



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

Posgrado en Filosofía de la Ciencia
Facultad de Filosofía y Letras
Facultad de Ciencias
Instituto de Investigaciones Filosóficas

El Problema de la Probabilidad en la
Interpretación de Muchos Mundos

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

Maestro en Filosofía de la Ciencia

PRESENTA:

Carlos Gerardo Sanjuán Ciepielewski

TUTOR:

Dr. Elias Okon Gurvich
Instituto de Investigaciones Filosóficas



Ciudad Universitaria, Ciudad de México,
septiembre de 2016



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice general

Introducción	3
1. El Problema de la Medición de la Mecánica Cuántica	5
1.1. Mecánica cuántica estándar	6
1.2. Funciones de onda	8
1.3. Superposiciones	9
1.4. El problema de la medición	10
1.4.1. Algunos malentendidos con el problema de la medición	11
1.5. Estrategias para resolver el problema de la medición	12
2. Interpretaciones Everettianas	17
2.1. Lo común a toda interpretación everettiana	17
2.2. La interpretación de la relatividad de los estados	21
2.3. Everett-DeWitt-Graham	25
2.4. Una y muchas mentes	27
2.4.1. La interpretación de una mente	27
2.4.2. La interpretación de muchas mentes	29

ÍNDICE GENERAL

2.5. Muchos mundos más decoherencia	31
3. Teoría de la Decisión Racional Everettiana	37
3.1. Teoría de la decisión racional clásica	38
3.2. Teoría de la decisión racional a la Everett	40
3.3. Axiomas de riqueza	43
3.4. Los dictados de la razón	45
3.5. Lemas de la prueba	47
3.6. El teorema de representación everettiano	48
4. Evaluando la teoría de la decisión racional everettiana	51
4.1. Estrategias para solucionar el problema de la probabilidad	52
4.2. La objeción de las frecuencias	54
4.3. La objeción de los descendientes elitistas	60
Conclusiones	69
Agradecimientos	73
Bibliografía	78

Introducción

A inicios del siglo XX, Werner Heisenberg, Niels Bohr y Erwin Schrödinger, entre otros, discutieron cómo entender la imagen del mundo que surgía de la recién creada mecánica cuántica. Desafortunadamente, esta discusión fue interrumpida debido al clima instrumentalista que reinaba por esos años y por la creciente hegemonía de lo que hoy conocemos como la interpretación de Copenhage (basada, principalmente, en las ideas de Bohr). Durante más de la mitad de ese siglo, casi toda pregunta sobre la naturaleza de la mecánica cuántica fue ignorada o fácilmente descartada por razones ideológicas.

Hace poco más de cuarenta años la situación cambió —notablemente como resultado del trabajo de John Stewart Bell. Una vez más, tuvo sentido preguntarse por la imagen del mundo de la mecánica cuántica. En el centro de estas discusiones está el llamado *problema de la medición* y las denominadas *interpretaciones de la mecánica cuántica* — diferentes versiones de la teoría cuántica que intentan resolver el problema de la medición.

Entre estas interpretaciones está la mecánica cuántica everettiana —también conocida como interpretación de muchos mundos. En años recientes, sus defensores han sostenido que se han resuelto los problemas que esta interpretación, desde su creación en 1956, ha enfrentado. En particular, el llamado *problema de la probabilidad* se considera resuelto por medio de, asombrosamente, la teoría de la decisión racional.

El propósito de esta tesis es evaluar la estrategia everettiana de usar la teoría de la decisión racional para dar solución al problema de la probabilidad. En especial, examinaré dos objeciones que se han hecho en su contra y las respuestas a las mismas que se han propuesto.

Este texto está dividido en cuatro capítulos. El primero tiene como objetivo introducir al lector a las discusiones en torno al problema de la medición. Se presenta la mecánica cuántica estándar, el problema de la medición y las estrategias comunes que se han propuesto para solucionarlo. El segundo capítulo ofrece una presentación de las distintas formas que la mecánica cuántica everettiana ha tenido. El tercero es una introducción al esquema de la teoría de la decisión racional usado por los everettianos para resolver el problema de la probabilidad. El último capítulo es una evaluación de dos objeciones que se han hecho contra la idea de que el esquema que presento en el capítulo tres resuelve el problema de la probabilidad de la mecánica cuántica everettiana.

Capítulo 1

El Problema de la Medición de la Mecánica Cuántica

La mecánica cuántica surgió gradualmente a inicios del siglo XX. Su gestación fue lenta y, a diferencia de lo que ocurre con la relatividad general o la mecánica newtoniana, su creación es atribuible a más de una persona. Se trata de una teoría sorprendente tanto por sus aplicaciones como por la evidencia empírica que la respalda. El láser, los transistores, el microchip y las imágenes por resonancia magnética son algunos de los ejemplos más importantes de sus aplicaciones. Aun más importante, la mecánica cuántica es un pilar sobre el cual descansan otras áreas de nuestro conocimiento. Sin ella, por ejemplo, la química dejaría de ser un conjunto ordenado de ideas y pasaría a ser un montón de reglas del pulgar. Por si esto fuera poco, la teoría ha superado todo intento de ser falseada.

Pese a todas estas virtudes, la mecánica cuántica no es una teoría bien definida. Explicar por qué no lo es, qué significa que no lo sea y, más importante, convencer al lector de que este es un problema *genuino*, es el propósito de este capítulo.

La estructura del capítulo es la siguiente. En la siguiente sección presentaré a la mecánica cuántica en su *formulación estándar*. Después, hablaré sobre dos conceptos de la teoría que serán importantes en discusiones posteriores: funciones de onda y superposiciones. Posteriormente, describiré en qué consiste el llamado *problema de la medición*. Por último, introduciré tres estrategias generales que se han propuesto para solucionarlo.

1.1. Mecánica cuántica estándar

Una forma de presentar a la mecánica cuántica estándar es a partir de los siguientes postulados (esta forma de exponer la teoría fue introducida por Dirac [1958] y von Neumann [1955]; aquí sigo la presentación de Albert [1992]):

- 1 *Estados físicos*. A todo sistema físico se le asocia un espacio de Hilbert.¹ Los estados físicos del sistema son representados por vectores de norma uno dentro de dicho espacio (vectores estado, denotados por $|\psi\rangle$). Cada vector de norma uno del espacio de Hilbert corresponde a un estado físico posible del sistema.²
- 2 *Propiedades*. Las propiedades de los sistemas físicos son representadas por operadores hermitianos asociados al espacio de Hilbert del sistema.³ Además, los valores de las propiedades del sistema están dados por la regla *eigenvalor-eigenvector*: Un

¹Un espacio de Hilbert es un espacio vectorial con producto interno, lo cual permite medir ángulos y longitudes, y completo. Un espacio vectorial es un conjunto de vectores cerrados bajo la suma y la multiplicación por escalares. El producto escalar es una operación entre vectores que da como resultado un escalar.

²La relación entre estados físicos y vectores del espacio de Hilbert no es uno a uno, pues dos vectores en el mismo ‘rayo’ dentro del espacio de Hilbert corresponden al mismo estado físico. Por ejemplo, $|\psi\rangle$ representa al mismo estado que el vector $-|\psi\rangle$.

³Los operadores hermitianos son mapeos lineales del espacio vectorial que coinciden con su propio operador adjunto. Una propiedad importante que poseen está en que sus eigenvalores son siempre números reales.

sistema posee el valor λ de una propiedad representada por el operador A si y solo si su estado es un eigenvector de A con eigenvalor λ .⁴ La importancia de esta regla radica en que es un puente entre el formalismo matemático (básicamente en términos de vectores y operadores) y el mundo físico (sistemas físicos y sus propiedades).

3 *Dinámica.* Dado cualquier estado de un sistema físico a cualquier tiempo, es posible determinar su estado futuro a partir de la ecuación de Schrödinger.⁵ Esta ecuación es determinista y lineal.⁶

4 *Regla de Born.* Cuando en un sistema se ha llevado a cabo una medición, la probabilidad de obtener el valor λ_i de la propiedad A es igual a $\langle \psi | P_i | \psi \rangle$ donde P_i es el operador de proyección correspondiente al eigenvector de A con eigenvalor λ_i . Como consecuencia, al realizar una medición a un sistema que se encuentra en un eigenestado del operador que representa a la propiedad medida, la probabilidad de obtener el eigenvalor correspondiente a dicho eigenvector es 1. Por otro lado, cuando el estado medido no es eigenestado del operador que representa a la propiedad medida, la probabilidad de obtener a cualquiera de sus eigenvalores es < 1 . Esta regla nos permite hacer predicciones al hacer experimentos

5 *Postulado del colapso.* Después de una medición, el estado del sistema ‘colapsa’ al estado obtenido como resultado de dicha medición. Esto es, si se realiza una medición A a un sistema originalmente en el estado $|\psi\rangle$ y el resultado es a_i , el estado

⁴Los eigenvectores de un operador A son los vectores diferentes de cero que satisfacen $A|\psi\rangle = \lambda|\psi\rangle$. λ es el eigenvalor correspondiente.

⁵ La ecuación de Schrödinger está dada por $i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}|\Psi\rangle$ donde \hbar es la constante de Planck dividida por 2π y \hat{H} es el operador Hamiltoniano que representa la energía total del sistema.

⁶Que sea lineal significa que tiene la propiedad matemática de que la combinación lineal de dos de sus soluciones también es una de sus soluciones. Esto es, si el estado $|a\rangle$ evoluciona a $|a'\rangle$ y el estado $|b\rangle$ evoluciona a $|b'\rangle$, entonces todo estado de la forma $\alpha|a\rangle + \beta|b\rangle$ evoluciona a $\alpha|a'\rangle + \beta|b'\rangle$.

del sistema deja de ser $|\psi\rangle$ y pasa a ser $|a_i\rangle$, el eigenestado de A que corresponde al valor a_i . Esta forma de evolucionar a un sistema es indeterminista, no lineal y discontinua.

1.2. Funciones de onda

El postulado uno de la mecánica cuántica estándar nos dice que los estados físicos de un sistema están representados por vectores de norma uno en el espacio de Hilbert asociado al sistema. En ocasiones, sin embargo, resulta útil hablar del estado de un sistema no en términos de un vector sino en términos de su *función de onda*. Para poder explicar lo que es una función de onda será necesario hacer algunas observaciones.

El conjunto de eigenestados $\{|a_1\rangle, |a_2\rangle \dots |a_n\rangle\}$ de un operador A representa una base ortogonal en el espacio de Hilbert.⁷ Debido a esto, podemos escribir a todo vector del espacio en términos de dicha base:

$$|\psi\rangle = \psi_1|a_1\rangle + \psi_2|a_2\rangle + \dots \psi_n|a_n\rangle$$

donde los ψ_i son llamados *coeficientes de expansión* y están definidos por: $\psi_i = \langle a_i|\psi\rangle$ —i.e., el producto interno entre los vectores $|\psi\rangle$ y $|a_i\rangle$.

La función de onda es otra forma de escribir al estado $|\psi\rangle$ usando la base de posiciones o de momento, las cuales son continuas.⁸ Así, la función de onda de $|\psi\rangle$ representa al estado en términos de los coeficientes de dicho estado en la base de posición.

⁷Para un espacio de n dimensiones, una base ortogonal es un conjunto de n vectores del espacio que son mutuamente ortogonales.

⁸Estrictamente hablando, no hay base de momento ni posición en este caso.

1.3. Superposiciones

Otro concepto que aparecerá constantemente en discusiones posteriores es el de *superposición*. Matemáticamente, las superposiciones son algo simple: son vectores que resultan de la suma de otros vectores. Siendo el espacio de Hilbert un espacio vectorial, podemos tomar a dos vectores que representen a estados físicos de un sistema, sumarlos y obtener a otro vector que represente a otro estado del sistema. Más todavía, como podemos escribir a todo vector del espacio de Hilbert en términos de la suma de otros vectores, todo vector es una superposición de los vectores que, al ser sumados, dan como resultado a dicho vector. Otra vez, matemáticamente, no hay nada misterioso. Los vectores pueden ser sumados y el resultado es otro vector.

La situación se complica cuando queremos dar una descripción física de lo que son las superposiciones. El famoso gato de Schrödinger es una manera de hacer manifiesta esta dificultad. Tenemos a un gato (el sistema) que puede estar en dos estados de la propiedad ‘vivir’: o está vivo ($|vivo\rangle$) o muerto ($|muerto\rangle$). Como consecuencia de los postulados de la mecánica cuántica, también el estado que resulta de sumarlos representa un estado posible del sistema —*i.e.*, la superposición $\alpha|vivo\rangle + \beta|muerto\rangle$ también es un estado posible del sistema.⁹ El problema está en que, mientras entendemos qué significa que un gato esté ‘vivo’ o ‘muerto’, no es claro qué puede significar que esté en una superposición de dichos estados (nunca *vemos* que el gato está en una suma de estar vivo y muerto). Pese a esta dificultad, las superposiciones cuánticas son elementos importantes en la teoría sin los cuales ésta no tendría el poder predictivo que posee. Por ejemplo, son cruciales en la explicación de los experimentos de doble rendija [Feynman, 2010].

⁹Donde $\alpha^2 + \beta^2 = 1$.

En mecánica cuántica estándar, y debido a la regla eigenvector-eigenvalor, un estado como $\alpha|vivo\rangle + \beta|muerto\rangle$ no posee una propiedad bien definida para la propiedad ‘vivir’. Lo más que podemos decir de dicho estado, en el contexto, otra vez, de la mecánica cuántica estándar, es que, después de una medición de la propiedad ‘vivir’, los posibles estados después de la medición son $|vivo\rangle$ y $|muerto\rangle$ y que la probabilidad de obtener a cada uno de ellos depende del coeficiente de expansión, α y β asociado a cada uno.

Al final, lo que digamos sobre lo que son físicamente las superposiciones dependerá de la interpretación de la mecánica cuántica que adoptemos. Más adelante veremos qué son estas interpretaciones, pero por ahora basta con decir que son distintas formas de leer la mecánica cuántica que tienen implicaciones profundas sobre la imagen del mundo que nos da la teoría.

1.4. El problema de la medición

De los postulados de la mecánica cuántica se desprende que hay dos formas de *evolucionar* un sistema. Una es por el tercer postulado, la dinámica de la teoría, la ecuación de Schrödinger, y la otra es por el postulado cuatro y cinco, la regla de Born y el postulado del colapso. La teoría nos dice que debemos usar la regla de Born y el postulado del colapso al hacer mediciones y que, en ausencia de mediciones, la evolución está dada por la ecuación de Schrödinger.

Si bien no hay problema con el hecho de que una teoría posea dos reglas de evolución, la dificultad aparece cuando notamos que *dentro de los postulados nunca se define qué constituye hacer una medición*. Esta falta de definición implica que la teoría no puede decirnos, sin ambigüedad, cuándo usar una de las evoluciones y cuándo la otra. La

consecuencia es una teoría que no está bien definida. Este es el problema de la medición.

