribución de Maxwell-Boltzmann, Boltzmann usó el teorema H para concluir esto.

2.2. Boltzmann

En este contexto histórico fue que Boltzmann dio las primeras ideas sobre la mecánica estadística. En esta sección, se expondrán los conceptos más importantes que Boltzmann postuló y ayudaron a formar la física estadística. La meta de Boltzmann era poder dar un enfoque microscópico de la termodinámica, y poder explicar las propiedades térmicas que se observaban en los sistemas macroscópicos. Como primera conexión intentó darle un significado microscópico a la segunda ley de la termodinámica. Esta conexión viene demostrada por el teorema H.

Para hablar sobre teorema H primero se debe explicar la ecuación de transporte de Boltzmann. La ecuación de transporte de Boltzmann puede ser derivada desde la jerarquía de Bogoliubov-Born-Green-Kirkwood-Yvon o mirando el problema del número de partículas en un rango del estado de una molécula en el espacio de fase [6]. Para ambas rutas se debe hacer supuestos para poder seguir con el análisis, esto no se mostrará acá dado que no es pertinente seguir todos los pasos. La ecuación de transporte de Boltzmann describe cómo la distribución de moléculas $f(\vec{r}, \vec{p}, t)$, donde \vec{r} es la posición y \vec{p} es el momento, evoluciona en el tiempo. La ecuación es

$$\left(\frac{\partial}{\partial t} - \frac{\partial U}{\partial \vec{q}_1} \cdot \frac{\partial}{\partial \vec{p}_1} + \frac{\vec{p}_1}{m} \cdot \frac{\partial}{\partial \vec{q}_1}\right)f = \\ - \int d^3\vec{p}_2 d^2\Omega \left| \frac{d\sigma}{d\Omega} \right| |\vec{v}_1 - \vec{v}_2| [f(\vec{p}_1, \vec{q}_1, t)f(\vec{p}_2, \vec{q}_1, t) - f(\vec{p}'_1, \vec{q}_1, t)f(\vec{p}'_2, \vec{q}_1, t)].$$

Esta ecuación se puede leer como [4]: el término a la izquierda describe el movimiento de una partícula en un potencial U , el término de la derecha es la probabilidad de encontrar una partícula con momento \vec{p}_1 en \vec{q}_1 ser alterada por una colisión con una partícula con momento \vec{p}_2 . La probabilidad de esta colisión está dada por el producto de los factores cinéticos proporcionado por la sección transversal diferencial $\left| \frac{d\sigma}{d\Omega} \right|$, el flujo de partículas incidentes $|\vec{v}_1 - \vec{v}_2|$ y la probabilidad de encontrar dos partículas $[f(\vec{p}_1, \vec{q}_1, t)f(\vec{p}_2, \vec{q}_1, t)]$. El primer término subtrae la probabilidad e integra

sobre todos los momentos posibles y el ángulo sólido. El segundo término es la adición del proceso inverso.

Hablar de la ecuación de transporte de Boltzmann es importante para seguir con el teorema H; ya que este dice que: si f satisface la ecuación de transporte de Boltzmann, entonces $\frac{dH}{dt} \leq 0$, donde

$$H(t) = \int d^3\vec{p} d^3\vec{q} f(\vec{p}, \vec{q}, t) \ln(f(\vec{p}, \vec{q}, t)). \quad (2.3)$$

Con este teorema se puede demostrar que la distribución en el equilibrio es la distribución de Maxwell-Boltzmann o sea cuando $t \rightarrow \infty$ la distribución $f(\vec{p}, t) \rightarrow f_0(\vec{p}) = \alpha e^{-\beta E}$ [6]. Pero hay un punto importante aquí y es algo que también en su época desconcertó a los contemporáneos de Boltzmann, ¿cómo es posible llegar a una descripción de fenómenos irreversibles que muestran una dirección temporal pero se empezó con una teoría de un gas reversible? Esto viene dado por las suposiciones hechas al derivar la ecuación de transporte de Boltzmann.

Se usa la hipótesis de caos molecular (Stosszahlansatz). Esta dice que, después de una colisión las partículas, que interactuaron durante el choque, quedan descorrelacionadas. Entonces, por el caos molecular $f(\vec{r}, \vec{p}_1, \vec{p}_2, t) = f(\vec{r}, \vec{p}_1, t) f(\vec{r}, \vec{p}_2, t)$. Es por esta hipótesis que la simetría temporal se rompe y hace que haya una dirección particular para el tiempo. Porque, al descorrelacionar las partículas después de un choque, se rompe la simetría que las ecuaciones de movimiento clásicas tienen. El teorema H fue usado por Boltzmann para poder darle una base microscópica a la segunda ley de la termodinámica. La relación es

$$S(t) = -\frac{H(t)}{k_B} \quad (2.4)$$

donde S es la entropía y k_B la constante de Boltzmann. Pero también Boltzmann recibió muchas críticas sobre la posibilidad de asociar a H con la entropía.

Una de las objeciones que recibió fue sobre la posibilidad de poner todas las velocidades de las moléculas de forma opuesta a las dadas inicialmente, esto daría resultados opuestos a los que se esperaban. Si inicialmente se encontraba que H decrecía ahora al invertir las velocidades se halla que H aumenta, esto significa que hay un caso para el cual la entropía disminuye

lo cual va en contra de la segunda ley.

Otro problema dado por E. Zermelo es que Poincaré había demostrado que en ese modelo cinético de un gas aislado, el sistema se comportaba quasiperiódico. Esto quiere decir que la función H puede ir asumiendo valores más grandes después de un periodo.

2.3. Gibbs

Ya más adelante W. Gibbs (1839-1903) con su libro *Elementary Principles of Statistical Mechanics* en 1901 propuso la forma actual de hacer física estadística. Gibbs desarrolla las ideas de ensambles, el estado microcanónico y canónico. Las ideas de Gibbs eran más perspicaces que las de Boltzmann, porque introdujo conceptos no muy intuitivos a la mecánica estadística. Aunque Gibbs se volvió a formular la física estadística con ideas nuevas, también se encontraron dificultades en ellas.

Para mostrar las nociones dadas por Gibbs se supone un sistema compuesto de N partículas, la suposición atomista de la materia. Gracias a la mecánica clásica se conocen las ecuaciones de movimiento para un sistema de varias partículas, las ecuaciones de Hamilton:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad , \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} . \quad (2.5)$$

Donde el índice $i = 1, \dots, 3N$ y H es el Hamiltoniano del sistema. Se podría pensar que el problema ya se encuentra solucionado pero hay dificultades al tener estas ecuaciones; la cantidad de ecuaciones de movimiento son muchas pues existirán el mismo número de ecuaciones que grados de libertad, y en un volumen típico hay en el orden de 10^{23} partículas. En la actualidad, las herramientas que proporciona la computación podrían ayudar a resolver estas ecuaciones, suponiendo cierta simplicidad en ellas, pero aún así existe el problema de las condiciones iniciales del sistema; en la práctica es imposible de encontrar la posición y el momento de todas las partículas para el mismo tiempo, esto es en el caso clásico [3]. La mecánica cuántica impone restricciones aún mayores.

Para poder superar este dilema se preve un enfoque estadístico, donde se ve que a mayor número de partículas empiezan a aparecer regularidades. Lo

interesante de la mecánica estadística es que las leyes dadas solo funcionan cuando los grados de libertad son muchos.

Sea un sistema mecánico macroscópico en un instante de tiempo t , se dice que el microestado del sistema está dado por cada $q_i(t)$ y $p_i(t)$. El microestado se puede representar por un punto $\lambda(t)$ en un espacio abstracto de $6N$ dimensiones, el espacio de fase $\Gamma = \Pi_i^{3N}\{q_i, p_i\}$. Con esto claro, la propuesta de Gibbs del ensamble es: Se supone N copias de un sistema macroscópico, cada uno de ellos viene representado en el espacio de fase Γ por un punto $\lambda(t)$. Nótese que esto no significa que todos los puntos son el mismo, un sistema macroscópico puede tener diferentes microestados, cada microestado puede cambiar con respecto a otro, por lo menos por una posición o momento. Esto no afectaría a las propiedades macroscópicas que se miden, y para términos prácticos sería el mismo sistema macroscópico [4].

El ensamble de copias permite crear una densidad de probabilidad en el espacio de fase. Sea $d\mathcal{N}(p, q, t)$ el número representativo de puntos en el volumen infinitesimal $d\Gamma = \Pi_i^{3N} dp_i dq_i$ alrededor del punto (p, q) . Entonces se define una densidad en el espacio de fase:

$$\rho(p, q, t) d\Gamma = \lim_{\mathcal{N} \rightarrow \infty} \frac{d\mathcal{N}(p, q, t)}{\mathcal{N}}. \quad (2.6)$$

Es más conveniente escribir la evolución del sistema por esta densidad, por su sencillez matemática. Ahora que se pudo definir una densidad de probabilidad en el espacio de fase, se puede hallar los promedios de ensamble así:

$$\langle f \rangle = \int d\Gamma \rho(p, q, t) f(p, q). \quad (2.7)$$

Los promedios se encuentran gracias al ensamble compuesto de varias copias idénticas del sistema; pero en microestados diferentes que son admisibles por los parámetros macroscópicos. Pero no significa que el sistema que se está estudiando este en todos esos estados al tiempo, es solo un método para poder lograr describir una densidad en el espacio de fase, que luego se usa para poder calcular propiedades del sistema. Entonces entra la pregunta de que tan posible es hablar de copias imaginarias del sistema, sabiendo que solo existe uno.

La mecánica estadística que introdujo Gibbs tenía como base la idea de la distribución microcanónica, esta es básicamente que: $\rho(q, p)$ es una distribución de densidad en el espacio de fase, la cual es cero en todo el espacio excepto en una superficie de energía, entre E y $E + \Delta E$, donde ΔE es bastante pequeña. En esta superficie de energía ρ tiene un valor constante.

La idea física detrás de esta distribución es bastante simple de seguir. Desde la perspectiva de la teoría de probabilidad cuando no se tiene conocimiento del problema que se quiere tratar se supone que no hay razón para darle un mayor peso a algún resultado que a otro, debido a la ignorancia subjetiva que se tiene. Entonces al no conocer nada del sistema se dice que todos los resultados tienen la misma probabilidad de salir, luego la distribución de probabilidad es constante.

Entonces si tengo un sistema físico del que conozco solo su estado macroscópico, estos parámetros me pueden dar varios posibles estados microscópicos. Ya que, no se tiene más información suponemos que, cada posible estado microscópico tiene la misma probabilidad de ser el estado en el que el sistema se encuentra en realidad. Este es uno de los problemas al fundamentar la mecánica estadística, porque se basa en la ignorancia subjetiva que se tiene del sistema. Muchos argumentan que las teorías físicas deben ser edificadas en ideas objetivas, y que una rama tan importante de la física se asiente en algo como la ignorancia propia deja mucho que desear [31].

Aunque esta idea se base en la teoría de la probabilidad también ha sido muy criticada en las matemáticas. Hay preguntas como: si se llegará a conocer un poco más del sistema, ¿aún sería válida la distribución microcanónica? ¿si las probabilidades realmente tiene un peso que se les debería asignar? En busca de evitar estas preguntas pero no volver a construir toda una teoría lo que se propone es intercambiar el principio de probabilidades iguales por una idea más estable que tenga las mismas implicaciones que ese principio.

En este punto se puede incorporar la entropía de Gibbs. Usando la notación, $d\Gamma \equiv d^3x_1...d^3p_N$, $d\Gamma_1 \equiv d^3x_1d^3p_1$, $d\Gamma_{-1} \equiv d^3x_2...d^3p_N$; la función de distribución de puntos representativos es

$$\rho_N(x_1, x_2, \dots, x_N; p_1, p_2, \dots, p_N; t), \quad (2.8)$$

la cual da la densidad de probabilidad del sistema en todo el espacio de fase.

La función H de Gibbs esta dada por

$$H_G = \int \rho_N \log \rho_N d\Gamma. \quad (2.9)$$

Este punto es preciso para comparar la función H de Gibbs con la de Boltzmann. recordando que la función H de Boltzman está dada por

$$H_B = N \int f_1 \log f_1 d\Gamma_1. \quad (2.10)$$

Donde $f_1(x_1, p_1, t)$ es la probabilidad de densidad de una sola partícula,

$$f_1(x_1, p_1, t) = \int \rho_N d\Gamma_{-1}. \quad (2.11)$$

En el artículo escrito por Jaynes [17], se demuestra que

$$H_B \leq H_G, \quad (2.12)$$

y la igualdad se cumple solo cuando

$$\rho_N(x_1, x_2, \dots, x_N; p_1, p_2, \dots, p_N) = f_1(x_1, p_1) \dots f_1(x_N, p_N), \quad (2.13)$$

es decir, cuando hay independencia entre las partículas. Este es el caso para un gas que no tiene interacción entre partículas, gas ideal. Gracias a esto Jaynes prosigue mostrando que la entropía de Boltzmann, $S_B = k_B H_B$, solo es cierta cuando se habla de un fluido con la misma densidad y temperatura en todo el espacio pero sin fuerzas entre partículas. Mientras la entropía de Gibbs, $S_G = k_B H_G$, es válida para cualquier sistema porque da la entropía fenomenológica de la termodinámica.

Concluye que esta diferencia no puede ser pequeña porque hay interacciones entre partículas muy importantes que afectan ampliamente el resultado. Esto muestra que aunque la idea de Gibbs sobre un ensamble es exótica da los resultados esperados por la termodinámica. Mientras que Boltzmann usando ideas más intuitivas no llega a hacer una conexión con la termodinámica.

2.3.1. Ensamble Canónico

Para explorar un poco más la idea de Gibbs, se presentará el ensamble canónico desde un punto de vista cuántico. Este ensamble será útil como referencia para los próximos capítulos. Para encontrar este operador se extremizará la entropía de Gibbs bajo dos restricciones. La primera restricción es que el operador de densidad, ρ , esté normalizado.

$$\text{Tr}(\rho) = 1, \quad (2.14)$$

además se impone la restricción de tener un valor de energía fijo,

$$\text{Tr}(H\rho) = \langle E \rangle, \quad (2.15)$$

donde H es el operador Hamiltoniano del sistema [9]. El promedio de energía se escribe así, porque cuánticamente la traza permite escribir promedios de esta forma. La entropía de Gibbs para un sistema cuántico se escribe de la siguiente forma

$$S = -k_B \text{Tr}[\rho \ln(\rho)]. \quad (2.16)$$

Se ve que mantiene la forma de la entropía clásica, teniendo como diferencia que se cambia la integral por una traza.

Con ayuda de los multiplicadores de Lagrange, α y β , se encuentra la distribución de probabilidad que maximiza la entropía bajo las restricciones dadas. Entonces se tiene:

$$\delta\{\text{Tr}[\alpha\rho + \beta H\rho - k_B\rho \ln(\rho)]\} = \text{Tr}\{[(\alpha - k_B)I + \beta H - k_B \ln(\rho)]\delta\rho\} = 0. \quad (2.17)$$

Donde I es el operador unitario. Como $\delta\rho$ es arbitrario, lo que está dentro de los corchetes cuadrados debe ser igual a 0. Organizando se llega al operador

$$\rho = \exp\left[\left(\frac{\alpha}{k_B} - 1\right)I + \frac{\beta}{k_B}H\right]. \quad (2.18)$$

Por la condición de normalización que se tiene sobre el operador se encuentra

$$Z(\beta) \equiv \exp\left(1 - \frac{\alpha}{k_B}\right) = \text{Tr}\left(e^{\frac{\beta H}{k_B}}\right). \quad (2.19)$$

Aquí se ha definido la función de partición $Z(\beta)$ como la normalización del operador de densidad de probabilidad. Por medio de las restricciones se puede encontrar que el multiplicador β es igual a $\frac{1}{T}$. Entonces se puede reescribir el operador de densidad como

$$\rho = \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})}. \quad (2.20)$$

Históricamente se escogió la entropía de Gibbs porque da este resultado para un sistema a temperatura T .

2.4. Teorma de Liouville

Construyendo sobre las ideas de Gibbs, se puede encontrar el teorema de Liouville. Este teorema tiene mucha importancia en la mecánica estadística, para ver sus consecuencias en esta área primero se verá qué declara este teorema. La siguiente demostración del teorema sigue los pasos dados por [5].

Se considera un volumen Γ que encierra una región que se quiere estudiar en el espacio de fase, este volumen tiene una superficie σ . El cambio del número de puntos, microestados posibles del sistema, dentro de este volumen está dado por

$$\frac{\partial}{\partial t} \int_{\Gamma} \rho d\Gamma \quad (2.21)$$

donde ρ es la densidad en el espacio de fase definida anteriormente. En la ecuación anterior $d^{3N}q d^{3N}p = d\Gamma$. El cambio neto de puntos que salen de Γ por la superficie σ es

$$\int_{\sigma} \rho \mathbf{v} \cdot \hat{\mathbf{n}} d\sigma; \quad (2.22)$$

donde \mathbf{v} es el vector de velocidad del punto representativo en la región de superficie $d\sigma$. $\hat{\mathbf{n}}$ es el vector perpendicular a esta superficie con dirección de salida. Por el teorema de la divergencia se tiene:

$$\int_{\Gamma} \nabla \cdot (\rho \mathbf{v}) d\Gamma = \int_{\Gamma} \sum_{i=1}^{3N} \left(\frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right) d\Gamma \quad (2.23)$$

Debido a que la cantidad de puntos se conserva en el espacio de fase esto quiere decir que el ensamble que se considera no agrega nuevos miembros ni elimina los que ya se encuentran en este, esto permite concluir:

$$\int_{\Gamma} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right) d\Gamma = 0. \quad (2.24)$$

La condición para que esta integral sea cierta para cualquier Γ es que el integrando sea cero. Esto nos da la ecuación de continuidad para el espacio de fase,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (2.25)$$

usando la forma explícita de la divergencia,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) + \rho \sum_{i=1}^{3N} \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0. \quad (2.26)$$

Por las ecuaciones de Hamilton 2.5 el último término se cancela. Como ρ depende de p, q y t con los dos primeros términos se pueden organizar para que la ecuación quede de la siguiente forma:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + [\rho, H] = 0. \quad (2.27)$$

La ecuación 2.27 es el teorema de Liouville. Esto lo que dice es que la densidad "local" de puntos vista desde un observador que se mueve con el punto, se mantiene constante en el tiempo. Luego los puntos en el espacio de fase se mueven de la misma manera que un fluido incompresible en el espacio físico.

