



UNIVERSIDAD DE ANTIOQUIA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
INSTITUTO DE FÍSICA

¿(A)SIMETRÍA TEMPORAL EN TEORÍA CUÁNTICA?

TESIS PARA OPTAR AL GRADO DE
FÍSICO

FEDERICO ÚSUGA AGUDELO

ASESOR:
BORIS ANGHELO RODRÍGUEZ REY

MEDELLÍN
2021

Resumen

La esencia de este trabajo es examinar los fundamentos de la física a través del concepto de *simetría*. ¿Qué define una simetría en física? Esa es la pregunta inicial. El Cap. 1 está dedicado a desarrollar esta pregunta en el contexto de la mecánica clásica, y la respuesta se logra a través del *teorema de Noether*: las simetrías se representan en mecánica clásica por medio de *transformaciones canónicas*. Se establece pues la profunda relación entre el concepto de simetría y las leyes de conservación.

El Cap. 2 está dedicado a desarrollar esta misma pregunta, pero esta vez en el contexto de la teoría cuántica, y además, bajo el influjo de las ideas de Asher Peres. «Los fenómenos cuánticos no suceden en un espacio de Hilbert, suceden en un laboratorio», dice en el prefacio de su libro. Tal como se propone allí, en el Cap. 2 se desarrolla la pregunta en teoría cuántica con un enfoque pragmático y estrictamente *instrumentalista*. Con la ayuda de experimentos conceptuales se intenta adivinar la respuesta, y al fin, a través del *teorema de Wigner*, se logra una en que se acentúa el valor de la observación: las simetrías se representan en teoría cuántica por medio de *transformaciones unitarias* y *antiunitarias*, i.e., aquellas que preservan las *probabilidades* de transición. Sin embargo, en particular, ¿por qué la evolución temporal tiene que ser unitaria? Esta será una pregunta recurrente.

En el Cap. 3 se relaja la severidad de aquel enfoque instrumentalista y se procura plantear la primera pregunta en el contexto de los sistemas abiertos y en relación con los conceptos de *información* y *entropía*. Como sabemos, el concepto de entropía surgió originalmente en la termodinámica clásica y luego fue introducido en la teoría de la información por Shannon. Sin embargo, debido a su estrecha relación, es interesante que la *asimetría* entre pasado y futuro, intrínseca en el procesamiento de información, es equivalente a la flecha de tiempo que surge en los fenómenos termodinámicos irreversibles [8]. En este contexto, se propone un teorema que sugiere que un sistema es informacionalmente aislado si y solo si da lo mismo si el tiempo fluye en un sentido o en otro: un “mundo sin fricción” en que la entropía es constante, la de los sistemas y la del universo. Pero faltan ingredientes, la *evolución temporal generalizada* para sistemas abiertos y el análisis del proceso de *medición*. ¿Porqué se asume unitaria la evolución temporal en teoría cuántica? ¿Qué es una medición? ¿Es la medición un proceso *irreversible*? Estas son preguntas esenciales. La parte final del Cap. 3 se dedica justamente a sazonar aquel teorema y a examinar estas preguntas a través de la reinterpretación del *teorema de Kraus* y del *teorema de Neumark*. Así que la discusión final es sobre asimetría temporal.

Para Tatiana, desde luego.

Agradecimientos

Siempre consideré esta la parte más importante del trabajo de grado. De hecho fue la primera que comencé a escribir, cuando conmovido, mientras las alistaba, me iba acordando —una por una— de aquellas personas importantes de mi vida, que desde luego influyen en cada cosa que realizo, por ejemplo en esta tesis. Por supuesto la lista la encabezaba mi familia, y la lista de mi familia, mujeres: mi mamá, mis abuelas, mis tías, que con su profundidad y dulzura siempre supieron demostrarme su amor. Me acordé también de mis profesores. Recordé su generosidad y paciencia. Pensé en Boris, Óscar, Chucho, Edgar, Arnache, El cubano, El español, John Fredy, Mazo, Nelson, Johan, Daniel, Chepe, Jaime, Alejandro... Y aunque me sienta impedido para hablar del profe Boris, tengo que decir que al pensar en él no pude evitar recordar un fragmento de una autobiografía del gran Borges: «Una tradición repite que en los exámenes no formuló jamás una pregunta y que invitaba a los alumnos a elegir y considerar un aspecto cualquiera del tema. No exigía fechas, alegando que él mismo las ignoraba. Abominaba de la bibliografía, que aleja de las fuentes al estudiante.» Y no es que esto sea exacto en su caso, pero quien haya sido alumno suyo sabe a lo que me refiero. Luego me acordé de mis amigos, especialmente de Sebas, Shippo, Dani, Hugo, Guti, El Ruso, Chigui, El mono, Santi, Mateo, Nati, Cami, Fuan Je, Deivid, Lucho, Pablo, Ochoa, Sofi, Ramón, Zuleta, Pueblo... Aquí me detuve en dos, en mis grandes amigos Hugo Restrepo y Mateo Carmona, pues son de aquellas personas que comprenden de verdad el valor de los sueños y del ocio, son de esos que al parecer te preguntan “¿qué piensas?”, cuando en realidad te están diciendo: “¿Qué sueñas?” Decía Hugo en uno de sus cuentos: «Tal vez sólo el ocio nos permite el juego impredecible de lo inesperado, perderle el miedo a soñar y a ser ilusos.» ¡Claro que necesité dedicación!, pero hay que ser bastante iluso para creerse el papel de quien discute sobre simetría y asimetría temporal en física. Después de mis familiares, profesores y amigos, vinieron a mi mente más personajes. Cómo olvidar los personajes imaginarios de los libros, ¡los propios autores sobre todo! Desde Voltaire que me inspiró para pensar en asimetría, hasta Estanislao Zuleta, de quien se dice no pasó de primaria, pero me enseñó a leer. Y por supuesto pensé en Tati, a quien dedico con amor este “pastiche”, y a quien no supe alistar porque es todo al mismo tiempo: familiar, amiga y personaje imaginario.

Alguna vez alguien me dijo: «Quien no cree no siente, quien no siente no quiere, quien no quiere no vive y quien no vive no muere; no aprende a morir.» Pues bien: ¡muchísimas gracias a todos por permitirme creer!

Tabla de Contenido

1. SIMETRÍA Y LEYES DE CONSERVACIÓN EN MECÁNICA CLÁSICA	1
1.1. Principio de mínima acción	1
1.2. Conservación del momento lineal y del momento angular	3
1.2.1. ¿La ley de conservación del momento angular se deriva de las leyes de Newton?	6
1.3. Conservación de la energía	7
1.4. Simetría y leyes de conservación	10
1.4.1. La forma canónica de las ecuaciones de Euler-Lagrange (ÉL)	10
1.4.2. Conservación: primeras integrales de las ecuaciones de ÉL	12
1.4.3. Simetría: transformaciones canónicas	12
1.4.4. <i>Teorema de Noether</i>	14
1.4.5. Leyes de conservación	18
2. SIMETRÍA EN TEORÍA CUÁNTICA	23
2.1. Necesitamos un nuevo lenguaje	23
2.1.1. ¿Qué es la teoría cuántica?	28
2.1.2. ¿Qué es una medición?	30
2.1.3. ¿Qué es un estado cuántico?	31
2.1.4. Hacia la evolución unitaria	34
2.2. ¿Qué define una simetría en teoría cuántica?	39
2.2.1. <i>Teorema de Wigner</i>	41
2.2.2. ¿Dinámica cuántica?	43

2.2.3. Simetría y leyes de conservación	45
3. ¿ASIMETRÍA TEMPORAL EN TEORÍA CUÁNTICA?	49
3.1. <i>Teorema de aislamiento</i> ¹	49
3.2. Entropía de un sistema informacionalmente aislado	53
3.2.1. Entropía de una preparación	53
3.2.2. Un mundo sin fricción	55
3.3. Dinámica y mediciones generalizadas: ¿la simetría temporal es fundamental? ²	57
3.3.1. <i>Teorema de Kraus</i> : teorema de representación para dinámica generalizada	58
3.3.2. <i>Teorema de Kraus y entropía</i> : ¿un mundo sin fricción?	60
3.3.3. Medición generalizada: ¿un mundo con fricción?	63
4. CONCLUSIONES	69
Bibliografía	71

¹Gran parte de la discusión está basada en la Sec. 9.2 de [10].

²La discusión de esta sección consiste en proponer reinterpretaciones del *teorema de representación* que se desarrolla en el capítulo 9 y en el apéndice D de [8] y en los capítulos 14-16 de [1], y del *teorema de Neumark* que se desarrolla en la Sec. 9.6 de [8].

Capítulo 1

SIMETRÍA Y LEYES DE CONSERVACIÓN EN MECÁNICA CLÁSICA

La única justificación para nuestros conceptos y sistemas de conceptos es que ellos sirven para representar la complejidad de nuestras experiencias; fuera de esto no tienen legitimidad.

A. Einstein

1.1. Principio de mínima acción

En mecánica clásica existe una forma de unificar las leyes *fundamentales* en un único principio.¹ Sea cual sea el sistema físico², sea un sistema de partículas libres o un sistema de partículas sometidas a un campo gravitacional, sea un sistema de partículas sometidas a fuerzas restauradoras, o uno de partículas cargadas sometidas a un campo electromagnético, incluso un sistema de partículas relativistas, sus leyes de movimiento se obtienen al exigir que una cantidad llamada *acción* sea mínima.

Supóngase que en un tiempo t_1 un sistema de n partículas se encuentra en una posición fija. La evolución del sistema en el tiempo está descrita por una curva

$$x_i(t), y_i(t), z_i(t) \quad (i = 1, \dots, n),$$

en un espacio de configuración $3n$ -dimensional. En un intervalo de tiempo $[t_1, t_2]$, la cuestión es ¿cuál de las infinitas curvas que unen el par de puntos $(t_1, x_i(t_1), y_i(t_1), z_i(t_1))$ y

¹Ver [7], Vol. II, Cap. 19.

²Siempre y cuando todas las fuerzas se puedan obtener de una función potencial

$$U = U(t, x_1, y_1, z_1, \dots, x_n, y_n, z_n).$$

$(t_2, x_i(t_2), y_i(t_2), z_i(t_2))$ en aquel espacio?

El **principio de mínima acción** establece que el movimiento del sistema de partículas en un intervalo de tiempo $[t_1, t_2]$ está descrito por la curva que minimiza la integral

$$S = \int_{t_1}^{t_2} L(t, x_1, y_1, z_1, \dots, x_n, y_n, z_n, \dot{x}_1, \dot{y}_1, \dot{z}_1, \dots, \dot{x}_n, \dot{y}_n, \dot{z}_n) dt, \quad (1.1)$$

llamada acción, donde L es la función Lagrangiana o Lagrangiano. Una condición necesaria para que $S[x_i, y_i, z_i]$ tenga un *extremo* para una cierta curva $x_i(t), y_i(t), z_i(t), 1 \leq i \leq n$, es que esta curva satisfaga las ecuaciones de Euler-Lagrange (1.15) [3], es decir, las ecuaciones de movimiento del sistema físico.

Muy bien, entonces el problema de formular las ecuaciones de movimiento de un sistema se reduce a determinar su Lagrangiano, y ¿cómo lo hacemos? Lamentablemente no existe un método general. Por ejemplo, para una partícula de masa m y carga eléctrica q en un campo electromagnético \mathbf{E} y \mathbf{B} ,

$$L = \frac{1}{2}mv^2 - q\phi - q\mathbf{v} \cdot \mathbf{A},$$

donde ϕ y \mathbf{A} son los potenciales electromagnéticos. Aplicando las ecuaciones de Euler-Lagrange se obtiene pues la fuerza de Lorentz $\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$. En el caso de una partícula relativista, casi que por prueba y error [7], se sabe que

$$L = -m_0c^2\sqrt{1 - \frac{v^2}{c^2}} - q\phi - q\mathbf{v} \cdot \mathbf{A},$$

donde m_0 es la masa de la partícula en reposo y c es la velocidad de la luz en el vacío, función que de nuevo conduce a la fuerza de Lorentz.

Sin embargo, si en un sistema (no relativista) todas las fuerzas pueden ser obtenidas de una función potencial

$$U = U(t, x_1, y_1, z_1, \dots, x_n, y_n, z_n), \quad (1.2)$$

el Lagrangiano toma la siguiente forma:

$$L = T - U, \quad (1.3)$$

donde T es la energía cinética del sistema. En un sistema de n partículas de este tipo, donde la i -ésima partícula tiene masa m_i y coordenadas x_i, y_i, z_i ($i = 1, \dots, n$),

$$T = \frac{1}{2} \sum_{i=1}^n m_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \quad (1.4)$$

y la fuerza que actúa sobre la i -ésima partícula tiene componentes

$$X_i = -\frac{\partial U}{\partial x_i}, \quad Y_i = -\frac{\partial U}{\partial y_i}, \quad Z_i = -\frac{\partial U}{\partial z_i}. \quad (1.5)$$

Según el principio de mínima acción, si (1.1) tiene un mínimo, la curva que describe su movimiento debe satisfacer las ecuaciones de Euler-Lagrange,

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0, \quad \frac{\partial L}{\partial y_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{y}_i} = 0, \quad \frac{\partial L}{\partial z_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{z}_i} = 0, \quad (1.6)$$

o equivalentemente,

$$-\frac{\partial U}{\partial x_i} - \frac{d}{dt}m_i\dot{x}_i = 0, \quad -\frac{\partial U}{\partial y_i} - \frac{d}{dt}m_i\dot{y}_i = 0, \quad -\frac{\partial U}{\partial z_i} - \frac{d}{dt}m_i\dot{z}_i = 0, \quad (1.7)$$

que de acuerdo con (1.5), significa que debe satisfacer las ecuaciones de Newton:

$$X_i = m_i\ddot{x}_i, \quad Y_i = m_i\ddot{y}_i, \quad Z_i = m_i\ddot{z}_i. \quad (1.8)$$

Pero ¿por qué la acción tiene que ser *mínima*? ¿Tiene que ver con que “la Naturaleza no hace por muchos medios lo que puede ser hecho por pocos”? En realidad el principio de mínima acción solo se cumple para intervalos de tiempo suficientemente pequeños; es más, para sistemas con infinitos grados de libertad, como por ejemplo una cuerda vibrante o el campo electromagnético, este principio se reemplaza por el principio de acción *estacionaria*, lo que quiere decir que para arribar a sus ecuaciones de movimiento no es necesario que la acción sea mínima, solo que su primera variación sea igual a cero: *¿puede ser máxima!*³

1.2. Conservación del momento lineal y del momento angular

Es claro por qué en el contexto de la mecánica Newtoniana se toma por fundamental la tercera ley de Newton. Aquí se trata sobre todo de “aguzar los sentidos” para advertir las fuerzas, al menos las que se conocen, puesto que en muchos casos no conocemos las leyes correctas. De ahí que la tercera ley cobre tanta importancia, pues esta expresa una propiedad general de las fuerzas que nos permite conocer el estado del sistema sin conocer todos los detalles.

Según esta ley, las fuerzas entre un par de partículas son iguales y opuestas, más aún, actúan sobre la misma línea, luego, de la segunda ley de Newton:

$$\begin{aligned} \frac{d\mathbf{p}_1}{dt} &= -\frac{d\mathbf{p}_2}{dt} \\ \frac{d(\mathbf{p}_1 + \mathbf{p}_2)}{dt} &= 0, \end{aligned} \quad (1.9)$$

donde \mathbf{p}_1 es el momento lineal de la primera partícula y \mathbf{p}_2 el de la segunda. Es decir, mientras las fuerzas sean *internas*, el momento lineal del sistema se conserva. Este argumento se puede extender fácilmente a un sistema de n partículas, y así, debido a la interacción entre pares de partículas, el aumento del momento lineal de una partícula siempre se compensará por la disminución del momento de la otra.

La ley de conservación del momento lineal (en colisiones) también se puede construir a partir de una versión particular de la tercera ley de Newton y el principio de relatividad Galileana [7]. ¿Qué sucede si dos cuerpos con masas iguales, inicialmente en reposo, son separados por una pequeña explosión sobre un riel de aire? (Figura 1.1). Puesto que los cuerpos

³Ver [3], Sec. 36.2.

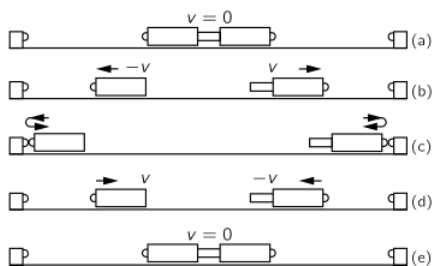


Figura 1.1: Experimento de acción-reacción con masas iguales. (Tomada de [7]).

son iguales, la situación debería ser simétrica, y ciertamente esto se confirma experimentalmente: *cuerpos iguales tendrán la misma rapidez*. Igualmente, sin preferencia entre izquierda y derecha, si los cuerpos se aproximan con la misma rapidez y chocan inelásticamente por medio de algún tipo de fijador, vuelven al reposo. Esto no es más que otra forma de enunciar, en el caso de las colisiones, la aguda observación de Newton expresada en su tercera ley. ¿Y qué hay de la explosión? En la pólvora, debido a la gran atracción entre la infinidad de átomos de oxígeno de las moléculas de nitrato de potasio (KNO_3) y la infinidad de átomos de carbono (C) y sodio (S), se produce una enorme conmoción en forma de energía cinética en la formación de las nuevas moléculas (e.g. CO), lo que finalmente se traduce en calor, incluso en ¡una llama! (luz), si el movimiento de las moléculas es suficientemente violento. Estas moléculas golpean contra las paredes del pequeño cilindro que las contiene y este a su vez golpea los cuerpos que flotan sobre el riel. Lo increíble de la observación de Newton es que nos permite conocer el estado de los cuerpos sin necesidad de entrar en estos detalles, sin conocer las leyes para las fuerzas intermoleculares y las fuerzas entre átomos ni los mecanismos de transformación de una forma de energía en otra, pues aquella propiedad de las fuerzas es *general*.

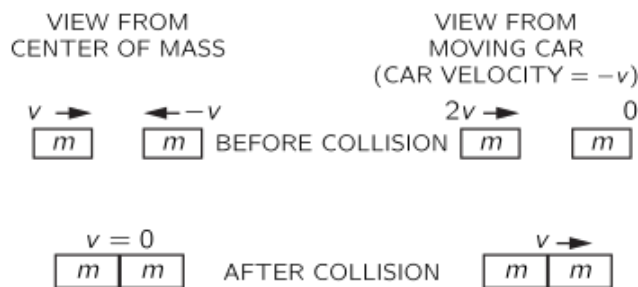


Figura 1.2: Colisión inelástica entre masas iguales vista desde diferentes sistemas de referencia. (Tomada de [7]).

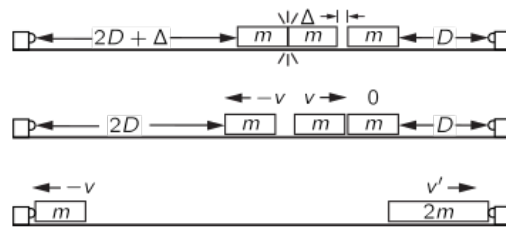


Figura 1.3: Experimento para verificar la situación de la Figura 1.2. (Tomada de [7]).

Analicemos situaciones nuevas a partir de lo que ya sabemos. ¿Qué sucedería si uno de los cuerpos está en reposo y el otro lo choca inelásticamente con cierta velocidad? De acuerdo al principio de relatividad de Galileo, las situaciones de la Figura 1.2 son físicamente equivalentes, respondiendo a la pregunta: los dos objetos pegados se moverían a la mitad de la velocidad inicial. Esto se verifica experimentalmente, como se muestra en la Figura 1.3, donde, en el viaje de ida y vuelta hasta colisionar inelásticamente y alcanzar de nuevo el reposo, el objeto de masa m recorre una distancia de casi $4D$ en el mismo tiempo en que el de masa $2m$ recorre $2D$. Lo interesante es notar que ya sea que se observe desde el sistema

de referencia del centro de masa o desde el sistema de referencia en movimiento, el momento lineal se conserva:

$$mv + m(-v) = 2m(0)$$

$$mv - mv = 0 \text{ (desde el CM),}$$

o

$$m(2v) + m(0) = 2m(v)$$

$$mv - mv = 0 \text{ (desde el carro).}$$

Para generalizar este resultado podemos continuar del mismo modo. Del resultado del experimento de la Figura 1.3 y aplicando de nuevo el principio de relatividad de Galileo podemos averiguar qué sucede cuando un cuerpo de masa m colisiona inelásticamente con un cuerpo de masa $2m$ en reposo. Y así sucesivamente: uno de masa m con uno de masa $3m$, uno de masa $2m$ con otro de masa $3m$, m versus $4m$, $3m$ versus $4m$... Así pues, sirviéndonos de la *simetría* del sistema de un par de objetos de masas diferentes —*cualesquiera*— que al chocar inelásticamente permanecen en reposo, sabremos qué sucede cuando uno de estos cuerpos está en reposo y el otro lo choca a cierta velocidad, ¡conservándose siempre el momento lineal! [7]

Cuando la interacción es por contacto parece claro que la tercera ley de Newton y la ley de conservación del momento lineal estén a un mismo nivel, pero ¿qué cambia si la acción es a distancia? Al parecer nada, pues según Newton las interacciones a distancia son instantáneas. Sin embargo, si una carga eléctrica se mueve repentinamente sintiendo una cierta fuerza de reacción, la interacción con otra partícula cargada —a una cierta distancia— no será instantánea, y por muy poco tiempo que tome, durante este tiempo el momento lineal no se conservará. Para salvar esta aparente contradicción, basta iluminar una superficie y medir la presión mecánica que se ejerce sobre esta producto del intercambio de momento entre el objeto y el campo electromagnético: el campo también posee momento lineal, momento angular y energía. Si se consideran los momentos mecánicos de las cargas junto con el momento del campo⁴, ¡el momento siempre se conserva!

Suponiendo que en relatividad especial también se conserva el momento lineal en un sistema donde solo actúan fuerzas internas, debe conservarse para todo sistema de referencia inercial, lo cual generaliza el principio de relatividad de Galileo para sistemas de referencia inerciales que se mueven a velocidades comparables con c . En este caso, la equivalencia entre sistemas físicos se da bajo transformaciones de Lorentz, y el momento lineal se conserva siempre y cuando tome la siguiente forma:

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (1.10)$$

En teoría cuántica también se cumple la ley de conservación del momento lineal, siempre que el sistema sea informacionalmente aislado [10], aunque, como veremos, es necesario introducir un nuevo lenguaje (Sec. 2.1).

Aunque en un principio nos dio la impresión de que la ley de conservación del momento lineal era una consecuencia natural de la mecánica Newtoniana, sabemos que también se

⁴Ver [7] Vol. 2, Ch. 27, *Field momentum*.

cumple en electromagnetismo, relatividad y mecánica cuántica, es por esto que esta ley es generalmente considerada más fundamental que las leyes de Newton [5].

Algo similar sucede con la ley de conservación del momento angular para un sistema de partículas aislado. Consideremos un sistema de n partículas, donde la i -ésima partícula tiene masa m_i ($i = 1, \dots, n$) y las fuerzas se deben exclusivamente a las interacciones entre las partículas. Sean \mathbf{f}_{jl} y $\boldsymbol{\tau}_{jl}$ la fuerza y el torque que ejerce la partícula l sobre la partícula j , respectivamente. Respecto al origen de un sistema de referencia conveniente, el torque sobre la partícula j es

$$\boldsymbol{\tau}_j = \mathbf{r}_j \times \sum_{l=1}^n \mathbf{f}_{jl}.$$

Lo que nos interesa comprobar es que efectivamente los torques internos se anulan por pares, de modo que el torque interno total se anule, o lo que es lo mismo, el momento angular del sistema se conserve:

$$\boldsymbol{\tau}_{jl} + \boldsymbol{\tau}_{lj} = \mathbf{r}_j \times \mathbf{f}_{jl} + \mathbf{r}_l \times \mathbf{f}_{lj},$$

o, según la tercera ley de Newton:

$$\begin{aligned} \boldsymbol{\tau}_{jl} + \boldsymbol{\tau}_{lj} &= (\mathbf{r}_l - \mathbf{r}_j) \times \mathbf{f}_{lj} \\ &= \mathbf{r}_{jl} \times \mathbf{f}_{lj}. \end{aligned} \tag{1.11}$$

Entonces, el momento angular de un sistema aislado de partículas se conserva siempre que \mathbf{r}_{jl} sea paralelo a \mathbf{f}_{lj} : en el caso particular del movimiento debido a *fuerzas centrales*, se concluye pues, que la ley de conservación del momento angular se sigue de las leyes de Newton [5].

1.2.1. ¿La ley de conservación del momento angular se deriva de las leyes de Newton?

¡Pero las fuerzas no tienen que ser necesariamente centrales según las leyes de Newton!

Es claro que existen fuerzas no centrales y que el momento angular, sin embargo, debe conservarse; en la electrodinámica, por ejemplo. En la Figura 1.5 se muestran dos protones que se mueven paralelos al eje x en sentidos opuestos con la misma rapidez. Así como el movimiento de un protón ejerce una fuerza $\mathbf{F}_E + \mathbf{F}_B$ sobre el otro, este ejerce una fuerza $-(\mathbf{F}_E + \mathbf{F}_B)$ sobre aquel. Ignorando el *delay* que se genera mientras una partícula siente el efecto del movimiento de la otra, el momento lineal se conserva siempre, aunque las fuerzas *no* sean centrales en todo momento. ¿Entonces el momento angular solo se conserva en el instante mostrado en la Figura 1.5? Claro que no, como se sabe, el campo electromagnético también posee momento angular, que si se considera junto al de las partículas, explica la conservación del momento angular en cada instante.

Sin embargo, la objeción recae directamente sobre las ideas en las que se funda la propia mecánica de Newton. Imaginemos por un momento que solo contamos con las leyes de Newton para describir la interacción entre un par de partículas aisladas con idéntico espín, e.g. un

par de electrones con espín *up*. Hasta donde se sabe, un electrón, a pesar de tener una masa bien definida, es una partícula con radio cero, una masa extraña que no tiene tamaño ni estructura, justamente la definición de una partícula Newtoniana. Inmediatamente surge un problema: los espines \mathbf{S}_i ($i = 1, 2$) definen una dirección espacial adicional a la definida por el vector \mathbf{r}_{12} , de la misma forma que en la Figura 1.5 las velocidades de los protones definen al eje x , así que la fuerza entre los electrones no tiene que ser necesariamente central. Si por ejemplo esta fuerza fuera

$$\begin{aligned}\mathbf{F}_{12} &= k\mathbf{r}_{12} \times (\mathbf{S}_1 + \mathbf{S}_2) \\ &= -\mathbf{F}_{21},\end{aligned}$$

donde k es una constante, tendría que concluirse que el momento angular del sistema no se conserva (Figura 1.6).

