

LENGUAJE SUBYACENTE A LA NOCIÓN DE MÁQUINA CUÁNTICA

Informe final de investigación

Andrés Sicard Ramírez

María Eugenia Puerta Yepes

Mario Elkin Vélez Ruiz

Departamento de Ciencias Básicas
Escuela de Ciencias y Humanidades
Universidad EAFIT
Medellín, Colombia, S.A.

1998

Índice general

1. Contexto y planteamiento del problema	7
1.1. Contexto del problema	7
1.2. Descripción de la investigación	12
1.3. Objetivos	12
1.3.1. Objetivo general	12
1.3.2. Objetivos específicos	12
2. Algunos preliminares matemáticos	13
2.1. Algunos elementos de la lógica	13
2.2. Algunos elementos del campo de los números complejos	14
2.3. Algunos elementos del análisis funcional	18
2.3.1. Espacios métricos	18
2.3.2. Espacios vectoriales	24
2.3.3. Espacios normados y espacios de Banach	28
2.3.4. Espacios con producto interior y espacios de Hilbert	29
2.3.5. Operadores	32
3. Algunos elementos de la mecánica cuántica	35
3.1. Observaciones iniciales	35
3.2. Comentarios al libro de Dirac: “Principios de M.C.”	36
3.3. Definiciones adicionales	61
4. Máquina cuántica	63
4.1. Elementos estáticos	63
4.1.1. Estados	64
4.1.2. Memoria	65
4.1.3. Posición actual	65
4.2. Elementos dinámicos	66
4.3. Observación final	69
5. Conclusiones	71
Bibliografía	77

Introducción

Es imposible definir donde comienza e inclusive donde termina un proyecto de investigación. En nuestro caso hemos decidido encuadrar nuestro proyecto de investigación *Lenguaje subyacente a la noción de máquina cuántica* entre dos proyectos; el primero de ellos *Máquinas de Turing dinámicas*, anterior a éste y el otro de ellos *¿Máquina de Turing cuántica autorrefencial: Una posibilidad?*, posterior a éste.

La investigación que realizamos en el año 1997 *Máquinas de Turing Dinámicas* estaba inscrita un campo problemático que hemos denominado *la ampliación de la noción de computabilidad*. En este proyecto, se acometió la posibilidad de ampliar la noción de computabilidad al interior de la teoría de las máquinas de Turing. La idea era representar la propiedad de *autorrefencia* en una máquina de Turing, es decir, formalizar la capacidad de que una máquina de Turing automodifique su comportamiento. Los resultados de imposibilidad obtenidos [35] no nos desanimaron, por el contrario nos han permitido madurar el problema y buscar otras alternativas y otras teorías donde abordarlo.

La primera vez que escuchamos acerca de la computación cuántica y en particular acerca de las máquinas cuánticas, fue por intermedio de Roger Penrose [27]. Aunque esta lectura inicial no fue muy prometedora, nos referenció la obra de David Deutsch [8]. La primera lectura de Deutsch y las posteriores relecturas del mismo, sirvieron para dejar en claro una cosa, no conocíamos ni entendíamos el *código* de la máquina cuántica. Por código de la máquina cuántica entendemos el lenguaje en el que ella es desarrollada y este lenguaje es el lenguaje de la mecánica cuántica. En un instante nos tropezamos con conceptos tales como “observable”, “autovector”, “espacio de Hilbert”, entre otros; para los cuales no teníamos ningún marco de referencia. Este proyecto de investigación está dirigido a construir esta competencia conceptual.

La investigación que hemos propuesto para el año 1999 *¿Máquina de Turing Cuántica Autorrefencial: Una Posibilidad?* es la continuación y el producto natural de nuestros dos proyectos de investigación anteriores. Conjeturamos que una máquina de Turing cuántica, puede automodificar su comportamiento, con base en los postulados establecidos por la mecánica cuántica, en relación a la alteración del sistema, cuando se realiza una observación sobre éste.

En relación a los contenidos de este informe, tenemos los siguientes: El primer capítulo introduce el marco de referencia sobre el cual se trabaja; el segundo capítulo expone algunos conceptos matemáticos sobre los cuales se construye la mecánica cuántica; el tercer capítulo presenta algunos conceptos fundamentales de la mecánica cuántica; finalmente, el cuarto capítulo desarrolla parcialmente la noción de máquina cuántica.

Es importante subrayar la naturaleza elemental de este proyecto. La construcción de este informe está pensada en un lector (que inicialmente somos nosotros) sin la enciclopedia conceptual (física y matemática) necesaria para acercarse a la noción de la máquina cuántica. A excepción hecha del primer capítulo, de algunos comentarios del tercero y de las conclusiones que sugerimos leer a cualquier lector, proponemos al lector hojear cada uno de los capítulos y comenzar a leer a partir desde donde considere necesario, esto por supuesto, pensando en las diferentes competencias de nuestros posibles lectores.

Finalmente deseamos agradecer a la Universidad EAFIT el apoyo para la realización de este proyecto

de investigación. Además deseamos agradecer a Jorge Iván García, Raúl Gómez, Fernando Gúzman, Juan Guillermo Lalinde, Nancy López, Jesús Mira, Daniel Velásquez y otros que en estos momentos se nos escapan, por la conversaciones sostenidas sobre diferentes aspectos de este proyecto.

Capítulo 1

Contexto y planteamiento del problema

1.1. Contexto del problema

Hemos decidido comenzar la presentación de nuestro proyecto de investigación señalando el contexto en el cual está inscrito, convencidos que cualquier problema de investigación está de alguna forma ampliado-delimitado, potenciado-disminuido, definido-indefinido por el contexto filosófico-teórico-científico en el cual se inserta y que una lectura de éste sin aquel, sería una lectura en el mejor de los casos técnica, pero definitivamente sería una lectura muy reducida del problema en cuestión.

Para nuestro problema en particular, nuestro contexto es la *posibilidad de ampliar la noción de computabilidad*. Presentamos a continuación algunas reflexiones realizadas entorno a esta problemática (publicadas en [34]).

A partir de la definición (informal) de computabilidad ofrecida por Soare: *"A computation is a process whereby we proceed from initially given objects, called inputs, according to a fixed set of rules, called a program, procedure, or algorithm, through a series of steps and arrive at the end of these steps with a final result, called the output. The algorithm, as a set of rules proceeding from inputs to output, must be precise and definite, with each successive step clearly determined. The concept of computability concerns those objects which may be specified in principle by computations, ..."* ([38], pàg. 286)

Es plausible interpretar la noción de computabilidad como una propiedad atribuible o no a cierta clase de objetos —se hablará entonces, de objetos computables y objetos no computables—. La clase de objetos que permite la pregunta por la computabilidad o no de sus miembros es muy heterogénea; por citar algunos ejemplos: es posible hablar de funciones computables o no computables (este es el objeto utilizado por la teoría de la computabilidad), de números computables o no computables (este fue el objeto seleccionado por Turing en el artículo fundacional de las máquinas de Turing [39]) y en una instancia de mayor cobertura, de procesos computables o no computables (objeto seleccionado por Penrose y tratado exhaustivamente en [27] y [28]). Una particularidad muy significativa de la propiedad de computabilidad, es su carácter binario excluyente, es decir, un objeto es o no es computable, esta característica genera una partición del universo de los objetos.