Como veremos más adelante, hay más de una forma de presentar el problema de la medición.¹⁰ Poner al problema en términos de la falta de una regla bien definida tiene la ventaja de hacer al problema claro y conciso. En términos más generales, pero poco informativos, el problema de la medición consiste en encontrar un formalismo matemático, en términos de espacios de Hilbert, vectores y operadores, (o uno equivalente o suficientemente similar), que pueda dar cuenta, sin ambigüedades, de varios aspectos experimentales para los cuales tenemos evidencia empírica (como puede ser que al hacer mediciones obtengamos resultados definidos, que dichos resultados tengan determinadas frecuencias, y que los estados después de una medición parezcan colapsar).

1.4.1. Algunos malentendidos con el problema de la medición

¿No hay un problema de la medición también en mecánica clásica al no poseer esta una noción bien definida de medición? No, no lo hay, al menos no en el mismo sentido. Es cierto que en mecánica clásica no hay una noción clara de lo que significa medir. Sin embargo, como en esta teoría dicho concepto no aparece en su formulación, que no lo defina no es un problema que haga que la teoría no esté bien definida. Según la mecánica clásica, el mundo es un conjunto de partículas gobernadas por las leyes de Newton. No hay referencia alguna a mediciones. Por otro lado, la mecánica cuántica estándar hace uso explícito de la noción de medición en su formulación.

¿El problema de la medición consiste en definir lo que es una ‘medición’? No, al

¹⁰Para introducciones más detalladas al problema de la medición, véase Krips [2013] Redhead [1989] Bell [1987] y Albert [1992].

menos no exactamente. Como dije arriba, el problema de la medición, en sentido estricto, consiste en conectar, sin ambigüedades, un formalismo matemático con algunos aspectos de nuestra experiencia.

La confusión surge porque el problema, *en el contexto de la interpretación estándar*, consiste en que no hay una definición de lo que significa medir. En el contexto de otras interpretaciones, sin embargo, el problema adquiere distintas formas. Por ejemplo, en las llamadas interpretaciones everettianas, que veremos en breve, el problema resurge, al menos parcialmente, como un problema relacionado con la probabilidad.

1.5. Estrategias para resolver el problema de la medición

A las propuestas que se han hecho para solucionar el problema de la medición se les conoce como *interpretaciones de la mecánica cuántica*. Debe notarse, no obstante, que hablar de ‘interpretaciones’ es un tanto desafortunado en este contexto, pues evoca la idea de una teoría bien establecida sobre la cual se discute cómo entenderla. Esto es incorrecto. Algunas de las interpretaciones de la mecánica cuántica son teorías distintas a ella, aunque obviamente relacionadas (deben estarlo para poder reproducir las predicciones exitosas que la mecánica cuántica estándar hace). Por ejemplo, las interpretaciones de colapso objetivo, como GRW, proponen una dinámica distinta a la ecuación de Schrödinger. Por esta razón, quizá sería más apropiado hablar de distintas versiones de la mecánica cuántica. Sin embargo, como dentro de las interpretaciones de la mecánica cuántica hay teorías (como la interpretación de muchos mundos) que buscan dejar al formalismo tal y como está, también sería incorrecto hablar de distintas versiones de la mecánica cuántica. Siendo esta la situación, lo más conveniente es seguir hablando de ‘interpretaciones’ con

la aclaración de que algunas de ellas buscan modificar al formalismo cuántico.

En lo que sigue presentaré tres estrategias generales para solucionar el problema de la medición. Para esto utilizaré una versión ligeramente modificada de un trilema que aparece en [Maudlin, 1995, p. 11].¹¹ Este trilema es una forma de presentar al problema de la medición, de modo que nos servirá tanto para introducir a las interpretaciones de la mecánica cuántica como para ofrecer al lector una forma alternativa de pensar sobre el problema de la medición. El trilema es el siguiente:

1. La función de onda de un sistema es completa —*i.e.*, la función específica, directa o indirectamente, todas las propiedades físicas del sistema.
2. La función de onda de un sistema evoluciona siempre siguiendo la ecuación de Schrödinger.
3. Mismas mediciones realizadas a sistemas representados por la misma función de onda frecuentemente tienen resultados diferentes, y la probabilidad de cada posible resultado es dada, al menos de forma aproximada, por la regla de Born.

Estas tres proposiciones son inconsistentes. Primero, si la función de onda es completa (1) y la función de onda evoluciona siempre de acuerdo a la ecuación de Schrödinger (2), entonces las mediciones de una misma propiedad a una misma función de onda deberían dar el mismo resultado ($\neg 3$). Segundo, si la función de onda es completa (1) y la medición

¹¹En su artículo, Maudlin divide al problema de la medición en tres distintos problemas, cada uno relacionado con un aspecto experimental distinto. El primero de estos aspectos es que al hacer mediciones en mecánica cuántica obtenemos resultados definidos. El segundo que al hacer una misma medición a una misma función de onda solemos obtener resultados distintos y estos parecen seguir la regla de Born. El tercero está en que al medir a un estado que no es eigenestado del operador que representa a la propiedad medida, el estado del sistema es alterado y posteriores mediciones de la misma propiedad al mismo sistema dan el mismo resultado. El trilema que aquí es usado representa al segundo aspecto experimental que Maudlin considera.

de una misma función de onda puede dar resultados distintos (3), entonces la función de onda no evoluciona de acuerdo con la ecuación de Schrödinger ($\neg 2$). Por último, si la función de onda evoluciona siempre siguiendo la ecuación de Schrödinger (2) y las mediciones de una misma función de onda pueden dar resultados distintos (3), entonces la función de onda no es completa ($\neg 1$).

La inconsistencia sólo surge cuando aceptamos las tres proposiciones; cada par posible es consistente. Así, aparecen naturalmente tres formas de solucionar el trilema a partir de negar una de las proposiciones (aunque también es posible negar más de una). Estas soluciones corresponden a tres estrategias distintas de solucionar el problema de la medición:

Si se niega 1, y entonces la función de onda de un sistema no especifica todas las propiedades físicas del sistema, uno llega a *interpretaciones de variables ocultas*. La idea de estas teorías es que hay más en el sistema físico considerado que lo que aparece representado por su función de onda. En otras palabras, la función de onda sólo es parte de la descripción física completa del sistema. La mecánica de Bohm es el ejemplo más significativo de interpretación que sigue esta estrategia. En esta interpretación, la descripción completa de un sistema físico incluye, además de la función de onda, las posiciones de las partículas de dicho sistema.

Por otro lado, si se niega 2, de modo que la evolución de la función de onda no siempre sigue la ecuación de Schrödinger, uno llega a *interpretaciones estocásticas*. Estas interpretaciones proponen solucionar el problema de la medición a partir de modificar la dinámica de la teoría. Un subconjunto importante de estas interpretaciones son las llamadas *interpretaciones de colapso objetivo*, como GRW (Ghirardi-Rimini-Weber). Este

subconjunto de interpretaciones busca dar una dinámica al colapso de la función de onda.

Por último, uno puede negar 3, que funciones de onda idénticas pueden dar resultados distintos cuando se les mide una misma propiedad y que estos resultados siguen, al menos de forma aproximada, la regla de Born. Seguir esta última estrategia nos lleva a interpretaciones everettianas (también llamadas de muchos mundos), el tema central del próximo capítulo.

Debe notarse que negar 3 es radicalmente distinto a negar 1 o 2. Mientras 1 y 2 son afirmaciones acerca de la naturaleza de la teoría cuántica —*i.e.*, si es completa o si sigue la ecuación de Schrödinger— 3 es un dato que obtenemos de manera empírica. Que mismas funciones de onda arrojen distintos resultados cuando se les hace una misma medición es algo que sabemos en el laboratorio. Así, cualquier interpretación que siga esta última estrategia debe dar algún tipo de teoría del error que explique por qué, *aparentemente*, 3 es verdad.

CAPÍTULO 1. EL PROBLEMA DE LA MEDICIÓN DE LA MECÁNICA CUÁNTICA

Capítulo 2

Interpretaciones Everettianas

Hacia el final del capítulo anterior encontramos que uno llega a una interpretación de muchos mundos al mantener tres proposiciones. Uno, la función de onda es completa. Dos, la función de onda siempre evoluciona de acuerdo con la ecuación de Schrödinger. Tres, contrario a lo que nos dice la experiencia, hacer mediciones a una misma función de onda nunca conduce a resultados distintos.

Este capítulo tiene como principal objetivo presentar las distintas versiones de la interpretación de muchos mundos que se han propuesto. La estructura del capítulo es la siguiente. Comenzaré describiendo lo común a toda versión de la mecánica cuántica everettiana. Después, presentaré cada una ellas. El capítulo termina con una introducción a la versión de la interpretación alcuualmente más defendida.

2.1. Lo común a toda interpretación everettiana

¿Qué es lo que hace a una interpretación everettiana ser una interpretación everettiana?

Para responder esta pregunta, usaré el famoso experimento mental del gato de Schrödinger

[1980] (que también usamos en el capítulo anterior para hablar de superposiciones).

Una versión del experimento es la siguiente: supongamos que tenemos un electrón en el estado $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$, que corresponde a la superposición de los estados espín arriba, $|\uparrow\rangle$, y espín abajo, $|\downarrow\rangle$, del electrón en el eje x . Ahora, imaginemos una caja cerrada con un gato dentro, adicionada con un dispositivo por el que pasará el electrón y que funciona del siguiente modo: cuando un electrón en el estado $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$ pasa por él, observamos que α^2 de las veces el espín del electrón es arriba y β^2 que es abajo (básicamente un Stern–Gerlach). Además, si resulta que el electrón termina en el estado $|\uparrow\rangle$, al gato dentro de la caja nada le ocurre; pero, si el estado del electrón es $|\downarrow\rangle$, un gas venenoso dentro de la caja es liberado y el gato muere.

Podemos usar el formalismo cuántico que vimos en el capítulo anterior para representar el estado del sistema —caja más electrón— antes de la interacción:

$$|listo\rangle_G (\alpha|\uparrow\rangle + \beta|\downarrow\rangle) \tag{2.1}$$

donde $|listo\rangle_G$ representa el estado del gato y la caja antes de la interacción. Después de esta —y si asumimos que la evolución del sistema está gobernada por la ecuación de Schrödinger (esta es la suposición central del experimento)— el estado pasa a ser:

$$\alpha|vivo\rangle_G|\uparrow\rangle + \beta|muerto\rangle_G|\downarrow\rangle \tag{2.2}$$

donde $|vivo\rangle_G$ y $|muerto\rangle_G$ representan los posibles estados del gato. Que este sea el estado final de la interacción se sigue de la linealidad de la ecuación de Schrödinger.

Así, al final de la interacción, la superposición del electrón, que en un inicio era

inofensiva por pertenecer al reino de lo microscópico, se filtra al mundo macroscópico, planteando una situación incómoda. Lo incómodo de la situación está en que, mientras entendemos qué significa que un gato esté vivo o muerto, que esté en una suma de dichos estados parece ridículo.

La descripción de lo que ocurre con el gato de Schrödinger depende de la interpretación de la mecánica cuántica que uno adopte. Por esta razón, el experimento sirve como punto de comparación entre ellas. Por ejemplo, en la interpretación estándar —el contexto en el que Schrödinger creó el experimento— todo sistema debe seguir la ecuación de Schrödinger mientras no haya una medición. Por consiguiente, el estado del sistema después de la interacción *es* 2.2; el gato *está* en la superposición de los estados $|vivo\rangle$ y $|muerto\rangle$ hasta que no se realice una medición (sea lo que sea que eso signifique).¹

La descripción del experimento mental que nos da la interpretación de muchos mundos es particularmente reveladora. Recordemos, que esta interpretación mantiene que la función de onda es completa y que la evolución de ésta siempre sigue la ecuación de Schrödinger. Como consecuencia, esta interpretación mantiene que, después de la interacción, el estado en 2.2 es el estado *real y completo* del sistema. La cuestión, como ya hemos visto, está en que nunca observamos (o al menos eso nos parece) estados macroscópicos en una superposición. El problema, dicho de otra forma, está en que si hiciéramos el experimento encontraríamos siempre al gato vivo o muerto. Así, la interpretación parece ir en contra de nuestras observaciones.

¹Podría argumentarse que el dispositivo Stern-Gerlach que posee la caja en mi versión del experimento *es* un aparato de medición. Si esto es así, parece que sí hubo una medición en la interacción entre el electrón y la caja. Sin embargo, uno puede modificar al experimento de modo que no incluya un Stern-Gerlach o a cualquier otro dispositivo que pueda pensarse como un aparato de medición. En el experimento original que Schrödinger [1980] imaginó, en lugar de un electrón y la propiedad del espín, se usa una sustancia radioactiva que tiene un 50% de probabilidad de decaer durante el transcurso de una hora. En ese escenario, no es viable objetar que en la interacción entre la caja y la sustancia radioactiva hubo algún tipo de medición.

Sin embargo, hay algo, o mejor dicho, alguien, a quien no hemos considerado hasta ahora y que es crucial para entender la motivación detrás de las interpretaciones everettianas. ¿Cómo representamos a los agentes dentro de la teoría? Si es cierto que los estados físicos siempre están gobernados por la ecuación de Schrödinger, los agentes (sistemas también físicos) igualmente deberían regirse por ella. Así, regresemos al estado de 2.2 (nuestro gato en superposición) y supongamos que un agente está a punto de abrir la caja y ver lo que hay dentro de ella. Representamos al sistema completo —electrón más gato más agente— antes de la interacción como:

$$|listo\rangle_A \left(\alpha |\uparrow\rangle_G |\uparrow\rangle_E + \beta |\downarrow\rangle_G |\downarrow\rangle_E \right) \quad (2.3)$$

donde $|listo\rangle_A$ representa el estado del agente antes de la interacción. Los estados del gato $|vivo\rangle$ y $|muerto\rangle$ fueron cambiados, por simplicidad, por $|\uparrow\rangle_G$ y $|\downarrow\rangle_G$, respectivamente, y añadí el subíndice “E” a los estados del electrón. Evolucionando el sistema por la ecuación de Schrödinger, obtenemos:

$$\alpha |\uparrow\rangle_A |\uparrow\rangle_G |\uparrow\rangle_E + \beta |\downarrow\rangle_A |\downarrow\rangle_G |\downarrow\rangle_E \quad (2.4)$$

donde $|\uparrow\rangle_A$ y $|\downarrow\rangle_A$ representan, respectivamente, al agente percibiendo que el electrón tiene espín arriba (y que el gato está vivo) y al agente percibiendo que el electrón tiene espín abajo (y que el gato está muerto).

Así, después de la interacción, *el agente también entra en la superposición* al ver dentro de la caja (y lo hace, en general, cuando observa cualquier resultado de un experimento cuántico). De forma que, el everettiano nos dice, el hecho de que nunca encontremos un

gato en una superposición —y en general el hecho de que *creamos* que mismas funciones de onda producen diferentes resultados cuando se les hace la misma medición— se debe a que nosotros, como sistemas físicos, *entramos* en una superposición al interactuar con el sistema.