Como aún no se conoce experimentalmente ningún caso en el que el momento angular total de un sistema aislado no se conserve, debe concluirse que la ley de conservación del momento angular es una ley física independiente [5].

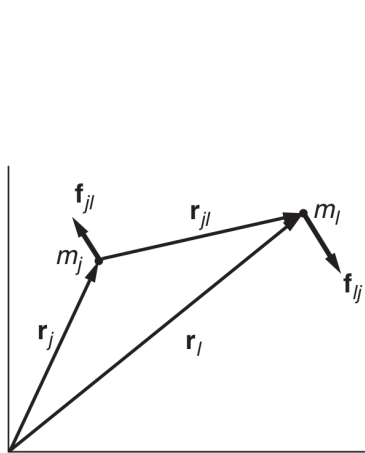


Figura 1.4: Caso de fuerzas no centrales. (Tomada de [5]).

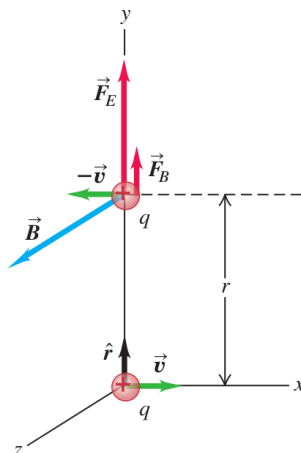


Figura 1.5: Fuerza entre dos protones en movimiento.⁶

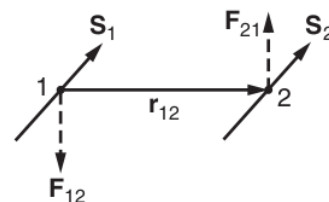


Figura 1.6: Ejemplo de fuerzas no centrales entre dos electrones con espín *up*. (Tomada de [5]).

1.3. Conservación de la energía

Como hemos visto, nada que se derive directamente de las leyes de Newton nos dice algo nuevo. La tercera ley de Newton es equivalente a la ley de conservación del momento lineal, lo mismo sucede con la ley de conservación del momento angular para fuerzas centrales. Este es el caso también de la ley de conservación de la energía mecánica. El teorema del trabajo y la energía para un sistema de partículas se deduce de la segunda ley de Newton, y expresa —para el caso de fuerzas conservativas— la conservación de la energía mecánica:

$$T_a + U_a = T_b + U_b = E_{mec}, \quad (1.12)$$

⁶Tomada de F. Sears, M. Zemansky, H.D. Young, R.A. Freedman, *Física Universitaria con Física Moderna Vol. 2*, Décimo tercera edición, Pearson, México (2013).

donde, en coordenadas del centro de masa, U_i es la energía potencial (gravitacional o elástica) del sistema y T_i es su energía cinética, ambas en la posición \mathbf{R}_i ($i = a, b$). E_{mec} es la energía mecánica del sistema, a este nivel de la discusión, la misma en todo momento, y que en principio puede ser cualquier número real.

Pero como sabemos, los resortes se detienen, los columpios dejan de pendular, las bolas de billar paran..., aunque también los resortes, los pivotes y las bolas de billar se calientan al cesar su movimiento, y aunque para nuestros toscos sentidos la energía simplemente desaparece, basta usar un termómetro suficientemente sensible para encontrarla. Las leyes de Newton incluyen también fuerzas *no* conservativas (e.g. la fricción), y así las cosas, la energía mecánica no se conserva y tampoco se pueden deducir las ecuaciones de movimiento del sistema a partir del principio de mínima acción: ni nuestros sentidos ni las leyes de Newton son suficientes para *ver* a un nivel fundamental. A nivel microscópico no existen fuerzas no conservativas [7], fuerzas como la de fricción son útiles a nivel macroscópico justamente para evitar las complicaciones que surgen de considerar el movimiento aleatorio de un sinnúmero de átomos, movimiento que en definitiva es el que proporciona una medida de la energía térmica. La energía térmica es solo energía cinética [7].

La energía parece pues que se “esconde”. ¿Es esta una peculiaridad de la energía? Pensemos en otro tipo de cantidad física, e.g. el momento lineal. Resulta que esconderlo no es tan simple. Un exceso de momento lineal de los átomos del pivote de un columpio, se traduce en un vaivén, por imperceptible que sea, pues se suman vectores, velocidades en este caso. Mientras que al sumar las energías cinéticas de los átomos estamos sumando los cuadrados de las rapidezces, cantidades sin dirección: incluso si el columpio está quieto, el pivote puede estar caliente. La energía es un concepto físico más sutil, entre otras cosas, porque es un camaleón que puede cambiar de una forma a otra [5].

La energía del campo electromagnético es otra de las formas. Esta se debe al vaivén de cargas eléctricas: energía mecánica transformada en energía electromagnética o viceversa:

$$-\frac{\partial u}{\partial t} = \nabla \cdot \mathbf{S} + \frac{\partial u_{mec}}{\partial t}, \quad (1.13)$$

donde u es la densidad de energía del campo y u_{mec} es la densidad de energía mecánica de las cargas, ambas dentro de cierta región del espacio, y \mathbf{S} es el vector de flujo de energía del campo o vector de Poynting. O sea, el campo pierde energía solo si interactúa con la materia dentro de una cierta región o si su energía escapa de esta región. Por otra parte, para entender mejor el significado del vector de Poynting, la presión que se ejerce al iluminar una superficie, la presión de radiación de la que hemos hablado, se estudia justamente por medio de este vector, siendo proporcional a su promedio temporal $\langle \mathbf{S} \rangle$. Es claro pues que la energía de la luz es también energía electromagnética.

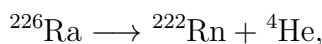
Otra forma de energía es la energía química, que es la que se libera en reacciones químicas, explosiones por ejemplo. Asimismo la energía nuclear, que se libera en reacciones nucleares, y aunque no se conozcan las leyes fundamentales [7], se tienen fórmulas para intentar comprenderla, como la de la energía total, que surge al remplazar la invariancia Galileana por la de Lorentz,

$$E^2 = (pc)^2 + (m_0c^2)^2.$$

De esta expresión es interesante notar que una partícula sin masa en reposo, e.g. un fotón,

con momento p , tiene energía, y que una partícula, ¡por el mero hecho de existir!, también: m_0c^2 .

En este punto la ley de conservación de la energía parece un acto de desesperación: fantasear con movimientos aleatorios de átomos en el vacío, imaginar sopas abstractas con energía que llenan el espacio, partículas sin masa que viajan a la velocidad de la luz... ¿Por qué creer que cuando aparece o desaparece energía mecánica debemos inventar una serie de cantidades y procesos abstractos para explicar la conservación de la energía? Porque *aún* no se conoce experimentalmente ningún proceso de la Naturaleza en el que la energía no se conserve. El experimento del cervecero inglés y físico aficionado, James P. Joule, para medir el equivalente mecánico del calor es uno de los ejemplos más importantes.⁷ Otro ejemplo es la emisión espontánea de rayos α (núcleos de ${}^4\text{He}$), pues cuando se pudieron medir con suficiente precisión masas nucleares se comprobó que el exceso de masa Δm en la reacción nuclear



era igual a $\Delta E/c^2$, donde ΔE era la energía cinética del retroceso del átomo de ${}^{222}\text{Rn}$ más la energía cinética del núcleo de ${}^4\text{He}$ [5]. Esto solo por mencionar algunos ejemplos.

A pesar de todo, aún no se ha discutido qué es la energía, y la razón es muy simple: no se sabe. Incluso la ley de conservación de la energía no se entiende profundamente. Por más que nos esforcemos por verificar experimentalmente este *principio matemático*, ¿cómo se explican las transformaciones? Esta ley no describe ningún tipo de mecanismo. Además, la energía que a los humanos nos resulta *útil* no se conserva tan fácilmente [7], complicándolo todo aún más. Pensemos solo en la conservación de la energía al escribir en *Overleaf* este trabajo de grado. ¿Cómo es que al teclear estas palabras la energía mecánica de mis dedos se transforma en energía eléctrica y energía térmica en un circuito electrónico, y finalmente la luz de la pantalla del computador me las muestra? Obviamente está dentro de las posibilidades, la siguiente pregunta sería, ¿qué tan probable es? Construir el computador es imposible sin tomar energía de algún otro lugar. Además de energía, es necesario pues, hablar de *entropía*, i.e., de las *probabilidades* de estas *posibilidades*.⁸

De la relación entre energía y entropía se ocupan la termodinámica y la mecánica estadística de Boltzmann-Gibbs. Sin embargo, en las siguientes secciones nos ocuparemos de la relación entre la energía y otro concepto fundamental de la física: la simetría. Esta relación es la que nos ayudará justamente a entender mucho mejor las leyes de conservación. Y a pesar de que no se discutirá profundamente, el triángulo lo completa la relación entre entropía y simetría, uno de los temas que más genera debate actualmente en el mundo de la física.⁸

⁷Ver Albert Einstein, Leopold Infeld, *La física, aventura del pensamiento*, Octava edición, Losada S.A., Buenos Aires (1939), p. 47-50.

⁸Ver Constantino Tsallis, *Introduction to Nonextensive Statistical Mechanics, Approaching a Complex World*, Springer, New York (2009), 1.3 *Symmetry, Energy and Entropy*.

1.4. Simetría y leyes de conservación

Aunque el principio de mínima acción unifica las leyes fundamentales de la mecánica clásica en un único principio, como se mostró, para deducir las ecuaciones de movimiento es necesario conocer el Lagrangiano del sistema, y no existe un método general para calcularlo. Como también se mostró, la ley de conservación del momento angular no se deriva directamente de las leyes de Newton, y aunque las leyes de conservación de la energía mecánica y del momento lineal sí, esto no nos dice nada nuevo. Todo esto nos deja con una mecánica incompleta y con un principio bastante elegante, pero solo para sistemas conservativos.

El propósito de esta sección es ir más allá de las leyes de Newton y, por medio de métodos analíticos generales, expresados en los formalismos Lagrangiano y Hamiltoniano, lograr más flexibilidad y una mayor profundidad en la discusión sobre las leyes fundamentales de la física. En especial, comprender mejor las leyes de conservación de la mecánica clásica a partir de su relación con las simetrías de los sistemas físicos.

Cuando hablamos de métodos analíticos generales nos referimos a las ideas de una de las grandes obras de Lagrange, el *cálculo de variaciones*, considerada por él mismo como su obra maestra, y que le permitió unificar la mecánica en su obra *Mécanique analytique*, y como dijera Hamilton, hizo de ella “una especie de poema científico”.⁹

1.4.1. La forma canónica de las ecuaciones de Euler-Lagrange (ÉL)

Las ecuaciones de ÉL correspondientes a la acción

$$S[q_1, \dots, q_n] = \int_{t_1}^{t_2} L(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) dt, \quad (1.14)$$

forman un sistema de n ecuaciones diferenciales de segundo orden

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad (i = 1, \dots, n). \quad (1.15)$$

Nuestro propósito es remplazar el conjunto de variables $t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, L$ por otro conjunto de variables que nos permita obtener una forma más conveniente y simétrica de las ecuaciones de ÉL. El truco está en usar las ecuaciones

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (i = 1, \dots, n) \quad (1.16)$$

para escribir $\dot{q}_1, \dots, \dot{q}_n$ como funciones de las variables

$$t, q_1, \dots, q_n, p_1, \dots, p_n, \quad (1.17)$$

⁹E. T. Bell, *Los grandes matemáticos (desde Zenón a Poincaré)*, Editorial Losada, S. A., Buenos Aires (1948), Cap. X: Una inmensa pirámide, Lagrange.

y escribir el Lagrangiano $L(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$ en términos de la función Hamiltoniana H :

$$H(t, q_1, \dots, q_n, p_1, \dots, p_n) = -L + \sum_{i=1}^n \dot{q}_i p_i, \quad (1.18)$$

donde las \dot{q}_i se consideran funciones de las variables (1.17).¹⁰

Las nuevas variables

$$t, q_1, \dots, q_n, p_1, \dots, p_n, H$$

son llamadas las *variables canónicas* correspondientes a la acción $S[q_1, \dots, q_n]$.

El siguiente paso es reescribir las ecuaciones de ÉL en términos de estas nuevas variables. El cambio consiste básicamente en expresar las derivadas parciales $\frac{\partial L}{\partial q_i}$ en términos de las derivadas parciales $\frac{\partial H}{\partial q_i}$, sin embargo, para evitar cálculos engorrosos, se usa la expresión para el diferencial de H , dH , y la propiedad de invariancia de este bajo cambios de variables independientes [3]:

$$dH = -\frac{\partial L}{\partial x} dx - \sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \sum_{i=1}^n p_i d\dot{q}_i + \sum_{i=1}^n \dot{q}_i dp_i. \quad (1.19)$$

Imaginen tener que expresar $d\dot{q}_i$ en términos de t, q_i y p_i para obtener las derivadas parciales de H . Una de las características más importantes de las variables canónicas es que justamente las ecuaciones (1.16) permiten cancelar los términos que contienen $d\dot{q}_i$:

$$dH = -\frac{\partial L}{\partial x} dx - \sum_{i=1}^n \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^n \dot{q}_i dp_i.$$

Se tiene entonces que

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}, \quad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i}, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i,$$

o, de (1.15) y (1.16),

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (i = 1, \dots, n). \quad (1.20)$$

Estas $2n$ ecuaciones diferenciales de primer orden reciben el nombre de *sistema canónico de ecuaciones de ÉL* correspondientes a la acción (1.14).

¹⁰Siempre que sea posible. H es justamente la *transformada de Legendre* de L , y esta transformación es posible siempre y cuando las ecuaciones (1.16) puedan ser resueltas para \dot{q}_i , o equivalentemente, que el determinante de la matriz Hessiana de L ,

$$\frac{\partial(p_1, \dots, p_n)}{\partial(\dot{q}_1, \dots, \dot{q}_n)} = \det \|L_{\dot{q}_i \dot{q}_j}\|,$$

sea diferente de cero. Y esto no es solo una condición matemática, pues hay sistemas físicos que la violan, como es el caso del Lagrangiano del sistema del ejercicio 4 del capítulo 8 de [4]

$$L = ax^2 + b\frac{\dot{y}}{x} + c\dot{x}\dot{y} + fy^2\dot{x}\dot{z} + g\dot{y} - k\sqrt{x^2 + y^2},$$

con a, b, c, f, g y k constantes.

1.4.2. Conservación: primeras integrales de las ecuaciones de ÉL

¿Y qué se gana con este cambio de variables? Antes de hablar de la estrecha relación entre simetría y leyes de conservación, es importante poder reconocer cantidades que no cambien en el tiempo a lo largo de cada una de las curvas integrales o soluciones de las ecuaciones de ÉL, y esto ciertamente se facilita en términos de las variables canónicas. Debemos entender primero qué significa *conservación* en este lenguaje.

Que una cierta cantidad se conserve significa que una función

$$\Lambda = \Lambda(t, q_1, \dots, q_n, p_1, \dots, p_n),$$

sea constante a lo largo de cada solución del sistema de ecuaciones (1.20), es decir, a lo largo de cada solución de las ecuaciones de movimiento del sistema físico. Se dice pues que $\Lambda(t, q_1, \dots, q_n, p_1, \dots, p_n)$ es una *primera integral* del sistema de ecuaciones de ÉL (1.20).

Averigüemos entonces bajo qué condiciones se da:

$$\begin{aligned} \frac{d\Lambda}{dt} &= \frac{\partial\Lambda}{\partial t} + \sum_{i=1}^n \frac{\partial\Lambda}{\partial q_i} \frac{\partial q_i}{\partial t} + \frac{\partial\Lambda}{\partial p_i} \frac{\partial p_i}{\partial t} \\ &= \frac{\partial\Lambda}{\partial t} + \sum_{i=1}^n \frac{\partial\Lambda}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial\Lambda}{\partial p_i} \frac{\partial H}{\partial q_i} \\ &= \frac{\partial\Lambda}{\partial t} + [\Lambda, H]_{PB}, \end{aligned} \tag{1.21}$$

donde

$$[\Lambda, H]_{PB} = \sum_{i=1}^n \frac{\partial\Lambda}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial\Lambda}{\partial p_i} \frac{\partial H}{\partial q_i}$$

son los corchetes de Poisson de las funciones Λ y H . Concluimos de (1.21) que si Λ no depende explícitamente del tiempo, una condición necesaria y suficiente para que sea una primera integral del sistema de ecuaciones de ÉL (1.20) es que los corchetes de Poisson $[\Lambda, H]_{PB}$ se anulen idénticamente [3].

1.4.3. Simetría: transformaciones canónicas

¿Qué significa entonces simetría? Desde un punto de vista matemático, significa que la acción (1.14) es *invariante* bajo ciertas transformaciones, o lo que es lo mismo, que dos problemas variacionales sean equivalentes bajo ciertas transformaciones de variables, i.e., que tengan los mismos extremales o soluciones de las ecuaciones de ÉL.

Esta invariancia podemos entenderla de dos formas. Se dice que la acción es invariante bajo la transformación¹¹

$$\begin{aligned} t^* &= \Lambda(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = \Lambda(t, q, \dot{q}) \\ q_i^* &= \Gamma_i(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = \Gamma_i(t, q, \dot{q}) \quad (i = 1, \dots, n), \end{aligned} \tag{1.22}$$

¹¹Aunque matemáticamente $t^* = \Lambda(t, q, \dot{q})$ representa una transformación general, físicamente solo se tratan la identidad, la traslación temporal y la reversión temporal, $\Lambda(t, q, \dot{q}) = t$, $\Lambda(t, q, \dot{q}) = t + \delta t$ y $\Lambda(t, q, \dot{q}) = -t$, respectivamente.

si

$$\int_{t_1^*}^{t_2^*} L(t^*, q^*, \dot{q}^*) dt^* = \int_{t_1}^{t_2} L(t, q, \dot{q}) dt,$$

donde la ecuación vectorial¹² $q^* = q^*(t^*)$ ($t_1^* \leq t^* \leq t_2^*$), es el resultado de remplazar $q(t)$ y $\dot{q}(t)$ ($t_1 \leq t \leq t_2$) en (1.22) y eliminar t .

Asimismo, la equivalencia entre problemas variacionales puede entenderse a partir de la invariancia de la forma canónica de las ecuaciones de ÉL bajo la transformación de las variables t, q_i y p_i en las nuevas variables

$$\begin{aligned} t, \\ Q_i = Q_i(t, q_1, \dots, q_n, p_1, \dots, p_n), \\ P_i = P_i(t, q_1, \dots, q_n, p_1, \dots, p_n), \end{aligned} \quad (1.23)$$

es decir, si bajo este cambio de variables, las ecuaciones canónicas se transforman en las nuevas ecuaciones

$$\dot{Q}_i = \frac{\partial H^*}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial H^*}{\partial Q_i}, \quad (1.24)$$

donde $H^* = H^*(t, Q_1, \dots, Q_n, P_1, \dots, P_n)$. Este tipo de transformaciones reciben el nombre de *transformaciones canónicas*. Ciertamente, decir que es posible obtener (1.24) a partir de (1.20), por medio de (1.23), es lo mismo que decir que

$$S[q_1, \dots, q_n, p_1, \dots, p_n] = \int_{t_1}^{t_2} \left(\sum_{i=1}^n p_i \dot{q}_i - H \right) dt \quad (1.25)$$

es igual a

$$S^*[Q_1, \dots, Q_n, P_1, \dots, P_n] = \int_{t_1}^{t_2} \left(\sum_{i=1}^n P_i \dot{Q}_i - H^* \right) dt, \quad (1.26)$$

y dos problemas variacionales son equivalentes si los integrandos de sus acciones correspondientes difieren entre sí por un diferencial total [3]:

$$\sum_{i=1}^n p_i dq_i - H dt = \sum_{i=1}^n P_i dQ_i - H^* dt + d\Lambda(t, q_1, \dots, q_n, p_1, \dots, p_n). \quad (1.27)$$

La función Λ es llamada *función generatriz*, y es la clave, pues si reescribimos (1.27) en la forma

$$d\Lambda = \sum_{i=1}^n p_i dq_i - \sum_{i=1}^n P_i dQ_i + (H^* - H) dt,$$

obtenemos precisamente la transformación canónica

$$p_i = \frac{\partial \Lambda}{\partial q_i}, \quad P_i = -\frac{\partial \Lambda}{\partial Q_i}, \quad H^* = H + \frac{\partial \Lambda}{\partial t}, \quad (1.28)$$

¹²Por simplicidad se usa la notación vectorial: q y \dot{q} representan los vectores n -dimensionales (q_1, \dots, q_n) y $(\dot{q}_1, \dots, \dot{q}_n)$, respectivamente.

pues esta establece una relación entre las antiguas variables y las nuevas variables que obviamente satisface (1.27).¹³

En suma, por medio de estas transformaciones no es posible obtener ningún conocimiento físico, *distinguir* entre situaciones físicas diferentes, pues, al preservarse la forma de las ecuaciones de movimiento, ¡las soluciones son las mismas! Una simetría es una equivalencia entre situaciones físicas diferentes [8].

1.4.4. Teorema de Noether

Consideremos una familia de transformaciones uniparamétricas

$$\begin{aligned} t^* &= \Lambda(t, q, \dot{q}; \varepsilon), \\ q_i^* &= \Gamma_i(t, q, \dot{q}; \varepsilon) \quad (i = 1, \dots, n), \end{aligned} \quad (1.29)$$

donde ε es el parámetro, las funciones Λ y Γ_i son diferenciables respecto a ε y

$$\begin{aligned} \Lambda(t, q, \dot{q}; \varepsilon = 0) &= t, \\ \Gamma_i(t, q, \dot{q}; \varepsilon = 0) &= q_i \quad (i = 1, \dots, n), \end{aligned} \quad (1.30)$$

representa la transformación identidad.

Ya estamos listos para entender la conexión entre la existencia de ciertas primeras integrales de un sistema de ecuaciones de ÉL y la invariancia bajo ciertas transformaciones de su respectiva acción, esto es, la estrecha relación entre las leyes de conservación de la física y las simetrías:

Toda familia uniparamétrica de transformaciones que deje invariante a la acción $S[q]$ conduce a una primera integral de su sistema de ecuaciones de ÉL [3], i.e., ***toda simetría conduce a una cantidad conservada.***

Matemáticamente:

Teorema 1.1 *Si la acción*

$$S[q] = \int_{t_1}^{t_2} L(t, q, \dot{q}) dt, \quad (1.31)$$

es invariante bajo la familia de transformaciones (1.29) para t_1 y t_2 arbitrarios, entonces

$$\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \gamma_i + \left(L - \sum_{i=1}^n \dot{q}_i p_i \right) \lambda = \text{const} \quad (1.32)$$

¹³Es posible escribir Λ como función de t , q_i y Q_i por medio de (1.23). Del mismo modo, es posible expresar la función generatriz como función de t , q_i y P_i reescribiendo (1.27) [3]. En este caso, la función generatriz toma la forma $\Gamma = \Lambda + \sum_{i=1}^n P_i Q_i$ y la transformación canónica

$$p_i = \frac{\partial \Gamma}{\partial q_i}, \quad P_i = \frac{\partial \Gamma}{\partial P_i} \quad H^* = H + \frac{\partial \Gamma}{\partial t}.$$

a lo largo de cada extremal de $S[q]$, donde

$$\begin{aligned}\lambda(t, q, \dot{q}) &= \left. \frac{\partial \Lambda(t, q, \dot{q}; \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0}, \\ \gamma_i(t, q, \dot{q}) &= \left. \frac{\partial \Gamma_i(t, q, \dot{q}; \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0}.\end{aligned}\tag{1.33}$$

Es interesante notar que el teorema de Noether reúne en un solo resultado las leyes de conservación y el principio de mínima acción. Esto se debe a que la primera variación de la acción es

$$\delta S = \int_{t_1}^{t_2} \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i(t) dt + \left(\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \gamma_i - H \lambda \right) \Bigg|_{t=t_1}^{t=t_2},$$

donde $h_i(t)$ representa la variación de $q_i(t)$, de modo que si suponemos que $S[q]$ tiene un extremo para alguna curva

$$q_i = q_i(t) \quad (i = 1, \dots, n),$$

entonces esta curva es una solución del sistema de ecuaciones de ÉL correspondientes a $S[q]$, así que

$$\left[\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \gamma_i + \left(L - \sum_{i=1}^n \dot{q}_i p_i \right) \lambda \right]_{t_1} = \left[\sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \gamma_i + \left(L - \sum_{i=1}^n \dot{q}_i p_i \right) \lambda \right]_{t_2}.$$

Teorema de Noether en teoría de campos

Vale la pena generalizar el resultado anterior y revisar, desde el punto de vista de la física, las enormes consecuencias que esto supone.

Matemáticamente, aunque el problema variacional que hemos tratado puede decirse general, pues los extremos no son fijos y la acción depende de n funciones q_1, \dots, q_n , estas dependen de una única variable independiente, el tiempo. Para extender el resultado anterior a la *teoría de campos* es necesario tratar un problema variacional aún más general: la variación de la acción en una región Ω variable, i.e., en el caso donde las variables independientes x_0, x_1, \dots, x_n varían, así como las $m + 1$ *funciones de campo* $u_0(x_0, x_1, \dots, x_n), \dots, u_m(x_0, x_1, \dots, x_n)$ y sus derivadas

$$\frac{\partial u_j}{\partial x_i} \quad (i = 0, \dots, n; j = 0, \dots, m).\tag{1.34}$$

Si consideramos que $x_0 = t$, simplificamos la notación por medio de los vectores $x = (t, x_1, \dots, x_n)$, $u = (u_0, \dots, u_m)$ y se entienden ∇u como el tensor de componentes (1.34) y dx como $dt dx_1 \cdots dx_n$, la acción toma la siguiente forma:

$$S[u] = \int_{t_1}^{t_2} dt \int \cdots \int_R L(x, u, \nabla u) dx_1 \cdots dx_n = \int_{\Omega} \mathcal{L}(x, u, \nabla u) dx,\tag{1.35}$$

donde $L(x, u, \nabla u)$ es el Lagrangiano del campo, $\mathcal{L}(x, u, \nabla u)$ es llamada *densidad Lagrangiana* del campo, $\Omega = R \times [t_1, t_2]$ es una región $(n + 1)$ -dimensional y R es una hipersuperficie.