Esta conclusión es la más clara al obtener el teorema de Liouville pero también hay consecuencias profundas dadas por este. Para ver las consecuencias que brinda el teorema seguiremos a [4].

2.4.1. Consecuencias del Teorema de Liouville

La primera consecuencia es que al hacer una inversión temporal el corchete de poisson $[\rho, H]$ cambia de signo y esto predice que la densidad

revierte su evolución. Es decir que al hacer la transformación $(p, q, t) \rightarrow (p, q, -t)$ el teorema de Liouville implica $\rho(p, q, t) = \rho(-p, q, -t)$. La segunda es sobre la evolución del promedio de ensamble en el tiempo.

$$\frac{d\langle f \rangle}{dt} = \int d\Gamma \frac{\partial \rho(p, q, t)}{\partial t} f(p, q) = \sum_{i=1}^{3N} \int d\Gamma f(p, q) \left(\frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} - \frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} \right) \quad (2.28)$$

En la ecuación anterior se usó el teorema de Liouville poniendo explícitamente el corchete de Poisson. Integrando por partes y recordando que ρ tiende a cero en los límites de la integración se llega a:

$$\frac{d\langle f \rangle}{dt} = \langle [f, H] \rangle. \quad (2.29)$$

Con esta relación sobre los promedios se puede ver qué condiciones son necesarias para que el ensamble se encuentre en el estado de equilibrio. Dados unos parámetros macroscópicos se sabe que si el ensamble corresponde a uno en el equilibrio los promedios de ensamble deben ser independientes del tiempo. Esto puede lograrse por

$$[\rho_{eq}, H] = 0. \quad (2.30)$$

Una posible solución a la ecuación anterior es que $\rho_{eq}(p, q) = \rho(H(p, q))$ esto es una solución ya que $[\rho(H), H] = \rho'(H)[H, H] = 0$. Esto muestra el supuesto básico de la mecánica estadística, ρ es constante en las superficie de energías constantes H . El supuesto de la mecánica estadística dictamina que el macroestado esta dado por una densidad uniforme de microestados. Esto es el principio de probabilidades iguales a priori.

La consecuencia anterior respondería la pregunta de cómo definir equilibrio para partículas en movimiento. Pero para saber si todos los sistemas evolucionan naturalmente al equilibrio y justificar el supuesto básico de la mecánica estadística se debe mostrar que densidades no estacionarias se van acercando a densidades estacionarias ρ_{eq} . Pero esto entra en un problema con la primera consecuencia (inversión temporal), dado una densidad $\rho(t)$ que se acerque a ρ_{eq} habrá otra dada la inversión temporal que se estará alejando de esta densidad de equilibrio. Lo que se ha propuesto normalmente es mostrar que $\rho(t)$ se encontrará en el intervalo cercano a ρ_{eq} una gran parte del tiempo, esto significa que los promedios temporales

están dominados por la solución estacionaria. Estos nos lleva al problema de ergodicidad, ¿es válido igualar el promedio temporal con el promedio de ensamble?.

2.5. Ergodicidad

En esta sección se hablará un poco sobre el problema de ergodicidad siguiendo a [12], se planteará las dificultades y se verán algunas soluciones que se han dado para este problema. Se habla de este problema ya que es fundamental la hipótesis de ergodicidad para la mecánica estadística, además la nueva perspectiva que se mostrará en los siguientes capítulos [1], no hace énfasis en esta hipótesis es más ni siquiera llega a ser una pregunta que necesite ser resuelta. Para ver por qué facilita tanto evitar esta hipótesis se dará esta sección.

La mecánica estadística tiene como base el principio de probabilidades iguales generalmente en la literatura siempre se empieza hablando de cómo se acepta el principio desde la teoría de la probabilidad. Pero hay un concepto que no es profundizado en general por los textos usados regularmente; el porqué se debe introducir la probabilidad en estos casos.

Libros como [6],[3] empiezan explicando cómo se manejaría un problema de muchas partículas desde la forma clásica. La idea que muestra es la forma habitual de introducir la mecánica estadística, se presenta que describir avogadro número de partículas es complicado aunque se tenga las ecuaciones y se pueda de alguna forma (numérica o computacional) de resolver todas las ecuaciones que salen. Entonces se sigue explicando que para resolver el problema se tiene en cuenta que la complejidad de las interacciones que hay entre el sistema y el ambiente, hace que el sistema se encuentre en muchos estados varias veces. Esto otorga un comportamiento probabilístico al sistema que puede llegar a comportarse como uno aislado.

Estos argumentos introducen las probabilidades al sistema mecánico sin tener que entrar en un conflicto con la mecánica clásica, nótese que luego de dada esa explicación siguen a formalizar lo dicho y construye la mecánica estadística que se conoce. Pero hay otra forma de introducir las probabilidades a estos sistemas de muchos grados de libertad y es por hipótesis, el libro [11] sigue esta filosofía. Aunque hayan dos formas de establecer las

probabilidades de igual forma se llega al problema de ergodicidad.

La hipótesis de ergodicidad afirma que el promedio temporal de una cantidad mecánica en un sistema aislado es igual al promedio de ensamble. Dada una teoría que explique algún fenómeno que nos interese se querrá probar que tan buenos resultados se obtienen de ella. Para esto se comparará con el experimento pero las cantidades físicas en una teoría aparecen como funciones de las variables dinámicas. Pero ya se ve un problema claro en este procedimiento cómo se puede comprar los datos medidos en el laboratorio con las funciones de la teoría; las funciones de las variables dinámicas toman valores diferentes para varios estados del sistema pero para poder comparar los resultados experimentales se debería saber cuál es el estado del sistema o sea dar todas las variables dinámicas para poder estar seguros que se está comparando los datos con el estado específico, pero como ya se ha dicho esto es imposible.

Pero hay un detalle en este caso y es que al medir una cantidad física esta no se observa de manera instantánea, se demora algún tiempo τ para poder observarla. En el caso de una cantidad termodinámica, el tiempo en que se demora midiendo va desde τ_0 , que es el tiempo necesario para que no hayan correlaciones del movimiento molecular, y τ_m que es el tiempo máximo en el que la propiedad macroscópica que se mide no cambia. Con la temperatura al medirla se necesita esperar un tiempo τ_0 para que se establezca el sistema y el termómetro, y la medición se puede tomar hasta τ_m cuando el sistema se vuelva a perturbar. Cualquier cantidad que se mida durante este intervalo se supone que no debe depender de τ en algunos sistemas que llegan al equilibrio τ_m puede extenderse al infinito.

Entonces con lo dicho anteriormente se puede proponer la siguiente idea de medición. Sea P un punto en el espacio de fase, y sea $f(P, t)$ una función sobre el espacio de fase, si esta función corresponde a una cantidad termodinámica no debe depender de su estado inicial P . Porque sistemas termodinámicos con cantidades termodinámicas iguales se pueden encontrar en estados dinámicos diferentes.

La relación de esta función con los parámetros está dada por un promedio en este caso el promedio temporal luego:

$$\overline{f}_\tau(P) = \frac{1}{\tau} \int_0^\tau f(P, t) dt \quad (2.31)$$

si se está en el equilibrio

$$\bar{f}(P) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau f(P, t) dt. \quad (2.32)$$

Para poder mantener consistencia con la termodinámica los promedios $\bar{f}_\tau(P)$ y $\bar{f}(P)$ deben ser independientes de τ . Además en el equilibrio los dos promedios deben ser iguales, esto es para τ_m muy grande.

El promedio de ensamble está definido así:

$$\langle f \rangle = \frac{1}{\Omega} \int_\Gamma f(P, t) d\Gamma \quad ; \quad \Omega = \int_\Gamma d\Gamma, \quad (2.33)$$

la integral es sobre el espacio accesible en el espacio de fase, $d\Gamma$ es el elemento de volumen. Como el punto P evoluciona gracias a las ecuaciones de Hamilton se preserva la medida por el teorema de Liouville, luego $\langle f \rangle$ no depende del tiempo. Entonces

$$\langle f \rangle = \overline{\langle f \rangle}. \quad (2.34)$$

Como el promedio temporal conmuta con el promedio de ensamble, hay casos especiales en los que no ocurre esto. Entonces

$$\overline{\langle f \rangle} = \langle \bar{f} \rangle. \quad (2.35)$$

Como \bar{f} es independiente de P y constante sobre todo el espacio de fase, entonces $\langle \bar{f} \rangle = \bar{f}$ por lo tanto

$$\langle f \rangle = \bar{f}. \quad (2.36)$$

Aquí se ha mostrado por qué es válido el uso del promedio de ensamble en vez del temporal. Para esta demostración se ha usado dos suposiciones. Primero que $\bar{f}(P)$ tiene el mismo valor para cualquier P en la misma trayectoria que este punto traza, lo que llama [23] el G-Path. Birkhoff demostró esta aseveración [25]. El problema es saber si $\bar{f}(P)$ tiene un valor definido independiente de la condición inicial P y si esto se satisface para P en cualquier G-path.

Boltzmann propuso que para satisfacer la igualdad 2.36, el G-path debe pasar por todos los puntos de la superficie de energía. Como la trayectoria

del punto pasa por todos los puntos dado suficiente tiempo, el promedio temporal es igual al promedio de ensamble. Esto es lo que Boltzmann llamó la hipótesis de ergodicidad, esta idea sobre la trayectoria del punto P es llamado la condición de ergodicidad en el sentido de Boltzmann.

Como la trayectoria dinámica es un conjunto unidimensional de puntos continuos y nunca se intersecta así misma en una superficie multidimensional constante de energía, hace imposible una correspondencia uno a uno de un espacio de dimensión dos o mayor a un espacio unidimensional, luego la ergodicidad en el sentido de Boltzmann no puede concebirse. Ergodicidad, la igualdad 2.36, no siempre implica ergodicidad en el sentido de Boltzmann.

Aunque la hipótesis ergódica se encuentre en un punto importante en la construcción de la física estadística, hay sistemas que no son ergódicos pero aún así se pueden tratar con la mecánica estadística. El calor específico de los sólidos pueden ser modelado de manera correcta bajo el supuesto de una red de partículas cuánticas vibrando en una aproximación armónica. El tratamiento clásico también es satisfactoria para temperaturas altas. Como se sabe, cualquier modo normal en vibraciones armónicas es independiente de los otros modos normales y la energía de cada modo es constante. Para describir el calor específico se puede hacer por medio de la aproximación armónica pero si existen términos no lineales pequeños o anarmónicos habrá un intercambio de energía entre los modos normales. Existe un problema famoso que E. Fermi, J. Pasta y S. Ulam investigaron [24], un sistema de una cadena lineal de $N + 1$ partículas idénticas de masa m conectadas por resortes idénticos. Pero en este caso ellos incluyeron interacciones no lineales al Hamiltoniano de este sistema, que puede ser escrito

$$H = \sum_k \frac{1}{2m} p_k^2 + \frac{\kappa}{2} \sum_k (q_{k+1} - q_k)^2 + \frac{\lambda}{s} \sum_s (q_{k+1} - q_k)^s, \quad (2.37)$$

donde λ es la constante de acoplamiento y $s = 3, 4$, esto representa los potenciales no lineales cúbicos y cuadráticos. Ellos querían mostrar la intuición que cualquier interacción no lineal entre partículas de un sistema generaría un comportamiento ergódico o irreversible para el sistema. Las ecuaciones de movimiento del sistema son

$$m\ddot{q}_k = \kappa(q_{k+1} - 2q_k + q_{k-1}) + \lambda[(q_{k+1} - q_k)^{s-1} - (q_k - q_{k-1})^{s-1}], \quad (2.38)$$

o transformando a coordenadas normales: $q_k = (\frac{2}{N})^{\frac{1}{2}} \sum_{j=1}^{N-1} x_j \sin\left(\frac{jk\pi}{N}\right)$, $m\dot{x}_j = y_j$ la ecuación queda como

$$\ddot{x}_j = -\omega_j^2 x_j + \lambda F_j(x), \quad (2.39)$$

con $\omega_j = 2(\frac{\kappa}{m})^{\frac{1}{2}} \sin\left(\frac{j\pi}{2N}\right)$. Donde $F_j(x)$ son las fuerzas no lineales. El sistema que ellos estudiaron fue de $N = 32$ y $N = 64$ mirando su evolución luego de un largo tiempo. Las fuerzas que usaron fueron cuadráticas, cúbicas y lineales. Además la energía de los modos está dada por

$$E_k = \frac{1}{2}m(\dot{x}_k^2 + \omega_k^2 x_k^2). \quad (2.40)$$

Simularon cómo las energías cambiaban en el tiempo cuando inicialmente se le da al menor modo una cantidad de energía. Se esperaba que la energía dada al menor modo fuese siendo distribuida paulatinamente a los demás modos y al final se tuviese una cantidad de energía igualmente distribuida en cada modo. Pero este no fue el hallazgo, para $N = 32$ con un potencial cúbico y con las constantes $\kappa = m = 1$ y $\lambda = 1/4$. Las condiciones iniciales eran: las velocidades iguales a 0 y los desplazamientos dados por la curva sinusoidal del menor modo normal. Lo que se vio fue que la energía regresa al primer modo de manera periódica, la energía es intercambiada en los modos más bajos pero no mucho en los modos más grandes. Se llega a la conclusión que el sistema regresa de forma casi exacta a su estado inicial. Entonces la presencia de anarmonicidad no implica ergodicidad para este sistema. Esto es un ejemplo de un sistema que no es ergódico, esto deja mucha desconformidad dado que no se llega ver porqué la hipótesis que se usa para formar la mecánica estadística no se cumple para sistemas que comúnmente son usados en la física y son utilizados de manera frecuente. Este descontento es una de las partes que no deja al marco teórico de la física estadística muy estable en comparación con otras ramas de la física.

2.6. Jaynes

Los problemas mostrados hasta ahora han sido varios y de gran complejidad. Por esto varios intentos se han hecho para proponer una nueva perspectiva, que no tenga estas tensiones. Han existido muchas propuestas

nuevas, tratando de basar la mecánica estadística en otras ideas. En esta sección se mostrará la visión de E.T Jaynes. Jaynes propone una relación entre la mecánica estadística y la teoría clásica de la información [28],[29]. Esto muestra una primera luz a usar herramientas de la información en áreas de la física.

Jaynes impulsado por la falta de un argumento que conecte los fenómenos microscópicos y los macroscópicos. Procede a estipular que para tener una mecánica estadística libre de objeciones esta debería seguir los siguientes requerimientos: libre de incongruencias matemáticas, no debería tener suposiciones externas arbitrarias y debería ser capaz de describir procesos tanto fuera del equilibrio como en equilibrio.

Jaynes explica que la segunda condición parece extraña al hablar de una teoría física porque en general se proponen hipótesis que luego intentan ser explicadas. Pone el ejemplo de la unidad de volumen en el espacio de fase que inicialmente fue puesta para poder conseguir los valores correctos de la presión de vapor en equilibrio, luego con la ayuda de la teoría cuántica se vio que esta unidad salía naturalmente de las leyes físicas. Esto hace que Jaynes proponga que la mecánica estadística sea un ejemplo de inferencia estadística. Como la expresión $-\sum p_i \log p_i$ aparece en la teoría de la información, que es una teoría de inferencia estadística, y en la mecánica estadística, Jaynes expone como crear una conexión entre las dos teorías haciendo que esta fórmula tenga el mismo concepto y así poder aplicar la teoría de la información clásica a la mecánica estadística. Con esto Jaynes muestra una simplificación de la matemática y dice que la mecánica estadística puede afrontar problemas más generales y nuevos problemas físicos.

Toda la perspectiva de Jaynes se basa en el principio de máxima entropía. Para poder entender este principio se plantea un problema que ocurre en muchos casos no solo en la física. Sea x una cantidad discreta que puede tener valores x_i con $i = 1, 2, \dots, n$. No se saben las probabilidades correspondientes p_i . Lo único que se conoce del problema son los promedios de la función $f(x)$,

$$\langle f(x) \rangle = \sum_i^N p_i f(x_i). \quad (2.41)$$

Entonces la pregunta que se hace es: ¿ con esta información cuál es el valor esperado de la función $g(x)$?. Esta clase de problemas aparecen en muchas áreas, como la teoría de la información, porque este es un problema de la teoría de la probabilidad. Encontrar las probabilidades con poca información es un problema en el que trabajaron los padres de la probabilidad. El principio dado por Laplace es un intento de resolver el problema, este principio asigna probabilidades iguales a dos eventos. Jaynes dice que a menos que haya una simetría en el problema se encuentran paradojas en los casos continuos. Por eso este tipo de formulaciones se ha dejado de usar en los problemas actuales continua diciendo Jaynes.

Para poder seguir exponiendo el planteamiento de Jaynes se debe entender que la perspectiva que acepta sobre la teoría de la probabilidad es una subjetiva. Esto quiere decir que como J.M. Keynes o H. Jeffreys acepta que las probabilidades son una cuantificación de la ignorancia humana. La probabilidad es solo una medida basada en la información que se tiene sobre el posible resultado que puede salir. Entonces la distribución de probabilidad solo es la representación del estado actual de nuestro conocimiento.