¿Cómo identificar la propiedad de computabilidad o no computabilidad de un objeto? Es necesario contar con un criterio de demarcación que construya la frontera **continua** entre los objetos computables y los que

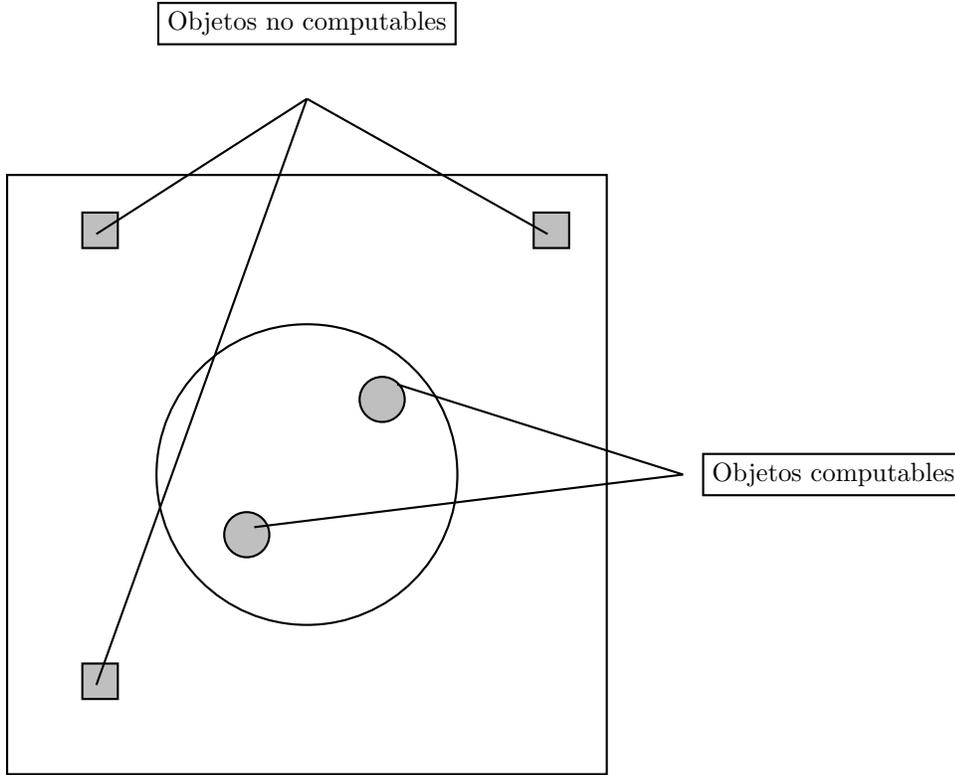


Figura 1.1: Partición del universo de los objetos: Objetos computables y objetos no computables.

no lo son, esta frontera es la definición misma de computabilidad, pero no una definición informal (como la ofrecida por Soare), sino una definición formal que permita distinguir sin equívoco un objeto computable de un objeto no computable. En la actualidad existen diferentes “versiones” para la definición formal de computabilidad —versiones que como es conocido, satisfacen la propiedad de ser coextensivas—; la versión con la cual se va a operar en este artículo es la correspondiente a las máquinas de Turing: un objeto es computable si es computable por una máquina de Turing, es decir, un objeto es computable si es Turing-computable.

La propiedad de computabilidad dota a los objetos *en principio* de una inteligibilidad completa, un objeto computable es un objeto conocido, es un objeto cuya aprehensión es plausible, es un objeto sintética y analíticamente descriptible, pero el precio que se debe retribuir por este “dominio” del objeto es muy elevado, un objeto computable es un objeto simple, es un objeto trivial (en el sentido de las máquinas triviales de von Foerster [41]).

En el mundo formal o en el mundo factual existen objetos que no son computables, por el lado formal se presentan por ejemplo el problema de la parada de una máquina de Turing [39] y el problema de la teselación [28]; por el lado factual, se mencionan procesos tales como la morfogénesis [40] y la cognición [28].

La clasificación binaria ejercida por la noción de computabilidad actúa como un paradigma-filtro que

clasifica los objetos en duplas: dominados y no dominados, simples y complejos, triviales y no triviales. Es frecuente que la ciencia realice grandes esfuerzos en “pulir” sus objetos para que éstos crucen el filtro y se conviertan así, en objetos aprehensibles; esta es la versión del paradigma de la simplicidad [24] observado desde la perspectiva de la computabilidad.

La definición formal de computabilidad ha sido puesta en entredicho desde sus primeras formulaciones. Con alguna frecuencia se han presentado personas que consideran que la definición de computabilidad no es completa ni definitiva, personas que han intentado romper el paradigma de la computabilidad por medio de definiciones más potentes de la misma. Se presentan apartes de la carta-respuesta enviada por Alonzo Church a József Pépés, con relación a la propuesta (implícita) de Pépés de una definición más poderosa de computabilidad. La carta cobra importancia en la medida que es escrita por Church, quien es el creador de una de las “versiones” formales de computabilidad (λ -cálculo). Por otra parte, la excelente explicación de Church de las consecuencias de contar con una definición más potente de computabilidad permite justificar el continuar con su búsqueda, aunque también ofrece argumentos bastante sólidos de la imposibilidad de encontrarla; pero de mayor importancia, es la posición de escepticismo adoptada por Church, muy diferente a la posición dogmática adoptada por algunos en la actualidad.

“Dear Mgr. [Monsignore] Pépés: ... In reply to your postal [card] I will say that I am very much interested in your results on general recursiveness, and hope that I may soon be able to see them in detail. In regard to your project to construct an example of a numerical function which is effectively calculable but not general recursive I must confess myself extremely skeptical - although this attitude is of course subject to the reservation that I may be induced to change my opinion after seeing your work.”

“...Therefore to discover a function which was effectively calculable but no general recursive would imply discovery of an utterly new principle of logic, not only never before formulated, but never before actually used in a mathematical proof - since all extant mathematics is formalizable within the system of Principia, or at least within one of its known extensions. Moreover this new principle of logic must be of so strange, and presumably complicated, a kind that its metamathematical expression as a rule of inference was not general recursive (for this reason, if such a proposal of a new principle of logic were ever actually made, I should be inclined to scrutinize the alleged effective applicability of the principle with considerable care).” ([36], págs. 175–176)¹

Este es el contexto en el cual se reflexiona sobre la posibilidad de ampliar la definición de computabilidad: Se presenta inicialmente una representación gráfica (figura 1.2) de las características que se considera debe poseer una nueva definición de computabilidad.

¹Algún lector podrá objetar que la propuesta de Pépés no es ampliar la noción de computabilidad, sino de refutar la tesis de Church-Turing. En algunos párrafos posteriores, se presenta la relación entre la tesis de Church-Turing y la propuesta de una nueva definición de computabilidad.

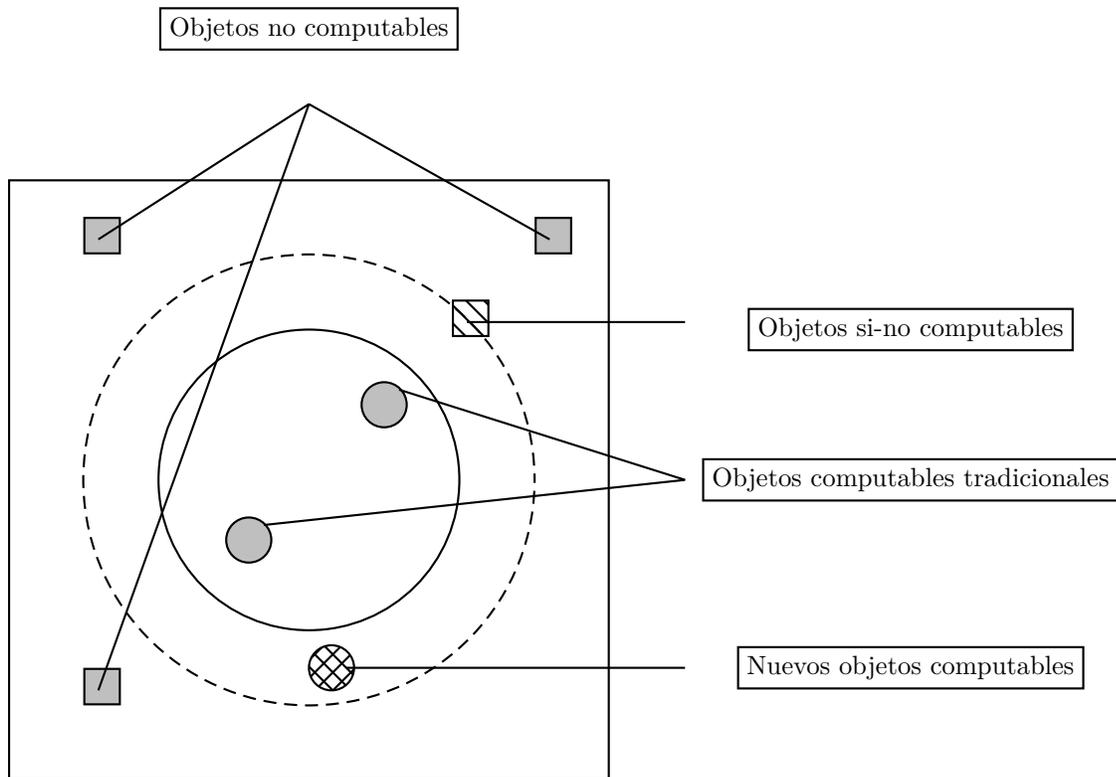


Figura 1.2: Ampliación de la definición de computabilidad.