Teorema 1.2¹⁴ Si la acción (1.35) es invariante bajo la familia de transformaciones r -paramétricas¹⁵

$$\begin{aligned} x_i^* &= \Lambda_i(x, u, \nabla u; \varepsilon) \sim x_i + \sum_{k=0}^r \varepsilon_k \lambda_i^{(k)}(x, u, \nabla u) \quad (i = 0, \dots, n), \\ u_j^* &= \Gamma_j(x, u, \nabla u; \varepsilon) \sim u_j + \sum_{k=0}^r \varepsilon_k \gamma_j^{(k)}(x, u, \nabla u) \quad (j = 0, \dots, m), \end{aligned} \quad (1.36)$$

para una región arbitraria Ω , entonces

$$\sum_{i=0}^n \frac{\partial}{\partial x_i} \left(\sum_{j=0}^m \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial u_j}{\partial x_i} \right)} \bar{\gamma}_j^{(k)} + \mathcal{L} \lambda_i^{(k)} \right) = 0 \quad (k = 0, \dots, r) \quad (1.37)$$

sobre cada superficie extremal de $S[u]$, donde

$$\bar{\gamma}_j^{(k)} = \gamma_j^{(k)} - \sum_{i=0}^n \frac{\partial u_j}{\partial x_i} \lambda_i^{(k)}. \quad (1.38)$$

Claramente, cuando $n = 0$ y $r = 0$ se obtiene el Teorema 1.1:

$$\begin{aligned} \frac{d}{dt} \left(\sum_{j=0}^m \frac{\partial L}{\partial \dot{u}_j} \bar{\gamma}_j + L \lambda \right) &= 0 \\ \sum_{j=0}^m \frac{\partial L}{\partial \dot{q}_i} \gamma_j + \left(L - \sum_{j=0}^m \dot{u}_j p_i \right) \lambda &= \text{const} \end{aligned}$$

a lo largo de cada extremal.

Pero detengámonos en este “simple” salto matemático entre el Teorema 1.1 y el Teorema 1.2 que aumenta en n la dimensión de la región en la que varía la acción S , porque en términos físicos tiene profundas consecuencias. Pensemos en el tipo de sistemas físicos representados en cada teorema. Por un lado, el primer teorema se ocupa de sistemas físicos *discretos*, e.g. un sistema de partículas que interactúan instantáneamente a distancia, como la Tierra y el Sol en el espacio vacío, donde lo que sucede *aquí* se debe a las condiciones de *allá* (y viceversa), a 151 millones de km; por otro lado, su generalización responde a sistemas *continuos*, como una cuerda vibrante, la membrana del timbal de Tito Puente, o más interesante aún, el *campo electromagnético*, donde, si sabemos lo que pasa *aquí y ahora*, se sabrá lo que sucederá un *poquito después* en el tiempo y un *poquito más allá* en el espacio. Sin embargo, en el ejemplo del campo electromagnético es necesario ir incluso más allá, pensar en la invariancia bajo transformaciones de Lorentz de las ecuaciones que describen la estructura del campo electromagnético, las *ecuaciones de Maxwell*. La geometría del espacio cobra importancia, y un simple espacio Cartesiano con coordenadas t, x, y, z no es adecuado para expresar su invariancia relativista. En particular, si se considera un espacio-tiempo plano, los eventos se localizan por medio del vector

$$x^i = (ct, x, y, z)$$

¹⁴La demostración detallada de este teorema se encuentra en la Sec. 37 de [3].

¹⁵ $\varepsilon = (\varepsilon_1, \dots, \varepsilon_r)$ y ‘ \sim ’ denota la igualdad para términos de orden 1 en ε : *transformaciones infinitesimales*.

y el tensor métrico g^{ij} nos dice cómo “medir distancias” en este nuevo “espacio”.¹⁶ En conclusión, mientras por un lado la escena de las leyes mecánicas de Newton se reduce solamente a los puntos del espacio tridimensional en los cuales la materia o las cargas están presentes, por el otro, *todo* el espacio-tiempo es la escena de las leyes de Maxwell [2]: *acción a distancia* versus *acción contigua*.

Asimismo, continuando con la comparación del sistema de partículas en movimiento en el espacio tridimensional y el campo electromagnético en el espacio-tiempo, esta mera generalización matemática supone otro abismo conceptual entre las leyes de la mecánica Newtoniana y las leyes de Maxwell que puso en aprietos a los mejores físicos y matemáticos del mundo a finales del siglo XIX, y que Einstein aclaró y enmendó en 1905 en su artículo *Sobre la electrodinámica de los cuerpos en movimiento*.¹⁷ Por un lado Newton, imperante, distinguiendo entre “lo absoluto y lo relativo, lo verdadero y lo aparente, lo matemático y lo común”, con ideas sospechosas sobre el espacio, el tiempo y el movimiento absolutos, inalcanzables por nuestros sentidos, e ideas sobre el espacio, el tiempo y el movimiento relativos, según él, los que como hombres comunes empleamos, “sin que esto represente ningún inconveniente en los asuntos comunes”; y sus leyes, invariantes bajo transformaciones de Galileo, transformaciones que en principio no plantean ningún límite para la velocidad de la luz y bajo las cuales las ecuaciones de Maxwell no son invariantes, reforzando así la idea de un sistema inercial privilegiado —el éter— en el cual, según se creía, estas toman su forma más simple y en el que la velocidad de la luz es justamente c . Por otro, las ecuaciones de Maxwell, invariantes bajo transformaciones de Lorentz, además, ningún experimento que pudiera probar que la velocidad de la luz dependiera del observador inercial, entre los más sofisticados, los de Michelson entre 1881 y 1929 que no pudieron medir el movimiento en el éter. Se planteó pues la disyuntiva: conforme al “hábito de pensamiento”, o existe un sistema privilegiado respecto al cual las leyes de la física toman formas especiales y la velocidad de la luz es c , y otros donde sus formas se complican y la velocidad de la luz puede aumentar o disminuir sin límite; o, como asumió Einstein, aceptando los *hechos* experimentales,

1. *las leyes de la física tienen la misma forma en todos los sistemas inerciales y*
2. *la velocidad de la luz en el vacío es la misma para todos los observadores inerciales.*

Aceptar estos *principios* supuso, como dijera Poincaré, “una mecánica completamente nueva”, con una nueva escena donde se desarrollan todos los eventos físicos, el espacio-tiempo, y con una nueva regla de transformación, la de Lorentz (equivalente a bajas velocidades a la de Galileo, satisfaciendo así el *principio de correspondencia*). Así pues, de aquel abismo conceptual nació la *teoría especial de la relatividad*, en un principio llamada por Planck *teoría de la relatividad*, aun por encima de un nombre que contaba con la anuencia de Einstein: *teoría de los invariantes*, el cual se ajustaba más a lo que postulaba sobre la invariancia de las

¹⁶En un espacio-tiempo plano su representación matricial es

$$(g^{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

¹⁷Ver la traducción de H. Quevedo en <http://webs.ftmc.uam.es/juancarlos.cuevas/Teaching/articulo-original.pdf>.

leyes físicas bajo transformaciones entre observadores inerciales. ¿Las leyes físicas deben ser pues simétricas? Fue la relatividad la que consiguió que esta clase de preguntas se volvieran comunes en física.¹⁸

1.4.5. Leyes de conservación

Sean campos clásicos, partículas relativistas, partículas Newtonianas..., inclusive sistemas cuánticos —como veremos en el Cap. 2—, toda simetría conduce a una cantidad conservada, y eso es lo que nos interesa discutir. Como se dijo, una simetría es una equivalencia entre situaciones físicas diferentes. Veamos en términos prácticos a qué nos referimos y aproximémonos de a poco al ideal expresado en el teorema de Noether.

Supongamos que un físico aficionado, ermitaño y neoludita tiene un grave accidente cerebral y desarrolla un síndrome nunca antes visto que le impide saber quién es, de dónde viene, para dónde va, ubicarse él mismo y ubicar todo lo que le rodea en el espacio y en el tiempo..., y que milagrosamente le permite conservar intactos todos sus conocimientos sobre física sin ningún tipo de relación con eventos históricos, sociales, políticos, culturales, etc., aislados por completo, en el mar de su mente, de todo lo demás que en algún momento supo de sí mismo y del mundo, como un único islote en tiempos del gran diluvio. De vez en cuando tiene pequeños episodios de lucidez en que le interesa saber en qué lugar está y en qué época. ¿Cómo podría aquel desgraciado por medio de experiencias científicas saberlo? En un golpe de gran lucidez recuerda el teorema de Noether, y se pregunta: —¿Quién será *el* tal Noether?”, pero no puede recordarlo. Sin embargo, iluminado, en seguida comprende que ninguno de sus rudimentarios y a la vez prolijos aparatos mecánicos, eléctricos, magnéticos, ópticos..., podrían ayudarle a averiguar en qué lugar de este planeta habitable y en qué época se encuentra, pues su funcionamiento en nada dependería del lugar donde estuviese o del momento de la historia en que se encontrase.

—Mis aparatos eléctricos y magnéticos no son lo suficientemente sensibles para que el campo magnético de este planeta me permita determinar en qué latitud me encuentre. Además, no importa en qué *momento* los encienda, ¡funcionarán *igualmente!*, pues alguien se los tuvo que haber inventado, y es un hecho que logré *reproducirlos*. ¡Un momento!, ¿y mi brújula? Sabría cuál es el norte, ¿o el sur?, ni recuerdo la convención... ¡Pero la paridad no se conserva! Aunque, ¿valdría la pena aquella parafernalia solo para distinguir entre izquierda y derecha?... Mucho menos podría servirme del campo gravitacional de este planeta, prácticamente constante sobre la superficie, ¿qué podría hacer con mi balanza?, aunque entre latitudes la única diferencia sea que esté *trasladada* y *rotada* en el espacio, para el mismo cuerpo ¡marcaría siempre el *mismo* valor!; un momento..., pero ¿trasladada y rotada respecto a qué? Puede ser respecto al Sol o las estrellas, ¡claro, podría observar el cielo! Pero ¿cómo estar seguro de que estoy en un lugar del universo y no en otro? Con un universo observable de 93 billones de años luz de diámetro, ¿por qué no podría existir un planeta habitable como este, con un sol joven y con este cielo?... Bueno, al menos conozco la edad del universo... —se dijo sin esperanza.

¹⁸Esta discusión se desarrolla con más detalle en Sepúlveda Soto, Alonso, *Los conceptos de la física: evolución histórica*, Editorial Universidad de Antioquia, Medellín (2012), Cap. 8: Electromagnetismo y éter y Cap. 9: El espacio-tiempo.

Pobre hombre, desubicado: ¿para qué saber únicamente física? ¿Se lo habría siquiera preguntado?! Con todo, el meollo del asunto es que la imposibilidad de distinguir entre situaciones físicas diferentes, i.e, el mismo fenómeno, en las mismas condiciones, pero en diferentes momentos o lugares del espacio, o en diferentes puntos del espacio-tiempo, redundan en una ley de conservación, como veremos.

A continuación, algunos ejemplos en el contexto de la mecánica de Newton y de la teoría clásica de campos.

1. **Conservación de la energía.** Primero veamos qué representa físicamente el Hamiltoniano en un sistema de n partículas. De acuerdo con la Ec. (1.18), la Ec. (1.3), y considerando que la energía cinética de este sistema está dada por (1.4) y la energía potencial por una función del tipo de (1.2),

$$\begin{aligned} H &= -L + \sum_{i=1}^n (\dot{x}_i p_{ix} + \dot{y}_i p_{iy} + \dot{z}_i p_{iz}) \\ &= -L + \sum_{i=1}^n (m_i \dot{x}_i^2 + m_i \dot{y}_i^2 + m_i \dot{z}_i^2) = -(T - U) + 2T = T + U. \end{aligned}$$

Habida cuenta de que el Hamiltoniano es la *energía total* del sistema, si su Lagrangiano no depende explícitamente del tiempo, y por lo tanto tampoco H ,

$$S[q] = \int_{t_1}^{t_2} (T - U) dt \quad (1.39)$$

es invariante bajo la transformación canónica

$$t^* = t + \varepsilon, \quad q_i^* = q_i, \quad (1.40)$$

i.e, es invariante bajo traslación temporal. Según (1.30) y (1.33), se tiene que

$$\lambda = 1, \quad \gamma_i = 0,$$

entonces, de (1.32), la energía total del sistema se conserva:

$$H = \text{const.}$$

Hay otra forma de verlo. Si H no depende explícitamente del tiempo, como

$$[H, H]_{PB} = 0,$$

H es una primera integral del sistema de ecuaciones de ÉL, o equivalentemente,

$$\frac{dH}{dt} = 0.$$

2. **Conservación del momento lineal.** Si la acción (1.39) es invariante bajo la transformación canónica

$$t^* = t, \quad x_i^* = x_i + \varepsilon, \quad y_i^* = y_i, \quad z_i^* = z_i, \quad (1.41)$$

i.e., si no se puede distinguir físicamente entre el sistema original de partículas y el mismo sistema trasladado en x , entonces el momento total del sistema en x se conserva:

$$P_x = \sum_{i=1}^n \frac{\partial L}{\partial \dot{x}_i} = \sum_{i=1}^n p_{ix} = \text{const},$$

pues $\lambda = 0$, $\gamma_{iy} = \gamma_{iz} = 0$ y $\gamma_{ix} = 1$.

Análogamente, la invariancia del sistema bajo traslaciones en y conduce a

$$P_y = \sum_{i=1}^n \frac{\partial L}{\partial \dot{y}_i} = \sum_{i=1}^n p_{iy} = \text{const},$$

y su invariancia bajo traslaciones en z a

$$P_z = \sum_{i=1}^n \frac{\partial L}{\partial \dot{z}_i} = \sum_{i=1}^n p_{iz} = \text{const}.$$

3. **Conservación del momento angular.** Del teorema de Noether, si el sistema de partículas es invariante bajo rotaciones alrededor del eje z , o sea, bajo la familia uniparamétrica de transformaciones canónicas

$$\begin{aligned} t^* &= t \\ x_i^* &= x_i \cos \varepsilon + y_i \sin \varepsilon, \\ y_i^* &= -x_i \sin \varepsilon + y_i \cos \varepsilon, \\ z_i^* &= z_i, \end{aligned} \tag{1.42}$$

como

$$\gamma_{ix} = \left. \frac{\partial x_i^*}{\partial \varepsilon} \right|_{\varepsilon=0} = y_i, \quad \gamma_{iy} = \left. \frac{\partial y_i^*}{\partial \varepsilon} \right|_{\varepsilon=0} = -x_i, \quad \gamma_{iz} = \left. \frac{\partial z_i^*}{\partial \varepsilon} \right|_{\varepsilon=0} = 0,$$

entonces el momento angular total en z se conserva:

$$L_z = \sum_{i=1}^n (p_{iy}x_i - p_{ix}y_i) = \text{const}.$$

Asimismo, la invariancia bajo rotaciones alrededor del eje y conduce a

$$L_y = - \sum_{i=1}^n (p_{iz}x_i - p_{ix}z_i) = \text{const}$$

y la invariancia bajo rotaciones alrededor del eje x a

$$L_x = \sum_{i=1}^n (p_{iz}y_i - p_{iy}z_i) = \text{const}.$$

4. **Conservación de la energía y del momento del campo electromagnético.** En consecuencia con la invariancia de Lorentz de las ecuaciones de campo, se usará la

notación “covariante”¹⁹ para expresar la invariancia relativista, y por simplicidad, la notación de Einstein.

De acuerdo con el teorema de Noether, la invariancia del campo bajo desplazamientos paralelos en el espacio-tiempo plano²⁰

$$\begin{aligned} x^{i*} &= x^i + \varepsilon^i \quad (i = 0, 1, 2, 3), \\ A_j^* &= A_j \quad (j = 0, 1, 2, 3), \end{aligned} \quad (1.43)$$

donde las funciones de campo no son más que las componentes del cuadripotencial²¹

$$A^j = \left(\frac{\phi}{c}, \mathbf{A} \right),$$

conduce a la “ecuación de conservación”

$$\frac{\partial}{\partial x^i} \left(\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial A_j}{\partial x^i} \right)} \frac{\partial A_j}{\partial x_k} - \mathcal{L} g^{ik} \right) = 0 \quad (k = 0, 1, 2, 3), \quad (1.44)$$

pues según (1.36), (1.43) y (1.38),

$$\lambda_i^{(k)} \rightarrow g^{ik}, \quad \bar{\gamma}_j^{(k)} \rightarrow -\frac{\partial A_j}{\partial x^i} g^{ik} = -\frac{\partial A_j}{\partial x_k}.$$

Justo esta ecuación define el *tensor de energía-momento* del campo²²

$$T^{ik} = \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial A_j}{\partial x^i} \right)} \frac{\partial A_j}{\partial x_k} - \mathcal{L} g^{ik},$$

de modo que

$$\frac{\partial T^{ik}}{\partial x^i} = 0 \quad (k = 0, 1, 2, 3). \quad (1.45)$$

Y como sabemos, la Ec. (1.45) es equivalente a decir que la integral del tensor sobre una hipersuperficie que contiene todo el espacio tridimensional,

$$P^i = \frac{1}{c} \int T^{ik} dS_k,$$

se conserva, i.e, el cuádrimomento del campo P^i se conserva [6].

¹⁹La palabra “covariante” usada de esta manera significa *invariante*. No significa que las componentes de los tensores estén expresadas en la base covariante.

²⁰i.e, la densidad Lagrangiana \mathcal{L} no depende explícitamente de las coordenadas espacio-temporales. $\mathcal{L}[u, \nabla u]$ representa pues un sistema cerrado, un campo sin interacciones.

²¹A partir de las cuales se obtienen \mathbf{E} y \mathbf{B} por medio de las ecuaciones

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

²²Sin embargo, según esta definición, el tensor T^{ik} no es único, pues cualquier tensor de la forma

$$T^{ik} + \frac{\partial \psi^{ikl}}{\partial x^l},$$

con $\psi^{ikl} = -\psi^{ilk}$, satisface (1.45), ya que $\frac{\partial^2 \psi^{ikl}}{\partial x^k \partial x^l} = 0$. Para definirlo de manera única se elige un ψ^{ikl} adecuado, de forma tal que el tensor sea *simétrico*, i.e, $T^{ik} = T^{ki}$, o equivalentemente, exigiendo la *conservación del momento angular* del campo [6].

Para finalizar, aterricemos en el mundo real con una pequeña reflexión sobre la importancia de este tipo de resultados.

Del mismo modo que sorprende la insólita desgracia de aquel hombre: no saber quién es, de dónde viene, para dónde va..., por haber sufrido un accidente con consecuencias super-inusuales que lo hacían casi un experto idiota, sorprende lo inusuales e ideales que resultan los objetos estudiados por los físicos: superficies perfectamente lisas, conductores perfectos, partículas y campos en el vacío, espacio-tiempo plano, sistemas inerciales, sistemas cerrados, etc. “Mundos sin fricción”. Cabría preguntarse: ¿y para qué teorías de mundos sin fricción en un mundo con fricción? En primer lugar, las teorías son útiles para *aproximar* cálculos en el mundo real. Teoremas como el de Noether, que “viven” en estos mundos, cobran importancia en la física en la medida en que los objetos abstractos se asocian con objetos familiares de nuestro mundo. Cuando experimentamos con los objetos del mundo, probamos sus propiedades físicas, y solo podemos verificar si efectivamente son realizaciones satisfactorias de los objetos abstractos de la teoría [8]. Además, porque no sabemos qué es el mundo real y las teorías físicas justamente nos permiten descubrirlo, entenderlo, y hasta *predecir* nuevos fenómenos. Por ejemplo, la teoría general de la relatividad fue desarrollada por Einstein de manera puramente deductiva y únicamente después fue sustentada por observaciones astronómicas [6]. O ¿fue necesario esperar la invención de la bomba de vacío o la llegada del hombre a la Luna para comprobar que en el campo gravitacional todos los cuerpos se mueven de la misma manera? No, el poder de la razón fue suficiente.

Capítulo 2

SIMETRÍA EN TEORÍA CUÁNTICA

Los fenómenos cuánticos no ocurren en un espacio de Hilbert, ocurren en un laboratorio.

A. Peres

2.1. Necesitamos un nuevo lenguaje

Conservación de la ignorancia

En tiempos de la Revolución francesa no saber cuántos huevos eran necesarios para preparar una tortilla normal podía ser un asunto de vida o muerte. Alguna vez un filósofo, matemático, científico y político francés de la época, Nicolas de Condorcet, amigo de Lagrange, antojado de una tortilla, pidió una docena. Inmediatamente el cocinero, receloso, le preguntó cuál era su oficio.

—Carpintero... —respondió Condorcet azorado.

—Muéstrame las manos —inquirió el cocinero. No eres carpintero.

Terminó en prisión. Tiempo después, fue envenenado u obligado a suicidarse.¹ En fin, las brutalidades del progreso.

Su amigo contó con mejor suerte. Lagrange, favorito de María Antonieta y respetado por Napoleón, llegó incluso a ser ídolo del pueblo que alguna vez pensó darle muerte. Incluso el ambicioso Laplace también escapó de la guillotina. El primero por su naturaleza sosegada y amable. El segundo por zorro. Ambos, según De Pastoret, porque resultaron útiles para calcular trayectorias en la artillería y para dirigir la preparación del nitro necesario para la

¹E. T. Bell, *Los grandes matemáticos (desde Zenón a Poincaré)*, Editorial Losada, S. A., Buenos Aires (1948), Cap. XI: De campesino a presumido, Laplace.

fabricación de pólvora.¹

En aquellos tiempos nadie calculaba trayectorias tan bien como Laplace. Aplicó detalladamente las leyes de la gravitación de Newton a todo el sistema solar, y en su obra, *Traité de mécanique céleste*, probó su estabilidad. La estabilidad de un sistema solar altamente idealizado, por supuesto. Laplace consideró el universo compuesto *única y exclusivamente* de partículas materiales en interacción, e inspirado en la necesidad de estudiar la astronomía matemática, contribuyó profundamente al desarrollo de la teoría de probabilidades, hallándola indispensable en todas las ciencias exactas. Sin embargo, consideró la teoría de probabilidades como un medio pragmático para aproximar las verdaderas leyes de la física. Decía Laplace:

Una inteligencia que en un momento dado conociese todas las fuerzas que animan la naturaleza y la posición respectiva de los seres que la componen, podría condensar en una sola fórmula el movimiento de los cuerpos más grandes del universo y el último de los átomos: para tal inteligencia nada sería indeterminado, el pasado y el futuro estaría ante sus ojos [Ullmo, 1959].

Pero el universo no solo está compuesto de materia. ¿Qué hay del campo electromagnético, por ejemplo? Si conocemos el campo en un cierto punto (ct_1, x_1, y_1, z_1) , las ecuaciones de Maxwell nos permiten saber cómo evolucionará en todo el espacio-tiempo, nos permiten estar al tanto de la historia y del destino del campo, su pasado y su futuro “están” ante nuestros “ojos”. Algo similar puede decirse de la geometría del espacio-tiempo, representada en el tensor métrico, cuya estructura está descrita por las ecuaciones de campo de Einstein. En realidad todo esto solo nos deja más cerca del sueño de Laplace. Es que, tácitamente,

las leyes de la física clásica suponen que si duplicamos exactamente todas las condiciones de un experimento, el resultado tiene que ser exactamente el mismo [8].

De esto hablaron Newton..., Lagrange, Laplace..., Maxwell..., Einstein, etc. Esta doctrina es llamada *determinismo*.

En nuestro tiempo, por “fortuna”, no conocer la receta de la tortilla, o no saber cuánto cuesta una docena de huevos, no es tan grave.² Incluso es *fundamentalmente imposible* calcular trayectorias, y se es un sabio por comprenderlo. Pensemos en los caminos que toma la luz polarizada cuando atraviesa un cristal birrefringente. De acuerdo con la teoría electromagnética *clásica*, la luz está formada por ondas electromagnéticas. Como se ilustra en el esquema de la Figura 2.1, parte de la luz proveniente de una fuente térmica S —la que se propaga en dirección z — es polarizada linealmente en dirección α respecto al eje x (Figura 2.2). Luego de ser colimada, atraviesa un cristal C con índice de refracción anisotrópico, cuyo eje óptico es perpendicular al eje x , de modo que las componentes del campo \mathbf{E}_x y \mathbf{E}_y se propagan con diferentes velocidades, i.e., la luz incidente se divide en dos haces con intensidades proporcionales a $\cos^2 \alpha$ y a $\sin^2 \alpha$, que es lo que finalmente vemos en una pantalla. De acuerdo con Maxwell, siempre que se repita este experimento, exactamente bajo las mismas condiciones, el resultado será exactamente el mismo.

Ahora supongamos que la luz se compone de fotones. Para esto, a diferencia del esquema clásico, imaginemos una fuente térmica lo suficientemente débil y un aparato muy pintoresco

²«Pero digamos que 1800 pesos la docena o algo así...», respondió el entonces Ministro de Hacienda colombiano Alberto Carrasquilla cuando en una entrevista de abril de 2021 le preguntaron cuánto costaba una docena de huevos.

imagenes/light_detection.png

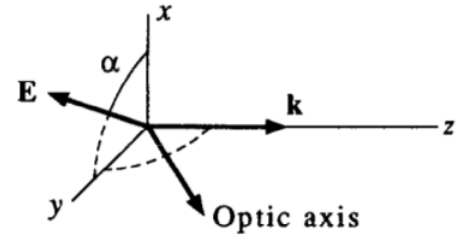


Figura 2.1: Experimento con una gran cantidad de fotones polarizados. (Tomada de [8] y editada).

Figura 2.2: Coordenadas usadas para describir la doble refringencia. (Tomada de [8]).

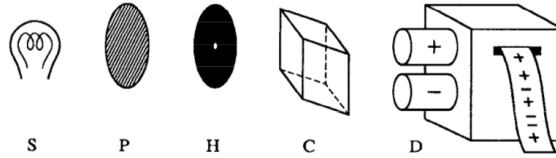


Figura 2.3: Esquema del experimento para la detección de fotones individuales. (Tomada de [8]).

con detectores lo suficientemente rápidos, a fin de que puedan registrarse los fotones individuales y, según se active un detector u otro, se imprima el resultado (Figura 2.3).³ Lo primero

³Es cierto que no hablamos estrictamente de fotones, sin embargo la idea a este nivel de la discusión es simplemente considerar *eventos individuales*. Según el efecto Hanbury Brown y Twiss o efecto *bunching*, prácticamente sin importar lo atenuada que pueda estar una fuente térmica, es sumamente improbable que en una posición fija y en un cierto instante no se detecten coincidencias (varios fotones). Nuestro aparato de detección debe ser tan rápido como para contrarrestar este efecto y al mismo tiempo imprimir ¡a la velocidad de la luz! ¡Así de pintoresco tiene que ser!