En contraposición con la perspectiva objetiva que dice que la probabilidad es la proporción de frecuencias de un experimento, calculando la distribución de probabilidad se está haciendo una predicción que se puede en principio comparar con el experimento. Esta perspectiva acepta las probabilidades como algo externo del humano. Jaynes dice que la visión subjetiva es más general porque las proporciones de frecuencia siempre pueden interpretarse desde esta perspectiva. Además manifiesta que la posición subjetiva acepta interrogantes que para el objetivismo no tienen importancia. Shannon en su teoría de la información propuso una función que describía de forma única y sin ambigüedad la cantidad de incertidumbre que tiene una distribución de probabilidad [30]. Esta cantidad es positiva, se incrementa con una incertidumbre decreciente y es aditiva para fuentes de incertidumbres independientes.

$$H(p_1, \dots, p_N) = -k \sum_i p_i \ln p_i, \quad (2.42)$$

k es una constante positiva. Aquí Jaynes hace la conexión y es un punto importante. Esta cantidad Shannon la llama entropía pero no es la entropía termodinámica, se dice que Von Neumann le propuso este nombre a Shan-

non ya que nadie sabía realmente qué era la entropía. Pero Jaynes hace el puente entre las dos teorías diciendo que la entropía termodinámica es la cantidad de incertidumbre que se tiene del sistema. Entonces esto es lo que quiere decir el principio de la máxima entropía, dadas unas restricciones o sea con una información parcial se debe encontrar la distribución que minimice la incertidumbre. Esto es equivalente a maximizar H bajo las condiciones conocidas. Esta es la forma de llegar a la distribución sin suposiciones arbitrarias que impondríamos bajo conjeturas que no estamos seguros de ser ciertas.

Jaynes prosigue usando multiplicadores de Lagrange y obtiene los siguientes resultados para un número m de funciones, dado $\langle f_r(x) \rangle = \sum_i p_i f_r(x_i)$:

$$p_i = \exp\{-[\lambda_0 + \lambda_1 f_1(x_i) + \dots + \lambda_m f_m(x_i)]\}. \quad (2.43)$$

Donde los λ 's son los multiplicadores de Lagrange. Además se define la función de partición como

$$Z(\lambda_1, \dots, \lambda_m) = \sum_i \exp\{-[\lambda_1 f_1(x_i) + \dots + \lambda_m f_m(x_i)]\}. \quad (2.44)$$

Los multiplicadores de Lagrange se pueden encontrar por

$$\begin{aligned} \langle f_r(x) \rangle &= -\frac{\partial}{\partial \lambda_r} \ln Z, \\ \lambda_0 &= \ln Z. \end{aligned}$$

Jaynes argumenta que siguiendo este principio se llega a una distribución de probabilidad imparcial correspondiente a la información del problema, de forma inequívoca y única. También permite la modificación de la distribución de probabilidad si se llega a conocer más información.

Gracias a esto se puede ver los resultados dados por la mecánica estadística fácilmente. Sea $E(\alpha_1, \alpha_2, \dots)$ los niveles de energía de un sistema, donde los α 's son parámetros externos como el volumen o potenciales gravitacionales. Si solo se conoce $\langle E \rangle$. Pero por la ecuación 2.43 ya se sabe las probabilidades de los niveles de energías E_i , esta es la distribución de Maxwell-Boltzmann. Entonces se puede seguir identificando los distintos parámetros de la ter-

modinámica.

$$\begin{aligned}\lambda_1 &= \frac{1}{kT}, \\ U - TS &= F(T, \alpha_1, \alpha_2, \dots) = -kT \ln Z(T, \alpha_1, \alpha_2, \dots), \\ S &= -\frac{\partial F}{\partial T} = -k \sum_i p_i \ln p_i, \\ \beta_i &= kT \frac{\partial}{\partial \alpha_i} \ln Z.\end{aligned}$$

Las fuerzas β_i pueden ser la presión, el tensor de estrés, los momentos magnéticos y/o eléctricos. Se puede también encontrar la distribución gran canónica solo poniendo una nueva restricción sobre el número de partículas. sea n_i el número de partículas del tipo i y no se conoce además para especificar el estado del sistema se necesita las n 's junto con un nivel de energía $E(\alpha_1, \alpha_2, \dots; n_1 n_2 \dots)$ y se conoce los valores esperados de la energía y cada número de partículas de todos los tipos

$$\langle E \rangle \quad , \quad \langle n_i \rangle \quad , \quad i = \{1, 2, \dots\}. \quad (2.45)$$

la función de partición es la de la gran canónica

$$Z(\alpha_1, \alpha_2, \dots | \lambda_1 \lambda_2, \dots, \beta) = \sum_{n_1 n_2 \dots} \sum_i \exp\{-[\lambda_1 n_1 + \lambda_2 n_2 + \dots + \beta E_i(\alpha_k | n_s)]\}. \quad (2.46)$$

La propuesta de Jaynes ha ganado popularidad y evita la pregunta de sobre cuál distribución se está sacando las probabilidades porque como su perspectiva es subjetiva lo único que se está teniendo en cuenta es que se está mirando la ignorancia subjetiva. No se debe crear un ensamble de copias para darle un significado a la probabilidad.

Capítulo 3

El Entrelazamiento y La Mecánica Estadística

3.1. Herramientas de la Teoría Cuántica

La mecánica cuántica en general se introduce desde la perspectiva del vector de estado, sus aplicaciones abarcan una gran cantidad de problemas. Sus herramientas son poderosas para algunos problemas, pero no siempre es sencillo poder seguir con este formalismo en casos particulares. Por esto se usa el formalismo de la matriz densidad u operador densidad cuando se habla de un sistema compuesto por un subsistema y su ambiente. Debido al entrelazamiento que existe ente el subsistema y su ambiente no se puede especificar un estado individual para el subsistema, las herramientas proporcionadas por el operador densidad permite representar las estadísticas de las medidas de uno de los componentes del sistema [7].

3.1.1. Operador de Densidad

Como se sabe cuánticamente el estado $|\psi\rangle$ da toda la información posible de un sistema, a esto se le llama un estado puro. Para especificar cuál es el operador densidad de un sistema se supone que el sistema cuántico puede estar en $|\psi_i\rangle$ para $i = 1, 2, \dots$. Cada estado posible del sistema tiene una probabilidad correspondiente, $p_i > 0$ y $\sum_i p_i$. Luego el operador densidad

para el sistema es :

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle \psi_i|. \quad (3.1)$$

En el caso en que solo haya un estado posible, se tiene el operador densidad para un estado puro $\rho = |\psi\rangle \langle \psi|$. Entonces se puede hablar de un ensamble de estados puros $\{p_i | \psi_i\rangle\}$ siendo, el operador densidad, la distribución clásica de probabilidad de matrices de estados puros. Nótese que se puede usar de manera similar el nombre de operador densidad y de matriz densidad, esta es la usanza de la literatura.

El operador densidad también puede ser especificado por las siguientes propiedades:

1. ρ tiene traza igual a uno, $\text{Tr}(\rho) = 1$.
2. ρ es un operador positivo, $|\phi\rangle$ es un vector arbitrario luego $\langle \psi | \rho | \psi \rangle \geq 0$.

Ya con esta breve introducción se pueden reescribir los postulados de la mecánica cuántica en este formalismo, siguiendo a [14]:

Postulado 1: Se puede asociar a un sistema físico aislado un espacio de Hilbert. El sistema está descrito completamente por el operador densidad que actúa sobre el espacio de Hilbert del sistema. Si un sistema cuántico está en el estado ρ_i con probabilidad p_i , el operador de densidad del sistema es $\sum_i p_i \rho_i$.

Postulado 2: La evolución de un sistema cuántico cerrado está descrita por una transformación unitaria. Esto es, el estado ρ del sistema en el tiempo t_0 está relacionado con el estado ρ' del sistema para el tiempo t por un operador unitario U el cual depende solo de los tiempos t_0 y t .

$$\rho' = U \rho U^\dagger. \quad (3.2)$$

Postulado 3: Las medidas cuánticas están descritas por un conjunto de operadores $\{M_m\}$. Estos operadores actúan en el espacio de Hilbert del sistema. El índice m habla de los posibles resultados de una medida al

hacer el experimento. Si se tiene el estado del sistema ρ inmediatamente antes de la medida la probabilidad que el resultado m ocurra es:

$$p(m) = \text{Tr}\left(M_m^\dagger M_m \rho\right), \quad (3.3)$$

y el estado del sistema luego de una medición es

$$\frac{M_m \rho M_m^\dagger}{\text{Tr}\left(M_m^\dagger M_m \rho\right)}. \quad (3.4)$$

Los operadores $\{M_m\}$ satisfacen,

$$\sum_m M_m^\dagger M_m = I. \quad (3.5)$$

Postulado 4: El espacio de un sistema físico compuesto es el producto tensorial de cada espacio de los componentes del sistema. Si hay N sistemas y cada sistema es preparado en el estado ρ_i , con el índice recorriendo cada sistema, el sistema total está representado por $\rho_1 \otimes \rho_2 \otimes \dots \otimes \rho_n$.

Estos son los postulados de la mecánica cuántica desde la perspectiva del operado densidad, estos postulados son análogos a la perspectiva del vector de estado. Pero se debe ver un poco más en detalle el postulado 3, este habla sobre las mediciones cuánticas pero con mayor generalidad. Habitualmente al introducir la mecánica cuántica solo se habla de medidas proyectivas, estas medidas tienen muchas restricciones y no abarcan todas las posibilidades de medición.

Se puede demostrar que las mediciones proyectivas junto con los demás axiomas de la cuántica son semejantes a las medidas generales [14], entonces si existe una equivalencia entre estos dos tipos de mediciones, ¿por qué preocuparse por un esquema más general?. Porque las mediciones generales no tienen tantas restricciones como las medidas proyectivas, esto da una variedad de propiedades que las medidas proyectivas no tienen. Además en algunos problemas que conciernen a la teoría cuántica de la información, es necesario usar mediciones generales, como por ejemplo la manera óptima de distinguir entre un conjunto de estados cuánticos.

Pero hay una ventaja que tienen las mediciones generales sobre las proyectivas, es lo que se nombra como *repetibilidad*. Las medidas proyectivas tienen

repetibilidad, puesto que si se tiene un estado inicial $|\psi\rangle$. Al hacer la primera medida se encuentra que $|\psi_m\rangle = \frac{P_m|\psi\rangle}{\sqrt{\langle\psi|P_m|\psi\rangle}}$. Operando otra vez a $|\psi_m\rangle$ con P_m , el estado no cambia entonces $\langle\psi_m|P_m|\psi_m\rangle = 1$, esto significa que si se repite la medición muchas veces se seguirá en el mismo estado [16]. Pero esto no es cierto para todos los experimentos. Si se está haciendo el experimento de la doble rendija y se quiere saber la posición de un fotón luego de pasar y eso se hace por medio de una placa, después de encontrar la posición del fotón no se puede volver a hacer una medición sobre ese fotón. Entonces por esto se debe recurrir en ciertos casos a las mediciones generales. De estas mediciones generales hay un caso particular que es usado en la teoría cuántica de la información, este caso se llama medidas POVM (Positive Operator-Valued Measure). Estas medidas sirven cuando solo se quiere saber las probabilidades de que un resultado salga sin darle importancia a cuál sea el estado del sistema después de la medida [14]. Por el postulado 3 se sabe que la probabilidad de que el resultado m salga viene dado por $p(m) = \text{Tr}(M_m^\dagger M_m \rho)$. Si se define :

$$E_m \equiv M_m^\dagger M_m. \quad (3.6)$$

Se puede demostrar fácilmente que E_m es un operador positivo y cumple con $\sum_m E_m = I$. El conjunto completo de $\{E_m\}$ se conoce como POVM, y cada E_m es un elemento asociado a POVM. Las mediciones proyectivas hacen parte de POVM, sea P_m un proyector tal que $P_m P_n = \delta_{mn} P_m$ y $\sum_m P_m = I$. Solo para este caso los elementos POVM son los mismos que los operadores de medición, ya que $E_m = P_m^\dagger P_m = P_m$.

3.1.2. Operador de Densidad Reducido

Ahora se puede plantear la siguiente situación: se tiene un sistema A que está correlacionado cuánticamente con el sistema B . Esto quiere decir que el sistema completo AB puede estar dado por un estado puro pero aún así no se puede describir individualmente el sistema A , entonces si solo se tiene acceso al sistema A pero no al B , cómo se puede modelar matemáticamente toda la información que es posible de extraer solo de mediciones locales en A . Aquí se usa el operador densidad reducido, este

se define así

$$\rho_A \equiv \text{Tr}_B(\rho_{AB}). \quad (3.7)$$

La traza es tomada únicamente en una base del espacio de Hilbert \mathcal{H}_B del sistema B . Esta operación puede ser interpretada como si se estuviese promediando sobre todos los grados de libertad del sistema que no se tiene acceso, esta operación se le llama la traza parcial sobre B . Para ver la utilidad que tiene este operador se verá un ejemplo sencillo, este ejemplo se puede generalizar fácilmente. Se tiene el estado de algún sistema compuesto por A y B dado por la función de onda Φ :

$$\Phi = \frac{1}{\sqrt{2}}(|a_1\rangle |b_1\rangle + |a_2\rangle |b_2\rangle), \quad (3.8)$$

este es un típico ejemplo de una función de onda donde los subsistemas A y B están entrelazados [26]. Los estados $|a_1\rangle, |b_1\rangle, |a_2\rangle, |b_2\rangle$ no son necesariamente ortogonales. La matriz de densidad es claramente:

$$|\Phi\rangle \langle \Phi| = \frac{1}{2} \sum_{ij=1}^2 |a_i\rangle \langle a_j| |b_i\rangle \langle b_j|. \quad (3.9)$$

Sea $\{|\psi_k\rangle\}$ una base ortonormal en \mathcal{H}_A y $\{|\phi_l\rangle\}$ una base ortonormal en \mathcal{H}_B . Considerando mediciones que solo afecten al sistema A se puede representar esta acción como $M = M_A \otimes I_B$, donde I_B es el operador identidad en el espacio de Hilbert de B . Siguiendo la idea de responder si se puede extraer información del sistema solo conociendo las mediciones locales en A , entonces se hace el siguiente razonamiento:

$$\langle M \rangle = \text{Tr}(\rho M) \quad (3.10)$$

$$= \sum_{kl} \langle \phi_l | \langle \psi_k | \rho (M_A \otimes I_B) | \psi_k \rangle | \phi_l \rangle \quad (3.11)$$

$$= \sum_k \langle \psi_k | \left(\sum_{l=1} \langle \phi_l | \rho | \phi_l \rangle \right) M_A | \psi_k \rangle \quad (3.12)$$

$$= \sum_k \langle \psi_k | (\text{Tr}_B(\rho)) M_A | \psi_k \rangle \quad (3.13)$$

$$= \sum_k \langle \psi_k | \rho_A M_A | \psi_k \rangle \quad (3.14)$$

$$= \text{Tr}_A(\rho_A M_A). \quad (3.15)$$

Aquí ρ_A es la matriz reducida definida al inicio de esta subsección. El ejemplo se puede expandir para sistemas con N subsistemas entrelazados. Se tiene ahora N componentes del sistema y solo se puede hacer mediciones en el componente i . El operador que describe estas mediciones sería $M = I_1 \otimes I_2 \dots \otimes I_{i-1} \otimes M_i \otimes I_{i+1} \otimes \dots \otimes I_N$. Ahora la matriz de densidad reducida para el sistema i es

$$\rho_i = \text{Tr}_{1,\dots,i-1,i+1,\dots,N}(\rho). \quad (3.16)$$

Con esto y siguiendo el ejemplo anterior se puede demostrar que [7]

$$\langle M \rangle = \text{Tr}(\rho M) = \text{Tr}_i(\rho_i M_i). \quad (3.17)$$

Entonces si se conoce el estado del sistema total y solo se puede hacer mediciones en un subsistema aún así gracias al formalismo del operador densidad se puede encontrar las estadísticas de las mediciones. Es importante aclarar que aunque se tenga un operador de densidad reducido para el sistema A no significa que se haya descrito el sistema de manera individual. Esto es solo una herramienta matemática para poder sacar las estadísticas de las mediciones.

3.1.3. Distancia de Traza

Las herramientas hasta ahora mostradas han sido impulsadas por una pregunta que se necesitaba responder pero que tenía una respuesta única. Ahora se encuentra con la pregunta ¿qué tan cercanos están dos estado cuánticos?. La respuesta a esta pregunta puede ser dada por varias definiciones de distancias pero la que se usará acá será la distancia de traza.

Una pregunta relacionada con la cercanía de los estados cuánticos es la de ¿qué quiere decir que dos informaciones sean similares?. En la teoría clásica de la información esto recae en poder diferenciar dos distribuciones de probabilidad. Pero como ya se había advertido esta respuesta no es única, existen diferentes medidas como la fidelidad o la distancia de Kullback–Leibler. Dependiendo del fin es deseable una a las otras.

Se tienen dos distribuciones de probabilidad $\{p_x\}$ y $\{q_x\}$. La distancia de traza clásica viene dada por:

$$D(p_x, q_x) \equiv \frac{1}{2} \sum_x |p_x - q_x|. \quad (3.18)$$

También se conoce como la distancia L_1 o la distancia de Kolmogorov. Esta distancia también se puede escribir de la siguiente manera

$$D(p_x, q_x) = \max_S |p(S) - q(S)| = \max_S \left| \sum_{x \in S} p_x - \sum_{x \in S} q_x \right|. \quad (3.19)$$

Donde la maximización es sobre todos los subconjuntos S de los índices $\{x\}$ [14]. Aunque estas igualdades puedan ser complicadas para hacer cálculos se conseguir una intuición de estas ecuaciones. Entonces la distancia de traza es la maximización de las diferencias entre la probabilidad que el evento S ocurra de acuerdo a la distribución $\{p_x\}$ y la probabilidad que el evento S ocurra de acuerdo a la distribución $\{q_x\}$. El evento S se podría decir que es el evento óptimo para examinar las diferencias entre $\{p_x\}$ y $\{q_x\}$.

Pasando a un contexto cuántico la distancia de traza entre dos estados cuánticos ρ y σ es,

$$D(\rho, \sigma) \equiv \frac{1}{2} \text{Tr} |\rho - \sigma|. \quad (3.20)$$

donde $|M| \equiv \sqrt{M^\dagger M}$ con la raíz positiva. Esta definición generaliza la distancia de traza clásica, si ρ y σ conmutan son diagonales en una misma base luego

$$\begin{aligned} D(\rho, \sigma) &= \frac{1}{2} \text{Tr} \left| \sum_i (r_i - s_i) |i\rangle \langle i| \right| \\ &= D(r_i, s_i). \end{aligned}$$

donde r_i y s_i son los elementos diagonales de ρ y σ respectivamente. Al igual que con la distancia de traza clásica se le puede dar un sentido más físico a esta definición. El siguiente teorema construirá la conexión [14].