Hasta aquí lo que el everettiano tiene es, a lo mucho, una idea interesante sobre cómo solucionar el problema de la medición. La sugerencia es tomar la mecánica cuántica estándar, quitar el postulado de la regla de Born y el postulado del colapso y tomarse en serio la idea de que los agentes son sistemas físicos que, como tales, también pueden entrar en superposiciones. Para tener una solución completa hace falta mostrar, entre otras cosas, que una interpretación de este tipo es consistente con la evidencia empírica, que no tiene ambigüedades (como sucedía en la interpretación estándar) y, en particular, debe explicar qué significa que agentes puedan entrar en superposición.

En lo que resta de este capítulo trataremos con las propuestas que se han hecho para hacer de esta sugerencia una solución al problema de la medición.

2.2. La interpretación de la relatividad de los estados

Hugh Everett desarrolló la primera interpretación de muchos mundos mientras era estudiante de doctorado en la universidad de Princeton. Su tesis, dirigida por John Wheeler y aceptada en 1957, fue una versión reducida de un texto que Everett había enviado a Wheeler un año antes. La reducción del texto se debió al distanciamiento entre las ideas de Everett sobre la mecánica cuántica y las ideas que por esos años imperaban dentro de la comunidad de físicos. En particular, el planteamiento de Everett era crítico de las ideas de Bohr y de la denominada interpretación de Copenhage. Después de egresar de Princeton,

Everett abandonó la academia y no participó activamente en las discusiones alrededor de su interpretación.

Si bien hay razones para creer que algo similar a la interpretación de muchos mundos ya había sido considerado antes, (por ejemplo, en algunas de las ideas de Schrödinger, véase Allori et al. [2009]), fue Everett el primero en desarrollar a detalle una propuesta de este tipo. Saber exactamente qué es lo que esta propuesta sostiene es difícil, en parte debido a la poca claridad del trabajo de Everett (como el lector podrá juzgar un poco más adelante). Aquí me apoyaré en el trabajo de Barrett [2014] y Albert [1992] para describir su postura.

Everett propuso tomar la mecánica cuántica estándar (1.1) reemplazando a los postulados de la regla de Born y del colapso por el *principio de la relatividad de los estados* [Everett, 1957, p. 78]. Este principio establece que hay una diferencia fundamental entre dos tipos de estados. Existen los estados absolutos, que representan las propiedades absolutas del sistema considerado; y existen los estados relativos, que representan las propiedades relativas de los subsistemas. Por ejemplo, un estado como 2.4, $\alpha|\uparrow\rangle_A|\uparrow\rangle_G|\uparrow\rangle_E + \beta|\downarrow\rangle_A|\downarrow\rangle_G|\downarrow\rangle_E$, es un estado absoluto que representa la totalidad del sistema considerado. Por otro lado, los subsistemas que lo componen (el electrón, el gato y el agente) están en estados relativos (el agente está en el estado de percibir que el espín del electrón es arriba *relativo* a que el electrón está en el estado de espín arriba). Siguiendo a Everett, no tiene sentido preguntarnos por el estado de un subsistema (*i.e.*, no tiene sentido preguntarnos por el espín del electrón) dado un estado como 2.4. Lo más que podemos decir sobre un subsistema es que tiene una propiedad *relativa* al resto de los subsistemas. Otra vez, lo que no está permitido es hablar —en nuestro ejemplo del

gato de Schrödinger— del estado del agente, el gato o el electrón independientemente del estado del resto del sistema. En el estado absoluto, $\alpha|\uparrow\rangle_A|\uparrow\rangle_G|\uparrow\rangle_E + \beta|\downarrow\rangle_A|\downarrow\rangle_G|\downarrow\rangle_E$, estos subsistemas no tienen una propiedad bien definida. Sólo en el contexto de un estado relativo tienen propiedades bien definidas.

Everett presentó su principio de la relatividad de los estados de la siguiente forma:

No existe, en general, nada como un estado único para un subsistema de un estado compuesto. Los subsistemas no poseen estados que sean independientes de los estados del resto del sistema, de modo que los estados de los subsistemas están generalmente correlacionados entre sí. Uno puede arbitrariamente elegir un estado para un subsistema, y ser conducido al estado relativo del resto. Así, estamos frente a una relatividad fundamental de los estados, implícita en el formalismo de sistemas compuestos. No tiene sentido preguntar por el estado absoluto de un subsistema —uno sólo puede preguntar por su estado relativo al estado del resto del subsistema [Everett, 1957, p. 180].²

Al dejar fuera el postulado del colapso y la regla de Born, Everett eliminó una de las dos formas de evolución que permite la mecánica cuántica estándar —dejando la ecuación de Schrödinger como única dinámica. Como consecuencia, su interpretación no necesita hablar de mediciones y, por lo tanto, no tiene la ambigüedad que posee la mecánica cuántica estándar.

²Traducción propia. Original: “There does not, in general, exist anything like a single state for one subsystem of a composite system. Subsystems do not possess states that are independent of the states of the remainder of the system, so that the subsystem states are generally correlated with one another. One can arbitrarily choose a state for one subsystem, and be led to the relative state for the remainder. Thus we are faced with a fundamental relativity of states, which is implied by the formalism of composite systems. It is meaningless to ask the absolute state of a subsystem —one can only ask the state relative to a given state of the remainder of the subsystem [Everett, 1957, p. 180].”

¿Es, sin embargo, el principio de la relatividad de los estados suficiente para tener una interpretación de la mecánica cuántica satisfactoria? Everett creía que sí, pues pensaba que lo más que podemos pedir de una teoría física es que satisfaga el criterio que él llamó *fidelidad empírica*. Este criterio básicamente establece que que todo *elemento de la realidad* con el que tenemos contacto debe ser representado por algún elemento presente en la teoría física que consideramos [Barrett, 2014, p. 4].³ Por ejemplo, si hiciéramos el experimento del gato de Schrödinger, encontraríamos al gato vivo o muerto. Estos estados corresponderían a distintos elementos de la realidad. Como estos elementos están presentes en el estado 2.4 (*i.e.*, dentro de dicho estado está tanto el estado del gato que corresponde a que esté vivo y el estado del gato donde está muerto), su propuesta satisface fidelidad empírica.

Independientemente de lo que Everett creyera, una interpretación satisfactoria de la mecánica cuántica debe ser más que fiel empíricamente. En particular, al menos debe ser capaz de reproducir las predicciones que la mecánica cuántica estándar hace (o unas muy similares). Al dejar fuera el postulado del colapso y la regla de Born, la interpretación de la relatividad de los estados no puede utilizar estos postulados para derivar predicciones cuánticas. Everett sabía esto y gran parte de su trabajo estuvo dedicado a mostrar que uno puede derivar los resultados estadísticos de la mecánica cuántica estándar desde su interpretación usando una noción de *tipicalidad* [Everett, 1957, pp. 187-194].⁴ Sin embargo, esta estrategia es circular: decir que un evento x es más típico que otro y no es más que decir que x es más probable que y .

³Everett nunca especificó qué es un *elemento de la realidad*.

⁴Básicamente, la propuesta consiste en derivar las probabilidades cuánticas a partir de los resultados que *típicamente* son obtenidos.

2.3. Everett-DeWitt-Graham

Siendo en un inicio crítico de las ideas de Everett, Bryce DeWitt cambió de parecer hacia 1970 con la publicación de *Quantum Mechanics and Reality*, artículo donde expresó su apego a lo que él creyó era la interpretación defendida por Everett. Presentando a Everett como creador, a sí mismo como su promotor y a Neil Graham —quien era estudiante de DeWitt al momento— como quien la clarificó, formuló la interpretación que, con el paso del tiempo, se convertiría en la versión más popular de las interpretaciones everettianas (conocida como EWG por obvias razones). En ella, cuando una interacción física genera un estado en superposición, como 2.4, el electrón, gato, agente, laboratorio y el universo entero, literalmente, se dividen en dos mundos. En uno el espín del electrón es arriba y en el otro es abajo. DeWitt lo expresó de forma más dramática:

Este universo está constantemente dividiéndose en un asombroso número de ramificaciones, todas resultado de interacciones tipo medición entre su gran número de componentes. Más aún, toda interacción cuántica teniendo lugar en cada estrella, cada galaxia, y en cualquier esquina remota del universo está dividiendo nuestro mundo local aquí en la Tierra en una inmensidad de copias de sí mismo [DeWitt, 1970, p. 161].⁵

EWG propone que pensemos en los estados en superposición como representando una multiplicidad de mundos, uno para cada elemento de la superposición. DeWitt pensó que uno llega a esta conclusión al asumir tres premisas. Uno, se debe interpretar la mecánica

⁵Traducción propia. Original: “This universe is constantly splitting into a stupendous number of branches, all resulting from the measurement like interactions between its myriads of components. Moreover, every quantum transition taking place on every star, in every galaxy, in every remote corner of the universe is splitting our local world on earth into myriads of copies of itself [DeWitt, 1970, p. 161].”

cuántica a partir, únicamente, de su formalismo. Dos, no existe un *reino clásico* ajeno al mundo cuántico. Tres, los colapsos nunca ocurren [DeWitt, 1970, p. 160].

La primera premisa —se debe interpretar la mecánica cuántica a partir, únicamente, de su formalismo— resulta poco razonable. Las matemáticas de una teoría, sin más, son estériles sobre lo que ocurre en el mundo. Sin una regla que conecte el formalismo con el mundo, la teoría carece de contenido *físico* y, por lo tanto, no podemos demandar que se deba interpretar una teoría únicamente a partir de su formalismo.

Independientemente de si se llega o no a EWG por las razones que DeWitt creyó, la interpretación enfrenta problemas serios. Primero, una solución satisfactoria al problema de la medición debe evitar hablar de interacciones tipo-medición, pues hacerlo equivale a recobrar la ambigüedad que la mecánica cuántica estándar tenía. El problema pasa de ser sobre definir lo que es una medición a definir lo que es una interacción tipo-medición. Como EWG habla de este tipo de interacciones, no es una solución satisfactoria. Otra forma de ver este problema es que EWG no nos precisa en qué momento ocurre la *ramificación* de mundos.

Segundo, EWG mantiene que estados en superposición representan una multiplicidad de *mundos*, tantos como elementos tiene una superposición. Sin embargo, como un mismo estado puede ser escrito en diferentes bases —recordemos que el estado del sistema es representado por medio de un vector en el espacio de Hilbert— ¿qué base es la que nos da la ramificación *correcta* de mundos?

Tercero, EWG parece incapaz, *en principio*, de poder recobrar las predicciones probabilistas de la mecánica cuántica estándar. Como sostiene que todo resultado cuántico ocurre con certeza, en ella no parece tener sentido hablar sobre la probabilidad de que tal

o cual resultado se obtenga.

Más adelante, hacia el final de este capítulo, volveré a decir algo sobre los últimos dos problemas que acabo de presentar. Como veremos en su momento, la versión actualmente más defendida de la interpretación everettiana propone una serie de soluciones a estos problemas.

2.4. Una y muchas mentes

En 1988, David Albert y Barry Loewer se propusieron desarrollar una interpretación que siguiera el espíritu de las ideas de Everett sin caer en los problemas que tiene EWG. Como resultado, crearon la interpretación de una mente (IUM) y la interpretación de muchas mentes (IMM). En esta sección trataremos a cada una por separado y nos centraremos en sus méritos y problemáticas.

2.4.1. La interpretación de una mente

Recordemos el estado 2.4:

$$\alpha |\uparrow\rangle_A |\uparrow\rangle_G |\uparrow\rangle_E + \beta |\downarrow\rangle_A |\downarrow\rangle_G |\downarrow\rangle_E$$

Como vimos antes, la interpretación de muchos mundos mantiene que 2.4 es el estado final y completo que resulta de la interacción, descrita en 2.1, entre un electrón, un gato y un agente. También hemos visto que lo extraño con esta postura está en que todo observador encontraría al gato vivo o muerto al abrir la caja, no a un gato en una superposición. Quizá, y esta es la idea central de la IUM, lo que ocurre es que *la mente, por alguna razón*

que desconocemos, es incapaz de entrar en una superposición.

El postulado central de la IUM es que la evolución de la mente no es determinista (como la evolución del electrón, la caja con el gato y el cuerpo del agente) sino probabilista [Albert and Loewer, 1988, p. 205]. La mente del agente antes de la interacción comienza en un estado en el que no tiene creencia alguna sobre el espín del electrón, $|liso\rangle_M$, y termina en $|\uparrow\rangle_M$ o $|\downarrow\rangle_M$ que, respectivamente, representan al agente creyendo que el espín del electrón es arriba o abajo —y que el gato está vivo o muerto. Además, la interpretación añade otro postulado que mantiene que la probabilidad de que la mente del agente termine pensando que el espín es arriba es igual a α^2 y la probabilidad de que termine pensando que el espín es abajo es igual a β^2 .⁶ Así, lo que tenemos es que la mente nunca se encuentra en superposiciones mientras mantenemos, al mismo tiempo, que el estado del mundo es 2.4.

La IUM está comprometida con un dualismo entre la mente y el cuerpo, pues mantiene que el cuerpo de un agente —y todo aquello que sea físico— puede entrar en una superposición, como 2.4, pero la mente no. Más aún, el dualismo puede verse en el hecho de que el estado de la mente es independiente del estado físico del agente.

Que la IUM sea dualista, no obstante, no debería ser un problema definitivo. Si bien las posturas dualistas no son populares en nuestros días, y probablemente no lo son por razones sólidas, esto no significa que una interpretación, por sólo postular una forma de dualismo, debe abandonarse. Sin embargo, el dualismo de la IUM es especialmente costoso. En esta interpretación, únicamente uno de los elementos de una superposición contiene una mente. Los otros elementos representan agentes sin mente y, aun peor, ningún elemento de la interpretación puede determinar cuál de esos elementos es el que tiene

⁶Por supuesto, el postulado se generaliza a cualquier medición.

mente. Este aspecto se vuelve más dramático si consideramos no una superposición de sólo dos estados, como 2.4, sino una de un exorbitante número de estados. En dicho caso sólo uno de ellos tendría una mente asociada y el resto tendría a agentes sin mente. De esto se seguiría que la mayor parte de la gente con la que hemos tenido contacto durante nuestras vidas posee, por así decirlo, un cerebro vacío (y, por lo tanto, no son, *realmente*, personas) [Albert, 1992, p. 130]. El problema, en pocas palabras, está en que en la IUM la mente no sólo no está relacionada con el cuerpo, sino que tampoco *superviene* del estado físico del agente, en particular, de su cerebro.⁷

2.4.2. La interpretación de muchas mentes

La interpretación de muchas mentes (IMM) surge a partir de los problemas de la IUM. En ella, cada agente está asociado con un continuo infinito de mentes. Además, suponemos que el conjunto infinito de esas mentes con un estado mental definido a cualquier tiempo es igual al cuadrado del valor absoluto del coeficiente del estado del cerebro asociado a ese estado mental [Albert, 1992, p. 130]. Así, en estados como 2.4, α^2 de la infinidad de mentes creerá que el electrón tiene espín arriba y β^2 que el electrón tiene espín abajo.