Sin embargo esta no es la única opción. La fuente térmica podría estar extremadamente atenuada, tanto como para tener un número promedio de fotones mucho menor que uno; aunque continuaríamos en el régimen clásico. También podríamos alejarnos decididamente del mundo clásico y considerar una fuente de fotones individuales, y así hablar rigurosamente de fotones. Se trata de procurar un efecto *antibunching*. Para ilustrarlo, consideremos un campo con un único modo en un estado de Fock $|n\rangle$. La función de coherencia cuántica de segundo orden para este estado es

$$g^{(2)}(\tau) = g^{(2)}(0) = \begin{cases} 0, & n = 0, 1 \\ 1 - \frac{1}{n}, & n \geq 2 \end{cases},$$

que choca con nuestra intuición clásica es que los fotones sean *indivisibles*: una partícula de energía $h\nu$ con polarización lineal —haciendo un ángulo α con el eje x — no es dividida por el cristal en dos de menor energía $h\nu \cos^2 \alpha$ y $h\nu \sin^2 \alpha$. De hecho, fotones como este de energía *completa* $h\nu$ algunas veces toman un camino, y se imprime un signo +, y otras veces toman el otro, y se imprime un signo -. Pero choca aún más con nuestra intuición clásica el hecho de que ¡no podemos predecir el camino que tomará un fotón! No podemos predecir cuál será el siguiente signo. ¿Por qué un fotón toma un camino y no el otro? No tenemos una explicación [8]. Claramente,

si duplicamos exactamente las condiciones del experimento, el resultado no necesariamente es el mismo.

Las secuencias aleatorias de signos no necesariamente serían las mismas. Las leyes de la teoría cuántica se desmarcan radicalmente de la leyes de la física clásica en la *irreproducibilidad* intrínseca de los resultados experimentales. ¿Está en juego entonces el método científico? Por supuesto que no. En realidad, si N_+ es el número de signos + y N_- es el número de signos - en la secuencia aleatoria que imprime el aparato de detección, a medida que $N_+ + N_-$ crece, las frecuencias relativas $N_+/(N_+ + N_-)$ y $N_-/(N_+ + N_-)$ tienden al resultado clásico de las intensidades de las componentes del campo:

$$\frac{N_+}{N_+ + N_-} \rightarrow \cos^2 \alpha, \quad \frac{N_-}{N_+ + N_-} \rightarrow \sin^2 \alpha. \quad (2.1)$$

*Repetiendo muchas veces un experimento, exactamente bajo las mismas condiciones, nos aproximamos a las mismas **probabilidades**.*

Así pues, más que indispensable, como señaló Laplace, la aleatoriedad se torna *fundamental*. No se trata de no contar con los medios para resolver analíticamente un problema de n cuerpos..., no se trata de usar una aproximación pragmática de las verdaderas leyes de la física, no se trata de conocimiento incompleto; se trata de una característica inherente a los fenómenos físicos. Se trata de *ignorancia total*. ¿Por qué un fotón toma un camino y no el otro? No existe una explicación, únicamente podemos pronunciarlos acerca de las probabilidades.

Superposición de fotones

¿Entonces la luz que incide en el cristal es una mezcla de fotones, unos con polarización lineal en x y otros con polarización lineal en y ? Supongamos que ubicamos entre el cristal y el detector otro cristal birrefringente idéntico, orientado exactamente del mismo modo, y que podemos disminuir el grosor del primer cristal tanto como queramos. Cuando el grosor de este cristal tiende a cero se verifica la misma tendencia de (2.1). Así pues, al disminuir el grosor del primer cristal, mientras el par de haces se superponen cada vez más, se va recuperando un único haz de fotones polarizados linealmente en la dirección original \mathbf{E} , no una mezcla de fotones, los unos con polarización lineal en x y los otros con polarización lineal en y . En general, al superponer un haz de fotones polarizados en x con otro de fotones

lo que sugiere que en una posición fija, independientemente del *delay* τ entre detecciones, cuando $n = 1$, muy probablemente no habrá coincidencias.

polarizados en y , se obtiene un haz de fotones polarizados elípticamente, no una mezcla de fotones linealmente polarizados [8].

¿Cuál es la polarización de un fotón?

Hasta ahora, al menos empíricamente, es posible conciliar los resultados de las teorías cuántica y clásica, pues para hablar de probabilidades necesitamos muchísimos eventos y una onda electromagnética clásica tiene una cantidad sumamente grande de fotones. Exploremos un poco más este acuerdo entre paradigmas.

Imaginemos dispositivos que nos permitan determinar la polarización de un “gran”⁴ número de fotones elípticamente polarizados de origen desconocido. Como en el caso clásico, donde el campo está representado por

$$\mathbf{E} = \mathbf{E}_x \cos(kz - \omega t + \delta_x) + \mathbf{E}_y \cos(kz - \omega t + \delta_y),$$

con $\delta = \delta_x - \delta_y$ arbitrario, determinar la polarización de un gran número de fotones equivale a medir sus parámetros de polarización δ y $|E_x/E_y|$. Supóngase que primero se intentan adivinar (o descartar) los casos más especiales —polarización lineal y polarización circular—. Suponemos entonces que los fotones están polarizados linealmente, i.e., $\delta = 0$, y diseñamos un dispositivo que nos permita aproximarnos al valor de $|E_x/E_y|$. En realidad ya lo tenemos. Este dispositivo consta de un cristal birrefringente con su eje óptico dispuesto como en la Figura 2.2 y de un detector muy especial alineado a la salida del cristal, como se muestra en la parte final de la Figura 2.3. Rotando el dispositivo respecto al eje que define la dirección de propagación del haz, es posible obtener una buena aproximación de $|E_x/E_y|$ a partir del ángulo en que el detector únicamente imprima signos $+$ o únicamente imprima signos $-$.⁵ Si para todo ángulo se obtuviera una secuencia aleatoria de signos $+$ y $-$, se descartaría la polarización lineal.

Por su parte, para adivinar o descartar polarización circular, se remplace el cristal por un selector de polarizaciones circulares. Los fotones polarizados circularmente salen de este dispositivo exactamente con la misma polarización, pero en dos haces separados, dependiendo de sus helicidades.⁶ Si los fotones de origen desconocido están polarizados circularmente, i.e., $|E_x| = |E_y|$, dependiendo de su helicidad, el detector debería imprimir únicamente signos $+$ o únicamente signos $-$. Si por el contrario se obtuviera una secuencia aleatoria, se descartaría la polarización circular.

En general, es posible diseñar un selector de polarizaciones elípticas, y con su ayuda, medir con buena precisión los parámetros de polarización δ y $|E_x/E_y|$ de un gran número de fotones polarizados elípticamente de origen desconocido.⁷

⁴“gran” significa una cantidad extremadamente grande, tan grande como para conciliar los resultados la teoría cuántica y la teoría clásica de la luz.

⁵Significa en realidad la prevalencia de una de las marcas, pues eventualmente algunos fotones habrán de desviarse hacia el otro detector y se imprimirá la otra marca.

⁶Un selector consta de una lámina de cuarto de onda (*qwp* por sus siglas en inglés); un cristal birrefringente alineado a la salida, con su eje óptico formando un ángulo de 45° respecto al eje óptico de la lámina, y una segunda lámina de cuarto de onda con su eje óptico perpendicular al de la primera. Ver **ejercicio 1.4** de [8].

⁷Ver **ejercicio 1.5** de [8].

Muy bien, ¿y si se trata de un *único* fotón? Como antes, podemos iniciar adivinando o descartando los casos más especiales. Comenzamos entonces suponiendo que el fotón está polarizado linealmente. Si el fotón hubiese sido efectivamente preparado con una polarización lineal haciendo un ángulo α con el eje x , ¿cómo podríamos adivinarlo? En el caso clásico que analizamos anteriormente es posible adivinar justamente por el hecho de que se trata de una profusión de fotones. Si al principio obtuviésemos una secuencia aleatoria, bastaría con rotar el dispositivo respecto al eje de propagación de la luz hasta alcanzar “uniformidad” en las impresiones. Pero ¿si solo tenemos *una* oportunidad? ¿Existe un dispositivo que nos permita medir los parámetros de polarización δ y $|E_x/E_y|$ de un único fotón individual elípticamente polarizado de origen desconocido? Este instrumento *no puede existir* [8]. Es *fundamentalmente imposible* determinar *cuál* es la polarización de un único fotón de origen desconocido. Lo que sí *es posible* responder es si el fotón tiene una polarización *específica*. Por ejemplo, ¿aquel fotón polarizado linealmente tiene polarización en y ? Existe una probabilidad de $\sin^2 \alpha$ de responder afirmativamente.

¿Cuál es la polarización de un fotón? Esta pregunta no tiene sentido.

En la teoría cuántica la aleatoriedad se torna fundamental y es fundamentalmente imposible calcular trayectorias; existen objetos sin una contraparte clásica; la intuición clásica falla rotundamente; suelen usarse las mismas palabras de la teoría clásica de un modo radicalmente diferente, hasta puede haber lugar a absurdos si se combinan mal... En fin, necesitamos un nuevo lenguaje para describir el mundo cuántico.

2.1.1. ¿Qué es la teoría cuántica?

Los estudiantes de física, tarde o temprano, nos vemos enfrentados a esta pregunta. Cuando sucede antes de tiempo, cuando confiamos tozudamente en nuestra intuición (clásica), solemos repetir las definiciones que alguna vez nos encontramos en los más famosos libros de divulgación o en las charlas de algún científico reputado, muy seguramente mal digeridas o malinterpretadas. Entre lo que repetimos es común hablar del universo microscópico, de la Naturaleza a escala atómica y subatómica, de probabilidades, etc. Cuando sucede “tarde”, o cuando se supone trastocada aquella intuición clásica por causa de los muchos cálculos de niveles de energía, tasas de transición, *cross sections*..., muchos nos damos cuenta, después de dos o más cursos de mecánica cuántica, de que tal vez nunca consideramos seriamente esta pregunta. Esta vez, confiando en nuestra experiencia y con un lenguaje más formal: espacios de Hilbert, interpretación probabilística, relaciones de incertidumbre («aunque deberían llamarse de ‘indeterminación’», decimos), etc., lo intentamos de nuevo: «La mecánica cuántica es...» Sin embargo, pocos quedamos satisfechos con la nueva definición. Incluso muchos regresamos nuevamente a nuestros preciados libros de divulgación y a las palabras de todos estos personajes que en algún momento nos inspiraron los más inocentes sueños, y, entre complacidos y apenados, nos sorprendemos de lo reveladoras que a la sazón resultan las mismas palabras que antes ni siquiera comprendíamos y, sobre todo, de lo mucho que aún ignoramos. «El gran Juan Maldacena en una charla sobre agujeros negros y entrelazamiento cuántico la define como “el conjunto de leyes que gobiernan la materia, *especialmente* átomos y partículas muy pequeñas” o como “el conjunto de leyes que usamos para describir

la Naturaleza”», podría pensar alguno buscando refinar su definición. Y quizá reflexionaría: «¿Por qué no incluye la luz!¿Especialmente?...» Nos gustaría por eso intentar responder esta pregunta desde el principio, antes de discutir sobre el nuevo lenguaje y sobre simetrías.

Según Asher Peres, en un sentido estricto,

*la teoría cuántica es un conjunto de reglas que permite el cálculo de **probabilidades** para los resultados de pruebas macroscópicas.*

Preparaciones y pruebas

«¿macroscópicas?!», más de un físico replicaría. Pero intentemos entender mejor esta definición. Según Peres, existen dos nociones primitivas (no definidas) en teoría cuántica, así como los puntos y las líneas rectas en los axiomas de la geometría Euclidiana, que son los dos tipos de tareas diferentes que realizan los físicos en los laboratorios: *preparaciones* y *pruebas*. «Una preparación es un procedimiento experimental que está completamente especificado, como una receta de un buen libro de cocina.» «Una prueba inicia como una preparación, pero esta también incluye un paso final en el cual *información*, previamente desconocida, es suministrada al físico que está realizando el experimento». Y nos dice: «Note que una preparación usualmente involucra pruebas, seguidas por una *selección* de resultados específicos.» En el ejemplo de la Figura 2.3, la información nos es dada por medio de los símbolos + y – impresos en la tira que suelta el aparato de detección, una secuencia de marcas nos permite *distinguir* en cada prueba cuál detector fue activado, i.e., cuál camino tomó el fotón. Esta secuencia de pruebas que pueden ser realizadas en un laboratorio, comienzan como un procedimiento experimental riguroso: «Defino un sistema de referencia adecuado; como quiero moverme en la frontera entre lo clásico y lo cuántico, usaré una fuente térmica, pero suficientemente atenuada para lograr un número promedio de fotones mucho menor que uno; luego alinearé un polaroid de forma tal que sus fibras absorban las ondas con un vector de campo eléctrico paralelo a ellas y que la luz resultante surja linealmente polarizada, haciendo un ángulo α con el eje x de mi sistema de referencia, eje definido perpendicular a la dirección de propagación de la luz; colimaré el haz, y finalmente usaré un cristal con un índice de refracción anisotrópico con su eje óptico perpendicular al eje x , de tal suerte que las componentes de la onda se propaguen independientemente dentro del cristal y cada uno de estos haces “busque” su detector correspondiente. Todo deberá estar perfectamente alineado.», diría un físico en su laboratorio. Y la información obtenida no es para nada trivial: las pruebas que siguen exactamente la misma receta del físico no tienen necesariamente el mismo resultado, como ya discutimos.

En términos de estas nociones primitivas, esta es la definición de Peres:

*En un sentido estricto, la teoría cuántica es un conjunto de reglas que permiten el cálculo de **probabilidades** para los resultados de las pruebas, las cuales siguen preparaciones especificadas.*

Si hablamos de probabilidades, en la práctica, tratamos con muchísimas pruebas, pues en

teoría deben ser infinitas. En nuestro ejemplo:

$$\frac{N_+}{N_+ + N_-} \xrightarrow{(N_+ + N_-) \rightarrow \infty} \cos^2 \alpha, \quad \frac{N_-}{N_+ + N_-} \xrightarrow{(N_+ + N_-) \rightarrow \infty} \sin^2 \alpha.$$

Pero ¿cuántas son “muchísimas”? Todo depende del nivel de precisión deseado por el experimentador.

¿Y qué hay sobre la descripción de la Naturaleza a escala atómica, subatómica...? En las siguientes secciones lo aclararemos. Sin embargo debe estar claro que estrictamente nos referimos a probabilidades de resultados de pruebas macroscópicas. Recordando el epígrafe: «Los fenómenos cuánticos no ocurren en un espacio de Hilbert, ocurren en un laboratorio.» En nuestro ejemplo, el resultado de la observación (un signo + o un signo -) es registrado por el aparato de forma clásica. Observamos los sistemas físicos (fotones, electrones, nucleones, átomos, etc.) y describimos sus propiedades justamente a través de su interacción con un aparato macroscópico.

Pero no debemos confundirnos, macroscópico no significa clásico. Aunque Peres distinga entre objetos macroscópicos y objetos microscópicos, no significa que estos obedezcan unas leyes de la física y aquellos otras. Lo que quiere decir es que deben usarse diferentes *modos de descripción*. Dejaremos para el último capítulo la discusión sobre la descripción cuántica del aparato de medición. En el presente capítulo usaremos un modelo más simple, pues lo que nos interesa discutir en este momento es cómo se representa matemáticamente una simetría.

2.1.2. ¿Qué es una medición?

Como se dijo, en teoría cuántica suelen usarse las mismas palabras de la mecánica clásica, sin embargo, sus significados pueden ser radicalmente opuestos. La medición puede ser uno de los casos más dramáticos. En principio parece que se hablara de lo mismo. Un aparato de medición necesariamente tiene que interactuar con el sistema físico que estudiamos. La medición es un proceso indirecto donde obtenemos información acerca de una cierta propiedad del sistema microscópico a través de su interacción con la propiedad correspondiente del dispositivo de medición, el cual suele ser macroscópico. Según Peres:

Tenemos que *interpretar* los resultados experimentales producidos por nuestro equipo. Hacemos esto construyendo un *modelo teórico* mediante el cual el comportamiento del equipo macroscópico es descrito por unos pocos grados de libertad que interactúan con aquellos del sistema microscópico bajo observación. A esto lo llamamos una “medición” de un sistema microscópico.

Un ejemplo concreto de modelo teórico es el Hamiltoniano de interacción en el experimento de Stern-Gerlach (SG):

$$H_{int} = -\boldsymbol{\mu} \cdot \mathbf{B}(\mathbf{r}),$$

donde el momento angular $\boldsymbol{\mu}$ del átomo, nuestro sistema de interés, está acoplado a la posición del centro de masa \mathbf{r} , el cual claramente es un grado de libertad macroscópico. En mecánica clásica el concepto de medición es completamente análogo, simplemente cambia el modo de descripción del sistema que se estudia. Ya no se habla de un sistema microscópico y el

Hamiltoniano y el momento angular se tratan en el contexto de la formulación Hamiltoniana de la mecánica clásica.

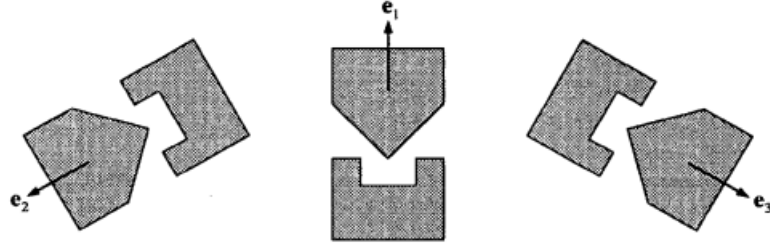


Figura 2.4: Tres posibles orientaciones para el imán del SG, cada una haciendo un ángulo de 120° con las otras. (Tomada de [8]).

Sin embargo lo que nos interesa es acentuar las diferencias, pues hemos expresado la necesidad de un nuevo lenguaje. La diferencia fundamental con el concepto de medición clásica es que en teoría cuántica no se asume que la propiedad que va a ser medida *existe* objetivamente antes de la interacción del aparato de medición con el sistema. Regresemos al SG. En la Figura 2.4 se muestran tres posibles orientaciones para el experimento de SG representadas por los vectores unitarios \mathbf{e}_i ($i = 1, 2, 3$), donde cada una hace un ángulo de 120° con una de las demás y de -120° con la otra. Es un hecho que cada uno de estos experimentos —realizado por separado— arroja el mismo resultado:

$$\mu_i = \boldsymbol{\mu} \cdot \mathbf{e}_i = \pm\mu \quad (i = 1, 2, 3).$$

Si asumimos que estas propiedades existen objetivamente, o sea, independientemente del proceso de medición, llegamos a una contradicción:

$$\mu_1 + \mu_2 + \mu_3 = \boldsymbol{\mu} \cdot (\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \equiv 0.$$

¿Cómo podrían estas componentes ser todas iguales a $\pm\mu$ y al mismo tiempo sumar cero? Pero lo resume de este modo:

*La contradicción es fundamental. una vez asociamos valores **discretos** con las componentes de un vector que puede ser **continuamente** rotado, el significado de estos valores discretos no puede ser el de componentes vectoriales “objetivas”, las cuales serían independientes del proceso de medición [8].*

En el Cap. 3, Subsec. 3.3.3, se amplía esta discusión.

2.1.3. ¿Qué es un estado cuántico?

Un sistema cuántico es cualquier cosa que admita una descripción dinámica cerrada dentro de la teoría cuántica [8]. Por ejemplo en el experimento de SG el sistema cuántico no es el átomo en sí mismo, tampoco en los experimentos ópticos que hemos discutido son los fotones. En realidad son el momento angular $\boldsymbol{\mu}$ del átomo y la polarización de los fotones, respectivamente. Este es el problema con el concepto de sistema cuántico, que depende de

la propiedad en que estemos interesados. En vista de que este concepto no parece tan firme, nos concentremos mejor en el de *estado cuántico*, pues tiene una definición *operacional* clara y puede ser definido a partir de las nociones primitivas de *preparación* y *prueba*:

Un estado está caracterizado por las probabilidades de los diferentes resultados de toda prueba concebible.

Consideremos una cierta preparación y un conjunto de pruebas. Si hablamos de probabilidades, necesariamente debemos considerar que cada una de las pruebas concebibles debe realizarse muchísimas veces siguiendo idénticas preparaciones. Pues bien, después de hacerlo, encontramos que las frecuencias relativas de los resultados de cada prueba tienden justamente a sus respectivas probabilidades. Según la definición anterior, cada distribución de probabilidad de cada prueba que nos podamos imaginar, y que en principio sea realizable, i.e., que no viole las leyes de la física, define un cierto estado. Pero esto resulta bastante redundante, veámoslo con un ejemplo. Consideremos de nuevo un experimento óptico con fotones individuales en que los fotones son idénticamente preparados en el estado de polarización circular derecha $|RHC\rangle$. Sabemos que si usamos un selector de polarización circular podremos predecir *con certeza* por dónde surgirá el fotón, y además, podremos *predecir* las *probabilidades* de los diferentes resultados de cualquier otra prueba concebible, e.g., una lámina de cuarto de onda seguida por un cristal birrefringente orientado arbitrariamente; incluso la prueba podría servirse únicamente del cristal, y testear por ejemplo polarización vertical; en fin, o cualquier otra prueba realizable que nos imaginemos. Es decir, la distribución de probabilidad de una única prueba puede ser suficiente para conocer las distribuciones de todas las demás. En general, las probabilidades no son independientes y uno puede especificar —de muchas maneras diferentes— un conjunto restringido de pruebas tales que, si las probabilidades de los resultados de esas pruebas son conocidas, es posible predecir las probabilidades de los resultados de cualquier otra prueba [8].

Consideremos ahora un par de definiciones que nos serán útiles para sintetizar la discusión de esta subsección en un postulado.

Pruebas repetibles

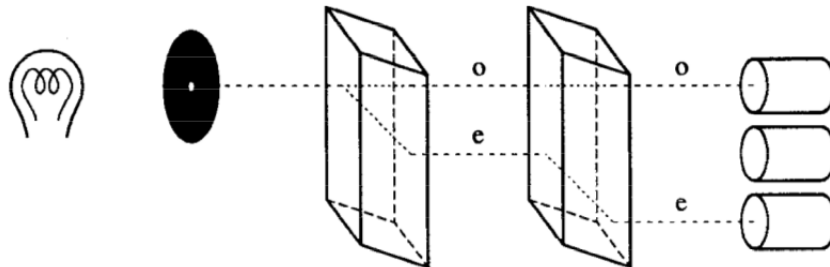


Figura 2.5: Ejemplo de una prueba repetible con luz polarizada. (Tomada de [8]).

En la Figura 2.5 se muestran dos pruebas consecutivas que arrojan el mismo resultado. La luz térmica es colimada y luego dividida por un cristal birrefringente en dos haces *o* (ordinario) y *e* (extraordinario). La novedad está en los tres detectores y en el segundo cristal,

idéntico al primero, dispuesto justo después, de forma que se realicen pruebas consecutivas idénticas y que el intervalo de tiempo entre las pruebas sea prácticamente despreciable. Empíricamente se sabe que el detector del medio nunca es excitado, como se ilustra. En este capítulo únicamente consideraremos pruebas de esta clase, i.e. pruebas que siempre arrojan el mismo resultado: pruebas *repetibles*.

Claramente no todas las pruebas son repetibles. Sirviéndonos del mismo esquema, podemos imaginarnos fácilmente un contraejemplo. Si la prueba constara del cristal y una lámina de cuarto de onda, y consideráramos dos pruebas idénticas, una justo después de la otra, evidentemente el resultado no sería el mismo, pues al final de la segunda prueba los tres detectores se activarían. Es que en realidad casi todas las pruebas son irrepetibles. Las pruebas repetibles son idealizaciones, tan solo una noción conveniente para el análisis teórico, como veremos.

Pruebas cuánticas maximales

En el experimento de SG, independientemente de la orientación del imán respecto a la dirección de movimiento de las partículas de momento angular μ , una de espín s siempre tendrá $(2s + 1)$ resultados diferentes. Cada orientación define una prueba diferente, sin embargo, todas tienen la misma cantidad de resultados distintos. En general, si N es el número de resultados diferentes que se obtienen al realizar una prueba a un cierto sistema cuántico, una prueba *maximal* o *completa* es cualquier prueba que tenga exactamente N resultados distintos. En nuestro ejemplo, cualquier orientación del imán define una prueba completa.

Deber estar claro que una prueba se dice *completa* respecto a un cierto grado de libertad, justamente el que define el sistema cuántico. Por ejemplo, en el SG la prueba se dice completa respecto al espín de la partícula; si nos ocupáramos de la polarización, sería completa respecto a esta. Los demás grados de libertad son ignorados. En la práctica, el resultado de cada una de estas pruebas es observado *correlacionando* el valor del grado de libertad interno —que está siendo testeado— con la *posición* de la partícula que sale, lo cual puede ser luego detectado por medios macroscópicos [8].

Todo lo que hemos discutido puede ser sintetizado de la siguiente manera:

Determinismo estadístico. Si un sistema cuántico es preparado de tal manera que arroje con certeza un resultado predecible en una prueba maximal especificada, los distintos resultados de **cualquier** otra prueba tienen **probabilidades** definidas. En particular, estas probabilidades no dependen de los detalles del procedimiento usado para preparar el sistema cuántico, para que este arroje un resultado específico en una determinada prueba maximal. Un sistema preparado de esa manera se dice que es un estado **puro**.

2.1.4. Hacia la evolución unitaria

El propósito de este capítulo es comprender —desde una perspectiva *instrumentalista*— cómo se representa una simetría en teoría cuántica. A continuación, guiados por las ideas de Peres, discutiremos algunos hechos *empíricos* que dan lugar a algunos *postulados* de la teoría cuántica. Sin embargo, también nos serviremos de la intuición matemática para establecer otros, cuyas consecuencias, por supuesto, deben ser verificadas experimentalmente. No pretendemos desarrollar todos los conceptos y métodos de la teoría ni discutir en toda su amplitud sus implicaciones, solamente los necesarios para ganar algo de claridad acerca de su estructura lógica, para familiarizarnos con nuestro nuevo lenguaje, con la fenomenología de las pruebas cuánticas, y al fin, para comprender el teorema de Wigner y sus profundas consecuencias; en definitiva, para entender cómo se representa una simetría en teoría cuántica.

Ley de reciprocidad

Antes de adentrarnos en el corazón de la teoría cuántica, es necesario entender un resultado que no tiene un análogo clásico y que —en general— está fundado en la experiencia.

Esta vez, en lugar de considerar pruebas consecutivas idénticas, consideremos pruebas consecutivas *diferentes*. En la Figura 2.6 se ilustra el resultado de someter partículas de espín 1 en un estado mezclado a dos pruebas de SG consecutivas diferentes, siempre que las separaciones angulares de los haces que surgen de la primera prueba no permitan que se superpongan, y que la separación de ambos imanes sea la adecuada para que estos haces prácticamente no se dispersen y el segundo imán actúe sobre cada haz esencialmente del mismo modo, tal como lo sugiere la disposición simétrica de los cúmulos. Al menos dos cosas llaman mucho la atención. En primer lugar, que todos los nueve caminos son perfectamente distinguibles y, más aún, la forma cuasi-clásica en que lo imaginamos, tal como en los experimentos con fotones (Figura 2.5), ¡por medio de trayectorias!, pues ¿cómo más podríamos interpretar los resultados experimentales?