Teorema 3.1.1. *Sea $\{E_m\}$ un POVM, con $p_m \equiv \text{Tr}(\rho E_m)$ y $q_m \equiv \text{Tr}(\sigma E_m)$ como las probabilidades de obtener un resultado de una medida m . Luego*

$$D(\rho, \sigma) = \max_{\{E_m\}} D(p_m, q_m), \quad (3.21)$$

donde la maximización es sobre todos los POVMs $\{E_m\}$.

Demostración. Véase que

$$D(p_m, q_m) = \frac{1}{2} \sum_m \left| \text{Tr}(E_m(\rho - \sigma)) \right|. \quad (3.22)$$

Usando la descomposición espectral se tiene que $\rho - \sigma = Q - S$, donde Q y S son operadores positivos con soporte ortogonal. Luego $|\rho - \sigma| = Q + S$, y

$$\left| \text{Tr}(E_m(\rho - \sigma)) \right| = \left| \text{Tr}(E_m(Q - S)) \right| \quad (3.23)$$

$$\leq \text{Tr}(E_m(Q + S)) \quad (3.24)$$

$$\leq \text{Tr}(E_m|\rho - \sigma|). \quad (3.25)$$

Luego

$$D(p_m, q_m) \leq \frac{1}{2} \sum_m \text{Tr}(E_m|\rho - \sigma|) \quad (3.26)$$

$$= \frac{1}{2} \text{Tr}(|\rho - \sigma|) \quad (3.27)$$

$$= D(\rho, \sigma), \quad (3.28)$$

se usó la relación de completitud de los elementos de POVM. Por otro lado si se escogen medidas donde los elementos POVM incluyen proyectores sobre el soporte de Q y S , se ve que existen mediciones que dan distribuciones de probabilidad tales que $D(p_m, q_m) = D(\rho, \sigma)$.

□

Entonces se puede decir que si dos operadores de densidad son cercanos, según la distancia de traza, cualquier medición hecha en este estado cuántico dará una distribución de probabilidad la cual estará cercana, en el sentido clásico de la distancia de traza. También se puede ver como el límite superior a la distancia de traza entre las distribuciones de probabilidad que se encuentran al medir esos estados cuánticos.

La mecánica clásica nos habla de un sistema físico definido que para todos los tiempos está especificado. Este sistema evolucionan de manera determinista dadas las ecuaciones de movimiento. Pero lo que sorprende

al tratar con sistemas termodinámicos es que aunque se hable de un sistema clásico este puede mostrar propiedades que dependan de promedios estadísticos [8]. La conexión que hay entre el determinismo y estas probabilidades es una discusión que no se ha podido solucionar dado a que han existido varias soluciones pero todas con sus fallas. Los métodos típicos requieren hablar de promedios de ensamble, promedios temporales, aleatoriedad. Todos estos conceptos siguen siendo muy debatidos y parecería que no solucionará ninguna pregunta seguir con ellos [12]. Por eso se ha estado buscando nuevos fundamentos para la mecánica estadística.

La mecánica estadística tiene como base el postulado de probabilidades iguales; este viene dado por la ignorancia subjetiva que el observador del sistema tiene. Aunque la mecánica estadística no tiene problemas al comparar sus resultados teóricos con los experimentales, el cimiento filosófico en el que reposa da mucho de que hablar dando así posturas diferentes que desde los inicios de su teoría no han podido ser unificados. En este capítulo se expondrá cómo Popescu et al. en [1] muestran una posible luz sobre el problema de unificar las perspectivas de la mecánica estadística. La idea principal que se quiere mostrar es cómo se puede reemplazar el postulado de probabilidades iguales por un principio canónico general basado en el entrelazamiento cuántico, al poner el entrelazamiento cuántico como nuevo cimiento ya no se tiene probabilidades subjetivas sino objetivas dadas por la teoría cuántica. Este nuevo enfoque permite evitar problemas con la ignorancia subjetiva pero también esquivo el problema de ergodicidad, gracias a esto se puede ver un fundamento más claro y sólido.

Este capítulo solo se enfocará en mostrar cómo la mayoría de los estados del universo están termalizados sin hablar en ningún momento de la evolución de estos estados. Si en general los estados del universo están termalizados se esperaría que cualquier evolución lleve los estados al equilibrio pero en el siguiente capítulo se darán especificaciones de esta conjetura junto con detalles que hacen ver las complicaciones de formalizar las ideas intuitivas que se tienen sobre el equilibrio.

3.2. Definiciones

Antes de seguir sería preciso dar la notación que se usará a lo largo de este capítulo y el siguiente. Cuando se tenga un sistema cuántico grande descrito por un espacio de Hilbert \mathcal{H} este se llamará el universo. Este universo será dividido en dos subsistemas. El primero se llamará el sistema S y el segundo se le dará el nombre de ambiente E , en ocasiones S se le dirá subsistema y E se le podrá llamar baño estos nombres son equivalentes a los anteriores. Esto lleva a descomponer \mathcal{H} como $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, con dimensiones d_S y d_E respectivamente, se supone que la dimensión del ambiente es mucho mayor que la del sistema. Para evitar espacios de dimensión infinita se introduce un tope para altas energías y así mantener la dimensión finita. Además se eliminarían términos de interacción del Hamiltoniano que lo lleven a subespacios no permitidos. Aún no se ha especificado nada sobre el subsistema o el ambiente (Excepto las proporciones de sus dimensiones) esto permite decir que cualquier descomposición del espacio de Hilbert especifica un subsistema y un baño. El subsistema S puede ser cualquier cosa desde una partícula hasta el conjunto de partículas distribuidas por todo el baño.

El estado global puro del universo se escribirá como $|\phi(t)\rangle$ en un tiempo t y su matriz de densidad se escribirá $\rho(t) = |\phi(t)\rangle \langle\phi(t)|$. El estado del sistema se encontrará al hacer una traza parcial del ambiente sobre el estado del universo $\rho_S = \text{Tr}_E \rho(t)$; similarmente el estado del ambiente está dado por $\rho_E = \text{Tr}_S \rho(t)$ [13]. El promedio temporal del universo es:

$$\omega = \langle\rho(t)\rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \rho(t) dt \quad (3.29)$$

de manera análoga ω_S y ω_E son el promedio temporal para el sistema y el ambiente respectivamente [12]. También se tiene el promedio $\langle.\rangle_\phi$ que es sobre todos los estados $|\phi\rangle \in \mathcal{H}_R$ de acuerdo a la medida estándar (unitariamente invariante). Esta medida se usa para hallar volúmenes de conjuntos de estados.

3.3. Idea Conceptual

Este nuevo tratamiento de los fundamentos de la mecánica estadística se tiene un universo. Dando como condición que el ambiente sea lo suficientemente grande. Este universo está descrito por un estado cuántico puro (se conoce el estado de manera exacta) que obedece una restricción global. Se plantea que el sistema alcanza el equilibrio térmico por medio de la interacción mutua (termalización) es producto del entrelazamiento del sistema y el ambiente [1]. Lo que se presentará más adelante es una definición más rigurosa que ayudará a dar cotas para la expresión “ambiente suficientemente grande”. Esta idea permite formular un principio canónico general: el sistema estará termalizado para casi todos los estados puros del universo. esto es soportado por límites cuantitativos. La restricción que se impone no es una específica, esto generaliza los resultados tradicionales dados en la literatura donde se toma por restricción la energía [4].

Ya poniendo lo dicho en un contexto un poco más matemático y siguiendo a Popescu et al. se supone tener un universo asilado y bastante grande este tiene dos partes el sistema S y el ambiente E . La dimensión del ambiente es mucho más grande que la del sistema. Además se le impone una restricción global al universo llamada R . Desde la mecánica cuántica esto puede ponerse como restricciones en el espacio de Hilbert, restricción de los estados posibles:

$$\mathcal{H}_R \subseteq \mathcal{H}_S \otimes \mathcal{H}_E, \quad (3.30)$$

Donde \mathcal{H}_S y \mathcal{H}_E son los espacios de Hilbert del sistema y el ambiente con dimension d_S y d_E respectivamente. Es bueno recalcar que R es una restricción arbitraria generalmente se toma como la energía del universo. Ahora se define el estado equiprobable del universo bajo R como:

$$\mathcal{E}_R = \frac{1}{d_R} \mathbb{1}_R, \quad (3.31)$$

Donde $\mathbb{1}_R$ es el operador identidad (proyección) sobre el espacio de Hilbert \mathcal{H}_R que tiene dimensión d_R . Esto se relaciona con el principio de probabilidades iguales porque este es el estado máximamente mezclado en \mathcal{H}_R [10] por ser así todos los estados bajo la restricción de R tienen la misma probabilidad de salir.

Definimos Ω_S como el estado canónico que está restringido por R cuando el universo se encuentra en el estado \mathcal{E}_R . Esto significa que si se hace una traza parcial del ambiente al universo da como resultado el estado canónico:

$$\Omega_S = \text{Tr}_E \mathcal{E}_R. \quad (3.32)$$

Para lo que sigue se hace un supuesto importante y es que el universo está en un estado puro $|\phi\rangle$ y no en un estado mixto \mathcal{E}_R , esto quiere decir que se conoce todo lo que es permitido por la mecánica cuántica del universo. Si estuviese en un estado mixto significaría que nosotros no tenemos toda la información que se pudiese tener [10]. Ahora lo que se quiere ver es que pese a que el estado del universo es puro el estado reducido del sistema

$$\rho_S = \text{Tr}_E |\phi\rangle \langle \phi|, \quad (3.33)$$

se acerca al estado canónico para la gran mayoría de los casos es decir:

$$\rho_S \approx \Omega_S. \quad (3.34)$$

Por consiguiente para casi todos los estados puros del universo $|\phi\rangle \in \mathcal{H}_R$ el sistema se comporta como si el universo estuviese en el estado mixto equiprobable \mathcal{E}_R . Este es el principio general canónico. Clarificando lo esbozado, el estado canónico del sistema Ω_S es el estado del sistema cuando el universo se encuentra en el estado equiprobable \mathcal{E}_R . Se puede interpretar el principio general canónico como un principio que estipula que las probabilidades iguales del sistema son aparentes porque para casi cualquier estado del universo, que sea puro, un subsistema de este universo que cumpla con ser lo suficientemente pequeño se comporta como si el universo estuviese en el estado equiprobable \mathcal{E}_R . Cabe recordar que aún no se ha especificado la restricción R entonces todo este análisis es general, la restricción no necesariamente debe ser la energía u otras cantidades que se conserven. Esto hace que Ω_S no deba ser obligatoriamente el estado canónico usual, puede ser el gran canónico o cualquier otro que sea acorde con la restricción impuesta[9]. Este principio puede ser de utilidad cuando la interacción entre el ambiente y el sistema no es débil o cuando las interacciones son complicadas el principio general canónico también aplica.

3.4. Formulación Matemática

Hasta ahora no se han entrado en los detalles ni en qué significa bastante pequeño o bastante grandes, ni tampoco se ha demostrado el principio general canónico. En esta sección se orientará en los detalles matemáticos, las herramientas usadas y la demostración de los teoremas que Popescu et. al. siguieron. En la siguiente sección se dará una perspectiva más física a lo hecho aquí.

Para empezar se debe decir cuál será la distancia que usaremos para darle un sentido de cercanía a los estados ρ_S y Ω_S . La distancia a usar es una bastante conocida en la teoría cuántica de la información, la distancia de traza [14]. Esta se define como:

$$D(\rho_S, \Omega_S) = \frac{1}{2} \text{Tr} |\rho_S - \Omega_S| = \frac{1}{2} \text{Tr} \sqrt{(\rho_S - \Omega_S)^\dagger (\rho_S - \Omega_S)}. \quad (3.35)$$

Esta distancia de traza se relaciona con la distancia de norma de manera sencilla

$$\|\rho_S - \Omega_S\|_1 = 2D(\rho_S, \Omega_S). \quad (3.36)$$

Teniendo ya una forma de darle sentido al concepto de que dos estados son cercanos entonces se puede plantear el teorema central de [1]:

Teorema 3.4.1. *Para un estado escogido de manera aleatoria $|\phi\rangle \in \mathcal{H}_R \subseteq \mathcal{H}_S \otimes \mathcal{H}_E$ y un $\epsilon > 0$ arbitrario, la distancia entre la matriz densidad reducida del sistema $\rho_S = \text{Tr}_E(|\phi\rangle \langle\phi|)$ y el estado canónico $\Omega_S = \text{Tr}_E(\mathcal{E}_R)$ esta dado probabilísticamente por*

$$Pr_\phi\{\|\rho_S - \Omega_S\|_1 \geq \eta\} \leq \eta', \quad (3.37)$$

Donde

$$\eta = \epsilon + \sqrt{\frac{d_S}{d_E^{eff}}}, \quad (3.38)$$

$$\eta' = 2 \exp(-Cd_R \epsilon^2). \quad (3.39)$$

y las constantes son: $C = (18\pi^3)^{-1}$, $d_R = \dim \mathcal{H}_R$, $d_S = \dim \mathcal{H}_S$. d_E^{eff} es la medida efectiva del tamaño del ambiente,

$$d_E^{eff} = \frac{1}{\text{Tr } \Omega_E^2} \geq \frac{d_R}{d_S}. \quad (3.40)$$

Donde $\Omega_E = \text{Tr}_S \mathcal{E}_R$. Ambas cantidades η y η' serán pequeñas esto implica que el estado estará cercano al estado canónico con alta probabilidad cuando la dimensión efectiva del ambiente sea mucho más grande que la del sistema (es decir $d_E^{eff} \gg d_S$) y $d_R \epsilon^2 \gg 1 \gg \epsilon$. Esta última condición se puede asegurar cuando el espacio accesible total es grande (es decir $d_R \gg 1$), escogiendo $\epsilon = d_R^{-\frac{1}{3}}$.

3.4.1. Lema de Levy

Para poder demostrar el teorema 3.4.1 se usará el lema de Levy [18]. Este dice que al seleccionar un punto ϕ aleatoriamente de una hiperesfera de dimensión alta y que $f(\phi)$ no cambie muy rápido, entonces $f(\phi) \approx \langle f \rangle$ con alta probabilidad, Más exactamente:

Lema 3.4.2. *Dada una función $f : \mathbb{S}^d \rightarrow \mathbb{R}$ definida en la hiperesfera d -dimensional \mathbb{S}^d , y un punto $\phi \in \mathbb{S}^d$ es escogido de manera uniformemente aleatoria,*

$$Pr_\phi\{|f(\phi) - \langle f \rangle| \geq \epsilon\} \leq 2 \exp\left(-\frac{2C(d+1)\epsilon^2}{\eta^2}\right) \quad (3.41)$$

donde η es la constante de Lipschitz de f , dado por $\eta = \sup |\nabla f|$ y $C = (18\pi^3)^{-1}$.

Los conceptos manejados por el teorema 3.4.2 son conocidos generalmente a excepción de la llamada constante de Lipschitz. Para poder entender qué es la constante de Lipschitz se debe ver primero qué significa que una función sea Lipschitz continua.

La definición de continuidad en el cálculo básico es:

Definición 3.4.1. continuidad Sea $f : I \rightarrow \mathbb{R}$ donde I puede ser un intervalo abierto (a, b) o uno cerrado $[a, b]$, además $C \in I$. Se dice que f es continua en C si y solo si para todo $\epsilon > 0$ existe un $\delta > 0$ tal que $|x - c| \rightarrow |f(x) - f(c)| < \epsilon$.

La definición anterior es la usada por lo general pero hay sutilezas en este concepto que no siempre son mostradas; como por ejemplo que δ depende de donde se ponga el punto C , esto se ve claramente en la siguiente función: $f : (0, 1) \rightarrow \mathbb{R}$, $f(x) = \frac{1}{x}$ al C estar más lejos del 0 permite un δ más grande pero al acercarse al 0 el δ debe ser más pequeño. La continuidad de Lipschitz permite que se defina un δ constante sin importar donde se encuentre el C . Para resolver este detalle se motiva la definición de Lipschitz continuo. Se es Lipschitz continuo con constante η si

$$|f(x) - f(y)| \leq \eta |x - y|, \quad (3.42)$$

esto permite decir que $|f(x) - f(y)| < \epsilon$ entonces $\delta < \frac{\epsilon}{\eta}$. Ahora si f es derivable y ∇f es acotado, para x y y dados existe ξ entre ellos tal que:

$$\implies f(x) - f(y) = \nabla f(\xi)(x - y) \quad (3.43)$$

$$\implies |f(x) - f(y)| \leq |\nabla f(\xi)| |x - y| \quad (3.44)$$

$$\implies |f(x) - f(y)| \leq \sup |\nabla f(\xi)| |x - y|, \quad (3.45)$$

como ∇f es acotado se tiene que $\sup |\nabla f(\xi)| = \eta$.