En la IMM, la evolución de cada mente individual es probabilista e igual a la descrita para la IUM. Sin embargo, y debido a que siempre hay una continua infinidad de mentes para todo estado mental, la evolución de las mentes, como conjunto, es invariablemente determinista [Albert, 1992, p. 130].

Con esto la IMM recobra las mentes que había dejado fuera la IUM. Sin embargo, la IMM resulta poco atractiva cuando consideramos su incapacidad de explicar algunas

⁷La relación de superveniencia es una relación entre propiedades que establece que un conjunto de propiedades *A* superviene de otro conjunto *B* cuando no puede haber diferencia entre las propiedades *A* sin haber también una diferencia en las propiedades *B*.

correlaciones. Para aclarar esto, no obstante, será necesario discutir algunas cosas antes.

Resulta que la IMM es local.⁸ Esto debe resultar extraño puesto que el teorema de Bell junto con los experimentos iniciados por Aspect et al. [1981] mostraron que nuestro mundo no es local. Supongamos que tenemos un par de partículas enredadas en el siguiente estado tipo EPR:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 - |\downarrow\rangle_1 \otimes |\uparrow\rangle_2) \quad (2.5)$$

También supongamos que la partícula 1 se encuentra en la región espaciotemporal \mathcal{A} y la 2 en \mathcal{B} y que \mathcal{A} y \mathcal{B} están separadas de forma espacial. Ahora imaginemos que un agente, Alicia, hace una medición de espín a la partícula uno y otro agente, Bob, hace una medición de espín a la partícula 2. En este contexto, el teorema de Bell nos dice que no puede haber ninguna forma local de dar cuenta de las correlaciones observadas entre los resultados de los experimentos hechos por Alicia y Bob. Esto es, el teorema descarta la posibilidad de poder explicar las correlaciones entre los resultados que obtienen Alicia y Bob en términos de lo que ocurre en el vecindario inmediato de las partículas medidas. Así, la medición que uno haga en la partícula 1 afectará, *inmediatamente*, al estado de la partícula 2, sin importar la distancia que las separe.

Hay, no obstante, que el teorema hace y que la IMM niega. El teorema asume que cuando Alicia o Bob miden el espín de las partículas hay *cuestiones de hecho* sobre los resultados obtenidos. Esto es, se asume que cuando Alicia, por ejemplo, mide el espín de la partícula 1, uno de los dos resultados posibles (espín arriba o abajo) *ocurre en el mundo*. Como en la IMM el estado del mundo después de la medición sigue siendo una

⁸En una teoría local, si dos regiones espacio-temporales, \mathcal{A} y \mathcal{B} , están separadas de forma espacial, entonces los eventos en \mathcal{A} no pueden afectar los eventos en \mathcal{B} .

superposición, nunca hubo, físicamente, tal cosa como un resultado para la medición que hizo Alice [Albert, 1992, p. 131].

Según la IMM, después de las mediciones que hacen Alicia y Bob, el estado del mundo sigue siendo una superposición. Además, la mitad de las mentes de Alicia creerán que se obtuvo el resultado $|\uparrow\rangle$ y la otra mitad que fue $|\downarrow\rangle$. Lo mismo sucederá con Bob. Lo extraño de la situación está en que las mentes de Alicia y Bob estarán relacionadas de forma tal que cuando una de ellas cree haber obtenido espín arriba, la otra cree haber obtenido espín abajo. Sin embargo, estas correlaciones son un completo misterio en la imagen que nos presenta la IMM. Según ésta, no hay ningún hecho físico en el mundo que determine esta correlación.

2.5. Muchos mundos más decoherencia

La versión actualmente defendida de las interpretaciones everettianas, que llamaremos muchos mundos más decoherencia (MMD), se compone de dos partes [Vaidman, 2014] [Wallace, 2012]:

1. Una teoría matemática que nos da la evolución del estado cuántico del universo, el cual es representado por un vector en el espacio de Hilbert.
2. Una explicación de cómo es que nuestras experiencias pueden ser obtenidas a partir de la función de onda del universo y su evolución.

El primer componente de MMD es básicamente la mecánica cuántica estándar (1.1) sin el postulado del colapso ni la regla de Born y con el añadido de que la función de onda del universo es la ontología fundamental de la teoría (*i.e.*, según MMD, lo único que *hay* a

nivel fundamental es la función de onda del universo, el resto son objetos emergentes). Al igual que EWG, esta interpretación mantiene que las interacciones físicas en el universo tienen como resultado el surgimiento de mundos donde cada resultado cuántico ocurre. Sin embargo, MMD, a diferencia de EWG, considera a estos mundos —y a todos los objetos en ellos— como entidades emergentes. Sólo la función de onda del universo existe a nivel fundamental.

El segundo componente responde a un requisito general que debe satisfacer toda teoría física. Toda teoría que pretenda decirnos algo sobre el mundo debe tener una forma de conectar su parte teórica con lo que observamos —debe, de algún modo, ser capaz de, en principio, decirnos algo sobre los objetos de nuestra experiencia. Si no lo hiciera, por más virtudes teóricas que poseyera, no podría ser corroborada ni usada para hacer predicciones. Este requisito, que quede claro, es distinto al de adecuación empírica. Éste demanda afinidad entre el contenido de la teoría y el mundo; aquél que, en principio, podamos ser capaces de decir algo del mundo a partir de la teoría. Llamemos a este requisito *conexión empírica mínima*.

Durante la mayor parte de la historia de la ciencia, la conexión empírica mínima ha sido un requisito trivial de las teorías científicas. Esto es, hay algo trivial en la demanda de que la mecánica newtoniana —o la teoría de la evolución— conecte su formalismo con nuestra experiencia cuando, por así decirlo, la mayoría de los *objetos* con los que trata la teoría (posiciones, velocidades, el sol, la gravedad, etc.) son los objetos que también constituyen nuestra experiencia.

La aparición de la mecánica cuántica cambió la situación de forma considerable. Esta nueva física trataba con sistemas con los que no tenemos contacto directo y, además,

su formalismo, en términos de vectores en el espacio de Hilbert y operadores, era tan lejano a nuestra experiencia que el puente entre ésta y la teoría debía ser construido de forma explícita con reglas interpretativas —como, por ejemplo, con la regla eigenvalor-eigenvector de la mecánica cuántica estándar.

Cuando pasamos de la mecánica cuántica estándar a sus interpretaciones, la conexión empírica mínima no está garantizada. Es necesario incluir reglas para extraer el contenido empírico de dichas interpretaciones. Esto es especialmente el caso con la mecánica cuántica everettiana. La interpretación inicia con la idea de eliminar a los postulados de la mecánica cuántica estándar que nos permitían conectar al formalismo con lo que observamos en el laboratorio —la regla de Born y el postulado del colapso. Como consecuencia, uno termina con una teoría que nos habla sobre la función de onda del universo —un objeto tan alejado de la experiencia como quizá ningún otro en la historia de la ciencia— y su evolución.

En este contexto, vuelven a aparecer dos problemas clásicos de las interpretaciones everettianas:

1. **El problema de la base privilegiada:** ¿cuál es la base *correcta* en la que debemos escribir la función de onda del universo?
2. **El problema de la probabilidad:** ¿cómo la probabilidad puede ser incorporada dentro de la teoría si, de acuerdo a ella, toda posibilidad física ocurre?

Ambos problemas son importantes para que la MMD pueda satisfacer al criterio de conexión empírica mínima. Sin una base, no es posible *leer* de la función de la onda del universo algo físicamente significativo. Sin la solución al problema de la probabilidad, la interpretación no puede decirnos por qué obtenemos ciertas frecuencias particulares al

realizar experimentos cuánticos. La solución a estos problemas es una condición necesaria, pero no suficiente, para que MMD satisfaga la conexión empírica mínima. Además se requiere, por ejemplo, que pueda hacerse un mapeo entre los elementos de la función de onda del universo a los objetos de nuestra experiencia. Wallace [2011] ha propuesto una forma de lograr esto adoptando una noción de funcionalismo que permita conectar aspectos de la función de onda con objetos de nuestra experiencia. Lo importante que uno debe notar aquí está en que la demanda de este tipo de mapeos no es exclusiva de MMD—esa es una cuestión presente en toda interpretación que sostenga que la función de onda es el único objeto fundamental de la teoría, como argumenta Maudlin [2010]— mientras que los dos problemas de arriba son exclusivos de muchos mundos.

La solución de MMD al problema de la base privilegiada es lo que hace a esta interpretación significativamente distinta a otras interpretaciones everettianas. Uno podría pensar en MMD como una versión actualizada de EWG. Después de todo, en ambas interpretaciones se mantiene que estados en superposición representan a una multiplicidad de mundos. Hay, no obstante, dos diferencias importantes. La primera consiste, como mencioné antes, en que MMD hace una distinción entre elementos fundamentales y emergentes de la teoría, sosteniendo que sólo la función de onda del universo es fundamental. La segunda es el uso que hace del fenómeno de decoherencia para solucionar al problema de la base privilegiada.

Decoherencia fue introducida por primera vez por Zeh [1970] y consiste en un proceso que ocurre cuando un sistema cuántico interactúa con su ambiente. Esta interacción tiene como resultado que los distintos elementos de una superposición cuántica dejan de interferir entre sí. MMD usa a este proceso para obtener una base en la cual escribir a la

función de onda que tiene características interpretativas favorables (escrita en esta base, la función de onda tiene elementos que se comportan, más o menos, de forma clásica). Aunque formalmente sigue siendo posible escribir a la función de onda del universo en otras bases, éstas no tendrán las características apropiadas para la emergencia de objetos macroscópicos.⁹

En el problema de la probabilidad podemos distinguir dos cuestiones separadas. La primera, que llamaremos *el problema de la coherencia*, tiene que ver con cómo hacer sentido de las nociones de probabilidad dentro de una interpretación everettiana. El segundo, que llamaremos *el problema cuantitativo*, consiste en derivar las predicciones que hace la mecánica cuántica estándar a partir del formalismo de la interpretación everettiana. Uno puede dar solución a uno de los problemas sin necesariamente solucionar al otro (*e.g.*, uno podría *narrar* alguna historia razonable sobre cómo hay incertidumbre en un mundo everettiano, logrando con esto que tenga sentido, al menos en principio, hablar de probabilidades, sin tener una forma de derivar el valor que éstas tienen).

Lo que debería seguir en este punto es una discusión sobre las propuestas disponibles para solucionar al problema de la probabilidad. Pero, será necesario interrumpir el flujo de esta discusión —flujo que será retomado en el capítulo cuatro— debido a que no es posible hacer justicia a las propuestas actuales para salir del problema de la probabilidad sin hacer alusión a lo que llamaré *teoría de la decisión racional everettiana*, el tema del próximo capítulo.

⁹Véase Wallace [2011, 2012] para una defensa de esta postura. Véase Janseen [2008] para una visión crítica.

Capítulo 3

Teoría de la Decisión Racional

Everettiana

En 1990, David Deutsch publicó *Quantum theory of probability and decisions*, artículo que ha influido significativamente a la interpretación de muchos mundos. En este artículo, Deutsch argumenta que las probabilidades de la mecánica cuántica estándar pueden ser derivadas —y por lo tanto no necesitan ser postuladas— de los otros postulados de la teoría. Esto es, Deutsch afirmó haber probado, a partir de la teoría de la decisión racional y la mecánica cuántica estándar (sin la regla de Born ni el postulado del colapso), que un agente racional actuaría *como si* hubieran probabilidades cuánticas y como si dichas probabilidades fueran iguales a las dadas por la regla de Born.

David Wallace [2003] llevó las ideas de Deutsch al contexto de la interpretación de muchos mundos; luego, en [2005b], presentó una serie de suposiciones que él creía implícitas en la prueba de Deutsch; y, en [2010, 2012], modificó aspectos importantes de la propuesta y la desarrolló a detalle. El objetivo de este capítulo es exponer esta propues-

ta, a la cual llamaremos *teoría de la decisión racional everettiana* (TDRE).

3.1. Teoría de la decisión racional clásica

Para entender la salida everettiana al problema de la probabilidad que apela a la teoría de la decisión racional, resulta útil comenzar con bosquejo de la teoría de la decisión racional clásica [Savage, 1972]. Ésta es un marco teórico que trata con decisiones en situaciones de incertidumbre. El objetivo es distinguir qué cursos de acción son mejores que otros. El tipo de situaciones que le interesan al teórico de la decisión racional están compuestas por los siguientes elementos:

- **Agente.** Es un individuo con ciertas preferencias que no conoce el estado del mundo de forma completa y que debe tomar un curso de acción en este contexto.
- **Estados del mundo.** Es un conjunto Γ de estados del mundo posibles, $\{\gamma_1, \gamma_2, \gamma_3, \dots\}$. El agente ignora qué elemento de este conjunto es, o será, aplicable a su mundo.
- **Actos.** Es el conjunto de cursos de acción, $\{A, B, C, \dots\}$, que el agente puede tomar en determinado momento. Formalmente, los actos pueden ser representados por funciones que van del conjunto Γ de estados del mundo al conjunto Ξ de consecuencias.
- **Consecuencias.** Es un conjunto Ξ de situaciones que resultan de que el agente haya elegido cierto curso de acción y de que el mundo resulte ser algún elemento del conjunto de estados del mundo Γ . Se asume que el agente prefiere unas consecuencias sobre otras y que estas preferencias pueden ser representadas por medio de una función de utilidad u .

Con estos elementos se pueden modelar una gran variedad de situaciones de toma de decisiones. Tomemos, por ejemplo, a Juan, una persona que debe decidir llevar o no su paraguas al trabajo. Él, por supuesto, es el agente; los posibles estados del mundo son que llueva o no llueva; los actos son que lleve el paraguas o no lo lleve y las consecuencias son lo que sucedería dado el acto que tomó y lo que haya ocurrido en el mundo (*e.g.*, que se moje porque llovió y no llevó su paraguas).

Para lograr su objetivo de distinguir qué cursos de acción son mejores que otros, la teoría de la decisión racional además impone una serie de constricciones de racionalidad sobre las preferencias del agente. Por ejemplo, se exige que las preferencias del agente sean transitivas: si el agente prefiere el acto A al acto B y prefiere B a C , entonces prefiere el acto A al acto C .

Con un conjunto lo suficientemente fuerte de constricciones de racionalidad, el teórico de la decisión racional puede probar lo que se conoce como un *teorema de representación* —*i.e.*, se puede probar que, para cualesquiera preferencias del agente que satisfagan las constricciones de racionalidad, hay una medida de probabilidad Pr , sobre el conjunto de estados Γ , y una función de utilidad u , sobre el conjunto de consecuencias Ξ , tales que, dados dos actos, A y B , el agente prefiere A a B si y sólo si el valor de expectación de A es mayor o igual al valor de expectación de B . Aquí el valor de expectación de un acto X está definido como:

$$EU(X) = \sum_{i=1}^N Pr(\gamma_i)u(X(\gamma_i)) \quad (3.1)$$

donde los γ_i son los posibles estados del mundo relevantes, $X(\gamma_i)$ es la función que tiene como rango al conjunto de consecuencias, $Pr(\gamma_i)$ es una medida de probabilidad sobre

el conjunto de posibles estados del mundo y $u(X(\gamma_i))$ es una función de utilidad sobre el conjunto de consecuencias. Entender qué representa cada uno de los elementos anteriores es la clave para entender la diferencia entre la teoría de la decisión clásica y la everettiana.