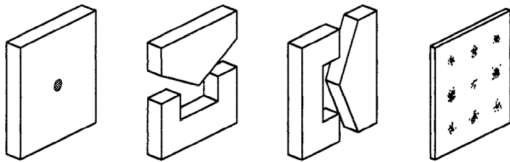


Figura 2.6: Dos experimentos de SG consecutivos para partículas de espín 1. (Tomada de [8]).

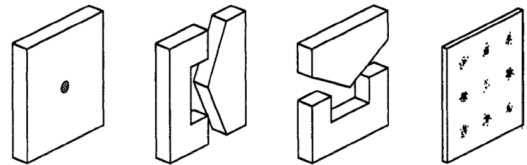


Figura 2.7: Las mismas pruebas de la Figura 2.6, pero en orden inverso. (Tomada de [8]).

En la figura 2.7 se ilustra un experimento con las mismas pruebas del anterior, pero intercambiadas, y en principio, con resultados similares. De hecho empíricamente se encuentra que los resultados son “iguales” [8], desde luego, después de realizar muchísimas pruebas. Sean $P_{\mu m}$ ($\mu, m = \pm 1, 0$) la matriz de probabilidades de transición del primer experimento y $\Pi_{m\mu}$ la del segundo, i.e., la “tabla” de probabilidades para observar el resultado μ a partir de la preparación del estado puro m y la tabla para observar el resultado m a partir de

la preparación del estado puro μ , respectivamente. La matriz $P_{\mu m}$ se define como $I_{\mu m}/I_m$, donde I_m representa las tres intensidades de los haces que surgen del primer imán e $I_{\mu m}$ las intensidades de los nueve haces que surgen del segundo. Idealmente, $I_m = \sum_{\mu} I_{\mu m}$, por lo cual se tiene que⁸

$$\sum_{\mu} P_{\mu m} = 1. \quad (2.2)$$

Matemáticamente, la ley de reciprocidad establece que

$$P_{\mu m} = \Pi_{m\mu}. \quad (2.3)$$

El principio de interferencia

Adentrémonos pues en el corazón de la teoría cuántica. Veamos varias formas de plantear la misma cuestión que nos ocupará: ¿por qué se asume unitaria la evolución temporal en teoría cuántica? ¿Cómo se suman las probabilidades cuánticas? Pese a ser observables, ¿son las probabilidades realmente lo fundamental?

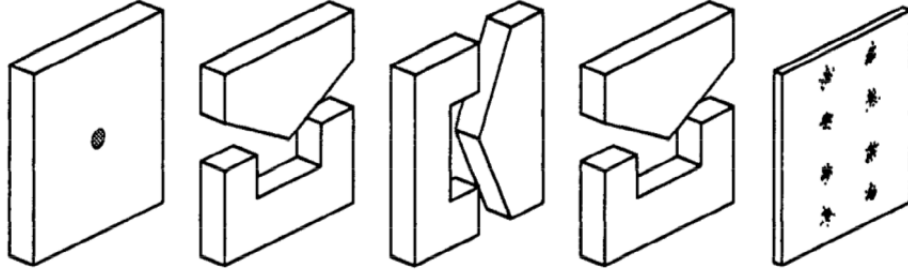


Figura 2.8: Tres experimentos de SG consecutivos para partículas de espín 1/2. (Tomada de [8]).

En la Figura 2.8 se muestran tres experimentos de SG consecutivos, dos de los cuales —los de los extremos— se asumen idénticos. La fuente de partículas de espín 1/2 prepara mezclas aleatorias, de modo que en cada SG la probabilidad de obtener alguno de los dos resultados posibles es exactamente la misma. Se asume también que cualquiera de los caminos que pueda tomar una partícula es completamente *distinguishible* de los demás, i.e, los haces no se superponen a la salida de cada SG, y tampoco se separan demasiado, pues el siguiente SG debe actuar aproximadamente del mismo modo sobre cada uno. Resultan pues ocho cúmulos y ocho caminos completamente distinguibles; es más, los caminos se pueden distinguir sin perturbar la dinámica de los sistemas cuánticos. Así pues, el rasgo característico de este experimento conceptual es la distinguibilidad. ¿Cómo se calculan las probabilidades? Por ejemplo, ¿cuál

⁸ $P_{\mu m}$ también satisface

$$\sum_m P_{\mu m} = 1.$$

Las matrices de elementos no negativos que satisfacen esta ecuación y la Ec. (2.2) son llamadas matrices doblemente estocásticas.

es la probabilidad de que una partícula arribe a alguno de los cúmulos superiores? Tenemos dos eventos mutuamente excluyentes, que la partícula resulte en el cúmulo superior derecho, o que arribe al cúmulo superior izquierdo. Al lanzar dos monedas, ¿cuál es la probabilidad de que al final ambas muestren la misma cara? Tenemos dos eventos mutuamente excluyentes, que ambas monedas muestren un busto, o que ambas lo oculten. En las dos situaciones, la cuántica y la clásica, las probabilidades se calculan del mismo modo, ¿por qué? Por la capacidad de distinguir con certeza cuando ocurren los eventos. Una vez la partícula arriba a la pantalla, sabemos cuál camino tomó; asimismo, una vez las monedas se detienen en el suelo, sabemos cuál evento ocurrió. ¿Cómo se calculan las probabilidades entonces? Por medio de la regla de adición de probabilidades:

$$P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}, \quad (2.4)$$

la cual es válida para la ocurrencia de dos eventos A y B . Generalicemos este resultado. Usemos los índices $m\mu n$ para rotular cada uno de los ocho cúmulos, m hace referencia a la preparación de los estados del primer SG, μ a la de los estados del segundo y n al posible resultado. Por tanto, la intensidad del cúmulo $m\mu n$ es proporcional a

$$\Pi_{n\mu} P_{\mu m} = P_{\mu n} P_{\mu m},$$

de acuerdo con la ley de reciprocidad. Ahora imaginemos que vamos debilitando de a poco el campo magnético que produce el SG del medio, de modo que gradualmente la separación horizontal de los haces disminuye, hasta que finalmente los haces terminan completamente *superpuestos* por pares (Figura 2.9). ¿Cómo se calcula ahora la probabilidad de que una partícula arribe a alguno de los cúmulos superiores? Ingenuamente podría pensarse que es cuestión de sumar las probabilidades de los posibles eventos, o sea, ¡que se obtienen cuatro cúmulos, cada uno con aproximadamente el doble de su antigua intensidad!, pero ¿cómo puede ser el paso por un camino *indeterminado* la ocurrencia de un evento? ¿Cómo así que un camino indeterminado? μ es el índice que nos indica los dos posibles caminos que puede tomar una partícula luego de salir del SG del medio. Justamente la separación horizontal de la que hemos hablado es signo de la intensidad del campo del medio. Si aquel campo es tal que no es posible distinguir entre un valor de μ u otro, entre un camino u otro, estamos ante una situación radicalmente diferente. ¿Y cuál es la diferencia? Justamente la capacidad de distinguir con certeza entre posibles alternativas: la *información*. Todo esto se resume en el siguiente postulado:

Principio de interferencia. Si un sistema cuántico puede seguir varios posibles caminos desde una preparación determinada a una prueba determinada, la probabilidad para cada resultado de esa prueba **no** es en general la suma de las probabilidades separadas correspondientes a los distintos caminos.

Además, si lo pensamos bien, se estaría violando la *repetibilidad* de las pruebas, o lo que es lo mismo, el postulado *determinismo estadístico*. Entonces ¿cómo se calculan las probabilidades cuando hay *superposición*, cuando la indistinguibilidad es esencial? En la Figura 2.9 vemos que de la superposición de los haces resultan dos que se refuerzan y dos que se “aniquilan”, dos que interfieren constructivamente y dos que interfieren destructivamente. Y esto ya tiene una explicación clásica a través de las ecuaciones de Maxwell: las *amplitudes* del campo son aditivas y satisfacen un conjunto de ecuaciones diferenciales lineales, mientras que sus intensidades no lo son, pues son cuadráticas en las amplitudes del campo. Por paradójico que

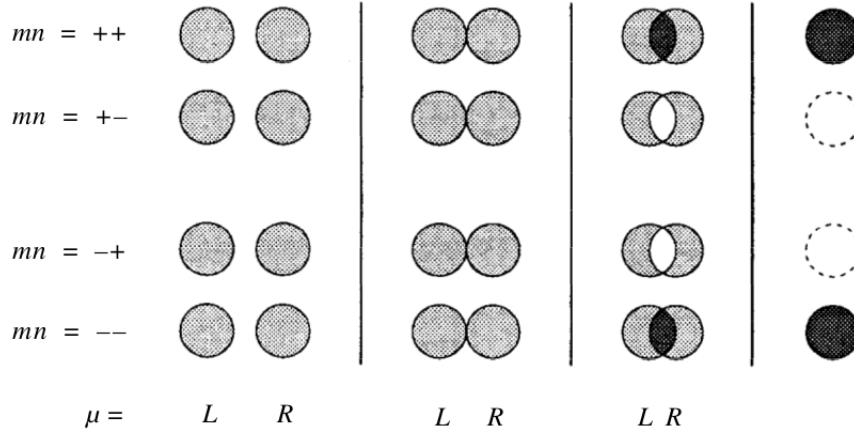


Figura 2.9: Comportamiento de los ocho haces en el experimento de la Figura 2.8 cuando el campo del imán del medio es gradualmente apagado. (Tomada de [8]).

pueda sonar, lo que sugiere esto es una teoría cuántica ondulatoria en que las intensidades, proporcionales al número de fotones por unidad de área, esto es, a la *probabilidades* de arribo de estos fotones, se deduzcan a partir de *amplitudes* de transición. Lo que sugiere esto es la existencia de amplitudes de transición *complejas*, donde

$$P_{\mu m} = |C_{\mu m}|^2 \quad \text{y} \quad \Pi_{m\mu} = |\Gamma_{m\mu}|^2.$$

Y ¿por qué complejas? De la ley de reciprocidad ya sabemos que $|C_{\mu m}| = |\Gamma_{\mu m}|$, así que son las fases la que se tornan fundamentales, las que justamente pueden dar lugar a los fenómenos de interferencia. En lugar de sumar intensidades, es decir, en lugar de sumar probabilidades de transición, se suman amplitudes de transición complejas que satisfagan una ecuación diferencial lineal. ¿Y cómo se suman? Por analogía con el electromagnetismo —i.e con una teoría *clásica*—, se postula lo siguiente:

Ley de composición de amplitudes de transición. *Las fases de las amplitudes de transición pueden ser elegidas de tal manera que, si se dispone de varios caminos desde el estado inicial hasta el resultado final, y si el proceso dinámico no deja huella que permita distinguir cuál camino fue tomado, la amplitud completa para el estado final es la **suma** de las amplitudes para los distintos caminos.*

De modo que en nuestro triple SG, para no violar el postulado de *determinismo estadístico* y de acuerdo con la *ley de composición de amplitudes de transición*, tenemos que

$$\sum_{\mu} \Gamma_{n\mu} C_{\mu m} = \delta_{nm}, \quad (2.5)$$

y ya que de la ley de reciprocidad sabemos que $|C_{\mu m}| = |\Gamma_{n\mu}|$, si elegimos las fases de modo que $\Gamma_{n\mu} = C_{\mu m}^*$, llegamos por fin, a partir de experimentos conceptuales, a una relación importantísima:

$$\sum_{\mu} C_{\mu m}^* C_{\mu m} = \delta_{nm}, \quad (2.6)$$

donde el Det $C_{\mu m} = 1$ y además

$$\sum_m C_{\mu m} C_{\nu m}^* = \delta_{\mu\nu}. \quad (2.7)$$

Las matrices que satisfacen esta relación son llamadas matrices *unitarias*.

Bien, entonces tenemos dos opciones. La primera es considerar fundamentales las probabilidades de transición, asumirlas conocidas (después de todo son observables), y a partir de ellas calcular las amplitudes de transición, su “raíz cuadrada”. O considerar fundamentales las amplitudes de transición, una cantidad que no es observable, y a partir de ellas calcular las probabilidades, i.e., su módulo cuadrado. Un enfoque estrictamente instrumentalista sugeriría tomar la primera opción, i.e., dada $P_{\mu m}$, encontrar una matriz unitaria $C_{\mu m}$ que satisfaga $P_{\mu m} = |C_{\mu m}|^2$. De hecho, siempre que $P_{\mu m}$ sea *doblemente estocástica*, este problema tiene solución para un conjunto *denso* de $P_{\mu m}$. Sin embargo también debe satisfacer desigualdades demasiado complejas.⁹ Esto sugiere entonces que el enfoque más conveniente es el segundo. Se toman pues por *fundamentales las amplitudes, no las probabilidades*. Esto no quiere decir que se abandone el enfoque instrumentalista. De todos modos cualquier consecuencia matemática que surja de la relación entre $P_{\mu m}$ y $C_{\mu m}$ tendrá que ser verificada experimentalmente.

Veamos una razón más para considerar fundamentales las amplitudes, no las probabilidades. Si modificamos nuestro triple experimento de SG (Figura 2.8), de modo que ya las pruebas de los extremos no sean idénticas, llegamos a un resultado muy alentador, que nos indica que vamos por el camino correcto. En general debemos hablar de la probabilidad de obtener el resultado \mathbf{r} a partir de la preparación m . En situación de indistinguibilidad, de acuerdo con la ley de composición de amplitudes, la amplitud completa para el estado final $C_{\mathbf{r}m}$ es

$$\sum_{\mu} \Gamma_{\mathbf{r}\mu} C_{\mu m} = C_{\mathbf{r}m}, \quad (2.8)$$

donde $|\Gamma_{\mathbf{r}\mu}|^2 = \Pi_{\mathbf{r}\mu}$, y en situación de distinguibilidad, o sea, cuando el campo del medio está encendido, la probabilidad de transición del camino $m\mu\mathbf{r}$ es $\Pi_{\mathbf{r}\mu} P_{\mu m}$ y su amplitud de transición $\Gamma_{\mathbf{r}\mu} C_{\mu m}$. Lo interesante es que ¡la matriz $C_{\mathbf{r}m}$ también es unitaria! Más aún, las matrices Γ y C están ordenadas *temporalmente*. Para conectar el presente con el futuro, i.e., para ir de la preparación m al resultado \mathbf{r} , primero actúa C y luego actúa Γ , lo cual evidentemente puede ser generalizado, pues C también es unitaria. Esto sugiere que la evolución dinámica de los sistemas cuánticos está representada por el producto de matrices unitarias [8].

¿Por qué se asume unitaria la evolución temporal en teoría cuántica? ¿Cómo se suman las probabilidades cuánticas? Pese a ser observables, ¿son las probabilidades realmente lo fundamental? Lo fundamental son las amplitudes de transición, no las probabilidades. En caso de que la ley de adición de probabilidad no sea aplicable, a partir de la amplitud completa $C_{\mathbf{r}m}$ para el estado final, la cual es la suma de las amplitudes para los distintos caminos, se calcula la probabilidad de obtener el resultado \mathbf{r} siguiendo la preparación m :

$$P_{\mathbf{r}m} = |C_{\mathbf{r}m}|^2.$$

⁹Quien desee profundizar en este problema algebraico puede consultar las páginas 41-43 de [8].

Y aunque de momento no podamos dar una respuesta definitiva a la primera pregunta, las ecuaciones (2.5) a (2.8) sugieren que la evolución unitaria se sigue de la ley de composición de amplitudes de transición, donde es crucial que el proceso dinámico no deje huella alguna que permita distinguir cuál camino fue tomado.

2.2. ¿Qué define una simetría en teoría cuántica?

Una simetría es una equivalencia de situaciones físicas diferentes [8]. Reconsideremos esta definición en mecánica clásica para advertir más claramente la diferencia con la definición en teoría cuántica. La invariancia de la acción bajo transformaciones infinitesimales de coordenadas y de las funciones de campo definen una simetría en mecánica clásica. Por medio del cambio de la acción —su primera variación— nos es imposible adquirir conocimiento físico, información, pues ¡la acción no cambia bajo transformaciones!, nos es imposible distinguir entre situaciones físicas diferentes. Eso es lo esencial. Si pensamos un poco mejor lo que se dijo en el Cap 1, puede resultar contradictorio que sirviera de ejemplo el caso de nuestro físico “desubicado” en el contexto de la física de los sistemas inerciales, pues su sistema de referencia era el planeta rotante, un sistema de referencia no inercial. De nuevo, debemos entender que el teorema de Noether en relatividad especial, y en general el cálculo de variaciones, nos sirve para aproximar cálculos en el mundo real. Además, la discusión sobre simetría va mucho más allá, tiene que ver con *indistinguibilidad*. Imaginemos nuevamente a nuestro físico aficionado con su brújula. Aunque aquel planeta rote respecto a su eje, nuestro personaje, brújula en mano y parado sobre la superficie mirando hacia el norte, rota junto con el planeta con ¡la aguja de la brújula apuntando siempre en la misma dirección! Es decir, por medio de la dirección del campo magnético del planeta no puede distinguirse entre situaciones físicas diferentes.

Dicho de otra forma, la existencia de una simetría implica la equivalencia de una transformación activa y una transformación pasiva [8]. Un ejemplo más simple es la rotación rígida de una superficie plana, como se ilustra en la Figura 2.10. Al realizar una transformación activa

$$\begin{aligned} F'_x &= F_x \cos \theta - F_y \sin \theta \\ F'_y &= F_x \sin \theta + F_y \cos \theta \end{aligned} \tag{2.9}$$

y una transformación pasiva

$$\begin{aligned} F_{x'} &= F_x \cos \theta + F_y \sin \theta \\ F_{y'} &= -F_x \sin \theta + F_y \cos \theta \end{aligned} \tag{2.10}$$

simultáneamente, el resultado es una situación final y una situación inicial indistinguibles (Figura 2.10(c)):

$$F'_{x'} = F_x, \quad F'_{y'} = F_y. \tag{2.11}$$

Una *simetría cuántica* también implica una equivalencia entre una transformación activa y una transformación pasiva. Debemos pues antes responder una pregunta: ¿de qué tipo de transformaciones hablamos en teoría cuántica? Del mismo modo que el estado clásico

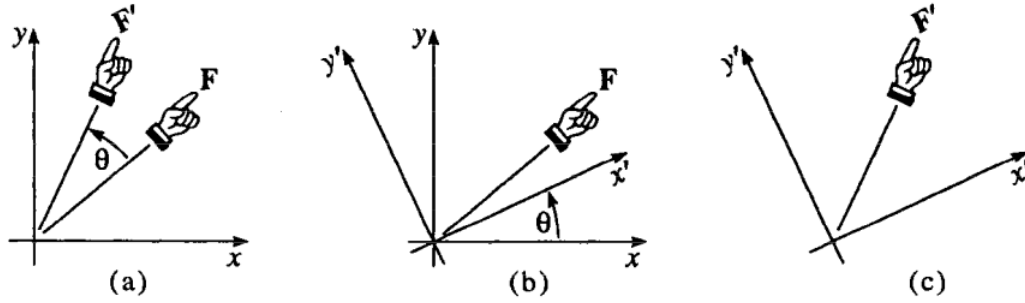


Figura 2.10: (a) Transformación activa. (b) Transformación pasiva. (c) Transformaciones activa y pasiva simultáneamente. (Tomada de [8]).

representado por \mathbf{F} cambia a \mathbf{F}' por medio de una transformación activa en \mathbb{R}^2 (Figura 2.10(a)), en teoría cuántica una transformación activa es un cambio de estado del sistema cuántico representado por el vector *complejo* $|\psi\rangle$:

$$|\psi\rangle \rightarrow |\psi'\rangle,$$

pero sobre un espacio de Hilbert \mathcal{H} . Mientras que una transformación pasiva, representada por un operador unitario, es una elección de una base diferente —i.e. una prueba maximal diferente— en \mathcal{H} , análogo a lo que se ilustra en la Figura 2.10(b) en \mathbb{R}^2 .

En este punto podríamos procurar una situación análoga a la expresada en la Ec. (2.11), donde, luego de realizar una transformación activa y una transformación pasiva simultáneamente, no es suficiente conocer las componentes de los vectores para desvelar una posible transformación. Sin embargo, ya no es tan simple “traducir” o razonar a partir de analogías, justamente porque el espacio vectorial de la teoría cuántica es complejo. En teoría cuántica no existe una relación uno a uno entre estados y vectores complejos, infinitos vectores *paralelos* entre sí representan el mismo estado. Es cierto que la condición de estar normalizados constituye un filtro, sin embargo, esto no impide que pueda introducirse una *arbitrariedad de fase*, ya que $|v\rangle$ y $e^{i\delta}|v\rangle$ tienen exactamente la misma norma para todo $|v\rangle$ en \mathcal{H} . Entonces ¿por qué no eliminamos aquella arbitrariedad? Porque las fases son fundamentales en teoría cuántica. Con arreglo al principio de superposición,¹⁰

$$|u\rangle + e^{i\delta}|v\rangle$$

y

$$|u\rangle + |v\rangle$$

representan estados puros realizables, y además, diferentes, pues como sabemos, según nuestro estudio de fotones polarizados, una diferencia de fase δ entre haces superpuestos da lugar, en general, a fotones polarizados elípticamente. De manera que el simple hecho de elegir una nueva base —e.g. $\{|u'\rangle, e^{i\delta}|v'\rangle\}$ — para expandir el vector que resulta de una transformación activa, i.e.,

$$|\psi'\rangle = \alpha|u'\rangle + \beta(e^{i\delta}|v'\rangle) \quad (\alpha, \beta \in \mathbb{C}),$$

¹⁰**Principio de superposición.** Cualquier vector complejo, excepto el vector nulo, representa un estado puro realizable.

y que conserva las mismas componentes del vector ‘antiguo’ respecto a la base ‘antigua’ $\{|u\rangle, |v\rangle\}$, o sea,

$$|\psi\rangle = \alpha |u\rangle + \beta |v\rangle,$$

claramente nos permite distinguir entre situaciones físicas diferentes.

¿Cómo lidiamos pues con esta arbitrariedad de fase? Aunque ya se sospecha el tipo de transformación de la que hablamos, hay que establecerlo formalmente, y para esto basta recordar qué es la teoría cuántica. La esencia de la teoría cuántica es proveer una representación matemática de los estados, junto con reglas para calcular las probabilidades de los distintos resultados de cualquier prueba [8]. Imaginemos un experimento óptico en que preparamos fotones con polarización vertical, i.e.,

$$|V\rangle = \cos \theta |\theta_+\rangle + \sin \theta |\theta_-\rangle,$$

y donde realizamos pruebas para averiguar si los fotones tienen polarización lineal haciendo un ángulo θ con el eje vertical. Conforme a la regla de Born,¹¹ la probabilidad de pasar una prueba exitosamente para el estado $|\theta_+\rangle$ es

$$|\langle V|\theta_+\rangle|^2 = \cos^2 \theta.$$

Ahora imaginemos que otro físico, en otra latitud de la Tierra, realiza exactamente la misma preparación e intenta averiguar si los fotones tienen polarización lineal haciendo un ángulo θ respecto al eje vertical de su laboratorio, i.e, el mismo experimento, pero rotado y trasladado rígidamente, y acaso en otro momento. A pesar de que —respecto a la base antigua— los fotones hayan sido preparados con diferente polarización

$$|V'\rangle = \cos \theta |\theta'_+\rangle + \sin \theta |\theta'_-\rangle,$$

y de que se realicen pruebas para encontrar una polarización diferente $|\theta'_+\rangle$, como sabemos, el resultado debe ser el mismo:

$$|\langle V'|\theta'_+\rangle|^2 = \cos^2 \theta.$$

Por medio de las *probabilidades* sí que es imposible distinguir entre situaciones físicas diferentes. En suma, las leyes de la física no dependen del lugar ni del momento en que suceden los fenómenos, y en teoría cuántica esto se funda en la preservación de las probabilidades de transición.

2.2.1. Teorema de Wigner

Si una cierta transformación U es una simetría del sistema físico, esta preserva las probabilidades de transición. Esto define una simetría en teoría cuántica.

Consideremos una transformación U sobre un espacio de Hilbert \mathcal{H} ,

$$\begin{aligned} U : \mathcal{H} &\rightarrow \mathcal{H} \\ |u\rangle &\mapsto |u'\rangle. \end{aligned}$$

Y eso es todo, no asumiremos nada más acerca de U .

¹¹**Regla de expectación cuántica (regla de Born).** Sean $|u\rangle$ y $|v\rangle$ dos vectores normalizados. La probabilidad de que un sistema cuántico preparado en el estado $|u\rangle$ pasará exitosamente una prueba para el estado $|v\rangle$ es $|\langle u|v\rangle|^2$.

Teorema 2.1¹² Si

$$|\langle u'|v'\rangle|^2 = |\langle u|v\rangle|^2, \quad \forall |u\rangle, |v\rangle \in \mathcal{H},$$

es posible redefinir las fases de los nuevos vectores $|u'\rangle$ y $|v'\rangle$ de tal forma que, para cualquier par de coeficientes complejos α y β , tenemos que la transformación \mathbf{U} es **lineal** y **unitaria**, i.e.

$$\mathbf{U}(\alpha |u\rangle + \beta |v\rangle) = \alpha |u'\rangle + \beta |v'\rangle \quad y \quad \langle u'|v'\rangle = \langle u|v\rangle,$$

o es **antilineal** y **antiunitaria**, i.e.

$$\mathbf{U}(\alpha |u\rangle + \beta |v\rangle) = \alpha^* |u'\rangle + \beta^* |v'\rangle \quad y \quad \langle u'|v'\rangle = \langle v|u\rangle.$$

Este teorema nos dice cómo se representa una simetría en teoría cuántica. Primero, la transformación debe respetar el principio de superposición. Si los estados $|u\rangle$ y $|v\rangle$ se transforman de acuerdo a

$$\begin{aligned} |u\rangle &\rightarrow |u'\rangle, \\ |v\rangle &\rightarrow |v'\rangle, \end{aligned}$$

entonces una superposición de estos dos estados se transforma según

$$\alpha |u\rangle + \beta |v\rangle \rightarrow \alpha |u'\rangle + \beta |v'\rangle,$$

i.e., un estado puro “evoluciona”¹³ en un estado puro. Segundo, debe preservar la normalización de los estados cuánticos y sus productos internos. En palabras de Schumacher, si \mathbf{U} es una simetría del sistema físico, el sistema cuántico es *informacionalmente aislado*, tal como veremos en el Cap. 3.