Gracias a la normalización, los estados puros en \mathcal{H}_R se pueden representar como puntos sobre la superficie de una hiperesfera de dimensión $2d_R - 1$, o sea \mathbb{S}^{2d_R-1} [10]. Luego se puede aplicar 3.4.2 a estado cuánticos ϕ aleatoriamente seleccionados. Para los estados seleccionados aleatoriamente $\phi \in \mathcal{H}_R$, se desea mostrar que $\|\rho_S - \Omega_S\|_1 \approx 0$ con alta probabilidad. Para poder usar 3.4.2 primero debe encontrarse la constante de Lipschitz ya teniendo una idea de qué es ser Lipschitz continuo se encontrará una cota para la constante η de la función $f(\phi) = \|\rho_S - \Omega_S\|_1$ que es la función que nos interesa para el problema físico. Entonces para poder lograr esto se procederá de la siguiente forma se definen dos estados reducidos $\rho_1 = \text{Tr}_E(|\phi_1\rangle \langle \phi_1|)$ y $\rho_2 = \text{Tr}_E(|\phi_2\rangle \langle \phi_2|)$, entonces

$$|f(\phi_1) - f(\phi_2)|^2 = \|\rho_1 - \Omega\|_1 - \|\rho_2 - \Omega\|_1^2. \quad (3.46)$$

como $\|M\|_1$ es una distancia (esto es $d(\rho_1, \Omega) = \|\rho_1 - \Omega\|_1$) es cierto para un espacio métrico que

$$|d(x, z) - d(y, z)| \leq d(x, y), \quad (3.47)$$

Por lo tanto

$$\left| \|\rho_1 - \Omega\|_1 - \|\rho_2 - \Omega\|_1 \right|^2 \leq \|\rho_1 - \rho_2\|_1^2 = \|\text{Tr}_E(|\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2|)\|_1^2. \quad (3.48)$$

Como existe una cota a la norma de una traza parcial dada por

$$\|\text{Tr}_{\mathcal{B}}(M)\|_p \leq [\dim(\mathcal{H}_{\mathcal{B}})]^{\frac{p-1}{p}} \|M\|_p, \quad (3.49)$$

entonces la cota sobre los estados reducidos queda

$$\|\text{Tr}_E(|\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2|)\|_1 \leq \| |\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2| \|_1. \quad (3.50)$$

Por lo tanto se tiene hasta ahora:

$$\|\rho_1 - \rho_2\|_1^2 \leq \| |\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2| \|_1^2, \quad (3.51)$$

Usando la hermiticidad de ρ y el teorema espectral se puede descomponer $\rho = UDU^\dagger$ donde U es un operador unitario y D es diagonal. Junto con las propiedades de la traza $\text{Tr}(\sqrt{UD^2U^\dagger}) = \text{Tr}(U\sqrt{D^2}U^\dagger) = \text{Tr}(\sqrt{D^2})$ se llega a

$$\| |\phi_1\rangle\langle\phi_1| - |\phi_2\rangle\langle\phi_2| \|_1^2 = 4(1 - |\langle\phi_1|\phi_2\rangle|^2) \quad (3.52)$$

entonces

$$4(1 - |\langle\phi_1|\phi_2\rangle|^2) \leq 4|\phi\rangle - |\phi\rangle|^2. \quad (3.53)$$

Uniando todos los pasos anteriores

$$|f(\phi_1) - f(\phi_2)|^2 \leq 4|\phi\rangle - |\phi\rangle|^2 \quad (3.54)$$

o sea

$$|f(\phi_1) - f(\phi_2)| \leq 2|\phi\rangle - |\phi\rangle|, \quad (3.55)$$

Con esto se muestra que $\eta \leq 2$.

3.4.2. Demostración del Principio General Canónico

En esta parte se dará una demostración matemática explícita del principio general canónico usando el lema de Levy se entrará en los detalles matemáticos de usar este lema y las cotas adicionales que se necesitan para poder llegar al teorema 3.4.1. Habiendo dado una cota para η ahora se puede usar por completo el lema 3.4.2 para la función $f(\phi) = \|\rho_S - \Omega_S\|_1$ recordando que se reemplazará d en el lema por $d = 2d_R - 1$. Tomando la parte derecha de la desigualdad de 3.4.2 y sustituyendo la dimensión se tiene:

$$2 \exp\left(-\frac{2C(d+1)\epsilon^2}{\eta^2}\right) = 2 \exp\left(-\frac{4Cd_R\epsilon^2}{\eta^2}\right) \quad (3.56)$$

como $\eta \leq 2$ entonces

$$2 \exp(-Cd_R\epsilon^2) \geq 2 \exp\left(-\frac{4Cd_R\epsilon^2}{\eta^2}\right) \geq \Pr_{\phi}[|f(\phi) - \langle f \rangle| \geq \epsilon]. \quad (3.57)$$

Mirando más atentamente $|f(\phi) - \langle f \rangle| \geq \epsilon$, como la norma de traza es un distancia se tiene que $\|\rho_S - \Omega_S\|_1 \geq 0$ entonces

$$\|\rho_S - \Omega_S\|_1 - \langle \|\rho_S - \Omega_S\|_1 \rangle_{\phi} \geq \epsilon. \quad (3.58)$$

Nombrando a $\mu = \epsilon + \langle \|\rho_S - \Omega_S\|_1 \rangle_{\phi}$ y $\mu' = 2 \exp(-Cd_R\epsilon^2)$ esto permite organizar el lema de Levy así:

$$\Pr_{\phi}[\|\rho_S - \Omega_S\|_1 \geq \mu] \leq \mu'. \quad (3.59)$$

Debido a que $d_R \gg 1$ Se asegura que ϵ y μ' son cantidades pequeñas al escoger $\epsilon = d_R^{-1/3}$. Para llegar a 3.4.1 falta acotar μ con las dimensiones de los espacios conocidos, se impondrá una cota a $\langle \|\rho_S - \Omega_S\|_1 \rangle_{\phi}$; lo primero para lograr esta empresa es acotar este promedio con trazas del estado del sistema y luego se calcularán estas trazas para poder dejar la cota en términos de las dimensiones del sistema y la dimensión efectiva del ambiente. Se procede a encontrar la relación entre $\|\rho_S - \Omega_S\|_1$ y $\|\rho_S - \Omega_S\|_2$ esto se hará para tener una facilidad de manejo ya que la norma de Hilbert-Schmidt ($\|\cdot\|_2$) es más sencilla para trabajar que la norma de traza y luego

se procederá con los planeado.

La relación entre estas dos normas se puede ver desde el manejo de matrices.

Sea M una matriz $n \times n$ se sabe que si M tiene λ_i valores propios entonces:

$$\text{Tr } M = \sum_i \lambda_i \quad (3.60)$$

con esto se puede escribir de manera explícita la norma de traza

$$\|M\|_1^2 = (\text{Tr } |M|)^2 = n^2 \left(\frac{1}{n} \sum_i |\lambda_i| \right)^2. \quad (3.61)$$

Como la función x^2 es convexa se puede usar la desigualdad de Jensen que dice: sean $a_1, a_2, \dots, a_n \leq 0$ constantes y $a_1 + \dots + a_n = 1$ sea $f : I \rightarrow \mathbb{R}$ donde I es un intervalo, $x_1, \dots, x_n \in I$ entonces

$$f(a_1 x_1 + \dots + a_n x_n) \leq a_1 f(x_1) + \dots + a_n f(x_n), \quad (3.62)$$

con esto se puede decir que

$$\left(\frac{1}{n} \sum_i |\lambda_i| \right)^2 \leq \frac{1}{n} \sum_i |\lambda_i|^2. \quad (3.63)$$

Pero se sabe que

$$\sum_i |\lambda_i|^2 = \text{Tr}(|M|^2) = \|M\|_2^2 \quad (3.64)$$

se llega entonces a la conclusión que

$$\|M\|_1^2 = n^2 \left(\frac{1}{n} \sum_i |\lambda_i| \right)^2 \leq n^2 \frac{1}{n} \sum_i |\lambda_i|^2 = n \|M\|_2^2 \quad (3.65)$$

Gracias a lo anterior la relación entre normas es:

$$\|\rho_S - \Omega_S\|_1 \leq \sqrt{d_S} \|\rho_S - \Omega_S\|_2 \quad (3.66)$$

Esta relación se usará un poco más adelante.

Volviendo al cálculo de $\langle \|\rho_S - \Omega_S\|_1 \rangle_\phi$ se acotará la norma de Hilbert-Schmidt y con la relación entre normas se dará la desigualdad que limite el

promedio de la norma de traza. Para empezar se recuerda que $\langle f^2 \rangle - \langle f \rangle^2 \geq 0$ donde $\langle f \rangle = \int_{\mathcal{M}} f(x)p(x)dx$. Tomando a f como $f = \|\rho_S - \Omega_S\|_2$ entonces

$$\langle \|\rho_S - \Omega_S\|_2 \rangle \leq \sqrt{\langle \|\rho_S - \Omega_S\|_2^2 \rangle} \quad (3.67)$$

acordándose que este promedio es tomado a los estados $|\phi\rangle$, se omitirá por ahora el subíndice indicando este promedio, esto hace que Ω_S se tome constante. Por hermiticidad de $\rho_S - \Omega_S$

$$\sqrt{\langle \|\rho_S - \Omega_S\|_2^2 \rangle} = \sqrt{\langle \text{Tr}(\rho_S - \Omega_S)^2 \rangle} \quad (3.68)$$

$$\sqrt{\langle \text{Tr}(\rho_S - \Omega_S)^2 \rangle} = \sqrt{\langle \text{Tr}(\rho_S)^2 \rangle - 2 \text{Tr}(\langle \rho_S \rangle \Omega_S) + \text{Tr}(\Omega_S^2)} \quad (3.69)$$

porque $\langle \rho_S \rangle = \Omega_S$. Luego se llega a

$$\langle \|\rho_S - \Omega_S\|_2 \rangle \leq \sqrt{\langle \|\rho_S - \Omega_S\|_2^2 \rangle} = \sqrt{\langle \text{Tr}(\rho_S)^2 \rangle - \text{Tr}(\Omega_S^2)}. \quad (3.70)$$

Por la relación entre la norma de traza y la norma de Hilbert-Schmidt se concluye lo que se quería

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{d_S(\langle \text{Tr}(\rho_S)^2 \rangle - \text{Tr}(\Omega_S^2))}. \quad (3.71)$$

Aunque ya se ha acotado $\langle \|\rho_S - \Omega_S\|_1 \rangle$ se quiere relacionar esta cota con las dimensiones del sistema para esto se procederá a demostrar la desigualdad

$$\langle \text{Tr} \rho_S^2 \rangle \leq \text{Tr} \langle \rho_S \rangle^2 + \text{Tr} \langle \rho_E \rangle^2, \quad (3.72)$$

recordando que el promedio es tomado con respecto a los estados $|\phi\rangle$, los métodos usados para encontrar esta desigualdad son usados también en destilación de entrelazamiento aleatoria y codificación de canal cuántico aleatorio [19]. Para poder hacer este cálculo se introduce una segunda copia del espacio de Hilbert. Ahora el problema se trabaja en $\mathcal{H}_R \otimes \mathcal{H}_{R'}$, donde $\mathcal{H}_{R'} \subseteq \mathcal{H}_{S'} \otimes \mathcal{H}_{E'}$. Percatándose de lo siguiente

$$\text{Tr}_S(\rho_S)^2 = \sum_k (\rho_{kk})^2 = \sum_{k,l,k',l'} (\rho_{kl})(\rho_{k'l'}) \langle kk'|ll' \rangle \langle l'l'|kk' \rangle. \quad (3.73)$$

Sea $F_{SS'}$ la operación "flip" $S \longleftrightarrow S'$ definida de esta manera:

$$F_{SS'} = \sum_{S,S'} |s'\rangle \langle s|_S \otimes |s\rangle \langle s'|_{S'} \quad (3.74)$$

Entonces

$$\begin{aligned} \sum_{k,l,k',l'} (\rho_{kl})(\rho_{k'l'}) \langle kk'|ll' \rangle \langle l'l'|kk' \rangle &= \text{Tr}_{SS'}((\rho_S \otimes \rho_{S'}) F_{SS'}) \\ &= \text{Tr}_{RR'}((|\phi\rangle \langle \phi| \otimes |\phi\rangle \langle \phi|)_{RR'} (F_{SS'} \otimes \mathbb{1}_{EE'})) \end{aligned}$$

Pero como se quiere $\langle \text{Tr}(\rho)^2 \rangle = \int \text{Tr}(\rho)^2 d\phi$. Entonces para resolver esto se requiere saber $V = \int (|\phi\rangle \langle \phi| \otimes |\phi\rangle \langle \phi|) d\phi$. V puede representarse como:

$$V = \alpha \Pi_{RR'}^{sim} + \beta \Pi_{RR'}^{anti}, \quad (3.75)$$

α y β son constantes y $\Pi_{RR'}^{sim/anti}$ son proyectores en el subespacio simétrico y antisimétrico de $\mathcal{H}_R \otimes \mathcal{H}_{R'}$ respectivamente, esto es posible por la invarianza unitaria de V. Debido a que

$$(|\phi\rangle \langle \phi| \otimes |\phi\rangle \langle \phi|) \frac{1}{\sqrt{2}}(|ab\rangle - |ba\rangle) = 0 \quad \forall a, b, \phi \quad (3.76)$$

la parte antisimétrica siempre debe ser 0 entonces $\beta = 0$. Por la normalización de V se llega a $\alpha = \frac{1}{\dim(RR'_{sim})}$, la dimensión del espacio RR'_{sim} está dada por el álgebra lineal $\dim(RR'_{sim}) = \frac{d_R(d_R+1)}{2}$. Entonces

$$V = \langle |\phi\rangle \langle \phi| \otimes |\phi\rangle \langle \phi| \rangle = \frac{2}{d_R(d_R+1)} \Pi_{RR'}^{sim}. \quad (3.77)$$

luego se tiene que

$$\langle \text{Tr}(\rho)^2 \rangle = \text{Tr}_{RR'} \left(\left(\frac{2}{d_R(d_R+1)} \Pi_{RR'}^{sim} \right) (F_{SS'} \otimes \mathbb{1}_{EE'}) \right) \quad (3.78)$$

Al ser $\Pi_{RR'}^{sim}$ un proyector simétrico se escribe

$$\Pi_{RR'}^{sim} = \frac{1}{2}(\mathbb{1}_{RR'} + (F_{RR'})) \quad (3.79)$$

donde $F_{RR'}$ es el operador "flip" $R \longleftrightarrow R'$. Como $F_{RR'}$ es un operador que actúa sobre RR' puede escribirse como $F_{RR'} = \mathbb{1}_{RR'}(F_{SS'} \otimes F_{EE'})$.

reuniendo todo lo anterior

$$\langle \text{Tr}_S \rho_S^2 \rangle = \text{Tr}_{RR'} \left(\frac{1}{d_R(d_R + 1)} (\mathbb{1}_{RR'} + \mathbb{1}_{RR'}(F_{SS'} \otimes F_{EE'})) (F_{SS'} \otimes \mathbb{1}_{EE'}) \right) \quad (3.80)$$

Distribuyendo y sabiendo que al hacer dos veces la operación "flip" no afecta nada se sigue lo siguiente

$$\text{Tr}_{RR'} \left(\frac{\mathbb{1}_{RR'}}{d_R(d_R + 1)} F_{SS'} \otimes \mathbb{1}_{EE'} + \frac{\mathbb{1}_{RR'}}{d_R(d_R + 1)} (\mathbb{1}_{SS'} \otimes F_{EE'}) \right). \quad (3.81)$$

Por las propiedades aditivas de la traza junto con $\mathbb{1}_R \otimes \mathbb{1}_{R'}$ y $\frac{1}{d_R(d_R+1)} \leq \frac{1}{d_R^2}$ se llega a:

$$\begin{aligned} \langle \text{Tr}_S \rho_S^2 \rangle &= \text{Tr}_{RR'} \left(\left(\frac{\mathbb{1}_{RR'}}{d_R(d_R + 1)} \right) (F_{SS'} \otimes \mathbb{1}_{EE'}) \right) \\ &+ \text{Tr}_{RR'} \left(\left(\frac{\mathbb{1}_{RR'}}{d_R(d_R + 1)} \right) (\mathbb{1}_{SS'} \otimes F_{EE'}) \right) \\ &\leq \text{Tr}_{RR'} \left(\left(\frac{\mathbb{1}_R}{d_R} \otimes \frac{\mathbb{1}_{R'}}{d_R} \right) (F_{SS'} \otimes \mathbb{1}_{EE'}) \right) \\ &+ \text{Tr}_{RR'} \left(\left(\frac{\mathbb{1}_R}{d_R} \otimes \frac{\mathbb{1}_{R'}}{d_R} \right) (\mathbb{1}_{SS'} \otimes F_{EE'}) \right) \end{aligned}$$

Recordando que $\frac{\mathbb{1}_R}{d_R} = \mathcal{E}_R$ y $\Omega_S = \text{Tr}_E(\mathcal{E}_R)$ se tiene que:

$$\begin{aligned} \text{Tr}_{RR'} \left(\frac{\mathbb{1}_R}{d_R} \otimes \frac{\mathbb{1}_{R'}}{d_R} (F_{SS'} \otimes \mathbb{1}_{EE'}) \right) + \text{Tr}_{RR'} \left(\frac{\mathbb{1}_R}{d_R} \otimes \frac{\mathbb{1}_{R'}}{d_R} (\mathbb{1}_{SS'} \otimes F_{EE'}) \right) \\ = \text{Tr}_{SS'}((\Omega_S \otimes \Omega_S) F_{SS'}) + \text{Tr}_{EE'}((\Omega_E \otimes \Omega_E) F_{EE'}). \end{aligned}$$

Entonces se llega a lo que se quería:

$$\langle \text{Tr}_S \rho_S^2 \rangle \leq \text{Tr}_S \Omega_S^2 + \text{Tr}_E \Omega_E^2, \quad (3.82)$$

y esto es lo mismo que

$$\langle \text{Tr}_S \rho_S^2 \rangle \leq \text{Tr}_S \langle \rho_S \rangle^2 + \text{Tr}_E \langle \rho_E \rangle^2 \quad (3.83)$$

Gracias al resultado 3.66 se obtiene

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{d_S (\text{Tr}_E \langle \rho_E \rangle^2)}. \quad (3.84)$$

Si se define $d_E^{eff} \equiv \frac{1}{\text{Tr}_E \Omega_E^2}$ como la dimensión efectiva del ambiente en el estado canónico, esto mide la dimensión del espacio en el que el ambiente es más probable de estar, como $\langle \rho_E \rangle_\phi = \Omega_E$ se concluye que

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{\frac{d_S}{d_E^{eff}}} \quad (3.85)$$

Ya con esto se tiene el teorema 3.4.1. Cuando el ambiente es mucho más grande que el sistema μ y μ' de la ecuación 3.59 serán pequeñas ($d_E^{eff} \gg d_S$) implicando $\|\rho_S - \Omega_S\|_1 \approx 0$ con alta probabilidad. Aunque ya se llegó a la desigualdad que se quería se puede notar lo siguiente, sean los valores propios de Ω_E iguales a λ_E^k con su máximo valor propio λ_E^{max} se ve que

$$\begin{aligned}
\text{Tr}_E \Omega_E^2 &= \sum_k (\lambda_E^k) \\
&\leq \lambda_E^{max} \sum_k \lambda_E^k \\
&= \max_{|\phi_E\rangle} \langle \phi_E | \text{Tr}_S \left(\frac{\mathbb{1}_R}{d_R} \right) | \phi_E \rangle \\
&= \max_{|\phi_E\rangle} \sum_s \langle s \phi_E | \frac{\mathbb{1}_R}{d_R} | s \phi_E \rangle \\
&\leq \frac{d_S}{d_E}
\end{aligned}$$

en conclusión $d_E^{eff} \geq d_R/d_S$ entonces se obtiene

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{\frac{d_S}{d_E^{eff}}} \leq \sqrt{\frac{d_S^2}{d_R}}. \quad (3.86)$$

Con esto se finaliza la demostración del teorema 3.4.1 en la siguiente sección se hablará de sus consecuencias físicas.