En la figura 1.2 se observa que de nuevo existen objetos que no son computables (representados por un cuadrado), objetos computables tradicionales —es decir aquellos que son computables bajo la definición actual de computabilidad (representados por un círculo)— y objetos que bajo la definición actual de computabilidad no son computables, pero que bajo la ampliación de la definición sí lo son (representados por una círculo con rayas). La transformación de objetos no computables en objetos computables exige *a-priori* una jerarquía de la no computabilidad, es decir, deben existir diferentes grados de no computabilidad, de lo contrario, si todos los objetos no computables comparten el mismo grado de no computabilidad, una vez se logre transformar uno de ellos en computable, se lograría transformarlos a todos.²

Pero de mayor importancia, la **discontinuidad** presente en la frontera construida por la nueva definición de computabilidad —a diferencia de la continuidad bajo la definición actual—, representa el compromiso con

²Una vez elaborada esta idea, el autor supo de la existencia de estos grados de no computabilidad a partir de [31], bibliografía que una vez revisada, permitirá complementar esta idea.

la propuesta moriniana: la ratificación de la complejidad. Se admite las limitaciones y mutilaciones causadas por los métodos simplificadores y, se reconoce la necesidad de contar con nuevos operadores-principios-definiciones complejos para construir una imagen más fiel de lo real. Los objetos no computables son objetos complejos por naturaleza, incluso se puede afirmar que la complejidad inherente a ellos (por lo menos en muchos casos), es la causa de su no computabilidad. Características tales como recursividad organizacional, emergencia de propiedades, inconsistencia, indeterminismo, etc., hacen parte de sus cualidades y por extensión de los obstáculos epistemológicos que emergen en el intento de aprehenderlos. La propuesta de la complejidad, quiebra los principios de la lógica clásica para permitir que se incorporen nuevas posibilidades de verdad. En el contexto de la nueva computabilidad se acepta y se espera la nueva categoría de objetos sí-no computables (representados por cuadrado con rayas). Por supuesto, esto es inadmisibles bajo la mirada consistente y binaria de la lógica-ciencia actual, pero esta es precisamente la propuesta de Edgar Morin, en sus palabras:

*“No se trata de retomar la ambición del pensamiento real simple de **controlar y dominar lo real**. Se trata de ejercitarse en un pensamiento capaz de tratar, de dialogar, de negociar, con lo real.”* (las negrillas son nuestras) ([24], pàg. 22).

Un magnifico ejemplo del fuerte tejido que existe entre la complejidad, la computabilidad y la ampliación de la computabilidad, lo ofrecen algunos trabajos realizados para construir modelos de los sistemas vivos [30, 15]. Kampis por una parte, a partir del hecho de que la evolución es uno de las principales características de los sistemas vivos y que ésta produce innovaciones en el sistema, las cuales aumentan la complejidad del mismo; y por otra parte, a partir de las limitaciones de los modelos computables actuales en donde es necesario conocer el futuro antes de que este pueda ser computado, es decir, una computación debe saber de antemano que va a computar; concluye acerca de la imposibilidad de utilizar la metáfora de la máquina computable para describir dichos sistemas. Además, dado que los sistemas vivos presentan la propiedad de **auto-modificar** su comportamiento (propiedad que escapa a ser modelada por reglas *a-priori*) es necesario contar con nuevos modelos de computabilidad, que sean capaces de **modificar su comportamiento en tiempo de ejecución**.

Por otra parte, con base en los comentarios realizados por Soare, en relación con la aceptación de la tesis de Church-Turing:

*“This may be viewed as roughly analogous to Euclidean geometry or Newtonian physics capturing a large part of everyday geometry or physics, but not necessarily all conceivable parts. Here, Turing has captured the notion of a function computable by a mechanical procedure, and as yet there is no evidence for any kind of computability which is not included under this concept. **If it existed, such evidence would not affect Turing’s Thesis about mechanical computability any more than hyperbolic geometry or Einsteinian physics refutes the laws of Euclidean geometry or Newtonian physics. Each simple describes a different part of the universe.***

*... Some have cast doubt on Turing’s Thesis on the grounds that there might be physical or biological processes which may processes, say, the characteristic function of the halting problem. **It is possible that these may exist (although there is presently no evidence) but if so, this will have absolutely no effect on Turing’s Thesis because they will not be algorithmic or mechanical procedures as required in ... and in Turing’s Thesis.***” ([38], pàgs. 294–295; las negrillas son nuestras)

Se evita incursionar en la relación entre la amplitud de la computabilidad y la tesis de Church-Turing. Esto no quiere decir que se ignore el campo problemático planteado por ella. Una vez ampliada la definición de computabilidad, sería necesario regresar a la tesis de Church-Turing y reflexionarla con respecto a la nueva definición, quizás para construir una tesis ampliada (como lo es la relación entre la física de Newton y la física de Einstein), quizás para construir una tesis alternativa (como lo es la relación entre la geometría euclidiana y la geometría hiperbólica).

Hecha la apuesta por la posibilidad de ampliación de la definición de computabilidad y por extensión directa, el fortalecimiento de los aparatos de captura de lo real, surgen las siguientes preguntas: ¿Cuál es la posibilidad de repetir este proceso?, ¿Se puede repetir *ad-infinitum* o tiene límite?, ¿Cuáles son los alcances de estas múltiples ampliaciones? En términos generales la pregunta es por los límites de la ciencia, con la complejidad a bordo por supuesto y con la noción de objeto computable —bajo nuevas y más potentes definiciones de computabilidad— como uno de sus elementos de base. Se defiende la infinitud del proceso, es decir existe optimismo en la capacidad humana para aumentar sus aparatos de cognición, pero aunque el proceso es infinito tiende a una asíntota insuperable, es decir se acepta la incompletitud final del proceso. Este es un juego contra lo real, en el cual no es posible triunfar. Este es el sentido trágico de esta aventura llamada ciencia, pero él mismo constituye su esencia.

1.2. Descripción de la investigación

Desde un punto de vista estrictamente matemático, el edificio conceptual utilizado por la mecánica cuántica es la aplicación a un caso en particular, de ciertos constructos matemáticos generales, bajo ciertas condiciones y restricciones. Esta investigación va a desarrollar simultáneamente ambas direcciones, por un lado, se van a presentar —vía definiciones— los conceptos y teoremas utilizados con el grado de generalidad usual en matemáticas, y por otro lado, se van a presentar las aplicaciones —vía construcciones— de dichos conceptos y teoremas en el contexto de la mecánica cuántica.

En términos generales la investigación propuesta va a discurrir sobre dos conceptos: El primero de ellos es el concepto de *estado de un sistema cuántico*, el cual se asocia con los vectores de un espacio de Hilbert denominado como $L^2[-\infty, +\infty]$ (técnicamente, los vectores pertenecen a un subespacio de $L^2[-\infty, +\infty]$); el segundo de ellos es el concepto de *observable de un sistema cuántico*, el cual se asocia con operadores lineales hermíticos en dicho subespacio.