Ahora sí es posible hablar formalmente de una equivalencia entre una transformación pasiva y una transformación activa, ya tenemos la transformación análoga a aquella transformación ortogonal (clásica) expresada en la Ec. (2.9), y que adivinamos en la sección anterior a partir de argumentos cualitativos y experimentos conceptuales no demasiado elaborados. Volvamos a nuestro montaje de tres experimentos de SG consecutivos (Figura 2.8), con el campo magnético del medio lo suficientemente débil para que haya lugar a interferencia. Si nos ponemos imaginativos y pensamos nuestro montaje en un laboratorio, con nuestro sistema de referencia definido en cualquiera de sus esquinas, una *transformación pasiva* corresponde por ejemplo a rotar y trasladar rígidamente todo el laboratorio con sus equipos y el sistema de referencia. Y una *transformación activa*, a trasladar y rotar rígidamente las partículas cuya dinámica describe una cierta ecuación de movimiento. Ya sabemos que aquel cambio de base —o la labor hérculea de mover el laboratorio con el montaje en su interior— se realizaría a través de una transformación unitaria en \mathbb{R}^3 , i.e., a través de un elemento del *grupo Euclidiano*, que a su vez es la transformación inversa de otro elemento de este mismo grupo que describe la labor de mover los átomos de plata, la transformación activa. Del mismo

¹²La demostración de este teorema se puede encontrar en las páginas 218-219 de [8].

¹³Es importante advertir que en el teorema de Wigner \mathbf{U} se usa de una manera más general para transformaciones unitarias y antiunitarias arbitrarias, no solamente para evolución temporal. Por ejemplo, para transformaciones que pertenecen a un grupo continuo, como las traslaciones y rotaciones; o para transformaciones discretas, como la reflexión espacial o la reversión temporal. Sin embargo, al hablar de *dinámica*, cuando decimos *evolución* nos referimos en particular a evolución temporal.

modo que la transformación ortogonal pasiva expresada en la Ec. (2.10) es la transformación inversa de aquella de la Ec. (2.9). El resultado de nuestra quimera es una situación final y una situación inicial *indistinguishables*: como Galileo al interior del buque de carga, realizando experimentos mecánicos mientras ignora que el navío se desliza con velocidad uniforme sobre el mar, sería imposible saber que el laboratorio se encuentra en otro sitio a través de la mera observación del fenómeno de interferencia de las partículas de espín 1/2 (Figura 2.9).

Nuestro sistema cuántico en este caso es el espín de las partículas, esta es la *variable dinámica* que obedece una ecuación de movimiento. Del mismo modo que para las variables dinámicas clásicas una simetría se representa por medio de una *transformación canónica* (Sec. 1.4.3), en teoría cuántica, para sus variables dinámicas, una simetría se representa mediante una *transformación unitaria*. Resultado después de realizar simultáneamente una transformación activa y una transformación pasiva: una ecuación o un sistema de ecuaciones de movimiento que preservan su forma, las ecuaciones de Hamilton bajo transformaciones canónicas en mecánica clásica y la ecuación de Schrödinger bajo transformaciones unitarias en teoría cuántica.

2.2.2. ¿Dinámica cuántica?

Ocupémonos ahora de la evolución temporal. Ocupémonos ahora de la simetría de traslación temporal, pues, como hemos mencionado, tampoco es posible conocer el día o la hora en que se realiza un experimento simplemente observando el comportamiento de las partículas cuánticas. Sin embargo, en este caso no es tan simple. El tiempo no es una variable dinámica, así que no hay razón para exigir que la traslación temporal sea representada por medio de una transformación unitaria.

Del mismo modo que en mecánica clásica no hay razón para asumir que existe un Hamiltoniano $H(q, p)$ tal que las ecuaciones de Hamilton

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

sean invariantes bajo traslación temporal, en teoría cuántica tampoco hay razón para asumir que existe un Hamiltoniano H tal que la ecuación de Schrödinger

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle$$

sea invariante bajo esta traslación. Por ejemplo, en el **ejercicio 8.23** de [8] se pide mostrar que en el caso de un *oscilador armónico amortiguado* es posible escribir ecuaciones de movimiento —diferentes a las ecuaciones de Hamilton— que son invariantes bajo traslación temporal. O sea, existen casos en que pueden formularse otras leyes de movimiento. Sin embargo, como discutimos en el Cap 1, esto se debe a que la fuerza de amortiguación $-\gamma p$ no es fundamental, es el resultado fenomenológico de ignorar un sinnúmero de grados de libertad. Al igual que en mecánica clásica, en teoría cuántica se asume que las leyes *fundamentales* surgen al incluir en el Hamiltoniano (o en el Lagrangiano, de ser posible) todos aquellos grados de libertad, y del mismo modo que clásicamente una traslación temporal es representada por una transformación canónica, en teoría cuántica esta se representa por medio de un operador

unitario. Es decir, se asume que la evolución es unitaria por simple analogía con la mecánica clásica. Cabe preguntarse: ¿qué hubiese sido de la teoría cuántica sin la mecánica clásica? ¿Es posible responder a esto, y así repensar los fundamentos de la evolución temporal en teoría cuántica?

Evolución unitaria

Se asume pues que

$$|\psi(t_0)\rangle \rightarrow |\psi(t)\rangle,$$

donde

$$|\psi(t)\rangle = \mathbf{U}(t, t_0) |\psi(t_0)\rangle \quad (2.12)$$

y $t > t_0$. Además, que \mathbf{U} satisface la propiedad de grupo

$$\mathbf{U}(t_3, t_1) = \mathbf{U}(t_3, t_2)\mathbf{U}(t_2, t_1),$$

donde $t_3 > t_2 > t_1$. Según se asume en mecánica clásica, a nivel fundamental no existen fuerzas no conservativas. Análogamente, en teoría cuántica, la evolución temporal se asume unitaria, o equivalentemente, como lo establece el teorema de Wigner, esta transformación respeta el principio de superposición y preserva la normalización de los estados cuánticos y sus productos internos. Hechos empíricos sugieren la siguiente regla [8]:

Determinismo cuántico. *En un entorno perfectamente reproducible, un estado puro evoluciona en un estado puro.*

Estamos pues ante un escenario altamente idealizado en que el entorno es perfectamente reproducible, un “mundo sin fricción”. Y de nuevo, verificar experimentalmente el determinismo cuántico, es verificar si efectivamente los objetos del mundo son realizaciones satisfactorias de estos objetos abstractos de la teoría.

Pero ¿por qué *determinismo*? Que un estado puro evolucione en un estado puro, significa, según el postulado de *determinismo estadístico*, que así como existe una prueba maximal en t_0 que permite descubrir con toda certeza la preparación inicial que da lugar a $|\psi(t_0)\rangle$, también existe —en general— otra prueba maximal que permite descubrir con toda certeza el estado $|\psi(t)\rangle$ que resulta de la evolución. Además, para cualquier otra prueba maximal en t , únicamente podemos predecir las probabilidades de los diferentes resultados.

El Hamiltoniano y la ecuación de Schrödinger

Ahora nuestro propósito es establecer formalmente cómo cambia $|\psi(t)\rangle$ en un cierto instante. Asumiendo —por analogía con la mecánica clásica— que la traslación temporal se representa por medio de una transformación unitaria, se define el Hamiltoniano, como el operador autoadjunto

$$\mathbf{H} := i\hbar \frac{d\mathbf{U}(t, t_0)}{dt} \mathbf{U}^\dagger(t, t_0) = -i\hbar \mathbf{U}(t, t_0) \frac{d\mathbf{U}^\dagger(t, t_0)}{dt}.$$

Claramente, a partir de esta ecuación, se tiene que $i\hbar\dot{U} = HU$ (o $-i\hbar\dot{U}^\dagger = U^\dagger H$). Derivando respecto al tiempo la Ec. (2.12) y usando este último resultado se deriva la ecuación de Schrödinger:

$$\begin{aligned}\frac{d|\psi(t)\rangle}{dt} &= \left(\frac{dU(t, t_0)}{dt}\right) |\psi(t_0)\rangle \\ &= \frac{1}{i\hbar} HU(t, t_0) |\psi(t_0)\rangle,\end{aligned}$$

así que

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H |\psi(t)\rangle. \quad (2.13)$$

Vale la pena aclarar que en general H depende del tiempo t y no del tiempo inicial t_0 :

$$\begin{aligned}\frac{dU(t, t_0)}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{U(t + \Delta t, t_0) - U(t, t_0)}{\Delta t} \\ &= \frac{1}{i\hbar} HU(t, t_0),\end{aligned}$$

así que, usando la propiedad de grupo que satisface U ,

$$\begin{aligned}H(t) &= i\hbar \lim_{\Delta t \rightarrow 0} \frac{U(t + \Delta t, t_0)U^\dagger(t, t_0) - U(t, t_0)U^\dagger(t, t_0)}{\Delta t} \\ H(t) &= i\hbar \lim_{\Delta t \rightarrow 0} \frac{U(t + \Delta t, t) - \mathbb{1}}{\Delta t}\end{aligned}$$

En suma, la ecuación de Schrödinger describe una situación ideal en que un estado puro evoluciona en un estado puro, sin embargo, como dice el epígrafe, los fenómenos cuánticos no suceden en un espacio de Hilbert, suceden en un laboratorio. En nuestro análisis de los fenómenos nos hace falta un ingrediente fundamental, el proceso de *medición*, que como sabemos, no es un proceso pasivo, y como discutiremos en el Cap. 3, puede dar lugar a que estados puros evolucionen en estados mezclados.

2.2.3. Simetría y leyes de conservación

En mecánica clásica el teorema de Noether establece que toda simetría conduce a una cantidad conservada. En teoría cuántica sucede algo similar, sin embargo, las leyes de conservación no solo están profundamente relacionadas con la simetría de los sistemas físicos bajos ciertos cambios, sino también con el *principio de superposición*, pues, como establece el teorema de Wigner, una simetría se representa por medio de un operador unitario. El propósito de esta última subsección es discutir al respecto.

Usualmente en los textos de mecánica cuántica se responde de otra manera —equivalente— a la pregunta que motivó esta sección: ¿qué define una simetría en teoría cuántica? Allí lo indispensable es asumir —“así sin más”— que la evolución temporal es unitaria. Además,

si se asume que la física es simétrica bajo una cierta transformación Q , se arriba a una expresión simple que define una simetría en teoría cuántica. Veamos. Si el estado $|\psi_2\rangle$ es el resultado de la evolución temporal del estado $|\psi_1\rangle$, i.e.,

$$|\psi_2\rangle = U |\psi_1\rangle,$$

luego de realizar sobre todo el sistema la transformación representada por Q , la evolución temporal debería preservarse, es decir,

$$|\psi'_2\rangle = U |\psi'_1\rangle,$$

donde $|\psi'_i\rangle = Q |\psi_i\rangle$ ($i = 1, 2$). Luego, reescribiendo la ecuación anterior en términos de la operación de simetría Q y de la preparación inicial representada por $|\psi_1\rangle$, se tiene que

$$\begin{aligned} Q |\psi_2\rangle &= UQ |\psi_1\rangle \\ QU |\psi_1\rangle &= UQ |\psi_1\rangle \end{aligned}$$

Debido a que esta relación se mantiene para cualquier estado inicial $|\psi_1\rangle$, se concluye que si la física es simétrica bajo la transformación Q , esta debe conmutar con la transformación de evolución temporal:

$$QU = UQ.$$

Y si el sistema no está sujeto a fuerzas externas dependientes del tiempo, es decir, si la transformación de evolución temporal está dada por

$$U(t, t_0) = e^{-iH(t-t_0)/\hbar},$$

y si se supone una transformación infinitesimal

$$U(t, t_0) = 1 - iH\Delta t/\hbar,$$

Q también debe conmutar con el Hamiltoniano del sistema:

$$QH = HQ.$$

Esto define una simetría en muchos libros de mecánica cuántica. Y aunque es equivalente a lo que se propuso en esta sección, este enfoque dista mucho de ser instrumentalista.

Leyes de conservación

La idea en la que se fundan las leyes de conservación en teoría cuántica es sencilla: si la operación de Q sobre un estado supone simplemente un cambio de fase, esto es verdad por siempre [7]. Supongamos que

$$U(t, t_0) |\psi_1\rangle = |\psi_2\rangle. \quad (2.14)$$

El enunciado anterior significa que si

$$Q |\psi_1\rangle = e^{i\delta} |\psi_1\rangle, \quad (2.15)$$

entonces también es cierto que

$$Q |\psi_2\rangle = e^{i\delta} |\psi_2\rangle. \quad (2.16)$$

¿Y bajo qué condiciones se cumple esto? Siempre y cuando Q conmute con U , justamente:

$$\begin{aligned} Q|\psi_2\rangle &= QU|\psi_1\rangle \\ &= UQ|\psi_1\rangle \\ &= Ue^{i\delta}|\psi_1\rangle \\ &= e^{i\delta}U|\psi_1\rangle \\ &= e^{i\delta}|\psi_2\rangle. \end{aligned}$$

En palabras de Feynman: «Incluso si ignoramos los detalles de la maquinaria por medio de la cual un sistema pasa de un estado a otro, podemos aún decir que si una cosa está originalmente en un estado con un cierto carácter de simetría, y si el Hamiltoniano para esta cosa es simétrico bajo esa operación de simetría, entonces el estado tendrá el mismo carácter de simetría en todo momento. Esa es la base de todas las leyes de conservación de la mecánica cuántica.»

Ilustremos esto a través de un elemento del grupo Euclidiano; en particular, a través de una rotación rígida alrededor del eje z . Consideraremos una situación en que ningún *parámetro externo* proporcione información sobre el tipo de transformación realizada sobre el sistema cuántico. Por ejemplo, si el sistema es una partícula sometida a un campo magnético o a un campo eléctrico en dirección z , una rotación alrededor de este eje no cambiará en nada las cosas. Muy bien, supongamos que la transformación unitaria $R_z(\phi)$, que representa una rotación de un ángulo ϕ alrededor del eje z , es tal que al actuar sobre un cierto estado simplemente cambia su fase:

$$R_z(\phi)|\psi_0\rangle = e^{i\delta(\phi)}|\psi_0\rangle.$$

Si consideramos rotaciones infinitesimales ε , es cierto que cualquier ángulo ϕ es la suma de una cantidad n adecuada de rotaciones infinitesimales: $\phi = n\varepsilon$. O sea,

$$R_z(\phi) = [R_z(\varepsilon)]^n.$$

Bajo estas circunstancias, necesariamente $\delta(\phi)$ debe que ser lineal en ϕ , i.e.,

$$\delta(\sum_i a_i \phi_i) = \sum_i a_i \delta(\phi_i) \quad (a_i \in \mathbb{R}),$$

pues de lo contrario no se satisfaría que $n\delta(\varepsilon) = \delta(n\varepsilon)$:

$$\begin{aligned} R_z(\phi)|\psi_0\rangle &= [R_z(\varepsilon)]^n|\psi_0\rangle = R_z(n\varepsilon)|\psi_0\rangle \\ e^{in\delta(\varepsilon)}|\psi_0\rangle &= e^{i\delta(n\varepsilon)}|\psi_0\rangle. \end{aligned}$$

De acuerdo, entonces

$$R_z(\phi)|\psi_0\rangle = e^{im\phi}|\psi_0\rangle \quad (m \in \mathbb{R}).$$

Por otro lado, podemos también escribir que

$$R_z(\phi) = e^{i\hbar J_z \phi / \hbar},$$

donde el operador J_z es la componente z del momento angular del sistema y es el generador de la transformación continua $R_z(\phi)$. Si consideramos rotaciones infinitesimales y comparamos las ecuaciones

$$R_z(\varepsilon)|\psi\rangle = (\mathbf{1} + i\hbar J_z \varepsilon / \hbar)|\psi\rangle$$

y

$$\mathbf{R}_z(\varepsilon) |\psi\rangle = e^{im\varepsilon} |\psi\rangle = (1 + im\varepsilon) |\psi\rangle,$$

encontramos que

$$\mathbf{J}_z |\psi\rangle = m\hbar |\psi\rangle.$$

Finalmente, según la idea que se enunció al principio, expresada matemáticamente en las ecuaciones (2.14), (2.15) y (2.16), si \mathbf{R}_z es una simetría del sistema, entonces la componente z de su momento angular \mathbf{J}_z se conserva.

Capítulo 3

¿ASIMETRÍA TEMPORAL EN TEORÍA CUÁNTICA?

Babuc hizo fundir, por el mejor fundidor de la ciudad, una estatuilla compuesta por todos los metales, tierras y piedras más preciosas y más viles, y la llevó a Ituriel —una de las deidades de rango más elevado—, a quien dijo: —¿Vais a destruir esta hermosa estatua porque no está hecha exclusivamente de oro y diamantes?

Ituriel entendió el significado de la pregunta y decidió no pensar más en querer mejorar Persépolis, y limitarse a dejar marchar *el mundo tal como va...*

Voltaire

3.1. Teorema de aislamiento¹

Hemos hablado de *determinismo cuántico* en el capítulo anterior, sin embargo, no precisamos qué es un *entorno perfectamente reproducible*. Para entenderlo, expresemos formalmente qué es un sistema informacionalmente aislado y mostremos su profunda relación con el concepto de simetría en teoría cuántica. De acuerdo con el principio de interferencia, la regla de adición de probabilidades, válida para la ocurrencia de dos eventos A y B ,

$$P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\},$$

no siempre aplica para las probabilidades cuánticas, pues en algunos casos hablamos de caminos indeterminados, y ¿cómo podría ser el paso por un camino indeterminado la ocurrencia de un evento? En otras ocasiones sin embargo es posible determinar los caminos seguidos por el sistema, incluso sin perturbar su dinámica, de modo que aquella regla recupera su validez. ¿Cuál es la diferencia? La *información*. informalmente se dice que la información es la capacidad de distinguir con seguridad entre posibles alternativas [10], conocer con certeza el camino seguido por el sistema cuántico a partir de una huella, un vestigio que el sistema deje en algún lugar del universo, en el laboratorio o acaso en otro sistema cuántico.

¹Gran parte de la discusión está basada en la Sec. 9.2 de [10].

Hablamos de la posible interacción entre el sistema cuántico y su entorno. Así las cosas, debemos tratar con *sistemas abiertos*.

Denominemos Q nuestro sistema y E el sistema externo. Supongamos que los sistemas Q y E inicialmente no tienen correlaciones entre sí, además, que Q inicialmente está en el estado $|\phi\rangle$ y que E está en el estado puro fijo $|0\rangle$, i.e., se suponen condiciones externas fijas para la evolución de Q.² Una vez el sistema evoluciona por medio del operador unitario U sobre $\mathcal{H}^{(QE)}$, puede haber lugar a interacciones entre Q y E. Se dice que el sistema Q es un sistema abierto y que E es su entorno.

¿Qué significa pues que un sistema sea informacionalmente aislado? Heurísticamente, que el sistema Q no deje huella en su entorno, o sea, que no transfiera información a E. Formalmente, que el estado final del entorno no dependa del estado inicial del sistema cuántico, i.e., que

$$\rho^{(E)} = \text{Tr}_{(Q)} U(|\phi, 0\rangle \langle\phi, 0|) U^\dagger \quad (3.1)$$

sea el mismo para todo $|\phi\rangle \in \mathcal{H}^{(Q)}$.

Todo esto se asemeja demasiado a lo que ya hemos discutido sobre simetrías. Recordemos que una simetría es una equivalencia entre situaciones físicas diferentes. En teoría cuántica que una cierta transformación sea una simetría implica que se preserven las probabilidades de transición; e.g., son simetrías la rotación o la traslación (rígidas) en el espacio de un cierto sistema óptico, pues las leyes de la óptica no pueden cambiar bajo este tipo de transformaciones, i.e., las probabilidades de los diferentes resultados de una prueba maximal se preservan. Y ¿cómo se relaciona esto con nuestra definición de sistema informacionalmente aislado? Veamos. La demostración del *teorema de aislamiento* nos ayuda a responder esta pregunta.

Teorema 3.1 *Supóngase que el sistema compuesto QE evoluciona de acuerdo al operador unitario U y que E está inicialmente en algún estado puro $|0\rangle$. La evolución del subsistema Q es unitaria si y solo si Q es informacionalmente aislado.*

DEMOSTRACIÓN. Supongamos que la evolución del sistema Q es unitaria, i.e., que el estado final del sistema compuesto QE, resultado de la evolución unitaria de $|\phi, 0\rangle$,

$$U|\phi, 0\rangle = V|\phi\rangle \otimes |e_\phi\rangle,$$

es una purificación de $V|\phi\rangle$,³ donde V es una transformación unitaria sobre $\mathcal{H}^{(Q)}$. Debe probarse que efectivamente $|e_\phi\rangle$ no depende de $|\phi\rangle$. Para esto, consideremos elementos diferentes

²Por supuesto que el entorno podría representarse por medio de un estado mezclado, sin embargo, por simplicidad, se supone un estado puro, pues de todos modos aquel estado mezclado para E podría obtenerse a partir de otro estado puro entrelazado que represente un sistema más grande E'.

³Sean $|\Psi\rangle \in \mathcal{H}^{(QE)}$ y $\rho^{(Q)} \in \mathcal{H}^{(Q)}$. Se dice que $|\Psi\rangle$ es una purificación de $\rho^{(Q)}$ en E si

$$\text{Tr}_{(E)} |\Psi\rangle \langle\Psi| = \rho^{(Q)}.$$

En nuestro caso es claro que

$$\text{Tr}_{(E)} V|\phi\rangle \otimes |e_\phi\rangle \langle\phi| V^\dagger \otimes \langle e_\phi| = V|\phi\rangle \langle\phi| V^\dagger.$$

de $\mathcal{H}^{(Q)}$ y veamos que $|e_\phi\rangle$ no cambia. Sean $|\phi\rangle$ y $|\phi'\rangle$ estados diferentes de Q , de modo que

$$\begin{aligned} \mathbf{U}|\phi, 0\rangle &= \mathbf{V}|\phi\rangle \otimes |e_\phi\rangle, \\ \mathbf{U}|\phi', 0\rangle &= \mathbf{V}|\phi'\rangle \otimes |e_{\phi'}\rangle \end{aligned}$$

y su producto interno es

$$\begin{aligned} \langle\phi, 0|\mathbf{U}^\dagger\mathbf{U}|\phi', 0\rangle &= \langle\phi|\mathbf{V}^\dagger\mathbf{V}|\phi'\rangle \langle e_\phi|e_{\phi'}\rangle \\ \langle\phi|\phi'\rangle &= \langle\phi|\phi'\rangle \langle e_\phi|e_{\phi'}\rangle. \end{aligned}$$

Si $|\phi\rangle$ y $|\phi'\rangle$ no son ortogonales, entonces $|e_{\phi'}\rangle = |e_\phi\rangle$.

Si $\langle\phi|\phi'\rangle = 0$, también se cumple lo anterior. Consideremos otro estado $|\phi''\rangle$ de Q que no sea ortogonal a $|\phi\rangle$ ni a $|\phi'\rangle$, así que

$$\begin{aligned} \langle\phi'', 0|\mathbf{U}^\dagger\mathbf{U}|\phi', 0\rangle &= \langle\phi''|\mathbf{V}^\dagger\mathbf{V}|\phi'\rangle \langle e_{\phi''}|e_{\phi'}\rangle \\ \langle\phi''|\phi'\rangle &= \langle\phi''|\phi'\rangle \langle e_{\phi''}|e_{\phi'}\rangle \end{aligned}$$

y

$$\begin{aligned} \langle\phi'', 0|\mathbf{U}^\dagger\mathbf{U}|\phi, 0\rangle &= \langle\phi''|\mathbf{V}^\dagger\mathbf{V}|\phi\rangle \langle e_{\phi''}|e_\phi\rangle \\ \langle\phi''|\phi\rangle &= \langle\phi''|\phi\rangle \langle e_{\phi''}|e_\phi\rangle, \end{aligned}$$

por tanto $|e_{\phi'}\rangle = |e_\phi\rangle$.

Probemos la otra la implicación, la que más nos interesa. Así que supongamos que Q es un sistema informacionalmente aislado. Esto significa que la base de autoestados $\{|e_k\rangle\}$ y el espectro de valores propios $\{\lambda_k\}$ de $\rho^{(E)}$ (Ec. (3.1)) no dependen de $|\phi\rangle$. Por otro lado, de acuerdo con la descomposición de Schmidt es posible expandir cualquier estado del sistema compuesto QE , entrelazado o no, en términos de la base de autoestados $\{|e_k\rangle\}$ de $\rho^{(E)}$ y de la base de autoestados $\{|q_k\rangle\}$ de $\rho^{(Q)}$, de modo que nuestro estado final $|\Psi\rangle = \mathbf{U}|\phi, 0\rangle$ puede expresarse como

$$|\Psi\rangle = \sum_k \sqrt{\lambda_k} |q_k, e_k\rangle. \quad (3.2)$$

Según nuestra hipótesis, el hecho de que $|\Psi\rangle$ dependa de $|\phi\rangle$ se debe única y exclusivamente a los $|q_k\rangle$. Y ciertamente la dependencia de los $|q_k\rangle$ puede verse a partir de los operadores \mathbf{V}_k sobre $\mathcal{H}^{(Q)}$, que difieren de los operadores de Kraus solo por un factor que asegura la ortonormalidad de los elementos de la base $\{|q_k\rangle\}$:

$$\mathbf{V}_k|\phi\rangle = \frac{1}{\sqrt{\lambda_k}} \langle e_k|\mathbf{U}|\phi, 0\rangle = \frac{1}{\sqrt{\lambda_k}} \langle e_k|\Psi\rangle = |q_k\rangle.$$

Como Q es informacionalmente aislado, se asegura que los operadores \mathbf{V}_k preserven el principio de superposición. Por otra parte, dado que los $|q_k\rangle$ están normalizados, se tiene que los operadores \mathbf{V}_k son unitarios:

$$\langle\phi|\mathbf{V}_k^\dagger\mathbf{V}_k|\phi\rangle = \langle q_k|q_k\rangle = 1 \quad \forall |\phi\rangle \in \mathcal{H}^{(Q)}.$$

Sin embargo,

$$\langle \phi | \mathbf{V}_k^\dagger \mathbf{V}_l | \phi \rangle = \langle q_k | q_l \rangle = 0 \quad \forall |\phi\rangle \in \mathcal{H}^{(Q)},$$

así que $\mathbf{V}_k^\dagger \mathbf{V}_l = 0$, por lo tanto, para garantizar la unitariedad de \mathbf{V}_k , k no puede ser diferente de l . Lo que nos permite concluir que hay un único \mathbf{V}_k , i.e., el único autovalor de $\rho^{(E)}$ es $\lambda = 1$, ya que $\{\lambda_k\}$ es una distribución de probabilidad. Entonces, de acuerdo con la Ec. (3.2), $|\Psi\rangle = |q, e\rangle$ y

$$\mathbf{U} |\phi, 0\rangle = |\Psi\rangle = |q, e\rangle = (\mathbf{V} |\phi\rangle) \otimes |e\rangle,$$

es decir, cada subsistema evoluciona independientemente, el sistema Q no interactúa con su entorno E, y su evolución —por medio de \mathbf{V} — respeta el principio de superposición y es unitaria.