3.5. Significado Físico

El teorema anterior ya permite hablar del concepto importante que se sigue del artículo de Popescu et al la idea general de la física es poder dar una relación uno a uno entre las propiedades de un objeto físico y su representación matemática con esta correspondencia se puede decir que la teoría esta completa cuando todas las propiedades que son posibles de medir tienen su semejante en la teoría [7]. Por ejemplo en la física clásica a las cantidades como velocidad y distancia se le asigna los símbolos matemáticos v y x , el cuerpo puede ser especificado dando estas dos cantidades en un tiempo determinado. La mecánica clásica nos da un ejemplo sencillo pero cuando se pasa a la mecánica estadística se encuentra perspectivas diferentes que entran en conflicto con esta sencilla idea. En los inicios de la mecánica estadística hubo muchas controversias dado a que ahora la idea determinista que Newton y otros habían mostrado como cierta (Se

podía describir la naturaleza de manera predecible) ahora se introducía la probabilidad a sistemas que según Newton eran deterministas. En este punto se debe hacer claro que la necesidad de la probabilidad era por la falta de conocimiento pero no por un indeterminismo intrínseco como en la mecánica cuántica. Dos propuestas importantes surgieron en busca de una base conceptual sobre estas probabilidades, la visión de Gibbs y la de Boltzmann. Gibbs propuso la idea del ensamble, esta es una idea que hasta el día de hoy se sigue usando porque da los resultados correctos para las propiedades de sistemas termodinámicos.

La idea de Gibbs es la siguiente: Se tiene un sistema macroscópico con ciertos parámetros que se pueden medir. Este sistema está compuesto por varias partículas que tienen posiciones y momentos, pero estas propiedades están restringidas por los parámetros macroscópicos. Entonces como no se puede conocer el microestado del sistema, no se puede medir cada posición y momento de cada partícula, se tiene una libertad sobre cuál es el microestado dado el macroestado. El ensamble es el conjunto de microestados que cumplen con el macroestado, mientras el microestado cumpla con los parámetros macroscópicos hará parte del ensamble así el sistema no se encuentre en ese microestado. En teoría no hay limitación sobre el número de microestados que pertenezcan a un ensamble, este puede tener infinitos microestados. Esta idea de base para la función de densidad en el espacio de fase, cada punto en este espacio es un microestado estos microestados no interactúan entre ellos, con esto se puede construir la mecánica estadística que se conoce [5].

Por el contrario Boltzmann da una concepción de la mecánica estadística más intuitiva; él propone N partículas que interactúan entre ellas y cumplen los parámetros macroscópicos. Las partículas no son imaginarias como en el caso de los ensambles de Gibbs, ellas son las partículas que están en el sistema que se está estudiando. Se esperaría que la solución de Boltzmann fuera la que diera los resultados correctos dada la sencillez de su propuesta pero es la formulación de Gibbs la que da la termodinámica correcta. E.T Jaynes en [17] muestra las diferencias matemáticas de cada perspectiva y llega a la conclusión que la formulación de Gibbs da la entropía correcta sin importar cual sea el tipo de interacción que tenga el sistema, además encuentra que la entropía que sale de la formulación Boltzmanniana no es la misma entropía dada por la termodinámica, la entropía de Boltzmann es

correcta cuando no hay interacción entre las partículas.

Ya con la comprobación experimental se debería aceptar la perspectiva de Gibbs como la correcta y dejar a un lado la de Boltzmann pero esto deja a la mecánica estadística en un contexto filosófico poco fuerte, aunque la física busque una correspondencia entre experimento y teoría no se puede aceptar cualquier teoría solo por tener congruencia con el experimento, la física debe buscar una estabilidad conceptual en sus teorías. ¿ Por qué la perspectiva de Gibbs tiene un fundamento congruente ? lo que propone Gibbs aunque matemáticamente aceptable se basa en todos los microestados en los que el macroestado no se encuentra es decir, que para poder determinar la propiedades de estado el cual se está estudiando se debe estudiar todos los estados en los que el sistema no se encuentra pero podría estar, pero por qué para analizar un sistema que se encuentra en un estado definido y exacto debería examinar los posibles infinitos estado en donde no se encuentra si se hace un promedio sobre estos microestados ¿ cuál es el significado de este promedio?. Este problema ontológico deja la perspectiva de Gibbs en duda, mientras la de Boltzmann habla del sistema que se examina sin ningún problema de este estilo.

La lucha entre las perspectiva ha sido una de aún no se resuelve lo que se ha mostrado hasta ahora ha sido la matemática de la tentativa de unir ambas perspectivas por Popescu et al. pero ¿ por qué estos resultado matemáticos intentan dar una luz a la resolución de estas perspectivas? se puede ver fácilmente que cuando se habla de del estado $\Omega_S = \text{Tr}_B \mathcal{E}_R$ este condensa la perspectiva de Gibbs mientras que $\rho_S = \text{Tr}_E |\phi\rangle \langle\phi|$ habla de un subsistema específico bajo una restricción esta es la perspectiva de Boltzmann por lo tanto el resultado mostrado acá dice: a menos que se tenga un estado bastante especial se puede asegurar que el subsistema de un universo muy grande, dado por un estado puro, se comporta de la misma manera que si el universo se encontrara en un estado máximamente mezclado. Esto corresponde a decir que para la mayoría de los estados en un espacio de Hilbert se puede reconciliar ambas perspectivas.

Hay otro punto que da más mérito a la perspectiva mostrada acá y es que el estado Ω_S es la forma de mostrar la ignorancia subjetiva que se tiene del sistema. Se sabe que este es el punto de partida para la mecánica estadística pero esto también es un fundamento que generalmente es muy cuestionado gracias al teorema 3.4.1 se cambia esa probabilidad subjetiva por una ob-

jetivad debido al entrelazamiento. Por este cambio a un fundamente más sólido no se debe entrar en el problema de ergodicidad ya que en general la solución del problema de esa probabilidad se dice que el sistema pasa por todos los estados,y eso solo complica la pregunta dado que hay sistemas que no pasan por todos los estados pero aún así son descritos por un estado canónico. Ya no debe entrarse en la pregunta sobre la ergodicidad del sistema dado que esta perspectiva muestra que los promedios no son necesarios en el momento de darle un fundamento a la mecánica estadística.

Capítulo 4

Evolución Hacia El Equilibrio

En el capítulo anterior se expuso la idea de Popescu et al. [1] para poder reconciliar las ideas fundamentales de la mecánica estadística. Se mostró que para estados genéricos de un universo se tiene que el estado del sistema va a estar muy cercano del estado canónico. Ya con esta idea la siguiente pregunta que se puede hacer es sobre la termalización, ¿ Este fenómeno cómo ocurre? ¿ se puede deducir se las ecuaciones básicas(Schrödinger, Newton , etc)?. Este capítulo mostrará lo que Linden et al. [20] proponen para poder responder estas preguntas siguiendo las ideas ya dadas en el capítulo anterior. La generalidad que se ha dado a los fundamentos gracias Popescu et al. permite una flexibilidad al momento de manejar nuevos problemas por eso se ve que además de tener una buena base filosófica las matemáticas dan un buen esqueleto para sostener las bases y seguir construyendo con ellas. Ahora se quiere tratar el caso en que el estado no se encuentra en un tiempo específico sino ya en un tiempo general. Entonces se deberá apuntar la maquinaria que ya se tiene para poder redondear el problema; esto será lo primero a precisar. Luego se mostrará que por lo general los estados del sistema llegan al equilibrio y duran allí gran parte de su tiempo, todo esto bajo las especificaciones dadas al comienzo. Se seguirá dando un teorema que muestra como para la gran mayoría de casos de estados genéricos se llegará al equilibrio, aquí también se dará un argumento sobre cómo los estados lejanos al equilibrio llegan a equilibrarse.

Además se mostrará la independencia del estado inicial y problemas que no se han llegado a solucionar desde la perspectiva de (referencia).

4.1. Especificación del Equilibrio

Las preguntas planteadas anteriormente quieren enfocarse en el equilibrio pero para poder seguir viendo qué ocurre con este fenómeno se deberá establecer lo que significa, basados en conceptos fenomenológicos se dará la definición del equilibrio y qué se espera de este estado.

La manera más intuitiva que se tiene sobre el equilibrio es que el sistema, que se está analizando, se mantiene en el mismo estado durante un largo periodo de tiempo con las mismas características; entonces se dice que un sistema se equilibra si este evoluciona a un estado específico (puede ser puro pero en general es mixto) y se mantiene allí por casi todo el tiempo. Aún no se dice nada sobre las dependencias que pueda tener, o sea se puede tener una condición de dependencia laxa esto permite decir que el estado de equilibrio depende o no del estado inicial del subsistema y/o del estado inicial del ambiente de forma arbitraria. Dado esto no es importante cuál sea el estado de equilibrio este puede ser la distribución de Boltzmann o cualquier otra. El concepto de equilibrio dicho aquí es una forma general de hablar sobre este fenómeno porque se da mucha libertad al estado y sus dependencias sobre el ambiente por eso se seguirá puliendo para no dejarlo tan general.

El estado de equilibrio del sistema no debería depender exactamente del estado inicial del baño. Esto quiere decir que el baño debe tener unos parámetros macroscópicos (como la temperatura) tales que al cuando se llega al equilibrio el estado dependa de la de estos parámetros, ya se restringe un poco más a lo que se llamará equilibrio. Esta idea también es proveniente de lo que normalmente se espera del equilibrio porque los parámetros macroscópicos son los que en general se tienen completamente especificados y estos establecen el equilibrio. El estado exacto del baño no debería jugar un papel tan importante porque para los mismos parámetros macroscópicos pueden haber varios estados del baño que concuerden con ellos, se supone que con ciertos parámetros macroscópicos dados un subsistema llegará al equilibrio sin importar cual ha sido su estado inicial. Pueden dos subsistemas preparados con los mismos parámetros macroscópicos haber sido

iniciados en un estado A y el otro en el estado B llegarán al mismo estado de equilibrio sin importar en cual de los dos hayan empezado. Esto motiva lo siguiente: Si el subsistema es pequeño en comparación con el ambiente el estado de equilibrio del subsistema debería ser independiente de su estado inicial. Pero para poder corroborar y unir estas ideas con los resultados conocidos se impone una última restricción. Bajo condiciones del estado inicial y el Hamiltoniano, el estado de equilibrio del subsistema puede ser escrito como $\rho_S = \frac{1}{Z} \exp\left(-\frac{H_S}{k_B T}\right)$ la forma familiar dada por la mecánica estadística.

Haber podido dividir el problema de la termalización de esta forma permite analizar cada uno de los aspectos por separado además de darle una generalidad a todo el tratamiento sin tener que restringirse a situaciones que usualmente se le asocian a la termalización. Por ejemplo no se debe quedar en el régimen de corta o débil interacción entre el sistema y el baño o decir que el baño es uno típico (Dado una temperatura o rango de energía). Pueden tomarse situaciones en los que el sistema no llegue a equilibrio.

Lo siguiente es que se mostrará con suposiciones no muy fuertes los primeros dos supuestos, la idea de que un sistema llegará al equilibrio y que el estado no depende del ambiente, son propiedades de sistemas cuánticos. La evolución del sistema tiene implicado al Hamiltoniano por eso se estudiará el siguiente:

$$H = \sum_k E_k |E_k\rangle \langle E_k|, \quad (4.1)$$

donde $|E_k\rangle$ es el estado propio con energía E_k . Al Hamiltoniano anterior se le dará la única restricción de que tenga brechas de energías no degeneradas. Esto quiere decir que para una diferencia de energías dada, solo existe un par posible de estados con esa diferencia. El ejemplo puesto por (Linden et al.) es: si se tiene 4 valores propios de energía E_k, E_l, E_m, E_n entonces $E_k - E_l = E_m - E_n$ implica $k = l$ y $m = n$, o $k = m$ y $l = n$. Esto implica que los niveles de energía no son iguales para diferentes estados (no son degenerados).

Esta restricción del Hamiltoniano comprende que no importa como se divida el subsistema y el baño siempre van a estar interactuando, esto excluye los Hamiltonianos no interactuantes ($H = H_S + H_E$) estos Hamiltonianos tienen muchas brechas de energía degeneradas. Véase que si no hay inter-

acción en el Hamiltoniano la energía es $E = E_S + E_B$ y sean E_1, E_2, E_3, E_4 tales que se satisfaga $E_i = E_i^S + E_i^E$ $i = 1, 2, 3, 4$, esto lleva a una brecha degenerada. Este supuesto no es tan fuerte como pueda llegar a parecer porque cualquier perturbación que se le haga al Hamiltoniano romperá las degeneraciones sin importar lo pequeña que sea la perturbación. Aunque estos cambios se tardan en hacer efecto sobre la evolución del sistema las escalas temporales no son importantes en este momento. Gracias a esta restricción se pueden incluir interacciones complejas que por lo general no son analizadas en la literatura como interacciones de larga distancia o interacciones entre todas las partículas esto hace que la energía no llegue a ser una cantidad extensiva.

4.2. Equilibración

La perspectiva principal que se quiere dar a entender siguiendo a [20] viene en forma del siguiente resultado: Para cada estado puro de un sistema cuántico que se compone por un número grande de estados de energía propios y el cual evoluciona bajo un Hamiltoniano que tiene brechas de energías no degeneradas y por lo demás arbitrario, es tal que cada pequeño subsistema llegará al equilibrio. Esto quiere decir que todos los subsistemas pequeños cumplirán con las ideas anteriores sobre equilibrio exactamente que el sistema evolucionará a un estado particular y se quedará cercano a él o en este durante la mayoría del tiempo.

Hay en este resultado un requerimiento que anteriormente no fue nombrado, el requerimiento de una cantidad grande de estados propios de energía. La necesidad de que el sistema tenga muchos estados propios de energía es equivalente a decir que el estado variará bastante durante su evolución temporal, por ejemplo el caso trivial de un solo estado propio de energía es claro que este no cambiará para nada. Este caso tan particular no llega a ser de mucho interés porque este estado no evolucionará a otro y se diría que ya se encuentra en equilibrio. Los resultados que se quieren mostrar toman estados fuera del equilibrio. En este punto se usan las otras suposiciones hechas anteriormente para que el subsistema no dependa del estado inicial el subsistema debe perder esta información; si el subsistema empieza lejano al equilibrio este pasará por muchos estados en su camino al equilibrio lo cual implica que el universo también evolucione en muchos

estados. El hecho de que el subsistema haya llegado al equilibrio no significa que el universo deje de evolucionar, debido a la unitariedad, debe seguir evolucionando con la misma proporción que antes. Para que los estados en los que el subsistema se encuentre en no equilibrio ocurran poco, los estados del universo en los que el subsistema se encuentre en esas condiciones deben ser una fracción muy pequeña del total de estados por donde pasa el universo. Por esto el universo debe pasar por muchos estados y el requerimiento de que tenga muchos estados propios de energía se valida. Otra forma de saber qué pasa con el universo es observar qué ocurre con el ambiente. Por unitariedad el universo debe seguir evolucionando aunque el subsistema y se encuentre en equilibrio y no cambie. Esta Evolución puede darse por el cambio de correlaciones entre el subsistema y el baño o por cambios en el estado del baño. Lo que se muestra es: cuando el estado del baño pasa por muchos estados diferentes, cualquier subsistema alcanza el equilibrio. También se muestra que cuando el estado del universo pasa por muchos estados diferentes cualquier estado pequeño subsistema alcanza el equilibrio. Estas dos ideas exponen que la equilibración ocurre en estados productos iniciales entre el subsistema y el baño, para casi todos los estados iniciales del baño.

Para poder empezar a construir estos conceptos matemáticamente se puede empezar con la concepción matemática de evolucionar por muchos estados diferentes, esta se comprime en la dimensión efectiva del estado promediado temporalmente $d^{eff}(\omega)$ donde $\omega = \langle \rho(t) \rangle_t$. Esta medida de evolución por muchos estados se puede relacionar con los estados propios de energía así:

Se toma el estado del universo

$$|\psi(t)\rangle = \sum_k c_k e^{-i \frac{E_k t}{\hbar}} |E_k\rangle \quad (4.2)$$

cuyo operador de densidad es

$$\rho(t) = \sum_{k,l} c_k c_l^* e^{\frac{-i(E_k - E_l)t}{\hbar}} |E_k\rangle \langle E_l|, \quad (4.3)$$

su promedio temporal es, recordando la condición de no degeneración de

los niveles de energía,

$$\omega = \sum_k |c_k|^2 |E_k\rangle \langle E_k| \quad (4.4)$$

y dando la relación se calcula la dimensión efectiva resultando

$$d^{eff}(\omega) = \frac{1}{Tr(\omega^2)} = \frac{1}{\sum_k |c_k|^4}. \quad (4.5)$$

Similarmente el hecho de que el baño pase por muchos estados diferentes viene dado por $d^{eff}(\omega_B)$ con $\omega_B = \langle \rho(t) \rangle_t$. Debido a que el baño al evolucionar pase por muchos más estados dado que el universo debe seguir evolucionando y el subsistema quede en un espacio de estados más pequeño se preve que d_S es mucho más pequeño que $d^{eff}(\omega_B)$. Para formular ya el primer teorema se quiere ver la distancia entre $\rho_S(t)$ y su promedio temporal $\omega_S = \langle \rho_S(t) \rangle_t$. Como se espera que $\rho_S(t)$ vaya fluctuando alrededor de ω_S se analizará el promedio temporal de su distancia $\langle D(\rho_S(t), \omega_S) \rangle_t$ cuando este sea muy pequeño el subsistema debe pasar gran parte del tiempo muy cerca a ω_S . Esto quiere decir que el subsistema se equilibrará (según la definición anterior) a ω_S .