El resultado nos asegura que todo agente racional actuará *como si* estuviera maximizando el valor de expectación dada una función de utilidad y con respecto a cierta medida de probabilidad.

3.2. Teoría de la decisión racional a la Everett

La teoría de la decisión racional everettiana (TDRE) es un esquema para representar situaciones en las que un agente debe apostar sobre el resultado de un experimento cuántico en el contexto de un mundo donde MMD es el caso. Más a detalle, la TDRE está interesada en modelar situaciones donde un estado cuántico será preparado en una superposición, el sistema será medido en alguna base y un agente apostará sobre el resultado del experimento. Llamaremos a estas situaciones *problemas de decisión cuánticos*. El propósito de la teoría es hacer apuestas óptimas. La exposición que aparece a continuación sigue el trabajo de Wallace [2012], quien ha sido el mayor expositor y defensor de la TDRE.

Un comentario sobre esta presentación. La exposición de la TDRE que aparece a continuación es un tanto informal en el sentido de que no presento de forma explícita las pruebas de la teoría —en su lugar, doy los razonamientos generales que usa. Esto se debe a dos razones. La primera es que el objetivo de este capítulo es preparar el terreno para el próximo, donde discutiremos dos argumentos que se han levantado contra la TDRE. Como veremos en su momento, ninguno de estos dos argumentos ataca la validez de las pruebas de la teoría, de modo que no es necesario presentarlas a detalle. La segunda razón

—y la más importante, me parece— está en que, si bien el rigor formal es importante y debe procurarse en muchas situaciones, en ocasiones puede nublar la imagen general del tema tratado.

Iniciamos con los elementos que componen un problema de decisión cuántico:¹

- **Agente.** Es un individuo que se encuentra ante la posibilidad de realizar una apuesta sobre el resultado de un experimento cuántico en un mundo donde MMD es el caso. Se asume que él sabe que MMD es verdad y que conoce el estado del mundo donde se encuentra —por lo menos, conoce los coeficientes de expansión asociados a cada ramificación que resulte del experimento que se realice.
- **Estado del mundo.** Es el estado del mundo antes de que se lleve a cabo una medición, representado por $|\psi\rangle$. Este incluye no sólo al estado en superposición que se medirá, sino también al agente —y al mundo entero— antes de la medición.
- **Actos.** Es el conjunto de posibles cursos de acción que puede seguir el agente dado $|\psi\rangle$. Un acto incluye no sólo la asociación de consecuencias a distintos resultados del experimento —lo que, en sentido estricto, sería una apuesta— sino también a la preparación del estado a medir y a la medición que se elija hacer. Esto es, cuando el agente elige un acto está eligiendo una forma de preparar el estado a medir, la medición que se hará y la asignación de consecuencias a los distintos resultados. Los actos son representados por operadores unitarios, $\{A, B, C, \dots\}$. Si un agente prefiere el acto A al B , lo representamos como $A \succeq^\psi B$, donde el superíndice ψ indica que la preferencia se hace desde el estado del mundo $|\psi\rangle$. La cuestión de qué

¹Wallace nunca presenta estos elementos de forma explícita. Aquí lo hago con el propósito de ofrecer al lector una imagen más comprensiva de la TDRE.

actos están disponibles dado un estado del mundo no es trivial y, para especificarlos, la TDRE asume una serie de axiomas que veremos un poco más adelante.

- **Ramificaciones.** Son los mundos emergentes, representados por $\{|x_1\rangle, |x_2\rangle, \dots, |x_n\rangle\}$, que son producto de seguir el acto X . Así, podemos representar a todas las ramificaciones resultantes de seguir al acto A en el estado del mundo $|\psi\rangle$ como $A|\psi\rangle$. Se asume que el agente tiene preferencias sobre las ramificaciones —prefiere, por ejemplo, la ramificación donde recibe 1000 pesos a aquella donde, *ceteris paribus*, no recibe nada. Por simplicidad, asumiremos que podemos dividir al espacio de ramificaciones en subconjuntos que representan a macroestados distintos, *i.e.*, en lugar de preocuparnos por todas las ramificaciones que emergen de que se tome un curso de acción, trataremos como una sola ramificación a todas aquellas que son compatibles con un macroestado particular.²
- **Premios.** Es el conjunto de aspectos dentro de las ramificaciones que son relevantes para el agente en su toma de decisiones. Se asume que el agente prefiere unos sobre otros. Por simplicidad, uno puede pensar en los premios como asignaciones de distintas cantidades de dinero a los distintos *yo*es futuros del agente.

Puede resultar extraño que en una apuesta cuántica se asuma que el agente sabe que MMD es el caso. Parece que asumir esto corresponde a una petición de principio por parte del everettiano. Sin embargo, uno debe recordar que el proyecto de la TDRE es el de mostrar

²Esta simplificación no es trivial en ningún sentido. Sin embargo, argumentar a favor de ella nos llevaría a cuestiones que no son de nuestro interés y, por lo tanto, se tomará como un hecho que esto puede hacerse. Las razones que Wallace ofrece a favor de considerar a los macroestados, y no a las ramificaciones *per se*, en el contexto de un problema de decisión cuántico se derivan de la estructura que impone la teoría de decoherencia y, en especial, de la idea de que ésta permite obtener superposiciones de elementos que se comportan de forma clásica [Wallace, 2012, pp. 164-166].

que un agente racional en un mundo con ramificaciones actuaría *como si* hubiera probabilidades y éstas fueran dadas por la regla de Born, no que MMD es verdadera. Uno podría, no obstante, insistir en que uno no puede solucionar el problema de la probabilidad (el cual está íntimamente ligado al tema de la comprobación empírica), a partir únicamente de consideraciones sobre lo que ocurriría si MMD es el caso. En el próximo capítulo habrá lugar para evaluar este tipo de consideraciones. Por lo pronto, mi punto aquí es el de señalar que el objetivo de la TDRE no es argumentar a favor de que MMD es verdadera.

3.3. Axiomas de riqueza

Hay dos tipos de axiomas que la TDRE asume: axiomas de riqueza y axiomas de racionalidad. Los primeros son suposiciones que especifican qué actos están disponibles para un agente dado cierto estado del mundo. Los segundos son constricciones de racionalidad las preferencias del agente. Los axiomas de riqueza que presenta Wallace [2012][p. 166] son los siguientes:

- **Disponibilidad de los premios.** Todos los premios están disponibles al agente en toda ramificación. Esto es, el conjunto de actos disponibles siempre incluye alguno donde el agente recibe cualquier premio para cualquier macroestado.
- **Disponibilidad de la ramificación.** Dado cualquier conjunto de números reales positivos p_1, p_2, \dots, p_n cuya suma es la unidad, un agente siempre puede elegir algún acto que tiene n diferentes ramificaciones como resultados posibles y que asignan el peso p_i al resultado i .³

³Para que este y el anterior axioma tengan sentido uno debe pensar en las ramificaciones no en términos de los eigenvectores del acto seguido, sino como conjuntos de dichos eigenvectores que se identifican a partir de ser compatibles con un mismo macroestado.

- **Borradura.** Dado un par de estados, $|\psi\rangle$ y $|\varphi\rangle$, que tienen el mismo premio asignado, siempre hay un acto U disponible en $|\psi\rangle$ y en otro V en $|\varphi\rangle$ tal que $U|\psi\rangle = V|\varphi\rangle$.

Disponibilidad de los premios y disponibilidad de las ramificaciones reflejan el hecho de que los sistemas cuánticos pueden ser preparados en estados arbitrarios y que las asignaciones de premios a ramificaciones siempre pueden hacerse (uno siempre puede dar x cantidad de dinero en toda ramificación).

Borradura asegura que el agente siempre puede olvidar todo hecho sobre su situación que no esté directamente ligado a la asignación de un premio [Wallace, 2012, p.167]. Intuitivamente, lo que borradura garantiza es que el agente siempre puede ignorar el resultado del experimento cuántico que se haya obtenido en su ramificación siempre y cuando el premio que haya obtenido sea el mismo.

Supongamos que Juan, nuestro agente del caso clásico, ahora vive en un mundo everettiano. Imaginemos que, después de una jornada difícil de trabajo y de haberse mojado por no llevar su paraguas, decide, molesto y agobiado, darse unas vacaciones en Las Vegas. Días más tarde, al llegar a la ciudad, se ve invadido por la publicidad de un nuevo casino que ofrece apuestas cuánticas. Movidio por la curiosidad, decide visitarlo. Al entrar, no encuentra ruletas ni mesas para póquer, sino gente alrededor de aparatos de medición realizando apuestas. En la mesa más cercana está a punto de realizarse la medición del espín de un electrón en el estado $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$. Al aproximarse, un crupier le ofrece una hoja en la que aparece una lista de todos los experimentos que se realizan en el casino y las apuestas disponibles en cada caso. En este caso, de nuevo, Juan es el agente; el estado del mundo es el vector que representa la situación en la que se encuentra Juan justo antes

de que se realice la medición (que, por simplicidad y en el contexto de una apuesta particular, podemos representar por $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$); los actos son las decisiones que puede tomar Juan sobre qué apuesta hacer y en qué experimento hacerla; los premios son las cantidades de dinero que los *yoes* futuros de Juan recibirán (o perderán) en cada ramificación como resultado del acto que se siguió; y, por último, los mundos que emergen del acto son las ramificaciones.

En el contexto del casino, los axiomas de riqueza garantizan que tenga sentido realizar apuestas —permiten, por así decirlo, que el casino pueda diseñar las apuestas que quiera. Disponibilidad de premios y disponibilidad de ramificaciones garantizan, respectivamente, que el casino siempre puede dar cualquier suma de dinero en cualquier ramificación y que el casino puede diseñar las apuestas que quiera. Borradora permite que el agente esté exclusivamente interesado en el premio que adquiera, no en la ramificación en la que se encuentra.

¿Cómo debe actuar Juan?

3.4. Los dictados de la razón

Para responder esta pregunta, el everretiano debe imponer sobre el conjunto de preferencias del agente una serie de constricciones de racionalidad. Wallace presenta las siguientes [Wallace, 2012, 167-170]:

- **Ordenamiento.** La relación de preferencia entre actos, representada por \succeq^Ψ , es transitiva (si $A \succeq^\Psi B$ y $B \succeq^\Psi C$, entonces $C \succeq^\Psi D$), reflexiva ($A \succeq^\Psi A$) y conectada (o $A \succeq^\Psi B$ o $B \succeq^\Psi A$).

- **Consistencia diacrónica.** En un sentido especial, las preferencias de los agentes se mantienen consistentes a lo largo del tiempo: si el acto A está disponible en $|\psi\rangle$ y, si en cada ramificación que surge después de A , $|a_i\rangle$, están disponibles los actos B_1 y B_2 y en todas ellas el agente prefiere B_1 sobre B_2 , entonces el agente prefiere el acto A seguido de B_1 que el acto A seguido de B_2 en $|\psi\rangle$.
- **Indiferencia del microestado.** El agente no está interesado en el microestado particular en el que está, sino sólo en el macroestado en el que se encuentra.
- **Indiferencia a la ramificación.** El agente no está interesado en el surgimiento de ramificaciones *per se*. Si cierta operación deja a sus *yo*s futuros en n diferentes ramificaciones pero no cambia ninguno de los premios que reciben, él es indiferente a si la operación se llevó a cabo o no. En otras palabras, si el experimento se llevó a cabo pero no hubo apuesta alguna, entonces el agente es indiferente a si ocurrió o no una ramificación.
- **Superveniencia del estado.** Las preferencias del agente dependen solamente del estado físico en el que dejan a su ramificación. Esto es, si $A|\psi\rangle = A'|\psi'\rangle$ y $B|\psi\rangle = B'|\psi'\rangle$, entonces un agente que prefiere A sobre B cuando el estado del mundo es $|\psi\rangle$, debe también preferir A' sobre B' cuando el estado del mundo es $|\psi'\rangle$.
- **Solución de continuidad.** Si para algún estado $|\psi\rangle$, $A \succ_{\psi} B$, entonces permutaciones lo suficientemente pequeñas de A y B no cambian esto.

3.5. Lemas de la prueba

Dados los axiomas de riqueza y racionalidad, el everettiano puede probar los siguientes lemas en el contexto de un problema de decisión cuántico [Wallace, 2012, pp. 182-186]:

- **Lema de equivalencia.** Si dos actos asignan el mismo peso a las mismas ramificaciones y cada una de éstas recibe el mismo premio, entonces el agente debe ser indiferente a seguir uno o el otro.

- **Lema de nulidad.** Un agente es indiferente a un acto si y solo si dicho acto asigna peso cero a toda ramificación.

- **Lema de dominio.** Si dos actos generan el mismo par de ramificaciones y una de ellas es preferible a la otra (en una de ellas el premio es mayor), entonces el agente debe preferir el acto que asigna un peso mayor a la ramificación con el mayor premio.

En el próximo capítulo habrá lugar para hablar sobre las razones detrás de la aceptación de este tipo de lemas. Por ahora, centrémonos en lo que estos lemas consiguen y en lo que implican. Los tres lemas, en conjunto, vienen a mostrar que los pesos asociados a las ramificaciones juegan el mismo rol funcional que la probabilidad juega en un contexto clásico. Veamos sus análogos clásicos. Si dos apuestas posibles asignan la misma consecuencia a cada estado del mundo y cada estado del mundo tiene asignada la misma medida probabilidad (el mismo peso), entonces el agente es indiferente a dichas apuestas (equivalencia). Si una apuesta es sobre estados del mundo con probabilidad igual a cero, entonces el agente es indiferente a dicha apuesta (nulidad). Por último, si dos apuestas

son sobre dos estados del mundo y uno de ellos es preferible al otro, entonces uno debe favorecer la apuesta que asigna mayor probabilidad al estado más favorable (dominio).

Hay una diferencia crucial entre la medida de probabilidad usada en el contexto de la teoría de la decisión racional clásica y los pesos de la TDRE. En el caso clásico, dos agentes racionales con las mismas preferencias pueden ser representados por distintas medidas de probabilidad. Esto responde al hecho de que los agentes pueden tener creencias distintas sobre qué estados del mundo son más probables que otros (siempre y cuando la probabilidad asignada no sea vulnerable a argumentos de succión financiera, el agente es libre de poner la medida que crea más conveniente). En la TDRE, por otro lado, todo agente racional con las mismas preferencias actuará del mismo modo ante un problema de decisión cuántico. Esto es una consecuencia de que los pesos jueguen el mismo rol funcional que las probabilidades juegan (aspecto garantizado por los lemas) y del hecho de que el peso asignado a cada ramificación no depende de las creencias del agente.