□

¿Qué nos recuerda esto? Sabemos que la transformación

$$\begin{aligned} \mathbf{V} : \mathcal{H}^{(Q)} &\rightarrow \mathcal{H}^{(Q)} \\ |\phi\rangle &\mapsto |q\rangle, \end{aligned}$$

respeto el principio de superposición y es unitaria. Asimismo, en general, podemos tener un estado mezclado para Q representado por $\rho^{(Q)}$ y su evolución unitaria dada por

$$\rho'^{(Q)} \rightarrow \mathbf{V} \rho^{(Q)} \mathbf{V}^\dagger.$$

De modo que la probabilidad de que un sistema cuántico preparado en $\rho^{(Q)}$ pase exitosamente una prueba para el estado $|\psi\rangle$ es

$$\text{Tr } \rho^{(Q)} |\psi\rangle \langle \psi|.$$

Si suponemos que

$$\mathbf{V} |\psi\rangle = |p\rangle,$$

se tiene que

$$|p\rangle \langle p| \rightarrow \mathbf{V} |\psi\rangle \langle \psi| \mathbf{V}^\dagger,$$

así que la probabilidad de que un sistema preparado en $\rho'^{(Q)}$ pase exitosamente una prueba para el estado $|p\rangle$ es

$$\begin{aligned} \text{Tr } \rho'^{(Q)} |p\rangle \langle p| &= \text{Tr } \mathbf{V} \rho^{(Q)} \mathbf{V}^\dagger \mathbf{V} |\psi\rangle \langle \psi| \mathbf{V}^\dagger \\ &= \text{Tr } \rho^{(Q)} |\psi\rangle \langle \psi|, \end{aligned}$$

i.e., se preservan las probabilidades de transición. En particular, si $\rho'^{(Q)} = |\phi\rangle \langle \phi|$

$$|\langle \phi | \psi \rangle|^2 = |\langle q | p \rangle|^2, \quad \forall |\phi\rangle, |\psi\rangle \in \mathcal{H}^{(Q)},$$

¡¡justamente la hipótesis del teorema de Wigner! Y con esto respondemos a la pregunta sobre la relación entre simetría e información. Recapitemos. Si un sistema es informacionalmente aislado, este y su entorno no interactúan durante la evolución del sistema compuesto, es decir, están *dinámicamente aislados* o *desacoplados*, de modo que esta evolución está representada por

$$\mathbf{U} = \mathbf{U}^{(Q)} \otimes \mathbf{U}^{(E)}.$$

O equivalentemente, si el sistema compuesto QE está inicialmente en un estado producto, este permanecerá en un estado producto después de la evolución. Así pues, que un sistema cuántico Q sea informacionalmente aislado equivale a que —localmente— $\mathbf{U}^{(Q)} = \mathbf{V}$ sea una simetría del sistema, i.e., se preserven las probabilidades de transición.

3.2. Entropía de un sistema informacionalmente aislado

Si hablamos de información, necesariamente tenemos que hablar de *entropía*. Es momento de servirnos del concepto de entropía de una preparación para entender la relación entre simetría e información en un sistema informacionalmente aislado.

3.2.1. Entropía de una preparación

Lo que naturalmente puede ser ignorancia, bien podría considerarse como la ocasión de ganar información. Como sabemos, previo a una prueba, únicamente podemos pronunciarnos acerca de las *probabilidades* de los distintos resultados, predecirlas. La aleatoriedad se torna fundamental. Sin embargo, una vez realizada una prueba, uno de los posibles resultados se materializa, y lo que antes fue ignorancia, se convierte en certeza, en información. Supongamos que p_1, \dots, p_N son las probabilidades (conocidas) de los N posibles resultados de una prueba, i.e., después de haber realizado muchísimas veces —exactamente— la misma preparación y la misma prueba, la probabilidad de obtener el resultado j es

$$p_j \xrightarrow{n \rightarrow \infty} \frac{n_j}{n}.$$

En otras palabras, si el mismo experimento se realiza n veces, con n muy grande, el resultado j se presentará aproximadamente $n_j = np_j$ veces. Como en el ejemplo de nuestro pintoresco aparato de detección que imprime signos $+$ y $-$, sabemos que luego de repetir muchísimas veces el mismo experimento las frecuencias relativas se aproximarán a las probabilidades p_j ; y eso es todo. Podemos predecir las probabilidades, pero nos resulta imposible predecir el siguiente tipeo, nos es imposible predecir el orden de la secuencia aleatoria de n signos, cuál de sus

$$\frac{n!}{n_1!n_2! \cdots n_N!}$$

posibles disposiciones diferentes lograremos al final del experimento. Pero entre tantas posibilidades, una vez realizada la prueba, ¡una se materializa! Usando la aproximación de Stirling se tiene que

$$\log \frac{n!}{n_1!n_2! \cdots n_N!} \simeq n \log n - n - \sum_j (n_j \log n_j - n_j) = -n \sum_j p_j \log p_j.$$

Se define pues la *entropía* de la distribución de probabilidad $\{p_j\}$ como

$$S := - \sum_{j=1}^N p_j \log p_j, \quad (3.3)$$

i.e., una medida de la cantidad promedio de información lograda al materializarse un resultado; o una medida de nuestra ignorancia antes de la prueba, como se prefiera.

Sin embargo, una distribución de probabilidad de un estado depende tanto de la preparación seguida, como de la prueba realizada. Lo mismo la entropía. Veámoslo con un ejemplo.⁴ Fotones son preparados por un proceso que tiene el 70 % de probabilidad de producir polarización circular derecha y un 30 % de probabilidad de producirlos linealmente polarizados en la dirección x (horizontal). ¿Cuál es la entropía de estos fotones en una prueba de polarización circular? El operador

$$\rho = 0,7 |RHC\rangle \langle RHC| + 0,3 |H\rangle \langle H| \quad (3.4)$$

especifica completamente todas las propiedades del ensamble. La representación matricial de ρ depende de la base que se utilice, i.e, de la prueba implementada.⁵ Si medimos polarización circular:

$$\rho = \begin{pmatrix} 0,85 & -0,15i \\ 0,15i & 0,15 \end{pmatrix},$$

así que, según la Ec. (3.3), la entropía de los fotones para esta prueba es $S \approx 0,42271$. ¿Y en una prueba para polarización lineal? En este caso

$$\rho = \begin{pmatrix} 0,65 & -0,35i \\ 0,35i & 0,35 \end{pmatrix},$$

de modo que $S \approx 0,64745$. Como era de esperarse, obtenemos más información acerca de la preparación del estado a través de la prueba de polarización circular. De hecho, *información parcial*, pues existen extremos de ignorancia.

Según el postulado de *mezclas aleatorias*, es posible preparar un sistema cuántico con N estados de tal forma que *toda* prueba maximal tenga la misma probabilidad N^{-1} .⁶ De este modo no solo la ignorancia se conserva, ¡es máxima para *toda* prueba maximal! Veamos:

$$S = - \sum_{j=1}^{N-1} p_j \log p_j - p_N \log p_N,$$

donde

$$p_N \equiv 1 - \sum_{j=1}^{N-1} p_j,$$

y además, todas las otras p_j son consideradas variables independientes, de modo que

$$\frac{\partial S}{\partial p_k} = - \log p_k + \log p_N.$$

Así, el único extremo de S se da cuando $p_k = p_N = 1/N$. Se tiene entonces para una mezcla aleatoria que

$$\rho_{mn} = \frac{\delta_{mn}}{N}$$

⁴**Ejercicio 9.1** de [8], pág. 261.

⁵La relación entre este par de pruebas maximales está dada por

$$\begin{aligned} |RHC\rangle &= \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle) \\ |LHC\rangle &= \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle) \end{aligned}.$$

⁶Ver Cap. 2, páginas 31 y 32 de [8].

Por el contrario, según el postulado de *determinismo estadístico*, la máxima cantidad de información que se puede proveer acerca de la preparación de un sistema cuántico corresponde a la de un estado puro.

¿En qué sentido podemos pues hablar de la entropía de una preparación, si esta depende de la prueba que se realice? ¿Por medio de cuál distribución de probabilidad $\{p_k\}$ deberíamos caracterizar una cierta preparación? Por medio de la más especial, la que —para aquella preparación— minimiza S en (3.3). Se define la *entropía de una preparación* como el menor valor que puede ser alcanzado por la Ec. (3.3) para cualquier prueba maximal realizada después de esa preparación [8]. Y la prueba que minimiza la entropía de Shannon para una cierta preparación es justamente la que corresponde a la base ortonormal formada por los autovectores $|v_\mu\rangle$ de su matriz densidad ρ ,⁷ i.e.,

$$S(\rho) = - \sum_{\mu} w_{\mu} \log w_{\mu} = -\text{Tr} (\rho \log \rho), \quad (3.5)$$

donde

$$\rho |v_{\mu}\rangle = w_{\mu} |v_{\mu}\rangle,$$

y los autovalores $\{w_{\mu}\}$ forman una distribución de probabilidad.

Por ejemplo, en la base en que la matriz de densidad de (3.4) es diagonal se tiene que

$$\rho = \begin{pmatrix} \frac{-\sqrt{58+10}}{20} & 0 \\ 0 & \frac{\sqrt{58+10}}{20} \end{pmatrix}.$$

Para aquella preparación, de (3.5), la mínima entropía es aproximadamente 0.36535.

3.2.2. Un mundo sin fricción

Volvamos al teorema de aislamiento; pero esta vez supongamos que inicialmente el sistema compuesto QE está en un estado producto

$$\rho \otimes |0\rangle \langle 0|,$$

de modo que, en general, el subsistema Q inicia en un estado mezclado. Sabemos que si el sistema es informacionalmente aislado, la evolución de cada subsistema es independiente de la del otro y que el sistema compuesto permanecerá en un estado producto después de la evolución. Para el sistema Q

$$\rho \rightarrow \rho' = V\rho V^{\dagger},$$

para el sistema E

$$|0\rangle \langle 0| \rightarrow U^{(E)} |0\rangle \langle 0| U^{(E)\dagger} = |e\rangle \langle e|$$

y para el sistema compuesto

$$\rho \otimes |0\rangle \langle 0| \rightarrow \rho' \otimes |e\rangle \langle e|.$$

⁷Una prueba de esto se encuentra en las páginas 262-264 de [8].

Muy bien, ahora calculemos la entropía del sistema compuesto, antes y después de la evolución. En su descomposición espectral

$$\rho = \sum_m w_m |v_m\rangle \langle v_m|,$$

donde su espectro de autovalores $\{w_m\}$ es una distribución de probabilidad. Así pues, la entropía inicial del sistema QE está dada por

$$S(\rho \otimes |0\rangle \langle 0|) = - \sum W_{m\mu} \log W_{m\mu},$$

donde w_m , ω_μ y $W_{m\mu} = w_m \omega_\mu$ son autovalores de ρ , $|0\rangle \langle 0|$ y $\rho \otimes |0\rangle \langle 0|$, respectivamente. Ya que hay un único ω_μ y es igual a uno, se tiene que

$$S(\rho \otimes |0\rangle \langle 0|) = S(\rho).$$

Para la entropía final de QE, basta saber que la entropía $S = -\text{Tr}(\rho \log \rho)$ se mantiene invariante bajo la evolución unitaria $\rho \rightarrow \mathbf{V}\rho\mathbf{V}^\dagger$, i.e.,

$$\begin{aligned} \rho |v_m\rangle &= w_m |v_m\rangle \\ \rho \mathbf{V}^\dagger |v_m\rangle &= w_m |v_m\rangle \\ \mathbf{V}\rho\mathbf{V}^\dagger |v'_m\rangle &= w_m |v'_m\rangle \\ \rho' |v'_m\rangle &= w_m |v'_m\rangle. \end{aligned}$$

Por lo tanto

$$S(\rho') = S(\rho)$$

y

$$S(\rho' \otimes |e\rangle \langle e|) = S(\rho \otimes |0\rangle \langle 0|).$$

La otra implicación también es cierta, de modo que se propone el siguiente teorema:

Teorema 3.2 *Supóngase que el sistema compuesto QE evoluciona de acuerdo al operador unitario \mathbf{U} y que \mathbf{E} está inicialmente en algún estado puro $|0\rangle$. El sistema Q es informacionalmente aislado si y solo si su entropía no cambia después de la evolución.*

Este resultado se torna especialmente interesante cuando hablamos de evolución temporal. Como sabemos, el concepto de entropía surgió originalmente en la termodinámica clásica y luego fue introducido en la teoría de la información por Shannon. Sin embargo, debido a su estrecha relación, es interesante que la *asimetría* entre pasado y futuro, intrínseca en el procesamiento de información, es equivalente a la flecha de tiempo que surge en los fenómenos termodinámicos irreversibles.⁸ Situado en este contexto, el teorema anterior *sugiere* que un sistema es informacionalmente aislado si y solo si da lo mismo si el tiempo fluye en un sentido o en otro: un “*mundo sin fricción*” en que la entropía es constante, la de los sistemas y la del universo. Pero faltan ingredientes, la *evolución temporal generalizada* para sistemas abiertos y el análisis del proceso de *medición*. Además, ¿porqué se asume unitaria la evolución temporal en teoría cuántica? ¿Qué es una medición? Estas son a su vez preguntas

⁸El Cap. 9 de [8] está dedicado en gran medida a desarrollar esta idea.

esenciales que en el Teorema 3.2 se soslayan. En lo que resta de este capítulo nos dedicaremos justamente a sazonar este teorema y a examinar estas preguntas. Como discutiremos, la interacción entre el sistema y su entorno da lugar a entrelazamiento. Asimismo, la medición no es un proceso pasivo de adquisición de conocimiento, es un proceso activo. Dependiendo de su definición, puede hablarse de un proceso activo en que se hace uso de operaciones cuánticas con una evolución de estado no determinística, o de uno, según A. Peres, donde se usan dispositivos extremadamente complejos que usualmente involucran mecanismos de amplificación *irreversibles*.

3.3. Dinámica y mediciones generalizadas: ¿la simetría temporal es fundamental?⁹

En el cierre de la sección anterior podríamos agregar que este corolario del teorema de aislamiento describe un “*mundo sin fricción*” en que la ecuación de Schrödinger conserva su validez, un mundo en que estados puros evolucionan en estados puros y estados mezclados evolucionan en estados mezaclados de manera *determinística*:

$$\rho \rightarrow \rho' = U\rho U^\dagger,$$

donde, como mostramos, la entropía se conserva. Y aunque entendemos el valor de las teorías sin fricción en un mundo con fricción como el nuestro (según se discutió en la parte final del Cap 1), ¿por qué no intentar una teoría con fricción, una teoría de evolución temporal *no* unitaria en que —en general— estados puros evolucionen en mezclas, en que el entorno no sea *perfectamente reproducible* y únicamente lo conozcamos estadísticamente, algo así como un modelo de evolución probabilística donde transformaciones U_i —no necesariamente unitarias— actúen sobre un sistema con probabilidad p_i :

$$\rho \rightarrow \rho' = \sum_i p_i U_i \rho U_i^\dagger, \quad \sum_i p_i = 1, \quad (3.6)$$

en que ¡la entropía no se conserve! en general:

$$S(\rho') \geq S(\rho),$$

y por supuesto, en que la *asimetría* entre pasado y futuro, propia del procesamiento de información, y equivalente a la flecha de tiempo que surge en los fenómenos termodinámicos *irreversibles*, sea también signo de la evolución temporal en teoría cuántica?

Intentémoslo. Debemos pues generalizar la dinámica y el proceso de medición, y ya tenemos una pista: si el sistema es abierto, debe interactuar con su entorno, ya sea en el proceso de evolución o en el de medición.

⁹La discusión de esta sección consiste en proponer reinterpretaciones del *teorema de representación* que se desarrolla en el capítulo 9 y en el apéndice D de [8] y en los capítulos 14-16 de [1], y del *teorema de Neumark* que se desarrolla en la Sec. 9.6 de [8].

3.3.1. *Teorema de Kraus: teorema de representación para dinámica generalizada*

Muchas de las transformaciones más importantes de la teoría cuántica (evolución temporal, traza parcial, medición proyectiva, etc.) pueden ser representadas por medio de un *superoperador* $\mathcal{E}(\rho)$ ¹⁰ tal que

$$\rho \rightarrow \tilde{\rho}' = \mathcal{E}(\rho),$$

y que satisface simplemente tres condiciones. Primero, ser lineal, i.e., que para toda mezcla estadística $\rho = \sum_i p_i \rho_i$,

$$\mathcal{E}(\rho) = \sum_i p_i \mathcal{E}(\rho_i).$$

Además, no aumentar la traza y ser *completamente positivo*. No aumentar la traza quiere decir que

$$\text{Tr}[\mathcal{E}(\rho)] \leq \text{Tr}(\rho) = 1$$

y ser completamente positivo no implica solamente preservar la positividad de los operadores de densidad, supone también que quepa la posibilidad de considerar nuestro sistema cuántico Q un sistema abierto, de modo que si \mathcal{E}^Q actúa en Q, el sistema compuesto QE puede ser influenciado. Es decir, si se extiende trivialmente el superoperador \mathcal{E}^Q a $\mathcal{E}^Q \otimes \mathbb{1}^E$ sobre $\mathcal{H}^Q \otimes \mathcal{H}^E$, este nuevo superoperador del sistema compuesto debe ser también positivo.¹¹ Y eso es todo. Schumacher llama a este tipo de superoperadores transformaciones *físicamente razonables* de un sistema cuántico, mientras Audretsch las llama simplemente *operaciones cuánticas*.

Teorema 3.3¹² *La transformación $\rho \rightarrow \tilde{\rho}' = \mathcal{E}(\rho)$ es una operación cuántica si y solo si esta tiene una descomposición*

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger \quad (3.7)$$

con operadores lineales K_i sobre \mathcal{H}^Q que satisfacen la condición

$$\sum_i K_i^\dagger K_i \leq \mathbb{1}. \quad (3.8)$$

Varios comentarios. Los operadores K_i son usualmente llamados *operadores de Kraus*. Si se cumple que

$$\sum_i K_i^\dagger K_i = \mathbb{1},$$

la operación cuántica $\mathcal{E}(\rho)$ preserva la traza y se dice que es una *operación cuántica completa*. De lo contrario, es decir, si

$$\sum_i K_i^\dagger K_i < \mathbb{1},$$

¹⁰La tilde encima del operador densidad significa que el estado puede no estar normalizado.

¹¹En el apéndice D, Sec. D.1 de [10] se discute al respecto.

¹²La demostración de este teorema puede encontrarse ya sea en el Cap. 16 de [1] o en el apéndice D de [10].

$\mathcal{E}(\rho)$ no preserva la traza y se dice que es una *operación cuántica incompleta*. Además, como veremos a continuación, la descomposición (3.7) no es única.

Con esto queda establecida la principal herramienta para desarrollar una dinámica y un proceso de medición generalizados.

Evolución unitaria de un sistema compuesto

Hasta aquí todo parece servido para intentar arribar a una evolución como la expresada en (3.6); pero no tan rápido. Recordemos la pista que nos da el Teorema 3.2. La entropía se conserva si y solo si el sistema es informacionalmente aislado, de modo que ¿por qué no considerar nuestro sistema como un sistema abierto que *interactúa* con su entorno? Esa podría ser la clave para arribar a una dinámica generalizada y lograr que la entropía no se conserve. De hecho esta es la clave, pero hay un pequeño problema: la evolución del sistema compuesto debe considerarse *unitaria*, ¡PLOP!. Veamos. Supongamos que nuestro sistema compuesto se encuentra inicialmente en el estado producto $\rho^{\text{QE}} = \rho \otimes |i\rangle\langle i|$ y que evoluciona unitariamente al estado ρ'^{QE} —en general *entrelazado*—, provocando a su vez que el subsistema Q evolucione de ρ a ρ' , i.e.

$$\rho'^{\text{QE}} = U^{\text{QE}} |i\rangle\langle i| \rho \langle i| U^{\text{QE}\dagger},$$

mientras que

$$\rho' = \text{Tr}_E(\rho'^{\text{QE}}) = \sum_n \langle e_n| U^{\text{QE}} |i\rangle\langle i| \rho \langle i| U^{\text{QE}\dagger} |e_n\rangle, \quad (3.9)$$

donde $\{|e_n\rangle\}$ es una base ortonormal del entorno. Es evidente que los operadores de Kraus son

$$K_i = \langle e_n| U^{\text{QE}} |i\rangle,$$

así que podemos verificar si satisfacen la Ec. (3.8):

$$\sum_n K_n^\dagger K_n = \sum_n \langle i| U^{\text{QE}\dagger} |e_n\rangle\langle e_n| U^{\text{QE}} |i\rangle = \mathbb{1}.$$

De acuerdo con el teorema de Kraus, podemos decir pues que

$$\rho \rightarrow \rho' = \mathcal{E}(\rho) = \sum_n K_n \rho K_n^\dagger, \quad (3.10)$$

donde $\mathcal{E}(\rho)$ es una operación cuántica *no* unitaria que preserva la traza. Este tipo de evolución suele llamarse también *determinística* debido a que la operación cuántica \mathcal{E} se lleva a cabo con certeza, pues

$$\text{Tr}(\rho) = \text{Tr}[\mathcal{E}(\rho)] = 1,$$

en oposición a las transformaciones en que la aleatoriedad se torna fundamental y no se preserva la traza, como veremos en la Subsec. 3.3.3. Además, es evidente que la descomposición en (3.9) no es única, pues depende de la base del entorno que se elija. Podemos entonces resumir el resultado de la siguiente manera:

La evolución unitaria del sistema QE representada por U^{QE} provoca al mismo tiempo una evolución no unitaria del subsistema Q

$$\rho \rightarrow \rho' = \mathcal{E}(\rho)$$

que preserva la traza.

De hecho la implicación inversa es también verdadera. Así que todas las operaciones cuánticas que preserven la traza pueden ser reducidas de esta manera a transformaciones unitarias [1]. De este modo podemos decir que logramos lo que pretendíamos, pero no sin venderle el alma a la evolución unitaria. El precio de una evolución no unitaria —local— del subsistema Q, es una evolución unitaria —global— del sistema *cerrado* QE.

3.3.2. Teorema de Kraus y entropía: ¿un mundo sin fricción?

Sin embargo, así como lo anterior representa una posible *implementación* de evolución *no* unitaria, puede procurarse también —aunque sea localmente— que la entropía no se conserve. El propósito de esta subsección es justamente revisar el teorema de Kraus desde el punto de vista de la entropía de una preparación. Propondremos un corolario del teorema de Kraus sirviéndonos de algunos resultados interesantes y del hecho de que toda operación cuántica determinística puede ser reducida a una transformación unitaria.

El primero de estos resultados interesantes es el teorema de *subaditividad*.¹³ Este teorema nos dice que un par de sistemas correlacionados supone *más información* que los dos sistemas considerados por separado [8]. Es decir, si ρ^{QE} es en general un estado entrelazado,

$$S(\rho^{\text{QE}}) \leq S(\rho^{\text{Q}}) + S(\rho^{\text{E}}), \quad (3.11)$$

donde $\rho^{\text{Q}} = \text{Tr}_{\text{E}}(\rho^{\text{QE}})$, $\rho^{\text{E}} = \text{Tr}_{\text{Q}}(\rho^{\text{QE}})$ y $S(\rho^{\text{Q}}) + S(\rho^{\text{E}}) \equiv S(\rho^{\text{Q}} \otimes \rho^{\text{E}})$. El otro resultado interesante expresa la desigualdad triangular que obedecen $S(\rho^{\text{Q}})$, $S(\rho^{\text{E}})$ y $S(\rho^{\text{QE}})$:¹⁴

$$S(\rho^{\text{QE}}) \geq |S(\rho^{\text{Q}}) - S(\rho^{\text{E}})|. \quad (3.12)$$

Por otra parte, ya que la evolución no unitaria

$$\rho^{\text{Q}} \rightarrow \rho'^{\text{Q}} = \mathcal{E}(\rho^{\text{Q}})$$

se debe a la evolución unitaria del sistema compuesto

$$\rho^{\text{QE}} \rightarrow \rho'^{\text{QE}} = \mathbf{U}^{\text{QE}} \rho^{\text{QE}} \mathbf{U}^{\text{QE}\dagger},$$

nos serviremos también del hecho de que la entropía del sistema QE se *conserva*. Si en general ρ^{QE} es una mezcla estadística de los estados producto $|\alpha_k, \beta_k\rangle$, i.e.,

$$\rho^{\text{QE}} = \sum_k p_k |\alpha_k, \beta_k\rangle \langle \alpha_k, \beta_k| \quad \left(\sum_k p_k = 1 \right),$$

se tiene que

$$\rho'^{\text{QE}}(\mathbf{U}^{\text{QE}} |\alpha_k, \beta_k\rangle) = p_k(\mathbf{U}^{\text{QE}} |\alpha_k, \beta_k\rangle),$$

y esto significa que la entropía del sistema cerrado QE se conserva:

$$S(\rho^{\text{QE}}) = S(\rho'^{\text{QE}}). \quad (3.13)$$

¹³A. Wehrl, *Rev. Mod. Phys.* 50 (1978) 221.

¹⁴H. Araki and E. H. Lieb, *Comm. Math. Phys.* 18 (1970) 160.

Esta es otra forma de entender el nombre de operación cuántica *determinística* que se le da a \mathcal{E} .

Lo que se demostrará es que *si $\mathcal{E}(\rho^Q)$ es que una operación cuántica determinística, entonces la entropía del sistema Q no se conserva en general.* Según lo que acabamos de mostrar, decir que $\mathcal{E}(\rho^Q)$ es que una operación cuántica determinística es equivalente a (3.13). Por otro lado, si —por simplicidad— suponemos que el sistema compuesto evoluciona desde un estado producto a un estado entrelazado ρ'^{QE} :

$$\rho^{QE} = \rho^Q \otimes |i\rangle \langle i| \rightarrow \rho'^{QE},$$

tenemos que

$$S(\rho^{QE}) = S(\rho^Q), \quad (3.14)$$

y en virtud del teorema de subaditividad:

$$S(\rho'^{QE}) \leq S(\rho'^Q) + S(\rho'^E).$$

Sirviéndonos de la Ec. (3.13), de la Ec. (3.14) y del resultado anterior, llegamos a la siguiente desigualdad entre las entropías locales:

$$S(\rho^Q) \leq S(\rho'^Q) + S(\rho'^E). \quad (3.15)$$

Asimismo de (3.13), (3.14) y la desigualdad triangular (Ec. (3.12)), arribamos a esta otra desigualdad entre las entropías locales:

$$S(\rho^Q) \geq |S(\rho'^Q) - S(\rho'^E)|. \quad (3.16)$$

De manera que para combinar las ecuaciones (3.15) y (3.16) debemos considerar dos casos. También vale la pena recordar que $S(\rho) \geq 0$ para todo ρ , así que se considerarán triviales los resultados de este tipo. Muy bien, considerando entonces el primer caso, i.e., cuando $S(\rho'^Q) < S(\rho'^E)$, concluimos que

$$S(\rho'^E) \geq S(\rho^Q). \quad (3.17)$$

Y si $S(\rho'^Q) \geq S(\rho'^E)$, alcanzamos un resultado muy interesante:

$$S(\rho'^Q) \geq S(\rho^Q). \quad (3.18)$$

De cualquier forma la entropía nunca decrece, tal como lo establece el siguiente teorema:

Teorema 3.4 *Sea \mathcal{E} una transformación tal que*

$$\rho^Q \rightarrow \rho'^Q = \mathcal{E}(\rho^Q).$$

Si \mathcal{E} es una operación cuántica determinística entonces alguna de las dos entropías, ya sea la del sistema o la del entorno, no decrece después de la evolución, i.e.

$$S(\rho'^Q) \geq S(\rho^Q) \quad \text{o} \quad S(\rho'^E) \geq S(\rho^Q).$$

Debemos decir que al menos localmente logramos nuestro cometido, pues la nueva evolución temporal *no* unitaria se da en la dirección del aumento de la entropía. Aunque no necesariamente la del sistema; de hecho esta puede disminuir durante la evolución, pero no sin el aumento de la entropía del entorno. En conclusión, este resultado *sugiere* tan solo *asimetría local* entre pasado y futuro. *Globalmente* sigue dando lo mismo si el tiempo fluye en un sentido o en otro.