Teorema 4.2.1. *Considere cualquier estado $|\psi(t)\rangle \in \mathcal{H}$ evolucionando bajo un Hamiltoniano con brechas de energía no-degeneradas. Luego la distancia promedio entre $\rho_S(t)$ y su promedio temporal ω_S está acotado por:*

$$\langle D(\rho_S(t), \omega_S) \rangle_t \leq \frac{1}{2} \sqrt{\frac{d_S}{d^{eff}(\omega_B)}} \leq \frac{1}{2} \sqrt{\frac{d_S^2}{d^{eff}(\omega)}}. \quad (4.6)$$

Demostración: Recordando la relación entre la distancia de traza y la norma de Hilbert-Schmidt que se usó en el capítulo anterior

$$\|M\|_1 \leq \sqrt{n} \|M\|_2, \quad (4.7)$$

se usa para el operador $D(\rho_1, \rho_2)$:

$$\frac{1}{2} \text{Tr}_S \sqrt{(\rho_1 - \rho_2)^2} \leq \frac{1}{2} \sqrt{d_S \text{Tr}_S (\rho_1 - \rho_2)^2}, \quad (4.8)$$

por la concavidad de la función raíz cuadrada se obtiene

$$\langle D(\rho_S(t), \omega_S) \rangle_t \leq \sqrt{d_S \langle \text{Tr}_S(\rho_S(t) - \omega_S)^2 \rangle_t}. \quad (4.9)$$

Usando las expansiones para ρ_S y ω_S

$$\rho_S(t) = \sum_{k,l} c_k c_l^* e^{\frac{-i(E_k - E_l)t}{\hbar}} \text{Tr}_B(|E_k\rangle \langle E_l|), \quad (4.10)$$

$$\omega_S = \sum |c_k|^2 \text{Tr}_B(|E_k\rangle \langle E_k|), \quad (4.11)$$

se puede escribir $\langle \text{Tr}_S(\rho_S(t) - \omega_S)^2 \rangle_t$ como

$$\langle \text{Tr}_S(\rho_S(t) - \omega_S)^2 \rangle_t = \sum_{k \neq l} \sum_{m \neq n} \mathcal{T}_{klmn} \text{Tr}_S[\text{Tr}_B |E_k\rangle \langle E_l| \text{Tr}_B |E_k\rangle \langle E_l|] \quad (4.12)$$

donde \mathcal{T}_{klmn} es :

$$\mathcal{T}_{klmn} = c_k c_l^* c_m c_n^* \left\langle e^{\frac{-i(E_k - E_l + E_m - E_n)t}{\hbar}} \right\rangle_t. \quad (4.13)$$

Debido a la restricción impuesta al Hamiltoniano de brechas de energías no degeneradas y como solo se toman elementos $k \neq l$ y $m \neq n$ los terminos que son diferentes de 0 son $k = n$ y $l = m$,

entonces

$$\begin{aligned} \langle \text{Tr}_S(\rho_S(t) - \omega_S)^2 \rangle_t &= \sum_{k \neq l} |c_k|^2 |c_l|^2 \text{Tr}_S[\text{Tr}_B(|E_k\rangle \langle E_l|) \text{Tr}_B(|E_k\rangle \langle E_l|)] \\ &= \sum_{k \neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_l|s'b\rangle \langle s'b'|E_l\rangle \langle E_k|sb'\rangle \\ &= \sum_{k \neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_k|sb'\rangle \langle s'b'|E_l\rangle \langle E_l|s'b\rangle \\ &= \sum_{k \neq l} |c_k|^2 |c_l|^2 \text{Tr}_B[\text{Tr}_S(|E_k\rangle \langle E_k|) \text{Tr}_S(|E_l\rangle \langle E_l|)] \\ &= \sum_{k \neq l} \text{Tr}_B[\text{Tr}_S(|c_k|^2 |E_k\rangle \langle E_k|) \text{Tr}_S(|c_l|^2 |E_l\rangle \langle E_l|)] \\ &= \text{Tr}_B \omega_B^2 - \sum_k |c_k|^4 \text{Tr}_S[\text{Tr}_B(|E_k\rangle \langle E_k|)]^2 \leq \text{Tr}_B(\omega_B^2). \end{aligned}$$

La última igualdad se encuentra recordando la definición de ω_B y observando la segunda igualdad. Por la subaditividad débil de la entropía de Rényi [21]:

$$\text{Tr}(\omega^2) \geq \frac{\text{Tr}_B(\omega_B^2)}{\text{rank}(\rho_S)} \geq \frac{\text{Tr}_B(\omega_B^2)}{d_S}. \quad (4.14)$$

Uniendo todos los resultados con la desigualdad inicial del promedio de la distancia

$$\langle D(\rho_S(t), \omega_S) \rangle_t \leq \frac{1}{2} \sqrt{d_S \text{Tr}_B(\omega_B^2)} \leq \frac{1}{2} \sqrt{d_S^2 \text{Tr}(\omega^2)} = \frac{1}{2} \sqrt{\frac{d_S^2}{d^{eff}(\omega)}}. \quad (4.15)$$

Este resultado da base para hablar de la termalización de una forma matemática, se ve que el subsistema se equilibra cuando la dimensión de $d^{eff}(\omega)$ sea mucho mayor que la dimensión de dos copias del subsistema (d_S^2) o cuando la dimensión efectiva explorada por el baño $d^{eff}(\omega_B)$ sea mucho más grande que la dimensión del subsistema.

EL resultado anterior tiene varias generalidades que se quieren recodar. La restricción impuesta sobre el Hamiltoniano es una que no excluye muchos Hamiltonianos [10]. Además de que esta ha sido la única restricción sobre todo el universo, no se ha especificado nada del baño ni del subsistema. El baño no está necesariamente en equilibrio no se le ha dado ninguna interpretación con respecto a la termodinámica usual a ningún objeto tratado hasta ahora. no se ha hablado tampoco de ninguna forma en la que el subsistema llega al equilibrio ni que esta en algún estado específico.

Los valores propios de energía tampoco son importantes en las cotas dadas anteriormente, en el teorema al ser demostrado fue encontrado algunos valores propios de energía que al ser promediados dan 0. La energía es importante al buscar las formas exactas en las que evoluciona el sistema pero aquí se demostró que para la equilibración en intervalos de tiempo muy grandes no son muy importantes, las cotas son independientes del tiempo. La forma en que se dividió el universo (subsistema y baño) no es importante, solo es importante para el teorema 4.2.1 la dimensión del subsistema y no la especificación de la forma o un subsistema particular. Esto permite decir que cualquier subsistema con dimensión d_S estará en equilibrio, los

varios subsistemas bastante pequeños de dimensión d_S también estarán en equilibrio.

El teorema 4.2.1 puso una cota a la fluctuación de $\rho_S(t)$ alrededor de ω_S esto ya es un inicio pero aunque esta cota exista no se ha hablado de cuales sistemas tendrán fluctuaciones suficientemente pequeñas para poder decir que se encuentra en equilibrio. Ahora se explorará cuáles son los estados que se equilibrarán, el siguiente teorema dirá cuales estados tienen fluctuaciones muy pequeñas.

Teorema 4.2.2. *1. El promedio de la dimensión efectiva $\langle d^{eff}(\omega) \rangle_\psi$, donde el promedio es sobre estados puros aleatorios uniformemente distribuidos $|\psi\rangle \in \mathcal{H}_R \subset \mathcal{H}$. Es tal que*

$$\langle d^{eff}(\omega) \rangle_\psi \geq \frac{d_R}{2}. \quad (4.16)$$

2. Para un estado aleatorio $|\psi\rangle \in \mathcal{H}_R \subset \mathcal{H}$, la probabilidad $\Pr_\psi\{d^{eff}(\omega) < \frac{d_R}{4}\}$ de que $d^{eff}(\omega)$ es más pequeña que $\frac{d_R}{4}$ es exponencialmente pequeña:

$$\Pr_\psi\{d^{eff}(\omega) < \frac{d_R}{4}\} \leq 2 \exp\{-C\sqrt{d_R}\}, \quad (4.17)$$

con constante $c = \frac{(\ln 2)^2}{72\pi^3} \approx 10^{-4}$.

Demostración

1. Primero se prueba la cota de la pureza esperada de ω . Para esta prueba se usarán varios resultados ya empleados en el capítulo anterior como: la operación "flip" F y la identidad $\langle |\psi\rangle \langle \psi| \otimes |\psi\rangle \langle \psi| \rangle_\psi = \frac{\Pi_{RR}(\mathbb{1}+F)}{d_R(d_R+1)}$. Además se usará la notación de $|E_k\rangle \equiv |k\rangle$ y $|E_k\rangle \otimes |E_l\rangle \equiv |kl\rangle$. Se introduce el mapeo de desfase como $\$[\rho] \equiv \sum_k |k\rangle \langle k| \rho |k\rangle \langle k|$, implica que $\omega = \langle |\psi\rangle \langle \psi| \rangle_t = \$[|\psi\rangle \langle \psi|]$. Entonces usando la operación "flip" se escribe $\langle \text{Tr}(\omega)^2 \rangle_\psi$ como

$$\begin{aligned}
\langle \text{Tr}(\omega)^2 \rangle_\psi &= \langle \text{Tr}(\omega \otimes \omega) F \rangle_\psi \\
&= \text{Tr}(\$ \otimes \$ [\langle \psi \rangle \langle \psi | \otimes |\psi \rangle \langle \psi |]_\psi F) \\
&= \text{Tr} \left(\$ \otimes \$ \left[\frac{\Pi_{RR}(\mathbb{1} + F)}{d_R(d_R + 1)} \right] F \right) \\
&= \sum_{kl} \text{Tr} \left(|kl\rangle \langle kl| \left(\frac{\Pi_{RR}(\mathbb{1} + F)}{d_R(d_R + 1)} \right) |kl\rangle \langle kl| F \right) \\
&= \sum_{kl} \text{Tr}(|kl\rangle \langle lk|) \left(\frac{\langle kl| \Pi_{RR}(|kl\rangle + |lk\rangle)}{d_R(d_R + 1)} \right) \\
&= \sum_k \frac{2 \langle kk| \Pi_{RR} |kk\rangle}{d_R(d_R + 1)} \\
&\leq \sum_k \frac{2 \langle k| \Pi_R |k\rangle}{d_R(d_R + 1)} < \frac{2}{d_R}
\end{aligned}$$

se sigue directamente que

$$\langle d^{eff}(\omega) \rangle_\psi = \left\langle \frac{1}{\text{Tr}(\omega)^2} \right\rangle_\psi \geq \frac{1}{\langle \text{Tr}(\omega)^2 \rangle_\psi} > \frac{d_R}{2} \quad (4.18)$$

concluyendo la prueba.

2. Para demostrar la segunda parte se usará el lema de levy [?] pero no directamente sino sobre la función

$$f(\psi) \equiv f(\vec{x}(\psi)) = \ln \left(\text{Tr} \left(\tilde{\$} [|\psi\rangle \langle \psi|]^2 \right) \right) \quad (4.19)$$

Donde el operador $\tilde{\$}$ actúa sobre el subespacio $\mathcal{H}_T \subseteq \mathcal{H}$ generado por los estados de energía con proyección diferente de cero sobre \mathcal{H}_R (estados que satisfagan $\langle k| \Pi_R |k\rangle \neq 0$). El subespacio \mathcal{H}_T contiene todos los estados que pueden aparecer durante la evolución temporal

del estado inicial en \mathcal{H}_R , y $\tilde{\$}$ mapea estos estados de regreso a \mathcal{H}_R de acuerdo a

$$\tilde{\$}[\rho] = \sum_k |\tilde{k}\rangle \langle k| \rho |k\rangle \langle \tilde{k}| \quad y \quad |\tilde{k}\rangle = \frac{1}{\sqrt{\langle k| \Pi_R |k\rangle}} \Pi_R |k\rangle. \quad (4.20)$$

nótese que cuando el Hamiltoniano conmuta con Π_R , $\tilde{\$}$ es idéntico al $\$$ en \mathcal{H}_T . Calculando el promedio de la función se encuentra

$$\begin{aligned} \left\langle \ln \left(\text{Tr} \left(\tilde{\$}[|\psi\rangle \langle \psi|]^2 \right) \right) \right\rangle_\psi &\leq \ln \left\langle \text{Tr} \left(\tilde{\$}[|\psi\rangle \langle \psi|]^2 \right) \right\rangle_\psi \\ &= \ln \text{Tr} \left(\tilde{\$} \otimes \tilde{\$} [|\psi\rangle \langle \psi| \otimes |\psi\rangle \langle \psi|] F \right) \\ &= \ln \left(\text{Tr} \left(\tilde{\$} \otimes \tilde{\$} \left[\frac{\Pi_{RR}(\mathbb{1} + F)}{d_R(d_R + 1)} \right] F \right) \right) \\ &= \ln \left(\sum_{kl} \text{Tr} \left(|\tilde{k}l\rangle \langle kl| \left(\frac{\Pi_{RR}(\mathbb{1} + F)}{d_R(d_R + 1)} \right) |kl\rangle \langle \tilde{l}k| \right) \right) \\ &= \ln \left(\sum_{kl} \langle \tilde{l}k| |\tilde{k}l\rangle \left(\frac{\langle kl| \Pi_{RR}(|kl\rangle + |lk\rangle)}{d_R(d_R + 1)} \right) \right) \\ &\leq \ln \left(\frac{2}{d_R(d_R + 1)} \sum_{kl} \langle lk| \Pi_{RR} |kl\rangle \right) \\ &= \ln \left(\frac{2}{d_R(d_R + 1)} \sum_k \langle k| \Pi_R |k\rangle \right) < \ln \left(\frac{2}{d_R} \right). \end{aligned}$$

Para acotar la constante de Lipchitz de la función $f(\psi)$ se usará otra función

$$g(\psi) = \ln \text{Tr} \left[\left(\sum_n |\hat{n}\rangle \langle \hat{n}| \tilde{\$}[|\psi\rangle \langle \psi|] |\hat{n}\rangle \langle \hat{n}| \right)^2 \right] \quad (4.21)$$

donde $|\hat{n}\rangle$ es una base ortonormal de \mathcal{H}_R . Escribiendo

$$t_{nk0} = \text{Re}[\langle \hat{n}| |\tilde{k}\rangle \langle k| \psi \rangle] \quad y \quad t_{nk1} = \text{Im}[\langle \hat{n}| |\tilde{k}\rangle \langle k| \psi \rangle] \quad (4.22)$$

se sigue que

$$g(\psi) = \ln \text{Tr} \left[\left(\sum_{nkz} t_{nkz}^2 |\hat{n}\rangle \langle \hat{n}| \right)^2 \right] = \ln \sum_n \left(\sum_{kz} t_{nkz}^2 \right)^2. \quad (4.23)$$

Para encontrar la constante de Lipschitz de g es suficiente con encontrar una cota superior al gradiente

$$\frac{\partial g}{\partial t_{nkz}} = \frac{1}{\sum_{n'} (\sum_{k'z'} t_{n'k'z'}^2)^2} 2,2.t_{nkz} \sum_{k'z'} t_{nk'z'}^2. \quad (4.24)$$

Introduciendo la notación $p_n = \sum_{kz} t_{nkz}^2$, y notando que $\sum_n p_n = 1$, se encuentra

$$\begin{aligned} |\nabla g|^2 &= \sum_{nkz} \left(\frac{\partial g}{\partial t_{nkz}} \right)^2 \\ &= \frac{16 \sum_n p_n^3}{(\sum_n p_n^2)^2} \\ &\leq \frac{16 (\sum_n p_n^2)^{3/2}}{(\sum_n p_n^2)^2} \\ &= \frac{16}{(\sum_n p_n^2)^{1/2}} \\ &\leq 16 \sqrt{d_R}, \end{aligned}$$

Por lo tanto la constante de Lipschitz de g llega hasta $4\sqrt[4]{d_R}$. Para obtener la constante de Lipschitz de f se nota que $g(\psi) \geq f(\psi)$ la igualdad se da si $\{|\hat{n}\rangle\}$ es una base propia de $\text{Tr}_B |\psi\rangle \langle\psi|$. Ahora para dos vectores cuales quiera, sin perdida de generalidad asumir $f(\psi_1) \leq f(\psi_2)$, y tomar $\{|\hat{n}\rangle\}$ como la base propia de $\text{Tr}_B |\psi\rangle \langle\psi|$. Entonces,

$$f(\psi_1) - f(\psi_2) \leq g(\psi_1) - g(\psi_2) \leq 4\sqrt[4]{d_R} \|\psi_1 - \psi_2\|_2, \quad (4.25)$$

entonces la constante de Lipschitz para f está acotada por $4\sqrt[4]{d_R}$. Aplicando el lema de Levy a $f(\psi)$ y observando que $\Pr\{x > a\} \leq b$ y $x \geq y$ implica $\Pr\{y > a\} \leq b$ sustituyendo la cota en $\langle f(\psi) \rangle_\psi$ obtenida arriba entonces

$$\ln \left(\text{Tr} \left(\mathbb{E} [|\psi\rangle \langle\psi|^2] \right) \right) \geq \ln \left(\text{Tr} (\mathbb{E} [|\psi\rangle \langle\psi|]^2) \right) \quad (4.26)$$

esto da

$$\Pr_{\psi} \left\{ \ln(\text{Tr}(\$[\psi] \langle \psi |]^2)) > \ln \frac{2e^{\epsilon}}{d_R} \right\} \leq 2 \exp \left(-\frac{\epsilon^2 \sqrt{d_R}}{72\pi^3} \right) \quad (4.27)$$

Tomando la desigualdad que está dentro de los corchetes multiplicándola por menos y tomando su exponencial se llega

$$\Pr_{\psi} \left\{ d^{eff}(\omega) < \frac{d_R}{2e^{\epsilon}} \right\} \leq 2 \exp \left(-\frac{\epsilon^2 \sqrt{d_R}}{72\pi^3} \right), \quad (4.28)$$

y haciendo que $\epsilon = \ln 2$ se llega al resultado esperado.