Es difícil hacer demasiado énfasis en estas diferencias, pues podría argumentarse que el resultado más importante de la TDRE no está en lo que nos falta por ver, sino en la idea de que los pesos asociados a ramificaciones juegan el mismo rol funcional para un agente racional en un mundo everettiano que las probabilidades dadas por la regla de Born juegan en un mundo regido por la interpretación estándar. De este resultado a mostrar que los agentes actuarán *como si* la regla de Born es el caso sólo hay un paso.

3.6. El teorema de representación everettiano

Con todo lo anterior, el everettiano es capaz de probar un teorema de representación análogo al clásico. Esto es, se puede mostrar que existe una función de utilidad sobre el con-

3.6. El teorema de representación everettiano

junto de premios tal que un acto es preferible a otro si y sólo si su valor de expectación, calculado con respecto a esta función de utilidad y a los pesos asignados por la mecánica cuántica para cada ramificación, es mayor [Wallace, 2012, p. 172].

Capítulo 4

Evaluando la teoría de la decisión racional everettiana

Si la prueba de la TDRE es correcta, agentes racionales actúan del mismo modo tanto en universos con ramificaciones *a la Everett* como en universos con probabilidades *genuinas*. Al menos, esto es, en el contexto más bien artificial de las apuestas cuánticas que vimos en el capítulo anterior. ¿Cómo es que este resultado es relevante para el problema de la probabilidad everettiano?

El everettiano puede usar este resultado para resolver el problema de la probabilidad de dos maneras, los cuales veremos en breve más a detalle, que consisten en sostener que, o la TDRE resuelve el problema de la probabilidad completamente, o lo resuelve de forma parcial. La diferencia entre estas dos respuestas radica, como veremos, en una visión distinta sobre cómo entender la probabilidad dentro de la interpretación de muchos mundos.

En este capítulo haré dos cosas. Primero, presentaré cómo uno puede incorporar la

TDRE en una solución completa al problema de la probabilidad. Segundo, discutiré y evaluaré dos argumentos en contra del uso de este esquema para dar solución al problema de la probabilidad.

4.1. Estrategias para solucionar el problema de la probabilidad

En términos generales, existen al día de hoy dos propuestas para solucionar el problema de la probabilidad everettiano.¹ La primera opción, bautizada por Greaves [2004] como la *imagen de fisión* y defendida por Deutsch [1999], Greaves and Myrvold [2008] y Wallace [2003, 2005b, 2007, 2010, 2012], mantiene que en un universo como el descrito por la interpretación de muchos mundos no hay lugar para la probabilidad. Según esta propuesta, en un mundo determinista donde no hay ignorancia sobre condiciones iniciales, no tiene sentido hablar de la probabilidad de tal o cual resultado de un experimento cuántico. No obstante, la imagen de fisión defiende, todo rol importante que la probabilidad cumple (o parece cumplir) —como su papel en las predicciones cuánticas o como guía para la toma de decisiones— puede recuperarse desde este esquema a partir de la TDRE o de consideraciones similares.²

La segunda opción para resolver el problema de la probabilidad, que llamaré la *imagen de incertidumbre*, siguiendo a Albert [2015], mantiene que es posible hablar con sen-

¹En el pasado han habido otras estrategias que fueron abandonadas por diversas razones. Un ejemplo es la propuesta del mismo Everett, que mencionamos en 2.2, de recuperar las predicciones cuánticas a partir de una noción de tipicalidad. Véase Barrett [2014] para una discusión de estas alternativas.

²La propuesta de Greaves and Myrvold [2008], que será mencionada brevemente más adelante, no hace uso de la TDRE, sino más bien de lo que los autores consideran es un esquema para la comprobación de teorías que es sensible tanto a teorías probabilistas como a teorías como la mecánica cuántica everettiana.

4.1. Estrategias para solucionar el problema de la probabilidad

tido de probabilidades en un mundo everettiano. Tiene sentido puesto que, se argumenta, agentes viviendo dentro de un universo everettiano, pese a que el universo es completamente determinista, no tendrían conocimiento sobre ciertos aspectos de su posición en él (un agente podría no saber en qué ralmificación se encuentra, por ejemplo). Se ha argumentado de dos formas a favor de esta propuesta. A la primera la llamaré la *vía semántica*, sostenida por Saunders and Wallace [2008], Saunders [2010] y Wilson [2013]. En esta vía se argumenta que un análisis semántico de locuciones como ‘tengo incertidumbre sobre el resultado del experimento que está a punto de ser realizado’ muestra que en un mundo everettiano hay incertidumbre. A la segunda la llamaré la *vía de la ignorancia sobre la ramificación*, defendida por Vaidman [1998, 2012, 2014] y Tappenden [2009]. Esta vía mantiene que agentes viviendo en un universo everettiano *a la MMD* no sabrían en qué ramificación se encuentran después de haber realizado un experimento y no haber visto el resultado del mismo.

Tanto la imagen de fisión como la imagen de incertidumbre pueden usar la TDRE. Sin embargo, el papel de ésta es distinto en cada una. En la imagen de incertidumbre, la TDRE puede ser invocada para solucionar el problema cuantitativo de la probabilidad — *i.e.*, se puede invocar para encontrar el valor correcto que debemos asignarle a las distintas ramificaciones que resulten de realizar una medición. En la estrategia de fisión, por el otro lado, la TDRE es usualmente utilizada como solución completa del problema de la probabilidad.

El propósito del resto de este capítulo es examinar dos objeciones que se han levantado contra el uso de la TDRE para solucionar el problema de la probabilidad. La primera, que llamaré la *objeción de las frecuencias*, es sólo aplicable en el contexto de la imagen de

fisión. Esta objeción apunta a que el problema de la probabilidad no puede ser resuelto exclusivamente en términos de lo que un agente racional haría en un mundo everettiano. Esta crítica, hasta donde sé, sólo ha sido levantada por Albert [2015].

La segunda objeción, la *objeción de los descendientes elitistas* [Kent, 2010], pertenece a un tipo de objeción que ataca la prueba central de la TDRE —ataca la idea de que un agente en un mundo everettiano, si es racional, está obligado a actuar del mismo modo que un agente que sigue la mecánica cuántica estándar— a partir de sugerir estrategias de acción racional distintas a las del resultado de la TDRE. La idea está en que es racional, en un mundo everettiano, actuar siguiendo una regla distinta a la regla de Born. Este tipo de crítica es quizá la más común que se ha hecho contra la TDRE (Price [2010], Hemmo and Pitowsky [2007], Kent [2010] y Albert [2015]).³ No es mi intención examinar toda versión de este tipo de crítica, sino sólo centrarme en la versión defendida por Albert y Kent, la objeción de los descendientes elitistas.

4.2. La objeción de las frecuencias

La imagen de fisión parte de la idea de que el problema de la probabilidad everettiano puede solucionarse a partir únicamente de consideraciones sobre lo que un agente racional haría en un universo con ramificaciones *a la Everett*. La idea es que, si se consigue mostrar que agentes racionales deben actuar del mismo modo tanto en universos everettianos como en universos con probabilidades *genuinas* —cosa que, supuestamente, consigue la TDRE— entonces todo problema al que hacíamos referencia al hablar del problema de probabilidad ha sido resuelto.

³En [Wallace, 2012, pp. 189-198], el autor presenta no menos de ocho versiones de este tipo de objeción.

Contra esta idea, David Albert ha argumentado que el proyecto entero de la imagen de fisión está dirigido a resolver el problema equivocado. El problema real consiste, nos dice Albert, en dar cuenta de las frecuencias que observamos al realizar experimentos (o la apariencia de las mismas), no en dar cuenta del comportamiento de agentes racionales, cosa que en última instancia es resultado de lo que observamos:

Son esas frecuencias, o (en cualquier caso) la apariencia de dichas frecuencias, y no el comportamiento al apostar al que dan resultado en última instancia, lo que principalmente y en principio necesita de una explicación científica. Y el pensamiento de que uno podría librarse de explicar esas frecuencias o sus apariencias, el pensamiento de que uno podría ser capaz de evadir explicar dichas frecuencias o su apariencia —el cual es el pensamiento central de la estrategia de la teoría de la decisión racional— es (cuando piensas al respecto) disparatado [Albert, 2015, p. 166].⁴

Podemos presentar la objeción de las frecuencias de la siguiente forma:

- 1 Si la imagen de fisión resuelve el problema de la probabilidad, entonces debe dar cuenta de las frecuencias con las que nos encontramos al realizar experimentos cuánticos.
- 2 La imagen de fisión no es capaz de dar cuenta de las frecuencias que observamos al realizar experimentos cuánticos.

⁴Traducción propia. Original: “It is those frequencies, or (at any rate) the appearance of those frequencies, and not the betting behaviors to which they ultimately give rise, which primarily and in the first instance stands in need of a scientific explanation. And the thought that one might be able to get away without explaining those frequencies or their appearances, the thought that one might be able to make some sort of an end run around explaining those frequencies or their appearances —which is the central thought of the decision-theoretic strategy— is (when you think about it) mad”[Albert, 2015, p. 166].

∴ La imagen de fisión no resuelve el problema de la probabilidad everettiano.

La crítica de Albert, puesta de este modo, tiene la estructura de un *modus tollens*. Siendo su estructura simple, lo que queda por ver aquí es la fortaleza de sus premisas.

Sobre la premisa uno. En la cita anterior, Albert hace referencia a la necesidad de una explicación científica de las frecuencias que observamos. Sin embargo, como “explicación” es un concepto demasiado cargado, y como no es necesario usarlo para levantar la crítica de Albert, aquí hablo en su lugar de *dar cuenta* de las frecuencias. Más allá de la expresión que uno prefiera, lo importante está en notar que una solución al problema de la probabilidad everettiano necesita establecer una conexión entre la mecánica cuántica everettiana y lo que observamos (sea aparente o no). Si la mecánica cuántica everettiana no consigue establecer esta conexión, entonces no tiene forma de ser confirmada o desmentida por la experiencia y, por lo tanto, deja de ser una teoría empírica en cualquier sentido.

Sobre la premisa dos. Primero. La imagen de fisión está comprometida con la idea de que el comportamiento de agentes racionales en un mundo everettiano es suficiente para dar cuenta de todo papel que la probabilidad juega en un mundo con probabilidades *genuinas*. Esto no debería ser polémico pues se desprende de cómo hemos caracterizado a la imagen de fisión. De esto se sigue que una justificación de la segunda premisa puede hacerse a partir de señalar que no hay forma de establecer una conexión desde la imagen de fisión entre el comportamiento de agentes racionales en un mundo everettiano y las frecuencias que observamos.

Segundo. El comportamiento de agentes racionales al que llega la TDRE en su conclusión no puede dar cuenta de las frecuencias que observamos pues es completamente

independiente de las mismas. Esto puede verse en el hecho de que la TDRE no hace uso en ningún lugar de consideraciones empíricas para determinar cómo un agente racional debería actuar. Como no hace alusión a frecuencias, ni a probabilidades (recordemos que aquí estamos considerando la imagen de fisión), su resultado es aplicable en cualquier universo *sin importar qué frecuencias se obtengan al realizar experimentos cuánticos*.

Otra forma de ver este mismo punto está en notar que, según la TDRE, un agente racional en un universo everettiano debe actuar siguiendo la regla de Born con independencia a las frecuencias que observa.

Alguien podría pensar que lo que acabo de escribir es de alguna forma incorrecto. Cómo actúa un agente y cómo es el mundo (*i.e.*, las frecuencias que observamos) *deben* tener algún tipo de relación. Si un agente racional es indiferente a apostar, en el contexto de un juego con una moneda, sobre si la moneda caerá en cara o cruz, entonces —si asumimos que está interesado en maximizar su utilidad— podemos inferir que dicho agente asigna la misma probabilidad subjetiva tanto a la moneda cayendo en cara como a la moneda cayendo en cruz. Si, además, suponemos que el agente tiene conocimiento completo sobre todo aspecto relevante de la moneda, entonces podemos inferir no sólo información sobre sus creencias sino también información sobre la probabilidad de la moneda: tanto cara como cruz son igualmente probables. En resumen. La objeción de las frecuencias parece incorrecta pues, al menos en ocasiones, el comportamiento de un agente racional puede hacernos inferir la presencia y el valor ciertas probabilidades.

El problema con este tipo de razonamiento está en que, en la imagen de fisión, un agente racional everettiano no actúa a partir de consideraciones sobre probabilidad: simplemente no hay probabilidades desde esta forma de ver las cosas. Como no hay proba-

bilidades que guíen su comportamiento, no hay forma de usar su comportamiento para recobrar dichas probabilidades.

Otra forma de ver este punto es la siguiente. Tanto la TDRE como la teoría de la decisión racional clásica buscan probar que un agente racional prefiere el acto *A* al acto *B* si y sólo si el valor de expectación de *A* es mayor al de *B*. Pero, hay una diferencia crucial en el significado del valor de expectación en estos contextos. En el caso clásico, el valor de expectación de un acto es la suma de todas las probabilidades de los distintos resultados multiplicadas por la utilidad asociada a cada uno de estos resultados. En la TDRE, por otro lado, el valor de expectación de un acto es la suma de las utilidades asociadas a cada ramificación multiplicadas por un número que no es, al menos desde la imagen de fisión, una probabilidad.⁵ Estos números, los pesos cuánticos, no son, recordemos, de nuevo, probabilidades. Incluso si pudiéramos inferir siempre el peso cuántico asociado a cada ramificación a partir de cómo un agente actúa, dicho número no tendría por qué estar asociado a las frecuencias que observamos.

Y este asunto —que en la TDRE no haya conexión alguna entre la forma de actuar de un agente racional y cómo es el mundo— implica que el everettino no sólo no nos explica las frecuencias que observamos, sino que tampoco explica por qué apostamos como apostamos. En nuestro mundo, las acciones de agentes racionales dependen de forma crucial de cómo es el mundo.

Lo que necesitamos es dar cuenta de nuestra experiencia empírica de frecuencias. Y lo que se nos promete (que cae enteramente por debajo de lo que necesitamos) es una explicación de por qué es que apostamos como aposta-

⁵¿Qué son, entonces, estos coeficientes que aparecen asociados a cada ramificación? Hilary Greaves considera que, desde la imagen de fisión, debemos considerarlos una *medida de cuidado* —una medida del nivel en el que nos interesamos en la situación descrita por cierta ramificación.

mos. Y lo que se nos da (que cae enteramente por debajo de lo que se nos prometió) es un argumento al efecto de que si nosotros mantuviéramos un conjunto completamente distinto de convicciones sobre el mundo de las que realmente tenemos, nosotros apostaríamos del mismo modo a como en la realidad apostamos [Albert, 2015, pp. 166-167].⁶

En resumen, la objeción de las frecuencias apunta a una falta de conexión entre la imagen de fisión y las frecuencias que observamos al realizar experimentos cuánticos. Lo que nos ofrece la TDRE, y que no es suficiente, es una prueba al efecto de que un agente racional viviendo en un mundo everettiano actuaría, en circunstancias distintas a las nuestra, del mismo modo a como nosotros lo hacemos.