1.3. Objetivos

1.3.1. Objetivo general

Apropiar el lenguaje matemático subyacente a la noción de máquina cuántica.

1.3.2. Objetivos específicos

1. Apropiar el concepto de espacio de Hilbert y sus propiedades.
2. Presentar las características y los elementos del espacio de Hilbert usado por la mecánica cuántica.
3. Apropiar los conceptos de operador y funcional, y sus propiedades.
4. Presentar las características de los operadores y funcionales usados por la mecánica cuántica.

Capítulo 2

Algunos preliminares matemáticos

Esperamos que el nombre de este capítulo “Algunos preliminares matemáticos” deje inmediatamente claro al lector el carácter parcial e incompleto de su contenido. Este capítulo presenta algunos conceptos matemáticos sobre los cuales opera la mecánica cuántica. Este capítulo implícitamente supone la relación matemáticas-física, en un sólo sentido; primero el concepto matemático y luego el concepto físico (aunque, reconocemos por supuesto la existencia de la relación en el otro sentido). Este sentido implícito ha evitado que realicemos referencia en este capítulo a los conceptos físicos, por el contrario, el próximo capítulo, hará referencia a los conceptos de este capítulo. Los conceptos en matemáticas usualmente presuponen otros conceptos y éstos a su vez otros y así, sucesivamente. Esto conduce a que la elección de los conceptos *a-priori* para comprender un nuevo concepto este sujeta a las preferencias de cada autor, esta situación no ha sido diferente en nuestro caso.

2.1. Algunos elementos de la lógica

Presentamos la definición de un lenguaje de la lógica de predicados de primer orden, para este lenguaje construimos la noción de modelo. Con base en estos dos conceptos presentaremos las definiciones de campo y espacio vectorial.

Definición 2.1. *Un lenguaje de la lógica de predicados de primer orden está definido por [21]:*

$$\mathbb{L} = \{\{P_i, i \in I\}, \{F_j, j \in J\}, \{C_k, k \in K\}\} \text{ donde,}$$

$\{P_i, i \in I\}$: Conjunto de símbolos de predicado
 $\{F_j, j \in J\}$: Conjunto de símbolos de funciones
 $\{C_k, k \in K\}$: Conjunto de símbolos de constantes

Definición 2.2. *Sea \mathbb{L} un lenguaje de la lógica de primer orden, un **modelo** para \mathbb{L} está definido por [21]:*

$$\mathcal{U} = \langle A, \mathcal{I} \rangle \text{ donde,}$$

A : Conjunto no vacío, llamado dominio del modelo
 \mathcal{I} : Es una función biyectiva de interpretación tal que:

\neg -1 Cada símbolo P_i^n , de predicado de aridez n , es interpretado por una relación n -ádica R , es decir:

$$\neg(P_i^n) = R \iff R \subseteq A^n \quad (2.1)$$

\neg -2 Cada símbolo F_j^m , de función de aridez m , es interpretado por una función m -ádica, es decir:

$$\neg(F_j^m) = f \quad y \quad f : A^m \rightarrow A \quad (2.2)$$

\neg -3 Cada símbolo de constante C_k es interpretado por un elemento fijo de A , es decir:

$$\neg(C_k) = t \quad y \quad t \in A, \quad t \text{ fijo} \quad (2.3)$$

2.2. Algunos elementos del campo de los números complejos

Definición 2.3. Sea $\mathbb{L} = \{F_1^2, F_2^2, 0, 1\}$ un lenguaje de la lógica de primer orden compuesto por dos símbolos de función de aridez dos F_1^2 y F_2^2 y dos símbolos de constantes 0 y 1. Un **campo**¹ es un modelo de \mathbb{L} denotado por

$$\mathcal{C} = \langle X, +, \cdot, 0, 1 \rangle, \text{ donde}$$

X : Es un conjunto no vacío.

$+$: Es una función $+$: $X \times X \rightarrow X$ llamada **adición**.

\cdot : Es una función \cdot : $X \times X \rightarrow X$ llamada **multiplicación**.

0: Es llamado el elemento neutro de \mathcal{C} bajo la adición. $0 \in X$.

1: Es llamado el elemento identidad de \mathcal{C} bajo la multiplicación. $1 \in X$.

El modelo \mathcal{C} debe satisfacer los siguientes axiomas:

Axiomas para la operación adición:

$$\forall x \forall y \in X (x + y \in X) \quad (\text{clausurativa } +) \quad (C-1)$$

$$\forall x \forall y \in X (x + y = y + x) \quad (\text{conmutativa } +) \quad (C-2)$$

$$\forall x \forall y \forall z \in X (x + (y + z) = (x + y) + z) \quad (\text{asociativa } +) \quad (C-3)$$

$$\forall x \in X (x + 0 = 0 + x = x) \quad (\text{elemento neutro para } +) \quad (C-4)$$

$$\forall x, \exists (-x) \in X (x + (-x) = 0) \quad (\text{existencia elementos inversos bajo } +) \quad (C-5)$$

Axiomas para la operación multiplicación:²

$$\forall x \forall y \in X (x \cdot y \in X) \quad (\text{clausurativa } \cdot) \quad (C-6)$$

$$\forall x \forall y \in X (x \cdot y = y \cdot x) \quad (\text{conmutativa } \cdot) \quad (C-7)$$

$$\forall x \forall y \forall z \in X (x \cdot (y \cdot z) = (x \cdot y) \cdot z) \quad (\text{asociativa } \cdot) \quad (C-8)$$

$$\forall x \in X (x \cdot 1 = 1 \cdot x = x) \quad (\text{elemento neutro para } \cdot) \quad (C-9)$$

$$\forall x \in X - \{0\}, \exists x' \in X (x \cdot x' = 1) \quad (\text{existencia elem. inversos para } \cdot) \quad (C-10)$$

¹La estructura o modelo de campo, usualmente se introduce a partir de clases especiales de estructuras más simples llamadas grupos y anillos. Remitimos el lector a alguna bibliografía de álgebra moderna si está interesado en la presentación de los campos a partir de dichas estructuras, sugerimos mirar el texto [11].

²Observe el lector que la existencia de elementos inversos bajo la operación de multiplicación (axioma C-10), excluye al elemento 0X de satisfacer esta propiedad.

Axiomas operaciones adición y multiplicación

$$\forall x \forall y \forall z \in X (x \cdot (y + z) = x \cdot y + x \cdot z) \quad (\text{distributiva derecha}) \quad (\text{C-11})$$

$$\forall x \forall y \forall z \in X ((y + z) \cdot x = y \cdot x + z \cdot x) \quad (\text{distributiva izquierda}) \quad (\text{C-12})$$

Uno de los elementos sobre el cual opera la mecánica cuántica es el campo de los números complejos. Presentamos a continuación algunas definiciones de dicho campo.

Definición 2.4. *Un **número complejo** z es un par ordenado de números reales (a, b) y cada par ordenado de números reales (a, b) es un número complejo ([19], pág. 225).*

*De acuerdo a lo anterior, un número z es un elemento de \mathbb{R}^2 . El conjunto \mathbb{R}^2 es llamado el **conjunto de los números complejos** y es designado usualmente por \mathbb{C} . Es decir, $\mathbb{C} = \mathbb{R}^2 = \{z = (a, b) \mid a, b \in \mathbb{R}\}$.*

Definición 2.5. *Sean $z_1 = (a, b)$ y $z_2 = (c, d)$ números complejos. La **suma** $z_1 + z_2$ se define por:*

$$z_1 + z_2 = (a, b) + (c, d) = (a + c, b + d) \quad (2.4)$$

Definición 2.6. *Sean $z_1 = (a, b)$ y $z_2 = (c, d)$ números complejos. La **mutiplicación** $z_1 z_2$ se define por:*

$$z_1 z_2 = (a, b)(c, d) = (ac - bd, ad + bc) \quad (2.5)$$

Teorema 2.1. *El conjunto de los números complejos, la operación de suma de números complejos, la operación de multiplicación de números complejos, el elemento neutro para la suma de números complejos definido por $(0, 0)$ y el elemento identidad para la multiplicación de números complejos definido por $(1, 0)$ forman un campo. Es decir el modelo $\mathcal{C} = \langle \mathbb{C}, +, \cdot, (0, 0), (1, 0) \rangle$ es un campo.*

Una vez establecido el campo de los números complejos, presentamos la definición de la unidad imaginaria i , además presentamos los conceptos de conjugado y módulo de un número complejo.