La primera parte del teorema 4.2.2 nos habla de cómo el promedio de la dimensión efectiva es más grande que la dimensión del subespacio de Hilbert esto significa que si se tiene estados de un subespacio bastante grande se puede asegurar un $d^{eff}(\omega)$ grade esto implica un $\langle D(\rho_S(t), \omega_S) \rangle_t$ pequeño. La segunda parte solo confirma de manera más estricta el hecho de encontrar un $d^{eff}(\omega)$ pequeño, mostrando que la probabilidad de encontrar una dimensión efectiva menor a $\frac{d_R}{4}$ es exponencialmente pequeña.

Usando en teorema 4.2.2 se puede ver qué ocurre con un estado escogido de forma aleatoria del espacio total \mathcal{H} , un estado genérico. El análisis se sigue simplemente poniendo $d_R = d$ gracias a esto se comprende que $d^{eff}(\omega) \sim d$ ya que hay una probabilidad exponencialmente baja para que se dé el caso en que $d^{eff} < \frac{d}{4}$, como $d = d_S d_B$ la cota para las fluctuaciones queda $\sqrt{\frac{d_S}{d_B}}$. Un sistema de muchas partículas la dimensión del espacio de Hilbert crece de manera exponencial [12], si el subsistema es una fracción constante del número de partículas del baño esta proporción caerá de manera exponencial con el número total de de partículas luego los subsistemas se equilibrarán.

Se podría suponer que con los teoremas presentados hasta el momento el problema de termalización se ha resulto mayoritariamente pero qué ocurre con los sistemas que están lejos del equilibrio. Con lo dicho arriba sobre

los estados genéricos se pensaría que cualquier sistema debe llegar al equilibrio pero esto no es cierto debido a que los estados lejos del equilibrio no son genéricos; los estados lejos del equilibrio no son típicos, por el capítulo anterior se sabe que con cotas exponenciales la mayoría de los estados en el espacio de Hilbert son tales que un subsistema pequeño está en un estado canónico.

Para poder sacar algo de esta pregunta se planteará la situación común, hay un baño que consiste de un número muy grande de partículas de las cuales se conoce unos parámetros macroscópicos, dentro de este se pone un subsistema con un estado inicial arbitrario pero descorrelacionado con el ambiente. Ahora la pregunta es ¿el subsistema se equilibra?, se verá que para cualquier estado inicial del subsistema y para casi todos los estados iniciales del baño el subsistema se equilibra. Esto incluye cuando el subsistema está lejos del equilibrio.

El estado inicial del sistema está dado por $|\Psi\rangle_{SB} = |\psi\rangle_S |\psi\rangle_B$. El estado del subsistema es uno arbitrario $|\psi\rangle_S$ en el espacio de Hilbert. Dado unos parámetros macroscópicos el baño debe cumplir con estos, luego el estado del baño $|\phi\rangle_B \in \mathcal{H}_B^R \subseteq \mathcal{H}_B$. Esta restricción mantiene la generalidad de seguir en cualquier espacio de Hilbert restringido. Pero la restricción es solo inicial al evolucionar el baño en el tiempo este puede moverse fuera de \mathcal{H}_B^R . Usando el teorema 4.2.2 para $\mathcal{H}_R = |\psi\rangle_S \otimes \mathcal{H}_B^R$ entonces $d_R = d_B^R$ esto da como resultado que para casi todos los estados iniciales del baño y cualquier estado del subsistema este se equilibrará para estas condiciones, mientras que $d_B^R \gg d_S^2$. El mecanismo en que el subsistema se equilibra puede ser bastante complicado ya que el baño pasa por muchos estados diferentes y no llega al equilibrio. Aunque el baño no llegue al equilibrio y se salga del subespacio \mathcal{H}_B^R el subsistema puede equilibrarse de todas formas.

En principio puede que la evolución del subsistema sea sensible a la forma precisa del baño. Para ver que el baño no se equilibra de manera genérica se verá que $d^{eff}(\omega_S)$ es mucho mayor que $d^{eff}(\rho_B(t))$ lo cual muestra que el baño sigue evolucionando y no se equilibra en ningún estado. Como los dos sistemas están en un estado puro $\text{rank}(\rho_B(t)) = \text{rank}(\rho_S(t)) \geq d_S$ como la dimensión efectiva de un estado es siempre menor a su rango se obtiene

$$d^{eff}(\rho_B(t)) \leq d_S \quad (4.29)$$

pero

$$d^{eff}(\omega_B) \geq \frac{d_{eff}(\omega)}{d_S} \quad (4.30)$$

Pero para un estado genérico la segunda parte del teorema 4.2.2 dice $d^{eff}(\omega) > \frac{d_R}{4}$ se tiene

$$d^{eff}(\omega_B) \geq \frac{d^R}{4d_S} = \frac{d_B^R}{4d_S} \gg d_S \geq d^{eff}(\rho_B(t)). \quad (4.31)$$

Lo trabajado anteriormente muestra cómo subsistemas de dimensión pequeña en comparación con el ambiente se equilibrarán. Ahora se verá cual sería la dependencia del estado de equilibrio del subsistema. Hasta ahora el estado inicial podría hacer que el equilibrio sea un estado diferente dependiendo del inicial. Sea el estado de equilibrio del subsistema ω_S^ψ . Como es conocido se desearía que el estado de equilibrio dependa solo de los parámetros macroscópicos y no del estado inicial microscópico. El teorema siguiente prueba que para casi todos los estados en un subsistema restringido llevan al mismo estado de equilibrio.

Teorema 4.2.3. *1. Casi todos los estados iniciales de un subespacio restringido bastante grande llevan al mismo estado de equilibrio de un subsistema pequeño. En particular, con $\langle \cdot \rangle_\psi$ siendo el promedio sobre estados puros aleatoriamente uniformes $|\psi(0)\rangle \in \mathcal{H}_R \subset \mathcal{H}_S \otimes \mathcal{H}_B$ y $\Omega_S = \langle \omega_S^\psi \rangle_\psi$:*

$$\langle D(\omega_S^\psi, \Omega_S) \rangle_\psi \leq \sqrt{\frac{d_S \delta}{4d_R}} \leq \sqrt{\frac{d_S}{4d_R}}. \quad (4.32)$$

La primera desigualdad es más estricta pero más complicada

$$\delta = \sum_k \langle E_k | \frac{\Pi_R}{d_R} | E_k \rangle \text{Tr}_S(\text{Tr}_B(|E_k\rangle \langle E_k|))^2 \leq 1, \quad (4.33)$$

donde Π_k es el proyector sobre \mathcal{H}_R .

2. Para un estado aleatorio $|\psi\rangle \in \mathcal{H}_R \subset \mathcal{H}$, la probabilidad que $D(\omega_S^\psi, \Omega_S) > \frac{1}{2} \sqrt{\frac{d_S \delta}{d_R}} + \epsilon$ caiga exponencialmente con $\epsilon^2 d_R$:

$$\Pr_\psi \left\{ D(\omega_S^\psi, \Omega_S) > \frac{1}{2} \sqrt{\frac{d_S \delta}{d_R}} + \epsilon \right\} \leq 2 \exp(-C' \epsilon^2 d_R), \quad (4.34)$$

con $C' = \frac{2}{9\pi^3}$. Si se pone $\epsilon = d_R^{-1/3}$ da una distancia promedio pequeña con alta probabilidad cuando $d_R \gg d_S$.

Demostración

1. Se usa la relación ya establecida entre la distancia de traza y la distancia de Hilbert-Schmidt

$$\begin{aligned} \langle D(\omega_S, \langle \omega_S \rangle_\psi) \rangle_\psi &\leq \left\langle \frac{1}{2} \sqrt{d_S \text{Tr}(\omega_S, \langle \omega_S \rangle_\psi)^2} \right\rangle_\psi \\ &\leq \frac{1}{2} \sqrt{d_S \left\langle \text{Tr}(\omega_S - \langle \omega_S \rangle_\psi)^2 \right\rangle_\psi}. \end{aligned}$$

Ahora se miran las cotas del término que es promediado

$$\begin{aligned} \left\langle \text{Tr} \left(\omega_S, \langle \omega_S \rangle_\psi \right) \right\rangle_\psi^2 &= \langle \text{Tr}(\omega_S^2)_\psi \rangle - \text{Tr}_S(\langle \omega_S^2 \rangle_\psi) \\ &= \text{Tr}((\langle \omega_S \otimes \omega_S \rangle_\psi - \langle \omega_S \rangle_\psi \otimes \langle \omega_S \rangle_\psi) F) \\ &= \text{Tr}_{SS} \left(\text{Tr}_{BB} \left(\$ \otimes \$ [\langle |\psi\rangle \langle \psi| \otimes |\psi\rangle \langle \psi| \rangle_\psi - \frac{\Pi_R}{d_R} \otimes \frac{\Pi_R}{d_R}] \right) F \right) \\ &= \text{Tr}_{SS} \left(\text{Tr}_{BB} \left(\$ \otimes \$ \left[\frac{\Pi_{RR}(\mathbb{1} + F)}{d_R(d_R + 1)} - \frac{\Pi_{RR}}{d_R^2} \right] \right) F \right) \\ &= \sum_{kl} \text{Tr}_{SS} \left(\text{Tr}_{BB} \left(|kl\rangle \langle kl| \frac{\Pi_{RR}}{d_R^2} |lk\rangle \langle kl| \right) F \right) \\ &= \sum_{kl} \frac{\langle kl| \Pi_{RR} |lk\rangle}{d_R^2} \text{Tr}_S(\text{Tr}_B(|k\rangle \langle k|) \text{Tr}_B(|l\rangle \langle l|)) \\ &\leq \sum_{kl} \frac{\langle k| \Pi_R |l\rangle \langle l| \Pi_R |k\rangle}{d_R^2} \text{Tr}_S \left(\frac{(\text{Tr}_B(|k\rangle \langle k|))^2 + (\text{Tr}_B(|l\rangle \langle l|))^2}{2} \right) \\ &= \frac{1}{d_R} \sum_k \langle k| \frac{\Pi_R}{d_R} |k\rangle \text{Tr}_S \left((\text{Tr}_B |k\rangle \langle k|)^2 \right) \\ &\leq \frac{1}{d_R} \sum_k \langle k| \frac{\Pi_R}{d_R} |k\rangle = \frac{1}{d_R} \end{aligned}$$

en la segunda desigualdad se usó el hecho que $\text{Tr}_S(\text{Tr}_B |k\rangle \langle k| - \text{Tr}_B |l\rangle \langle l|)^2 \geq 0$ es la traza de un operador positivo. Ahora insertando la tercera línea contando desde el final y con la relación entre

la distancia de traza y la distancia de Hilbert-Schmidt se demuestra el resultado.

2. Para demostrar la segunda parte que todos los estados $|\psi\rangle$ llevan al mismo estado de equilibrio, se usa el ya conocido lema de Levy. Se aplica este lema directamente a la función $f(\psi) \equiv D(\omega_S^\psi, \Omega_S)$ en la hipersfera de dimensión $2d_R - 1$ de estados cuánticos. Para hallar la constante de Lipschitz se sigue de la misma manera que en el capítulo anterior

$$\begin{aligned}
 |D(\omega_S^{\psi_1}, \Omega_S) - D(\omega_S^{\psi_2}, \Omega_S)| &\leq D(\omega_S^{\psi_1}, \omega_S^{\psi_2}) \\
 &\leq D(\omega^{\psi_1}, \omega^{\psi_2}) \\
 &= \sqrt{1 - |\langle \psi_1 | \psi_2 \rangle|^2} \\
 &\leq \|\psi_1 - \psi_2\|_2.
 \end{aligned}$$

entonces la constante de Lipschitz satisface $\eta \leq 1$. Sustituyendo en el lema de Levy junto con el valor promedio encontrado en la primera parte de esta demostración se llega al resultado deseado.

Para un estado inicial que sea el producto del estado del sistema y el del ambiente $|\psi\rangle_{SB} = |\psi\rangle_S |\phi\rangle_B$ en el espacio $\mathcal{H}_R = |\psi\rangle \otimes \mathcal{H}_R^B$ se mostró que para estados genéricos el ambiente causa que el subsistema se equilibre aunque el estado de equilibrio ω_S^ψ del subsistema podría depender del estado inicial del baño $|\phi\rangle_B$. Esto no es así y casi todos los estados del baño en \mathcal{H}_B^R llevan al mismo estado de equilibrio del subsistema simplemente usando el resultado anterior a $\mathcal{H}_B = |\psi\rangle \otimes \mathcal{H}_B^R$ y luego $d_R = d_R^B$, si $d_R^B \gg d_S$ para casi todos los estados iniciales del baño llevarán al subsistema al mismo estado de equilibrio ω_S . como se usó la cota menos estricta este resultado no depende de la forma explícita de los estados de energías propios.

El nuevo enfoque dado en el capítulo anterior fue explorado aún más siguiendo a [20] en este capítulo viendo qué ocurre con los sistemas al evolucionar. El enfoque dado por [1] es fructífero en resultados y se puede seguir profundizando esto genera un interés para que se siga explorando esta perspectiva. Este capítulo mostró cómo en general los estados de un espacio de Hilbert bastante grande que tiene muchos estados propios de energía llegan

a estar bastante cerca del estado promedio (promediado temporalmente) y se dice que se equilibra este. Luego se muestra que en general el estado de equilibrio es ese estado promedio. Linden et al. mostraron que hay equilibrio para estados genéricos y también para estados fuera del equilibrio además mostraron la independencia del estado del baño. Aunque ellos no mostraron la independencia del estado del subsistema no debería influenciar el equilibrio. Siguiendo lo dicho por ellos este problema muestra más dificultades que los otros dado a que es necesario una especificación de la energía porque ahora jugaría un papel más importante, en los teoremas anteriores no fue necesario hablar de la energía excepto en el teorema 4.2.3 donde hay una desigualdad dada por los estados propios de energía.

Bibliografía

- [1] Popescu, S., Short, A., & Winter, A. (2006). Entanglement and the foundations of statistical mechanics. *Nature Physics*, 2(11), 754-758. <http://dx.doi.org/10.1038/nphys444>.
- [2] Goldstein, S., Lebowitz, J., Tumulka, R., & Zanghì, N. (2006). Canonical Typicality. *Physical Review Letters*, 96(5). <http://dx.doi.org/10.1103/physrevlett.96.050403>
- [3] Landau, L., Lifshits, E., Pitaevskii, L., Sykes, J., & Kearsley, M. (1980). *Statistical physics*. Amsterdam: Elsevier Butterworth Heinemann.
- [4] Kardar, M. (2007). *Statistical physics of particles*. Cambridge: Cambridge University Press.
- [5] Pathria, R., & Beale, P. (2012). *Statistical mechanics*. Beijing: World Publishing Corporation.
- [6] Huang, K. (2005). *Statistical mechanics*. New York [u.a.]: Wiley.
- [7] Schlosshauer, M. (2010). *Decoherence and the quantum-to-classical transition*. Berlin: Springer.
- [8] Callen, H. (1975). *Thermodynamics*. New York [etc.]: Wiley.
- [9] Reichl, L. (2016). *A Modern Course in Statistical Physics*. Weinheim: Wiley.
- [10] Sakurai, J. (2013). *Modern Quantum Mechanics*. [S.l.]: Pearson Education India.

- [11] Khinchin, A., & Gamow, G. Mathematical foundations of statistical mechanics.
- [12] Toda, M., Kubo, R., & Saito, N. (1998). Statistical physics I. Berlin: Springer.
- [13] Wilde, M. Quantum information theory.
- [14] Nielsen, M., & Chuang, I. (2015). Quantum computation and quantum information. Cambridge: Cambridge University Press
- [15] Lang, S. (1972). Linear algebra. Reading, Mass.: Addison-Wesley.
- [16] Vedral, V. (2013). Introduction to quantum information science. Oxford: Oxford University Press.
- [17] Jaynes, E. (1965). Gibbs vs Boltzmann Entropies. American Journal Of Physics, 33(5), 391-398. <http://dx.doi.org/10.1119/1.1971557>
- [18] Vitali D. Milman., & Gideon Schechtman. (1986). Asymptotic Theory of Finite Dimensional Normed Spaces. Springer Verlag.
- [19] Horodecki, M., Oppenheim, J., & Winter, A. (2005). Partial quantum information. Nature, 436(7051), 673-676. <http://dx.doi.org/10.1038/nature03909>.
- [20] Linden, N., Popescu, S., Short, A., & Winter, A. (2009). Quantum mechanical evolution towards thermal equilibrium. Physical Review E, 79(6). <http://dx.doi.org/10.1103/physreve.79.061103>.
- [21] Hayden, P., van Dam, W.(2002). Renyi-entropic bounds on quantum communication”, arXiv:quant-ph/0204093.
- [22] Muller, I. (2007). A history of thermodynamics. Berlin: Springer-Verlag GmbH.
- [23] Ehrenfest, P., Ehrenfest, T., & Moravcsik, M. The conceptual foundations of the statistical approach in mechanics.
- [24] Fermi, E., Pasta, J., Ulam, S. (1955). Collected Papers of Enrico Fermi, Vol II, University of Chicago press, Chicago 1965.

- [25] Birkhoff, G. (1931). Proof of the Ergodic Theorem. Proceedings Of The National Academy Of Sciences, 17(12), 656-660. <http://dx.doi.org/10.1073/pnas.17.2.656>
- [26] Susskind, L., & Friedman, A. (2015). Quantum mechanics. UK: Penguin Books.
- [27] Zettili, N. (2009). Quantum mechanics : concepts and applications.
- [28] Jaynes, E. (1957). Information Theory and Statistical Mechanics. Physical Review, 106(4), 620-630. <http://dx.doi.org/10.1103/physrev.106.620>
- [29] Jaynes, E. (1957). Information Theory and Statistical Mechanics. II. Physical Review, 108(2), 171-190. <http://dx.doi.org/10.1103/physrev.108.171>
- [30] Shannon, C. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27(3), 379-423. <http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [31] Penrose, O. (2014). Foundations of Statistical Mechanics. Newburyport: Dover Publications.