Una objeción similar puede hacerse contra el proyecto de Greaves and Myrvold [2008] de generar un esquema bayesiano de comprobación de teorías que sea sensible a teorías científicas tanto probabilistas como con ramificaciones. La intuición de este proyecto es usar la regla de Bayes:

$$Pr(T | E) = \frac{Pr(E | T) Pr(T)}{Pr(E)}$$

como base para actualizar nuestra confianza en distintas teorías físicas, sean estas probabilistas o no.⁷

En el contexto de la propuesta de Greaves y Myrvold, uno lee la regla del Bayes de forma tal que se convierte en una manera de *actualizar* nuestro nivel de confianza en

⁶Traducción propia, original: “What we need is an account of our actual empirical experience of frequencies. And what we are promised (which falls entirely short of what we need) is an account of why it is that we bet as we do. And what we are given (which falls entirely short of what we were promised) is an argument to the effect that if we held an altogether different set of convictions about the world than the ones we actually hold, we would bet the same way as we actually do” [Albert, 2015, pp. 166-167].

⁷La regla de Bayes, también conocida como el teorema de Bayes, es un resultado matemáticamente

cierta teoría ante la aparición de nueva evidencia ($Pr(T | E)$). El teorema nos dice que esta confianza debe ser actualizada multiplicando la probabilidad que dicha teoría predice para la evidencia en cuestión ($Pr(E | T)$) por nuestro nivel de confianza previo en la teoría ($Pr(T)$) dividido sobre la probabilidad de dicha pieza de evidencia ($Pr(E)$).

El problema con esta idea está en que no hay forma de hacer sentido de la regla de Bayes en algo como la imagen de fisión: uno de los *inputs* de esta regla —la probabilidad que la teoría asigna a cierta evidencia, $P(E | T)$ — no tiene sentido desde esta propuesta. Esto se debe a que esta imagen no nos provee una conexión entre la teoría y lo que observamos (no hace, en efecto, ninguna predicción), no tiene sentido este input.

4.3. La objeción de los descendientes elitistas

El tipo de crítica más común contra la TDRE apunta a la existencia de otras alternativas de acción racional, distintas a la de seguir la regla de Born, en el contexto de un problema de decisión cuántico (Price [2010], Hemmo and Pitowsky [2007], Kent [2010] y Albert [2015]). La idea está en que, si hay otras formas racionales de actuar, entonces un agente no debe —so pena de ser irracional— comportarse como si siguiera la regla de Born. En este contexto nos centraremos únicamente en la forma de esta crítica que llamaré, siguiendo a Adrian Kent, *la objeción de los descendientes elitistas* y que ha sido levantada por Kent [2010] y Albert [2015].

simple y, en apariencia, poco importante:

$$\begin{aligned} Pr(A \wedge B) &= Pr(A | B)Pr(B) \\ &= Pr(B | A)Pr(A) \\ Pr(A | B) &= \frac{Pr(B | A)Pr(A)}{Pr(B)} \end{aligned}$$

Lo interesante, y problemático, de este teorema está en las distintas interpretaciones que puede tener.

Para presentar esta objeción, será útil comenzar recordando al lema de equivalencia, 3.5: Si dos actos asignan el mismo peso a las mismas ramificaciones y cada una de éstas recibe el mismo premio, entonces el agente debe ser indiferente a seguir uno o el otro.

Un ejemplo:

Sea $|\alpha\rangle \otimes |\text{premio}\rangle$ un estado donde un agente recibe cierta cantidad de dinero y el resto del mundo es $|\alpha\rangle$, y $|\beta\rangle \otimes |\text{no premio}\rangle$ un estado donde el agente no recibe dinero y el resto del mundo es $|\beta\rangle$. Ahora, supongamos que hay dos actos, A y B, que son disponibles para el agente y que producirán los siguientes resultados:

$$A : 1/\sqrt{2}|\alpha\rangle \otimes |\text{premio}\rangle + 1/\sqrt{2}|\beta\rangle \otimes |\text{no premio}\rangle$$

$$B : 1/\sqrt{2}|\alpha\rangle \otimes |\text{no premio}\rangle + 1/\sqrt{2}|\beta\rangle \otimes |\text{premio}\rangle$$

Asumiendo que el agente únicamente está interesado en maximizar su riqueza, en particular asumiendo que es indiferente a si $|\alpha\rangle$ o $|\beta\rangle$ se obtiene, equivalencia nos dice que el agente debe ser indiferente entre estas apuestas [Wallace, 2012, pp. 172-173].⁸

La objeción de los descendientes elitistas mantiene, contrario a equivalencia, que podría ser racional para un agente actuar conforme a lo que ocurre en $|\alpha\rangle$ y $|\beta\rangle$. Como ejemplo, Albert sugiere que alguien podría actuar siguiendo la siguiente regla:

Medida de la gordura: Sea el grado de mi cuidado por lo que ocurre en cierta ramificación proporcional a qué tan gordo soy en esa ramificación multiplicado por el cuadrado del valor absoluto de la amplitud asociada a esa

⁸En las primeras formas de la TDRE, equivalencia era introducida como una premisa [Wallace, 2003]. Posteriormente wallace2010 la presentó, justo como lo hicimos en el capítulo anterior, como un lema. No obstante esto, como veremos más adelante, la objeción de los descendientes elitistas corre en ambos casos.

ramificación [Albert, 2015, p. 168].⁹

La medida de la gordura es sólo una de las formas —en este caso, jocosa— de la objeción de los descendientes elitistas. Lo importante está en que alguien, de forma racional, podría preferir el acto A sobre el acto B debido a la diferencia entre $|\alpha\rangle$ y $|\beta\rangle$. Uno podría usar otros criterios para seleccionar entre $|\alpha\rangle$ y $|\beta\rangle$ distintos del peso del agente. Por ejemplo, un agente podría pensar que la medida debe tomar en cuenta el coeficiente intelectual de sus descendientes (caso en el que tendría más sentido la palabra ‘elitista’), o algo arbitrario como el número de calcetines en cada una de las ramificaciones.

Por supuesto, algunas de estas estrategias de acción no tienen nada de *natural*. ¿Por qué un agente estaría interesado en que a sus descendientes con más peso les vaya mejor? Sin embargo, la cuestión no es si estas estrategias son naturales —sea lo que sea que signifique esto— sino que, en principio, no hay nada de irracional en un agente que cree que el mejor modo de actuar es siguiendo una de estas estrategias.

La importancia de equivalencia para la TDRE es enorme. Como vimos en el capítulo anterior, equivalencia, junto con los lemas de nulidad y dominio, garantizan que un agente tomará los pesos asociados a ramificaciones *como si* fueran probabilidades *para todo fin práctico*. Albert lo pone del siguiente modo: “Resulta que suponer equivalencia equivale a suponer que cualquier cosa que juega el rol funcional de las probabilidades dependerá de ninguna otra cosa que de las amplitudes de la mecánica cuántica”[Albert, 2015, p. 167].¹⁰

Sin equivalencia, un agente racional no tiene por qué comportarse como si los pesos

⁹Traducción propia. Original: “Let the degree to which I care about what transpires on a certain branch be proportional to how fat I am on that branch multiplied by the absolute square of the amplitude associated with that branch.”

¹⁰Traducción propia. Original “It turns out that supposing equivalence amounts to supposing that whatever it is that plays the functional role of probabilities can depend on nothing other than the quantum-mechanical amplitudes.”

cuánticos fueran probabilidades.

David Wallace ha respondido a esta objeción, en particular al ejemplo de la medida de la gordura, señalando que una estrategia de este tipo viola, de hecho, uno de los axiomas de racionalidad de la TDRE: consistencia diacrónica.

(La medida de la gordura) viola consistencia diacrónica. El agente propuesto por Albert es (*ex hypothesi*) indiferente sobre si hace dieta o no. Pero él no es indiferente sobre si sus futuros ‘yoes’ hacen dieta: él quiere que aquellos en ramificaciones con buenos resultados ganen peso, y que aquellos en ramificaciones con malos resultados pierdan peso [Wallace, 2012, 190-191].¹¹

Consistencia diacrónica, 3.5, mantiene que las preferencias de un agente deben permanecer iguales a lo largo del tiempo en un sentido especial: si un agente prefiere algún acto *A* después de haber hecho el acto *B*, sin importar las consecuencias obtenidas después de hacer *B*, entonces el mismo agente prefiere, antes de que el acto *B* haya sido llevado a cabo, hacer *B* seguido por *A*. Wallace afirma que un agente siguiendo la medida de la gordura viola esta restricción: en general, y ante una situación donde no hay ramificación alguna, dicho agente es indiferente a hacer dieta; pero, ante un evento con ramificaciones, el agente prefiere que sus futuras copias ganen peso.

Si uno acepta que hay conflicto entre seguir algo como la medida de la gordura y consistencia diacrónica, entonces la justificación de este axioma de racionalidad se vuelve central. Si el axioma es justificable, entonces la respuesta de Wallace es sólida. A favor de consistencia diacrónica, Wallace dice “Las acciones de un agente toman tiempo para

¹¹Traducción propia. Original: “(the fatness measure) violates diachronic consistency. Albert’s agent is (*ex hypothesi*) indifferent to dieting. But he is not indifferent to whether his future selves diet: he wants the ones on branches with good outcomes to gain weight, and the ones on branches with bad outcomes to lose weight.”

llevarse a cabo; sus deseos y objetivos toman tiempo para ser completados. Si sus preferencias no se mantienen consistentes durante este periodo temporal, la acción deliberativa no es posible de ninguna manera.”[Wallace, 2012, p. 168]¹² También, Wallace sostiene que consistencia diacrónica, siendo razonable en un contexto clásico, es indispensable en un universo everettiano: “En la presencia de amplias violaciones de consistencia diacrónica, la agencia en un universo everettiano no es posible de ninguna forma”[Wallace, 2012, p. 168].¹³

Así, la razón que Wallace ofrece para aceptar consistencia diacrónica se basa en la supuesta necesidad de dicho axioma para que la acción deliberativa sea posible. En general, como la toma de decisiones toma tiempo, es necesario que los agentes mantengan sus deseos *más o menos* iguales. Además, la necesidad de consistencia diacrónica es patente en un contexto everettiano: como el universo se encuentra dividiéndose en incontables ramificaciones cada segundo, si las preferencias de un agente son inconsistentes entre ramificaciones —esto es, si dichas preferencias violan consistencia diacrónica— entonces la acción racional es imposible.

Contra esta forma de argumentar en favor de consistencia diacrónica, Adrian Kent ha señalado que este axioma, de hecho, es violado constantemente en el mundo real: Las preferencias de un agente cambian con el tiempo debido, entre otras cosas, a acciones que dicho agente ha tomado en su pasado y cuyos resultados afectaron al agente de formas que no pudo haber predicho. Lo mejor que podemos decir de consistencia diacrónica en el mundo real, por lo tanto, es que en muchas ocasiones, a corto plazo, aproximadamente es

¹²Traducción propia. Original: “An agent’s actions take time to carry out; his desires and goals take time to be realized. If his preferences do not remain consistent over this timescale, deliberative action is not possible at all”[Wallace, 2012, p. 168].

¹³Traducción propia. Original: “In the presence of widespread, generic violation of diachronic consistency, agency in the Everett universe is not possible at all”[Wallace, 2012, p. 168].

verdad —lo cual, claramente, no es suficiente para tratar a este principio como un axioma de racionalidad [Kent, 2010, pp. 342-343].

Por supuesto, el everettiano siempre puede responder que el axioma de consistencia diacrónica es una restricción normativa acerca de cómo debería comportarse un agente *si es que es racional*, no una descripción sobre cómo, de hecho, se actúa en el mundo. La cuestión no es si en el mundo real los agentes siguen dicho principio, sino si es racional seguirlo (otros axiomas de racionalidad, como transitividad, por ejemplo, son violados continuamente en mundo real. Sin embargo, pocas personas estarían dispuestas a concluir a partir de este tipo de consideraciones que uno debe abandonar transitividad).

Otra forma de argumentar en contra de la razón que Wallace ofrece para aceptar consistencia diacrónica está en notar que ese tipo de razón sólo es válida si asumimos que la acción racional es posible en un mundo everettiano. Su argumento es que consistencia diacrónica es necesario para que haya racionalidad. Sin embargo, ¿por qué debería ser posible la acción racional en un mundo everettiano? Otra forma de ver la situación en la siguiente: Quizá la acción racional no es posible en un contexto everettiano y la mejor evidencia que tenemos de esto está en que dicha acción requiere de axiomas de racionalidad que no son necesarios en el contexto clásico.

David Albert ha argumentado que consistencia diacrónica no es un principio de racionalidad razonable. . Consistencia diacrónica requiere que nuestras preferencias entre futuros donde no hay ramificaciones de alguna forma determinen nuestras preferencias entre futuros con ramificaciones [Albert, 2015, p. 170]. La toma de decisiones en un universo con ramificaciones es un contexto nuevo y distinto al clásico, de modo que no parece justificado suponer que las mismas consideraciones que aplicamos en el clásico

son aplicables al contexto con ramificaciones.

En otras palabras, consistencia diacrónica asume que la presencia de ramificaciones a la hora de tomar decisiones es estéril (o casi estéril) para las preferencias de un agente. Esto parece injustificado. Ante mi inminente ramificación en x copias, parece razonable, al menos en principio, que mis consideraciones y preferencias sean distintas a las que tengo en un contexto donde no hay ramificaciones.

Hasta aquí hemos visto dos respuestas a las críticas de Wallace que conceden que seguir algo como la medida de la gordura entra en conflicto con el axioma de consistencia diacrónica. Hay, no obstante, otra forma de defender esta objeción que no hace esta concesión. Según esta respuesta, seguir algo como la medida de la gordura no tiene que ver con las preferencias de un agente y, por lo tanto, no puede entrar en conflicto con un axioma de racionalidad como consistencia diacrónica que, por definición, sólo compete a las preferencias del agente. Esta vía parece poco razonable debido a cómo es redactada la medida de la gordura: “Sea el grado de mi cuidado por lo que ocurre en cierta ramificación proporcional a qué tan gordo soy en esa ramificación multiplicado por el cuadrado del valor absoluto de la amplitud asociada a esa ramificación” [Albert, 2015, p. 168].