Definición 2.7. *Los números complejos son asociados con la unidad imaginaria i . Se presenta la definición del número i , y se representa cada número complejo de la forma $z = (a, b)$ en terminos de la unidad imaginaria.*

Se representa el número complejo $(0, 1)$ por i . Entonces:

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1. \quad (2.6)$$

*De aquí, se observa que en el conjunto de los números complejos, la raíz cuadrada de -1 es i , es decir, $i = \sqrt{-1}$, el número i es llamado la **unidad imaginaria**.*

*Cada número complejo de la forma $z = (a, b)$ puede ser escrito como $a + ib$, llamada la **forma rectangular** del número complejo (a, b) ([19], pág. 236). La representación de un número complejo en la forma $z = a + ib$ es la universalmente preferida ([26], pág. 4) y será la adoptada en este informe.*

Las operaciones de suma y multiplicación de números complejos, pueden ser reescritas en términos de la forma rectangular de los mismos. Sean $z_1 = a + ib$ y $z_2 = c + id$ números complejos, entonces:

$$z_1 + z_2 = (a + ib) + (c + id) = (a + c) + i(b + d) \quad (2.7)$$

$$z_1 z_2 = (a + ib)(c + id) = (ac - bd) + i(ad + bc) \quad (2.8)$$

Definición 2.8. Sea $z = a + ib$ un número complejo. La **parte real** de z designada $Re(z)$ es el número real a y la **parte imaginaria** de z designada $Im(z)$ es el número real b ([26], pág. 4).

Definición 2.9. Sea $z = a + ib$ un número complejo. El **inverso aditivo** de z designado por $-z$ es el número complejo $-z = -(a + ib) = -a + i(-b)$ ([19], pág. 237).

Definición 2.10. Sean $z_1 = a + ib$ y $z_2 = c + id$ números complejos. La **resta** $z_1 - z_2$ se define por ([19], pág. 238):

$$\begin{aligned} z_1 - z_2 &= z_1 + (-z_2) \\ &= (a + ib) + [-(c + id)] \\ &= (a + ib) + (-c + i(-d)) \\ &= (a - c) + i(b - d) \quad (\text{por 2.7}) \end{aligned} \tag{2.9}$$

Definición 2.11. Sea $z = a + ib$ un número complejo. El **conjugado** de z designado \bar{z} es el número complejo definido por ([26], pág. 5):

$$\bar{z} = a - ib \tag{2.10}$$

Teorema 2.2. El conjugado de un número complejo satisface las siguientes propiedades ([26], pág. 5).

Sean $z = a + ib$, $z_1 = a + ib$, $z_2 = c + id$ números complejos, entonces:

$$\overline{\bar{z}} = z \tag{2.11}$$

$$\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2 \tag{2.12}$$

$$Re(z) = \frac{(z + \bar{z})}{2} \tag{2.13}$$

$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2 \tag{2.14}$$

Definición 2.12. Sea $z = a + ib$ un número complejo. El **módulo** de z designado $|z|$ se define por ([26], pág. 6; [10], pág. 11):

$$\begin{aligned} |z| &= \sqrt{\bar{z}z} \\ &= \sqrt{(a - ib)(a + ib)} \\ &= \sqrt{a^2 + b^2} \end{aligned} \tag{2.15}$$

Teorema 2.3. El módulo de un número complejo satisface las siguientes propiedades ([26], pág. 6; [10], pág. 11):

Sean $z = a + ib$, $z_1 = a + ib$, $z_2 = c + id$ números complejos, entonces:

$$|z|^2 = z\bar{z} = a^2 + b^2 \quad (|z|^2 \text{ es llamado el cuadrado del módulo}) \tag{2.16}$$

$$|Re(z)| \leq |z| \tag{2.17}$$

$$|z| = |\bar{z}| \tag{2.18}$$

$$|z_1 z_2| = |z_1| |z_2| \tag{2.19}$$

$$|z_1 + z_2| \leq |z_1| + |z_2| \quad (\text{desigualdad triangular}) \tag{2.20}$$

$$|z_1 - z_2|^2 = (a - c)^2 + (b - d)^2 \tag{2.21}$$

Demostración. Vamos a demostrar la propiedad descrita por la ecuación (2.20) debido a que ilustra de manera muy eficaz el uso de las propiedades del conjugado y del módulo de un número complejo.

La propiedad a demostrar es que $|z_1 + z_2| \leq |z_1| + |z_2|$. Vamos a realizar la demostración para $|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$ y tomando raíz cuadrada a ambos lados, se obtendrá la demostración pedida.

$$\begin{aligned}
|z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} && \text{(por 2.16)} \\
&= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) && \text{(por 2.12)} \\
&= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2 && \\
&= z_1\bar{z}_1 + z_1\bar{z}_2 + \overline{z_1\bar{z}_2} + z_2\bar{z}_2 && \text{(por 2.11)} \\
&= |z_1|^2 + z_1\bar{z}_2 + \overline{z_1\bar{z}_2} + |z_2|^2 && \text{(por 2.16)} \\
&= |z_1|^2 + 2\operatorname{Re}(z_1\bar{z}_2) + |z_2|^2 && \text{(por 2.13)} \\
&\leq |z_1|^2 + 2|\operatorname{Re}(z_1\bar{z}_2)| + |z_2|^2 && \\
&\leq |z_1|^2 + 2|z_1\bar{z}_2| + |z_2|^2 && \text{(por 2.17)} \\
&= |z_1|^2 + 2|z_1||\bar{z}_2| + |z_2|^2 && \text{(por 2.19)} \\
&= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 && \text{(por 2.18)} \\
&= (|z_1| + |z_2|)^2 &&
\end{aligned}$$

□

Uno de los principales “usos” que hace la mecánica cuántica del campo complejo, esta relacionado con las funciones de valor complejo de variable real. Presentamos entonces, dichas funciones y algunas de sus propiedades.

Definición 2.13. Una *función de valor complejo* $\psi(c)$ de una variable compleja, es una función $\psi: A \rightarrow \mathbb{C}$, donde $A \subset \mathbb{C}$.

La función $\psi(c)$ puede ser expresada de la forma $\psi(c) = u(c) + iv(c)$, donde $u(c) = \operatorname{Re}(\psi(c))$, $v(c) = \operatorname{Im}(\psi(c))$ e i es la unidad imaginaria. Las funciones $u(c)$ y $v(c)$ son funciones definidas por $u: A \rightarrow \mathbb{R}$ y $v: A \rightarrow \mathbb{R}$, es decir, son funciones de valor real de una variable compleja.

Para efectos de “eliminar” la variable compleja de las funciones $u(c), v(c)$, se identifica el número complejo c por $c = a + ib$ y se obtiene $u(c) = u(a, b)$ y $v(c) = v(a, b)$, entonces $u(a, b)$ y $v(a, b)$ son funciones de valor real de dos variables reales ([26], pág. 19).