Evolución unitaria probabilística

Este nombre es extraño, pues puede indicar que la evolución es unitaria sin serlo, pero de hecho es el adjetivo *probabilística* el que hace el trabajo de invertir el sentido. Veamos por qué.

Ya estamos en posición de estudiar casos en que solo conozcamos el entorno estadísticamente, i.e., casos en que el entorno no sea *perfectamente reproducible*. Consideremos el caso de un ensamble de transformaciones unitarias U_i donde a cada una le corresponde una probabilidad p_i . Nuestro sistema evoluciona de acuerdo a

$$\rho \rightarrow \rho' = \mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger, \quad \sum_i p_i = 1.$$

En palabras de Schumacher, \mathcal{E} representa una evolución temporal “físicamente razonable”, pues claramente los operadores de Kraus $K_i = \sqrt{p_i} U_i$ satisfacen la relación de completitud

$$\sum_i K_i^\dagger K_i = \mathbb{1},$$

y al mismo tiempo, conforme al Teorema 3.4, la entropía nunca decrece, siempre y cuando $S(\rho') \geq S(\rho^E)$:

$$S(\rho') \geq S(\rho).$$

¿Y cuándo se da la igualdad? Cuando se tiene una única transformación unitaria U_i , es decir, cuando la evolución del subsistema es también unitaria. O equivalentemente, según el Teorema 3.2, cuando el sistema es informacionalmente aislado.

¿Existe realmente la evolución *no* unitaria?

Ya tenemos bastantes elementos para entender por qué la evolución temporal se asume unitaria en teoría cuántica. Desde el Cap. 2 (Subsec. 2.1.4), antes de discutir sobre simetrías, la *ley de composición de amplitudes de transición* nos sugirió que la evolución dinámica de los sistemas cuánticos está representada por un producto de matrices unitarias. Y esta ley es enfática en algo: el proceso dinámico no puede dejar huella en el ambiente que permita distinguir cuál camino fue tomado. Luego, cuando se discutió el concepto de simetría, fuimos un poco más allá. Se dijo que si una cierta transformación preserva las probabilidades de transición, entonces esta respeta el principio de superposición y preserva el producto interno. Es decir, una simetría continua, sea cual fuere, se representa en teoría cuántica por medio de una transformación unitaria y lineal. Siendo la traslación temporal un caso particular. Más adelante, en el Cap. 3, nos servimos del formalismo de los sistemas abiertos para definir un sistema informacionalmente aislado, y —por medio del teorema de aislamiento— estos sistemas se identificaron precisamente con la evolución unitaria y con la conservación de la entropía. En suma, tres formas equivalentes de referirse a lo mismo: en un entorno perfectamente reproducible, un estado puro evoluciona en un estado puro.

Sin embargo, ahora contamos con una operación cuántica determinística en que localmente la evolución temporal *no* unitaria puede darse en la dirección del aumento de la entropía.

Pero ¿cuál fue el precio? Justamente la evolución unitaria del sistema compuesto. ¿Por qué la evolución temporal se asume unitaria en teoría cuántica? La respuesta ya es aparente: las transformaciones unitarias son las transformaciones fundamentales que preservan la traza.

¿Existe pues algún tipo de operación cuántica no determinística en que la evolución *realmente* se dé en la dirección del aumento de la entropía?

3.3.3. Medición generalizada: ¿un mundo con fricción?

En esta subsección explotaremos la generalidad del teorema de Kraus, explorando el caso en que los operadores de Kraus definen una operación cuántica *no determinística*, i.e., en que

$$\sum_i K_i^\dagger K_i < \mathbb{1}.$$

Este es justamente el caso de las mediciones generalizadas. Una operación cuántica \mathcal{E} suele asociarse con evolución temporal, tal como lo sugiere Schumacher en el apéndice D de su libro, pues cuando habla de una transformación unitaria “físicamente razonable”, se refiere exclusivamente a la evolución *temporal* de un sistema cuántico. Y con razón, de acuerdo a lo que recién se discutió en las subsecciones anteriores. Pero ¿acaso el proceso de medición no marca también un antes y un después? Esto es lo que se intentará aclarar.

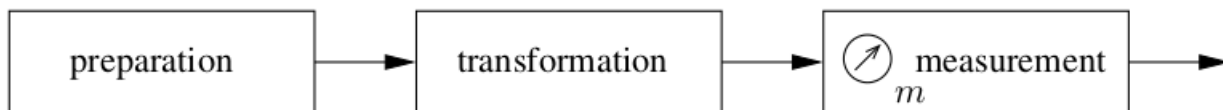


Figura 3.1: Las acciones de un aparato experimental sobre un sistema físico. (Tomada de [1]).

El teorema de Kraus es llamado también el teorema de *representación*, y la razón es muy simple: nos dice cómo representar cada una de las operaciones de un aparato experimental sobre un sistema cuántico (Figura 3.1). El proceso de *transformación* comprende precisamente las operaciones cuánticas \mathcal{E} que acabamos de discutir, aquellas en que el sistema cuántico evoluciona —en general *no* unitariamente— del estado ρ al estado ρ' ,

$$\rho \rightarrow \rho' = \mathcal{E}(\rho),$$

y en las que la operación \mathcal{E} se lleva a cabo con certeza, pues

$$\text{Tr}(\rho) = \text{Tr}[\mathcal{E}(\rho)] = 1,$$

de ahí que reciban el nombre de operaciones cuánticas *determinísticas*. Por el contrario, en el proceso de *medición* la aleatoridad se torna fundamental, y preguntarse por ejemplo por la polarización de un *único* fotón es un despropósito. En general, se considera un disparate creer que en teoría cuántica puede *determinarse* de antemano —a través de una prueba— el estado en que un único sistema fue preparado. Es por eso que en este contexto el proceso

de medición comprende las operaciones cuánticas \mathcal{M}_m en que el sistema cuántico evoluciona —no unitariamente— del estado ρ al estado $\tilde{\rho}'_m$,

$$\rho \rightarrow \tilde{\rho}'_m = \mathcal{M}_m(\rho),$$

y en las que la operación \mathcal{M}_m no se lleva a cabo con certeza, i.e.,

$$\text{Tr}[\mathcal{M}_m(\rho)] < 1.$$

Por eso reciben el nombre de operaciones cuánticas *no determinísticas*.

Pero ¿qué es una medición en este contexto? Para entender bien lo que acaba de decirse debe responderse antes esta pregunta, o lo que es lo mismo, deben establecerse los postulados de lo que en adelante llamaremos una *medición generalizada*. En este contexto una medición también supone la interacción de un aparato con el sistema de interés, donde justamente se procura información sobre el sistema. Esta es la esencia del estudio científico; la ciencia está basada en la observación de la Naturaleza [8]. Hasta aquí no hay desacuerdo entre este concepto de medición y el del Cap. 2. Usualmente se considera —“así sin más”— que el efecto de aquella interacción sobre el aparato puede ser precisado por medio de la lectura de un número real; mientras que Peres exige una *interpretación* parcialmente *clásica* de los resultados experimentales producidos por el dispositivo. Dice Peres al respecto:

Una medición no es un proceso pasivo de adquisición de conocimiento. Es un proceso activo que hace uso de equipo extremadamente complejo, y que usualmente involucra mecanismos de amplificación irreversibles. (La irreversibilidad no es accidental, es esencial, si queremos un *registro* indeleble, un registro objetivo[...]). Más aún, tenemos que *interpretar* los resultados experimentales producidos por nuestro equipo. Nosotros hacemos esto construyendo un *modelo teórico* mediante el cual el comportamiento del equipo macroscópico es descrito por unos pocos grados de libertad que interactúan con aquellos del sistema microscópico bajo observación. Llamamos a esto una “medición” del sistema microscópico.

Por lo pronto no profundizaremos en el concepto de medición de Peres,¹⁵ simplemente entenderemos el efecto sobre el dispositivo de medición a través de la lectura de un número entero m . Después de aclarar esto, podemos continuar con la definición de lo que es una medición generalizada. Otra característica fundamental es que no se asume que la propiedad¹⁶ medida *exista* objetivamente antes de la interacción entre el aparato y el sistema, tal como se discutió en el Cap. 2. Asimismo —como ya se mencionó— se considera que la aleatoriedad es fundamental. ¿Por qué *el* fotón toma un camino y no el otro? (Figura 2.3), no se sabe, simplemente podemos pronunciarlos acerca de las probabilidades de que tome uno o el otro. Muy bien, pero ¿qué tiene de “general” esta definición? Es que aún falta. Comúnmente se habla de la interacción de la propiedad del sistema —que se pretende medir— con la propiedad del dispositivo a la cual está *acoplada*. Sin embargo, el aspecto fundamental de las mediciones generalizadas es que *no* son necesariamente *repetibles*,¹⁷ es decir, si se consideran dos mediciones idénticas consecutivas, estas no siempre arrojarán idénticos resultados, de manera que este tipo general de medición no necesariamente se asocia con la adquisición de información acerca de una cierta propiedad del sistema (e.g. la energía, la polarización, el espín, etc.) Según esto, la medición es en sí misma un proceso de *preparación*. La medición

¹⁵Quien lo desee, puede encontrar en el Cap. 12 de [8] esta idea desarrollada.

¹⁶En el caso particular de medición proyectiva.

¹⁷En este contexto *medición* es sinónimo de *prueba*, el concepto que se definió en el Cap. 2.

proyectiva, que como sabemos supone la medición de esta clase de propiedades, es solamente un caso particular de este proceso más amplio. Por último, la medición puede ser *selectiva* o *no selectiva*. Aunque el que nos interesa es el caso de *medición selectiva general*, pues después de todo cuando medimos procesamos *información* (Figura 2.3), lo que nos interesa al final es *distinguir*. Medición selectiva significa simplemente que el estado $\mathcal{M}_m(\rho')$ que resulte del proceso de selección debe depender de un estado ρ' bien definido, como por ejemplo el que resulta de una evolución determinística $\mathcal{E}(\rho)$, donde ρ es el estado de un gran número de sistemas preparados idénticamente (Figura 3.2).

Ahora expresemos todo esto matemáticamente. La operación cuántica de medición prepara un estado ρ cualquiera (bien definido) en el estado $\tilde{\rho}'_m$ asociado con el resultado m de la medición, que ocurre con una probabilidad $p(m)$:¹⁸

$$\rho \rightarrow \tilde{\rho}'_m = \mathcal{M}_m(\rho) = \sum_i \mathbf{M}_{m,i} \rho \mathbf{M}_{m,i}^\dagger$$

y

$$p(m) = \text{Tr}[\mathcal{M}_m(\rho)] < 1,$$

donde los operadores $\mathbf{M}_{m,i}$ son los operadores de Kraus asociados con el resultado m de la medición, y por supuesto,

$$\sum_i \mathbf{M}_{m,i}^\dagger \mathbf{M}_{m,i} < \mathbb{1} \text{ y } \sum_{m,i} \mathbf{M}_{m,i}^\dagger \mathbf{M}_{m,i} = \mathbb{1},$$

lo primero porque $\mathcal{M}_m(\rho)$ es una operación cuántica no determinística y lo segundo porque $\sum_m p(m) = 1$. La operación cuántica \mathcal{M}_m es no determinística precisamente porque $\text{Tr}[\mathcal{M}_m(\rho)]$ es la probabilidad de obtener en la medición el resultado m .

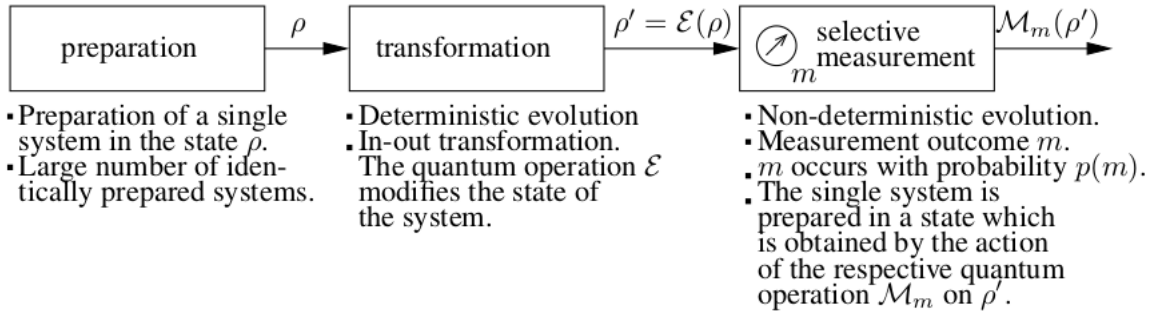


Figura 3.2: Descripción de las acciones de un aparato experimental sobre un sistema físico. (Tomada de [1]).

En adelante consideraremos por simplicidad el caso de $i = 1$ en que una medición selectiva general está descrita por un conjunto $\{\mathbf{M}_m\}$ de operadores de medición lineales:

$$\rho \rightarrow \tilde{\rho}'_m = \mathcal{M}_m(\rho) = \mathbf{M}_m \rho \mathbf{M}_m^\dagger$$

¹⁸Recordemos que la tilde en el estado $\tilde{\rho}'_m$ significa que el estado no está normalizado. En caso de usar la medición generalizada como uno proceso de *preparación*, bastaría agregarle a $\tilde{\rho}'_m$ una constante de normalización:

$$\rho'_m = \frac{1}{p(m)} \tilde{\rho}'_m.$$

y

$$p(m) = \text{Tr}[\mathcal{M}_m(\rho)] = \text{Tr}(\tilde{\rho}'_m) < 1,$$

donde cada resultado m tiene asociado un único operador M_m , y de nuevo,

$$M_m^\dagger M_m < \mathbb{1} \text{ y } \sum_m M_m^\dagger M_m = \mathbb{1}.$$

En general, los operadores M_m no son hermíticos. Tampoco son operadores de proyección P_m :

$$M_m M_{m'} \neq \delta_{mm'} M_m,$$

i.e., son ortogonales únicamente cuando $M_m = P_m$. Asimismo, el número de operadores M_m puede ser mayor que la dimensión del espacio de Hilbert en que actúa \mathcal{M}_m .¹⁹ Claramente, este tipo de mediciones no son repetibles, y en general, el conjunto $\{M_m\}$ no está relacionado con la medición de una cierta propiedad del sistema como la energía o el momento angular.

Reinterpretación del teorema de Neumark

Recordemos que toda operación cuántica determinística puede ser reducida a una transformación unitaria. Como dijimos, podemos considerar que nuestro sistema es abierto, que es un subsistema Q de un sistema más amplio QE, donde E hace las veces de entorno. La evolución no unitaria —local— de Q se debe justamente a la evolución unitaria —global— de QE que entrelaza a Q con su entorno. En este caso la clave está en la *interacción* de los subsistemas.

En el caso de las operaciones cuánticas no determinísticas sucede algo similar, pero hay un elemento adicional que se encarga de inducir la *aleatoriedad* en el resultado local de Q: *la medición proyectiva sobre QE*. Es decir, la evolución no determinística de Q debida a \mathcal{M}_m , puede reducirse a una transformación unitaria del sistema QE que entrelaza a Q y a E, y a una medición proyectiva sobre el sistema compuesto. Este resultado se conoce como el *teorema de Neumark*.²⁰ Sin embargo, para lo que se quiere expresar, es suficiente considerar una *implementación* particular de \mathcal{M}_m en donde se realiza simplemente una medición proyectiva sobre el entorno, y en la que el número de operadores M_m , es decir, el número de posibles

¹⁹La POVM (*positive operator valued measure*) es un tipo especial de medición generalizada en que justamente se evidencia esto. Resulta que todo operador lineal, como es el caso de M_m , puede ser expresado como

$$M_m = U_m \sqrt{M_m^\dagger M_m},$$

donde U_m es un operador unitario unívocamente determinado por M_m . Esta expresión suele llamarse *descomposición espectral izquierda* de M_m . De modo que ahora

$$\sum_m E_m = \mathbb{1} \text{ y } p(m) = \text{Tr}(\rho E_m),$$

donde $E_m = M_m^\dagger M_m$. En el **ejercicio 9.21** de [8] se pide comprobar que efectivamente el número de operadores E_m puede ser mayor que la dimensión del espacio de Hilbert en que actúa \mathcal{M}_m . Y la razón es muy simple. De nuevo, nuestro sistema de interés Q puede considerarse como un sistema abierto con un entorno E, así que lo que se dice una *medición generalizada* en Q bien puede considerarse una medición sobre el sistema compuesto QE.

²⁰En la Sec. 9.6 de [8] se encuentra la demostración de este teorema

resultados de la medición en Q, se considera igual a la $\dim(\mathcal{H}^E)$. Veamos. Sea L^{QE} una transformación lineal sobre un subespacio de $\mathcal{H}^Q \otimes \mathcal{H}^E$:

$$L^{QE} |\phi, 0\rangle = \sum_m M_m |\phi\rangle \otimes |m\rangle,$$

donde los operadores M_m sobre Q y la base ortonormal $\{|m\rangle\}$ del entorno definen L^{QE} y se asume que el sistema inicialmente se encuentra en un estado cualquiera $|\phi\rangle$ y el entorno en el estado fijo $|0\rangle$. Resulta que L^{QE} preserva el producto interno sobre este subespacio. Si $|\phi, 0\rangle$ y $|\phi', 0\rangle$ son dos elementos cualesquiera de aquel subespacio,

$$\begin{aligned} \langle \phi', 0 | L^{QE\dagger} L^{QE} |\phi, 0\rangle &= \sum_{mm'} \langle \phi' | M_m^\dagger M_m |\phi\rangle \langle m' | m \rangle \\ &= \langle \phi', 0 | \phi, 0 \rangle \end{aligned},$$

ya que $\sum_m M_m^\dagger M_m = \mathbb{1}$. Satisfecha esta condición, podemos servirnos de un resultado que nos permite considerar L^{QE} no solamente como una transformación lineal sobre aquel subespacio, sino también como una transformación unitaria U^{QE} sobre el espacio completo $\mathcal{H}^Q \otimes \mathcal{H}^E$ [1]. Ya tenemos entonces la transformación unitaria sobre QE que da lugar al entrelazamiento. Resta considerar la medición proyectiva sobre el entorno, i.e.

$$P_m = \mathbb{1} \otimes |m\rangle \langle m|.$$

Así pues, el resultado de una evolución no determinística de Q, puede entenderse a partir de la transformación sobre QE

$$P_m U^{QE} |\phi, 0\rangle = M_m |\phi\rangle |m\rangle,$$

o si el sistema Q antes de la medición se encuentra en un estado mezclado ρ , como

$$\begin{aligned} \rho^{QE} \rightarrow \rho'^{QE} &= P_m U^{QE} \rho^{QE} U^{QE\dagger} P_m \\ &= \tilde{\rho}'_m \otimes |m\rangle \langle m| \end{aligned},$$

donde $\rho^{QE} = \rho \otimes |0\rangle \langle 0|$ y $\tilde{\rho}'_m = M_m \rho M_m^\dagger$. Este resultado es interesante en términos de la entropía. Como las entropías de los sistemas compuestos son iguales a las entropías de los subsistemas, respectivamente, i.e.,

$$S(\rho^{QE}) = S(\rho) \quad \text{y} \quad S(\rho'^{QE}) = S(\tilde{\rho}'_m),$$

además

$$M_m^\dagger M_m < \mathbb{1}$$

y \mathcal{M}_m es una operación cuántica no determinística, i.e.,

$$\text{Tr}[\mathcal{M}_m(\rho)] = \text{Tr}(\tilde{\rho}'_m) < 1,$$

entonces

$$S(\tilde{\rho}'_m) > S(\rho), \tag{3.19}$$

o equivalentemente

$$S(\rho'^{QE}) > S(\rho^{QE}). \tag{3.20}$$

¿No era algo así lo que reclamábamos al final de la subsección anterior? Primero, la evolución se da en la dirección del aumento de la entropía. Segundo, la operación cuántica no determinística \mathcal{M}_m no solamente produce un aumento en la entropía *local*, justo como se consigue a través de la operación determinística \mathcal{E} , sino que también supone un aumento de la entropía *global*. Este resultado *sugiere* que el proceso de medición es *irreversible*. En palabras de Peres, que la distinción entre preparaciones y pruebas supone una dirección para el flujo del tiempo. ¿Sugiere esto que la *asimetría* entre pasado y futuro es fundamental en teoría cuántica? En realidad no, y la explicación es muy simple. Las poblaciones de $\tilde{\rho}'_m$ no definen una distribución de probabilidad, de modo que no tiene sentido hablar de la entropía de una preparación. Recordemos que la medición generalizada puede usarse también como un proceso de preparación. En ese caso, para preservar la interpretación probabilística, el estado $\tilde{\rho}'_m$ es normalizado:

$$\rho'_m = \frac{1}{p(m)} \tilde{\rho}'_m.$$

Las poblaciones de ρ'_m sí que definen una distribución de probabilidad, por lo cual vuelve a tener sentido hablar de la entropía $S(\rho'_m)$; sin embargo, en general, no siempre

$$S(\rho'_m) > S(\rho).$$

La entropía puede aumentar o disminuir durante la evolución. A través de este argumento no puede concluirse que la evolución del sistema se da en la dirección del aumento de las entropías local y global.

Capítulo 4

CONCLUSIONES

A lo largo de este trabajo dos preguntas son recurrentes: ¿qué define una simetría en física?, ¿por qué se asume unitaria la evolución temporal en teoría cuántica? Acerca de la primera, puede decirse que del mismo modo que en mecánica clásica las simetrías son representadas por medio de *transformaciones canónicas*, en teoría cuántica estas se representan a través de *transformaciones unitarias*. Mientras en mecánica clásica las unas dejan invariantes las ecuaciones de movimiento (e.g. las de Hamilton), en teoría cuántica las otras conducen a la invariancia de la ecuación de Schrödinger. Y en ambos casos, toda simetría conduce a una cantidad conservada.

En teoría cuántica resulta especialmente interesante la simetría de traslación temporal, pues el *teorema de Wigner*, el cual nos dice cómo se representa una simetría, no solo propone una representación por medio de transformaciones lineales y unitarias, sino que también incluye las transformaciones antilineales y antiunitarias. Una simetría particularmente llamativa que se representa de esta manera es la de *reversión temporal*. Lo interesante es que la evolución unitaria *parece* estar de acuerdo con esta simetría entre pasado y futuro en teoría cuántica, pues este tipo de evolución está estrechamente relacionada con la conservación de la entropía. En este contexto es que la segunda pregunta cobra importancia: ¿por qué se asume unitaria la evolución temporal en teoría cuántica? Porque las transformaciones unitarias son las transformaciones fundamentales que preservan la traza y, por medio del *teorema de Kraus*, es claro que están íntimamente relacionadas con las operaciones cuánticas *determinísticas*.

¿Existe entonces algún tipo de operación cuántica *no determinística* en que la evolución realmente se dé en la dirección del aumento de la entropía? Al parecer sí. Sirviéndonos del *teorema de Neumark*, mostramos que la *medición generalizada* se representa por medio de una operación cuántica de este tipo. Esto sugiere que la medición es un proceso *irreversible*, es decir, que la distinción entre preparaciones y pruebas supone una dirección para el flujo del tiempo en teoría cuántica. Bien puede ser, pero este argumento no es correcto. No tiene sentido hablar de la entropía de una preparación si antes no se normaliza el estado que resulta de una operación cuántica no determinística. Una vez normalizado, no puede decirse —en general— que la evolución del sistema se dé en la dirección del aumento de las entropías local y global.

Además aún quedan bastantes cuestiones por resolver. No se debe olvidar que el teorema

de Wigner reúne en un solo resultado operaciones cuánticas determinísticas y no determinísticas. Trata con transformaciones lineales y unitarias y con mediciones proyectivas —un caso particular de medición generalizada. ¿Entonces qué significa este teorema? ¿Habla únicamente de simetría? Más aún, ¿es también la simetría de reversión temporal una operación cuántica determinística? Esto podría cambiarlo todo. Aún queda mucha tela por cortar.

Bibliografía

- [1] Jürgen Audretsch. *Entangled Systems: New Directions in Quantum Physics*. Wiley-VCH Verlag GmbH & Co. KGaA, first edition, 2007.
- [2] Albert Einstein and Leopold Infeld. *The Evolution of Physics*. Cambridge University Press, London, N.W. I, first edition, 1967.
- [3] I. M. Gelfand. and S. V. Fomin. *Calculus of Variations*. Prentice-Hall, Englewood Cliffs, N.J, 1963.
- [4] Herbert Goldstein. *Classical Mechanics*. Addison-Wesley, second edition, 1980.
- [5] D. Kleppner. and R. Kolenkow. *An Introduction to Mechanics*. Cambridge University Press, second edition, 2014.
- [6] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields, Course of Theoretical Physics, Vol. 2*. Pergamon Press, fourth edition, 1975.
- [7] R. Feynman. R. Leighton. and M. Sands. *The Feynman Lectures on Physics Vol. I, Vol. II and Vol. III*. The Feynman Lectures Website: <https://www.feynmanlectures.caltech.edu>, 2013.
- [8] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 2002.
- [9] J. J. Sakurai. and Jim J. Napolitano. *Modern Quantum Mechanics*. Addison Wesley, 2nd edition, 2011.
- [10] B. Schumacher. and M. Westmoreland. *Quantum Processes, Systems and Information*. Cambridge University Press, 2010.
- [11] Alonso Sepúlveda Soto. *Los conceptos de la física: evolución histórica*. Editorial Universidad de Antioquia, tercera edición, 2012.