Si uno interpreta, en cita anterior, a “el grado de mi cuidado” como una cuestión acerca de las preferencias de un agente, entonces es fácil ver cómo esta medida y consistencia diacrónica entran en conflicto. Sin embargo, esta interpretación puede resistirse. Esto es, es perfectamente razonable suponer que un agente considere las diferencias entre $|\alpha\rangle$ y $|\beta\rangle$ como relevantes para sus decisiones en el contexto de una apuesta cuántica manteniendo, al mismo tiempo, que el agente no tiene preferencia alguna entre $|\alpha\rangle$ y $|\beta\rangle$. Para motivar este razonamiento, recordemos que tanto el teorema de representación

4.3. La objeción de los descendientes elitistas

clásico como el everettiano mantienen que un agente prefiere el acto A sobre el B si y sólo si el valor de expectación de A es mayor al de B . El valor de expectación de un acto X está definido como:

$$EU(X) = \sum_{i=1}^N Pr(\gamma_i)u(X(\gamma_i)) \quad (4.1)$$

donde, sin embargo, la interpretación de cada uno de los elementos de esta ecuación es distinta en el contexto clásico y el everettiano. En el caso clásico, el valor de expectación es la suma de todos los posibles estados del mundo multiplicados por la utilidad asociada a cada consecuencia (donde, recordemos, una consecuencias es el resultado de que el agente haya tomado cierto curso de acción y de que el mundo resulte ser de determinado modo). En el caso everettiano, por otro lado, el valor de expectación es la suma de las medidas asociadas a cada ramificación (los pesos cuántos, argumenta el everettiano; la medida de gordura, argumenta Albert) multiplicados por la utilidad asociada a cada premio (los cuales, recordemos, son las partes de las ramificaciones que son de interés para el agente). El cuadro 4.1 resume estas diferencias.

Cuadro 4.1: Diferentes formas de interpretar el valor de expectación de un acto.

Contexto	γ_i	$Pr(\gamma_i)$	$X(\gamma_i)$	$u(X(\gamma_i))$
Clásico	Estados del mundo	Probabilidad subjetiva	Consecuencias	Utilidad
Everettiano	Ramificaciones	Medida relevante	Premios	Utilidad

La diferencia que quiero resaltar aquí está es en relación a la interpretación de $Pr(\gamma_i)$. Según el contexto clásico, esta es una medida de probabilidad que representa las creencias del agente (si, por ejemplo, el agente cree que una moneda es justa, entonces esta medida asigna 50% tanto a que caiga cara como a que caiga cruz). Esta medida, por otro lado, es

todo un misterio en el caso everettiano: éste pretende argumentar que un agente racional *debería* tratar a los pesos cuánticos como si fueran una medida de probabilidad. Pero, la objeción de los descendientes elitistas, a partir de ejemplos como el de la medida de la gordura, nos muestra que siempre es racional para un agente utilizar una medida que dependa de las diferencias entre las ramificaciones relevantes en una apuesta cuántica. Y la razón por la cual es racional hacer esto no tiene que, necesariamente, depender de que el agente crea que una ramificación es preferible a otra. La preferencia del agente de $|\alpha\rangle$ sobre $|\beta\rangle$ puede depender, exclusivamente, de cosas que el agente cree (en contraste con cosas que el agente prefiere). La medida de gordura es un ejemplo: el agente podría considerar que esta medida debe ser sensible al peso de sus *yoés*, pese a que el mismo agente no prefiere ganar más peso.

En el caso everettiano, esta medida no está relacionada con las creencias del agente, ni a información que éste posee, ni a sus preferencias (pues, recordemos, para eso está la función de utilidad), ni a cuántos mundos emergentes hay (después de todo, recordemos, no hay forma de contarlos), *ni, en realidad, cuando uno lo piensa detenidamente, a nada*. Y este es el problema de la TDRE que la objeción de los descendientes elitistas expone: los pesos cuánticos, sin una regla para interpretarlos —como la regla de Born, en la mecánica cuántica estándar— no representan nada físicamente.

Conclusiones

La teoría de la decisión racional everettiana no resuelve el problema de la probabilidad que enfrenta la interpretación de muchos mundos. No es una solución completa (pues no puede dar cuenta de las frecuencias que observamos) ni parcial (pues no consigue probar que agentes racionales viviendo en mundos everettianos están obligados a actuar siguiendo la regla de Born). Esta es la idea central que he querido defender en esta tesis.

A lo largo de cada capítulo, sin embargo, hubo otras ideas que guiaron la redacción de este texto y que vale la pena mencionar aquí. El primer capítulo buscó convencer al lector de que el problema de la medición es un problema real e importante para todo aquel que piense en la mecánica cuántica como una teoría que aspira a tener un carácter fundamental. El segundo capítulo exploró las distintas formas que ha tomado la interpretación de muchos mundos; espero haber hecho esto de forma tal que estas interpretaciones no parezcan sólo una vía natural de resolver el problema de la medición sino también un área de investigación fértil. El tercer capítulo buscó introducir al lector la propuesta de solucionar el problema de la probabilidad everettiano a partir de la teoría de la decisión racional; espero haber logrado exponer de forma clara y correcta las motivaciones e ideas centrales de esta propuesta pese al tipo de exposición que escogí (ni muy formal ni completamente informal).

Es hasta el capítulo cuatro que presento mi caso a favor de la tesis principal de este texto —otra vez, que la teoría de la decisión racional no soluciona el problema de la probabilidad everettiano— a partir de clarificar y defender dos objeciones que, respectivamente, llamamos la *objeción de las frecuencias* y la objeción de los descendientes elitistas. Siendo que ambas objeciones no son de mi autoría, quisiera decir algo sobre lo que, creo, es mi aportación a la discusión en este tema. Lo que busqué hacer en el caso de la objeción de las frecuencias fue presentar esta objeción de forma distinta y, espero, más clara (dejando explícitas sus premisas) y argumentar a favor de las mismas. Con la objeción de los descendientes elitistas, por otro lado, ofrecí una defensa al tipo de críticas que se han hecho en su contra.

Quiero terminar haciendo algunas observaciones sobre la relación entre este texto y la interpretación de muchos mundos. Esta interpretación es una de las tres soluciones al problema de la medición más discutidas en la actualidad —siendo las otras dos la mecánica de Bohm y las interpretaciones de colapso objetivo. La probabilidad y los temas alrededor suyo —como la comprobación empírica— han sido una dificultad para las interpretaciones everettianas desde que éstas fueron creadas. Si es el caso, como he argumentado en este texto, que la probabilidad sigue siendo un problema para esta interpretación, ¿qué juicio es razonable hacer acerca de esta solución al problema de la medición?

Por un lado, el hecho de que esta interpretación no se haya librado del problema de la probabilidad pese a décadas de intentos por resolverlo es desalentador. Por el otro, no hay que olvidar que la última estrategia propuesta para resolver este problema (el uso de la teoría de la decisión racional) es relativamente nueva. Siendo esta la situación, no parece razonable ni abandonar el proyecto everettiano ni continuarlo de forma indefini-

4.3. La objeción de los descendientes elitistas

da (después de todo, sin una solución al problema de la probabilidad, difícilmente puede decirse que esta interpretación constituye un tipo de teoría científica). Quizá lo más razonable para los próximos años sea examinar las distintas formas mediante las cuales se puede argumentar que el problema de la probabilidad está resuelto usando la teoría de la decisión racional, siempre teniendo en mente que los recursos intelectuales destinados a esta interpretación son más difíciles de justificar mientras más y más tiempo pase sin que haya una forma de solucionar sus problemas centrales. Esa sería mi recomendación.

CAPÍTULO 4. EVALUANDO LA TEORÍA DE LA DECISIÓN RACIONAL EVERETTIANA

Agradecimientos

Primero, agradezco a mi tutor, Elias Okon, quien me introdujo a las discusiones sobre los fundamentos de la mecánica cuántica y me enseñó a pensar sobre estos temas. Su tiempo, dedicación y trabajo han sido fundamentales para este trabajo. Si alguna de las proposiciones en este texto es verdadera, se debe más a Elias que mí.

También agradezco a mis lectores: Cristian Alejandro Gutiérrez Ramírez, Fernanda Samaniego, Ana Rosa Pérez Ransanz, y a Daniel Sudarsky por sus observaciones y críticas a las primeras versiones de este trabajo.

A mis amigos Ian Quallenberg y Ricardo Santiago les agradezco en general el tiempo que dedicaron a escuchar sobre las ideas expuestas en esta tesis y, en particular, a sus demandas por claridad cuando mis explicaciones lo requerían.

Por último, agradezco el apoyo que recibí por parte del CONACYT durante el tiempo que estudié la maestría en filosofía de la ciencia en la UNAM (2012-2014).

CAPÍTULO 4. EVALUANDO LA TEORÍA DE LA DECISIÓN RACIONAL EVERETTIANA

Bibliografía

- David Albert. *Quantum Mechanics and Experience*. Harvard University Press, 1992.
- David Albert. Probability in the everett picture. In Simon Saunders, Jonathan Barret, Adrian Kent, and David Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*, pages 355–268. Oxford University Press, 2010.
- David Albert. *After Physics*, chapter Probability in the Everett Picture. Harvard University Press, 2015.
- David Albert and Barry Loewer. Interpreting the many worlds interpretation. *Synthese*, 1988.
- V. Allori, S. Goldstein, R. Tumulka, and N. Zanghi. Many-Worlds and Schroedinger’s First Quantum Theory. *ArXiv e-prints*, March 2009. URL <http://adsabs.harvard.edu/abs/2009arXiv0903.2211A>.
- Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental tests of realistic local theories via bell’s theorem. *Phys. Rev. Lett.*, 47:460–463, Aug 1981. doi: 10.1103/PhysRevLett.47.460. URL <http://link.aps.org/doi/10.1103/PhysRevLett.47.460>.
- J Barrett and P. (eds) Byrne, editors. *The Everett Interpretation of Quantum Mechanics: Collected Works 1955–1980 with Commentary*. Princeton University Press, 2012.
- Jeffrey Barrett. Everett’s relative-state formulation of quantum mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition, 2014. URL <http://plato.stanford.edu/archives/fall2014/entries/qm-everett/>.
- John Stewart Bell. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press, 1987.
- David Deutsch. *Comment on Lockwood*, volume 47. British Journal for the Philosophy of Science, 1996. pp. 222-8.
- David Deutsch. Quantum theory of probability and decisions. *Proceedings of the Royal Society of London*, A455, 1999.
- Bryce DeWitt. Quantum mechanics and reality. *Physics Today*, 23(9), 1970.
- P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 1958.

BIBLIOGRAFÍA

- Hugh Everett. On the foundations of quantum mechanics, 1957.
- Richard Feynman. *The Feynman Lectures on Physics*. California Institute of Technology, 2010. URL <http://www.feynmanlectures.caltech.edu/>.
- Hilary Greaves. Understanding deutsch's probability in a deterministic multiverse, April 2004. URL <http://philsci-archive.pitt.edu/1742/>. To appear in *Studies in the History and Philosophy of Modern Physics*, September 2004.
- Hilary Greaves. Probability in the everett interpretation, November 2006. URL <http://philsci-archive.pitt.edu/3103/>. Forthcoming in *Philosophy Compass*.
- Hilary Greaves and Wayne Myrvold. Everett and evidence, August 2008. URL <http://philsci-archive.pitt.edu/4222/>.
- M. Hemmo and I. Pitowsky. Quantum probability and many worlds. *Studies in History and Philosophy of Modern Physics*, 38(333-50), 2007.
- Hanneke Janseen. Reconstructing reality. Master's thesis, Radboud University Nijmegen, 2008.
- Adrian Kent. One world versus many: The inadequacy of everettian accounts of evolution, probability, and scientific confirmation. In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.
- Henry Krips. Measurement in quantum theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2013 edition, 2013. URL <http://plato.stanford.edu/archives/fall2013/entries/qt-measurement/>.
- Tim Maudlin. Three measurement problems. *Topoi*, 1995.
- Tim Maudlin. Can the world be only wavefunction? In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.
- Wayne C. Myrvold. What is a wavefunction?, June 2014. URL <http://philsci-archive.pitt.edu/10750/>.
- David Papineau. A fair deal for everettians. In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.
- Huw Price. Probability in the everett picture. In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.
- M.L.G. Redhead. *Incompleteness, Nonlocality and Realism*. Oxford: Clarendon Press, 1989.

- Simon Saunders. *Time, Quantum Mechanics, and Probability*, volume 114. Synthese, 1996. URL <http://users.ox.ac.uk/~lina0174/part3.pdf> for via TDNet. pp.373-404.
- Simon Saunders. Chance in the everett interpretation. In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. 2010.
- Simon Saunders and David Wallace. Branching and uncertainty, January 2008. URL <http://philsci-archive.pitt.edu/3811/>. This is a revised version of the paper of the same name deposited on 12 June 2007, as it will appear in the British Journal for the Philosophy of Science.
- Leonard Savage. *The foundations of statistics*. New York:Dover, 1972.
- Erwin Schrödinger. The present situation in quantum mechanics. *Proceedings of the American Philosophical Society*, (124):323–38, 1980. Traducción de John D. Trimmer.
- Paul Tappenden. Evidence and uncertainty in everett’s multiverse, December 2009. URL <http://philsci-archive.pitt.edu/5046/>. This is the final version of the paper which has been accepted by the British Journal for the Philosophy of Science for publication in 2010.
- Lev Vaidman. On schizophrenic experiences of the neutron or why we should believe in the many-worlds interpretation of quantum theory. *International Studies in the Philosophy of Science*, 1998.
- Lev Vaidman. Probability in the many-worlds interpretation of quantum mechanics. In Ben-Menahem et al, editor, *Probability in Physics*. 2012.
- Lev Vaidman. Many-worlds interpretation of quantum mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014. URL <http://plato.stanford.edu/archives/spr2014/entries/qm-manyworlds/>.
- Lev Vaidman. Many-worlds interpretation of quantum mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2015 edition, 2015. URL <http://plato.stanford.edu/archives/spr2015/entries/qm-manyworlds/>.
- von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.
- Daid Wallace. *The Emergent Multierse*. Oxford University Press, 2012.
- David Wallace. Epistemology quantized: Circumstances in which we should come to believe in the everett interpretation. *British Journal for the Philosophy of Science*, year = 2006.
- David Wallace. Everettian rationality: Defending deutsch’s approach to probability in the everett interpretation. *Studies in History and Philosophy of Modern Physics*, 2003.
- David Wallace. *Quantum Probability from Subjective Likelihood: improving on Deutsch’s proof of the probability rule*. unpublished, 2005a. URL <http://philsci-archive.pitt.edu/id/eprint/2302>.

BIBLIOGRAFÍA

- David Wallace. *Quantum Probability from Subjective Likelihood: improving on Deutsch's proof of the probability rule*. unpublished, 2005b. URL <http://philsci-archive.pitt.edu/id/eprint/2302>.
- David Wallace. Quantum probability from subjective likelihood: Improving on deutsch's proof of the probability rule. *Studies in History and Philosophy of Modern Physics*, 2007.
- David Wallace. How to prove the born rule. In A. Kent Saunders, J. Barrett and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.
- David Wallace. Decoherence and Ontology, or: How I Learned To Stop Worrying And Love FAPP. *ArXiv e-prints*, 2011.
- Alastair Wilson. Objective probability in everettian quantum mechanics. *The British Journal for the Philosophy of Science*, 64(4):709–737, 2013. doi: 10.1093/bjps/axs022. URL <http://bjps.oxfordjournals.org/content/64/4/709.abstract>.
- Heinz Dieter Zeh. On the interpretation of measurement in quantum theory. *Foundations of Physics*, 1(1):69–76, 1970. ISSN 1572-9516. doi: 10.1007/BF00708656. URL <http://dx.doi.org/10.1007/BF00708656>.