Definición 2.14. Si $\psi(x)$ es una función de valor complejo de una variable compleja ($\psi: A \rightarrow \mathbb{C}$) y $A = \mathbb{R}$, entonces $\psi(x)$ es llamada *función de valor complejo de una variable real* ($\psi: \mathbb{R} \rightarrow \mathbb{C}$). La función $\psi(x)$ puede entonces expresarse por:

$$\psi(x) = u(x) + iv(x) \tag{2.22}$$

donde $u(x), v(x)$ son funciones de valor real de una variable real, e i es la unidad imaginaria.³

Las funciones de valor complejo de una variable real, “heredan” las operaciones definidas sobre los números complejos. Entonces, sean $\psi(x) = u(x) + iv(x)$, $\psi_1(x) = u_1(x) + iv_1(x)$ y $\psi_2(x) = u_2(x) + iv_2(x)$ funciones de valor complejo de una variable real:

³El lector puede observar que en el caso de una función de valor complejo de una variable real $\psi(x)$, las funciones $u(x)$ y $v(x)$ son funciones de valor real de una variable compleja; no son funciones de valor real de dos variables reales, como hubiera sido el caso si $\psi(x)$ fuera una función de valor complejo de una variable compleja.

La suma $\psi_1(x) + \psi_2(x)$ está definida por:

$$\psi_1(x) + \psi_2(x) = (u_1(x) + u_2(x)) + i(v_1(x) + v_2(x)) \quad (2.23)$$

La compleja conjugada de $\psi(x)$ está definida por:

$$\psi^*(x) = u(x) - iv(x) \quad (2.24)$$

El cuadrado del módulo de $\psi(x)$ está definido por:

$$|\psi(x)|^2 = \psi^*(x)\psi(x) = u^2(x) + v^2(x) \quad (2.25)$$

La operación $|\psi_1(x) - \psi_2(x)|^2$ está definida por:

$$\begin{aligned} |\psi_1(x) - \psi_2(x)|^2 &= (\psi_1(x) - \psi_2(x))^*(\psi_1(x) - \psi_2(x)) \\ &= (u_1(x) - u_2(x))^2 + (v_1(x) - v_2(x))^2 \end{aligned} \quad (2.26)$$

La operación $|\psi_1(x) + \psi_2(x)|^2$ está definida por:

$$\begin{aligned} |\psi_1(x) + \psi_2(x)|^2 &= (\psi_1(x) + \psi_2(x))^*(\psi_1(x) + \psi_2(x)) \\ &= (u_1(x) + u_2(x))^2 + (v_1(x) + v_2(x))^2 \end{aligned} \quad (2.27)$$

Definición 2.15. Sea $\psi(x) = u(x) + iv(x)$ una función de valor complejo de una variable real, la **derivada** de $\psi(x)$ con respecto a x se define por ([1], pág. 231):

$$\frac{d}{dx}\psi(x) = \frac{d}{dx}u(x) + i\frac{d}{dx}v(x) \quad (2.28)$$

Definición 2.16. Sea $\psi(x) = u(x) + iv(x)$ una función de valor complejo de una variable real, la **integral** de $\psi(x)$ con respecto a x se define por ([1], pág. 231):

$$\int_a^b \psi(x) dx = \int_a^b u(x) dx + i \int_a^b v(x) dx \quad (2.29)$$

2.3. Algunos elementos del análisis funcional

Otro elemento importante sobre el cual opera la mecánica cuántica es llamado *espacio de Hilbert*. Un espacio de Hilbert es un concepto que pertenece a un área de las matemáticas llamada Análisis Funcional. Uno de los objetivos de esta sección será el de construir los elementos necesarios para llegar a la definición de espacio de Hilbert.

2.3.1. Espacios métricos

Dentro de una amplia taxonomía establecida por el análisis funcional para los conjuntos (espacios), hemos seleccionado comenzar por la noción de espacio métrico, nuestra construcción del espacio de Hilbert.

Definición 2.17. Un **espacio** X es un conjunto de elementos. Los elementos de X son llamados los **puntos** de X .

Definición 2.18. Una *métrica* o *función de distancia* d sobre un espacio X , es una función $d: X \times X \rightarrow \mathbb{R}$ que satisface los siguientes axiomas, para todo $x, y, z \in X$:

$$0 \leq d(x, y) < \infty \quad (\text{M-1})$$

$$d(x, y) = 0 \leftrightarrow x = y \quad (\text{M-2})$$

$$d(x, y) = d(y, x) \quad (\text{M-3})$$

$$d(x, y) \leq d(x, z) + d(z, y) \quad (\text{desigualdad triangular}) \quad (\text{M-4})$$

El axioma **M-1** nos dice que $d(x, y)$ es una función real, no negativa y finita.

Definición 2.19. Un *espacio métrico* es una dupla $\langle X, d \rangle$ donde X es un espacio y d es una métrica sobre X .

La noción de espacio métrico puede ser introducida a partir de modelos. Sea $\mathbb{L} = \{F^2\}$ un lenguaje de la lógica de primer orden con un símbolo de función de aridez dos F^2 . Un espacio métrico es un modelo de \mathbb{L} denotado por $\mathcal{X} = \langle X, d \rangle$ donde \mathcal{X} debe satisfacer los axiomas **M-1**, **M-2**, **M-3** y **M-4**.

Antes de presentar nuestro primer ejemplo de un espacio métrico, es necesario presentar la definición de valor absoluto e indicar algunas propiedades del mismo.

Definición 2.20. Sea $x \in \mathbb{R}$. El *valor absoluto* de x es un número real no negativo que se designa por $|x|$ y está definido por ([2], pág. 51):

$$|x| = \begin{cases} x & \text{si } x \geq 0, \\ -x & \text{si } x \leq 0. \end{cases} \quad (2.30)$$

De la definición del valor absoluto de x , se observa que:

$$-|x| \leq x \leq |x| \quad (2.31)$$

Teorema 2.4. Sean $x, a \in \mathbb{R}$ tal que $a \geq 0$, entonces:

$$|x| \leq a \leftrightarrow -a \leq x \leq a \quad (2.32)$$

Teorema 2.5. Sean $x, y \in \mathbb{R}$, entonces: ([2], pág. 51)

$$|x + y| \leq |x| + |y| \quad (\text{desigualdad triangular}) \quad (2.33)$$

Demostración. De la definición de valor absoluto se tiene que $-|x| \leq x \leq |x|$ y que $-|y| \leq y \leq |y|$, sumando estas desigualdades, se obtiene:

$$\begin{aligned} -|x| - |y| &\leq x + y \leq |x| + |y| \\ -(|x| + |y|) &\leq x + y \leq |x| + |y| \\ |x + y| &\leq |x| + |y| \quad (\text{por 2.32}) \end{aligned}$$

□

Ejemplo 2.1. La dupla $\langle \mathbb{R}, d \rangle$ donde \mathbb{R} es el conjunto de los números reales y d está definida por $d(x, y) = |x - y|$; es un espacio métrico. Es necesario demostrar que la función d satisface la definición de métrica, es decir, es necesario demostrar que d satisface los axiomas **M-1**, **M-2**, **M-3** y **M-4**. Soló presentaremos la demostración del axioma **M-4**.

Lema 2.1. $\langle \mathbb{R}, d \rangle \models d(x, y) \leq d(x, z) + d(z, y)$, para todo $x, y, z \in \mathbb{R}$.⁴

Demostración. Es necesario demostrar que:

$$|x - y| \leq |x - z| + |z - y|$$

Sea $a = x - z$ y $b = z - y$. Por el teorema 2.5, se tiene que:

$$|a + b| \leq |a| + |b|$$

luego

$$\begin{aligned} |(x - z) + (z - y)| &\leq |x - z| + |z - y| \\ |x - y| &\leq |x - z| + |z - y| \end{aligned}$$

□

Ejemplo 2.2. La dupla $\langle \mathbb{C}, d \rangle$ donde \mathbb{C} es el conjunto de los números complejos y d está definida por $d(z_1, z_2) = |z_1 - z_2|$; es un espacio métrico. De nuevo, es necesario demostrar que la función d satisface la definición de métrica, es decir, es necesario demostrar que d satisface los axiomas *M-1*, *M-2*, *M-3* y *M-4*. Soló presentaremos la demostración del axioma *M-4*.

Demostración.

Lema 2.2. $\langle \mathbb{C}, d \rangle \models d(z_1, z_2) \leq d(z_1, z_3) + d(z_3, z_2)$, para todo $z_1, z_2, z_3 \in \mathbb{C}$.

Demostración. Es necesario demostrar que:

$$|z_1 - z_2| \leq |z_1 - z_3| + |z_3 - z_2|$$

Sean $a = z_1 - z_3$ y $b = z_3 - z_2$. Por el teorema 2.3 (propiedad 2.20), se tiene que:

$$|a + b| \leq |a| + |b|$$

luego

$$\begin{aligned} |(z_1 - z_3) + (z_3 - z_2)| &\leq |z_1 - z_3| + |z_3 - z_2| \\ |z_1 - z_2| &\leq |z_1 - z_3| + |z_3 - z_2| \end{aligned}$$

□

□

Antes de presentar nuestro siguiente ejemplo de espacio métrico necesitamos la noción de continuidad de una función de valor complejo de una variable real, dicha noción será construida desde la perspectiva de los espacios métricos.

⁴El símbolo $\mathcal{U} \models \alpha$ es usado en la teoría de modelos, significando que el modelo \mathcal{U} satisface la propiedad α .

Definición 2.21. Sean $\langle X, d \rangle$ un espacio métrico, $x_0 \in X$ y $\epsilon > 0$.

Una **bola abierta** denotada por $B(x_0, \epsilon)$ es el conjunto formado por ([16], pág. 18):

$$B(x_0, \epsilon) = \{x \in X \mid d(x, x_0) < \epsilon\} \quad (2.34)$$

Una **bola abierta perforada** denotada por $B'(x_0, \epsilon)$ es el conjunto formado por ([26], pág. 34):

$$B'(x_0, \epsilon) = \{x \in X \mid 0 < d(x, x_0) < \epsilon\} = B(x, x_0) - \{x_0\} \quad (2.35)$$

Además, x_0 y ϵ son llamados el centro y el radio de la bola respectivamente.

Definición 2.22. Sean $\langle X, d \rangle$ un espacio métrico, $x_0 \in X$ y $\epsilon > 0$. Una bola abierta $B(x_0, \epsilon)$ de centro x_0 y radio ϵ es llamada un **entorno** de x_0 , denotado por $N(x_0, \epsilon)$. Si el radio no importa, el entorno de x_0 es denotado por $N(x_0)$. Si la bola abierta es perforada el entorno de x_0 es llamado un **entorno perforado** de x_0 y es denotado por $N'(x_0)$ ([1], pág. 47).

Ejemplo 2.3. Sea $x_0 \in \mathbb{R}$. Un entorno de x_0 de radio $\epsilon > 0$ es un intervalo abierto de centro x_0 definido por:

$$N(x_0, \epsilon) = (x_0 - \epsilon, x_0 + \epsilon)$$

Un entorno perforado de x_0 de radio $\epsilon > 0$ es un conjunto formado por:

$$N'(x_0, \epsilon) = (x_0 - \epsilon, x_0) \cup (x_0, x_0 + \epsilon)$$

Ejemplo 2.4. Sea $c_0 \in \mathbb{C}$. Un entorno de c_0 de radio $\epsilon > 0$ es un disco abierto de centro c_0 y radio ϵ definido por:

$$N(c_0, \epsilon) = \{c \in \mathbb{C} \mid |c - c_0| < \epsilon\}$$

Un entorno perforado de c_0 de radio $\epsilon > 0$ es un disco abierto perforado de centro c_0 y radio ϵ definido por:

$$N'(c_0, \epsilon) = \{c \in \mathbb{C} \mid 0 < |c - c_0| < \epsilon\}$$

Definición 2.23. Sea M un subconjunto de un espacio métrico X . Un punto x_0 de X es llamado un **punto de acumulación** de M si cada entorno de x_0 contiene al menos un punto $y \in M$ diferente de x_0 ([16], pág. 21).

Los conceptos de entorno de un punto y punto de acumulación de un conjunto, permiten construir definiciones más generales para los conceptos de límite de una función y continuidad de una función. Aunque es posible definir estos conceptos para funciones definidas entre espacios métricos, se haran las construcciones para dos clases de funciones: funciones de valor real de una variable real y funciones de valor complejo de una variable real. La razon de esto, es que se desea ilustrar la “conversion” de la definición usual de límite de una función de valor real de una variable real en términos de entornos y puntos de acumulación, además utilizar dicha definición para la definición del límite de una función de valor complejo de una variable real.

Definición 2.24. La conversión de la definición del límite de una función de valor real de una variable real está construida a partir de ([1], págs. 61-63).

Sea $f(x)$ una función de valor real de una variable real ($f : A \rightarrow B, A \subseteq \mathbb{R}, B \subseteq \mathbb{R}$). La definición “tradicional” del límite de $f(x)$ cuando x tiende a x_0 está dada por:

$\lim_{x \rightarrow x_0} f(x) = L$ significa que para cada $\epsilon > 0$, existe un $\delta > 0$ tal que si $0 < |x - x_0| < \delta$ entonces $|f(x) - L| < \epsilon$.

La definición anterior puede ser escrita en términos de entornos y de puntos de acumulación de la siguiente manera:

La desigualdad $0 < |x - x_0| < \delta$ significa:

$$\begin{aligned} & (0 < |x - x_0|) \quad \wedge \quad (|x - x_0| < \delta) \\ ((x - x_0 > 0) \quad \vee \quad (x - x_0 < 0)) \quad \wedge \quad & (-\delta < x - x_0 < \delta) \\ ((x > x_0) \quad \vee \quad (x < x_0)) \quad \wedge \quad & (x_0 - \delta < x < x_0 + \delta) \end{aligned}$$

entonces

$$\begin{aligned} x \in & ((x_0, +\infty) \cup (-\infty, x_0)) \quad \cap \quad (x_0 - \delta, x_0 + \delta) \\ x \in & (x_0 - \delta, x_0) \quad \cup \quad (x_0, x_0 + \delta) \end{aligned}$$

Es decir, $x \in N'(x_0)$ (x pertenece a un entorno perforado de x_0).

La desigualdad $|f(x) - L| < \epsilon$ significa:

$$\begin{aligned} -\epsilon & < f(x) - L < \epsilon \\ L - \epsilon & < f(x) < L + \epsilon \end{aligned}$$

entonces

$$\begin{aligned} f(x) \in & (L - \epsilon, +\infty) \quad \cap \quad (-\infty, L + \epsilon) \\ f(x) \in & (L - \epsilon, L + \epsilon) \end{aligned}$$

Es decir, $f(x) \in N(L)$ ($f(x)$ pertenece a un entorno de L).

Entonces, en forma general el principio invocado en la definición del límite significa que para cada entorno $N(L)$ existe otro entorno $N(x_0)$ tal que si $x \in N(x_0)$ entonces $f(x) \in N(L)$. Esto permite construir una definición formal del límite de una función de valor real de una variable real.

Sea $f(x)$ una función de valor real de una variable real, con dominio A y rango B ($f : A \rightarrow B, A \subseteq \mathbb{R}, B \subseteq \mathbb{R}$). Si x_0 es un punto de acumulación de A y $L \in \mathbb{R}$, el simbolismo $\lim_{x \rightarrow x_0} f(x) = L$, significa que para entorno $N(L) \in B$, existe un entorno $N(x_0) \in A$ tal que, si $x \in N'(x_0) \cap A$ implica que $f(x) \in N(L)$.

La definición exige que $x \in N'(x_0) \cap A$ en lugar de que $x \in N'(x_0)$ para asegurar que x está en el dominio de f , además, se requiere que x_0 sea un punto de acumulación de B para asegurar que la intersección $N'(x_0) \cap A$ no sea un conjunto vacío.

Definición 2.25. Si $f(x)$ está definida en x_0 y $\lim_{x \rightarrow x_0} f(x) = f(x_0)$, se dice que la función $f(x)$ es **continua en el punto** x_0 . Si $f(x)$ es continua en cada punto de A , entonces se dice que $f(x)$ es continua en el conjunto A .

Definición 2.26. A partir de la definición del límite de una función de valor real de una variable real, utilizando los conceptos de entorno y puntos de acumulación, es posible construir la definición del límite de una función de valor complejo de una variable real.

Sea $\psi(x)$ una función de valor complejo de una variable real, con dominio A y rango B ($f : A \rightarrow B, A \subseteq \mathbb{R}, B \subseteq \mathbb{C}$). Si x_0 es un punto de acumulación de A y $L \in \mathbb{C}$, el simbolismo $\lim_{x \rightarrow x_0} \psi(x) = L$, significa que para entorno $N(L) \in B$, existe un entorno $N(x_0) \in A$ tal que, si $x \in N'(x_0) \cap A$ implica que $f(x) \in N(L)$.

Definición 2.27. Si $\psi(x)$ está definida en x_0 y $\lim_{x \rightarrow x_0} \psi(x) = \psi(x_0)$, se dice que la función $\psi(x)$ es **continua en el punto** x_0 . Si $\psi(x)$ es continua en cada punto de A , entonces se dice que $\psi(x)$ es continua en el conjunto A .

Ejemplo 2.5. Como mencionamos anteriormente, el concepto de continuidad de una función de valor complejo de una variable real era necesario para presentar nuestro siguiente ejemplo de espacio métrico. Este espacio es la base a partir de la cual se va a construir el espacio de Hilbert sobre el cual opera la mecánica cuántica, y en particular el espacio sobre el cual opera la máquina cuántica.

Espacio: El conjunto de funciones de valor complejo de una variable real, continuas sobre el intervalo $(-\infty, +\infty)$. Este conjunto lo vamos a representar por $C(\mathbb{R}, \mathbb{C})$.

Métrica: Sean $\psi_1(x), \psi_2(x) \in C(\mathbb{R}, \mathbb{C})$, entonces se define la métrica d por:

$$d(\psi_1(x), \psi_2(x)) = \left(\int_{-\infty}^{+\infty} |\psi_1(x) - \psi_2(x)|^2 dx \right)^{\frac{1}{2}} \quad (2.36)$$

Entonces, es posible demostrar que la dupla $\langle C(\mathbb{R}, \mathbb{C}), d \rangle$ es un espacio métrico.

Un concepto importante para construir la noción de espacio de Hilbert es el relacionado con la “completitud” del espacio en cuestión. Presentamos a continuación algunos elementos para hablar de la completitud en relación a los espacios métricos.

Definición 2.28. Sea $\langle X, d \rangle$ un espacio métrico. Una **sucesión infinita** $\{x_n\}$ de X , es un conjunto de puntos x_1, x_2, \dots de X .⁵

Definición 2.29. Una sucesión $\{x_n\}$ de un espacio métrico $\langle X, d \rangle$ se dice **convergente** si existe un $x \in X$ tal que $\lim_{n \rightarrow \infty} d(x_n, x) = 0$; x es llamado el **límite** de $\{x_n\}$, es decir $\lim_{n \rightarrow \infty} x_n = x$.

Definición 2.30. Una sucesión $\{x_n\}$ de un espacio métrico $\langle X, d \rangle$ se dice que es una **sucesión de Cauchy** si para cada $\varepsilon > 0$ existe un $N = N(\varepsilon)$ tal que $d(x_m, x_n) < \varepsilon$ para todo $m, n > N$.

Definición 2.31. Sea $\langle X, d \rangle$ un espacio métrico. $\langle X, d \rangle$ es un **espacio métrico completo** si cada sucesión de Cauchy en $\langle X, d \rangle$ es convergente, es decir, cada sucesión de Cauchy tiene límite y éste es un elemento de X . Un espacio métrico $\langle X, d \rangle$ no completo, se denomina, **espacio métrico incompleto**.

Presentamos a continuación un teorema y algunos elementos relacionados con él para realizar la completación de un espacio métrico incompleto.

Definición 2.32. Sean $\langle X, d \rangle, \langle \tilde{X}, \tilde{d} \rangle$ espacios métricos. Una función $T: X \rightarrow \tilde{X}$ se dice que es **isométrica** o que es una **isometría** si T preserva las distancias, es decir, si para todo $x_1, x_2 \in X$, $\tilde{d}(Tx_1, Tx_2) = d(x_1, x_2)$, donde Tx_1 y Tx_2 son las imágenes de x_1 y x_2 respectivamente.

Definición 2.33. Sean $\langle X, d \rangle, \langle \tilde{X}, \tilde{d} \rangle$ espacios métricos. $\langle X, d \rangle, \langle \tilde{X}, \tilde{d} \rangle$ son llamado **espacios isométricos** si existe una función isométrica $T: X \rightarrow \tilde{X}$ tal que T sea biyectiva.

Teorema 2.6. Para cada espacio métrico $\langle X, d \rangle$ existe un espacio métrico completo $\langle \tilde{X}, \tilde{d} \rangle$ el cual tiene un subespacio $\langle W, \tilde{d} \rangle$ que es isométrico con $\langle X, d \rangle$ y es denso en $\langle \tilde{X}, \tilde{d} \rangle$. Este espacio \tilde{X} es único excepto para isometrías, es decir, si \hat{X} es cualquier espacio métrico completo con un subespacio denso \hat{W} isométrico con X , entonces \hat{X} y \tilde{X} son isométricos.

Intuitivamente podemos pensar que el teorema de completación para un espacio métrico incompleto X , produce un nuevo espacio métrico completo X' en cual se han adicionado los límites de las sucesiones de Cauchy que no pertenecían a X .

⁵El término sucesión hará referencia a una sucesión infinita.

2.3.2. Espacios vectoriales

Continuado con la taxonomía presentada por el análisis funcional para los espacios, presentamos el concepto de espacio vectorial y algunos elementos relacionados con él. Los espacios de las próximas secciones (espacios normados, espacios de Banach, espacios con producto interior y espacios de Hilbert) están contruídos sobre el concepto algebraico de espacio vectorial.

Definición 2.34. *Un espacio vectorial o espacio lineal sobre un campo K es una estructura algebraica, definida por*

$$\mathcal{E}\mathcal{L} = \langle X, K, +, \cdot, 0, 1 \rangle \text{ donde,}$$

X : Es un conjunto no vacío de elementos llamados **vectores**

K : Es un campo llamado el campo escalar y sus elementos son llamados **escalares**

$+$: Es una función $+: X \times X \rightarrow X$ llamada **adición vectorial**

\cdot : Es una función $\cdot: K \times X \rightarrow X$ llamada **multiplicación de vectores por escalares**⁶

0 : Es llamado el elemento neutro de X

1 : Es llamado el elemento unidad de K

$\mathcal{E}\mathcal{L}$ debe satisfacer los siguientes axiomas:

Axiomas de la adición vectorial:

$$\forall x \forall y \in X (x + y \in X) \quad (\text{clausurativa } +) \quad (\text{EV-1})$$

$$\forall x \forall y \in X (x + y = y + x) \quad (\text{conmutativa } +) \quad (\text{EV-2})$$

$$\forall x \forall y \forall z \in X (x + (y + z) = (x + y) + z) \quad (\text{asociativa } +) \quad (\text{EV-3})$$

$$\forall x \in X (x + 0 = 0 + x = x) \quad (\text{elemento neutro para } +) \quad (\text{EV-4})$$

$$\forall x \in X - \{0\}, \exists y \in X (x + y = 0) \quad (\text{existencia elementos inversos bajo } +) \quad (\text{EV-5})$$

Axiomas de la multiplicación de vectores por escalares:

$$\forall \alpha \in K, \forall x \in X (\alpha x \in X) \quad (\text{clausurativa } \cdot) \quad (\text{EV-6})$$

$$\forall \alpha \forall \beta \in K, \forall x \forall y \in X (\alpha(x + y) = \alpha x + \alpha y) \quad (\text{distrib. mult. vectorial}) \quad (\text{EV-7})$$

$$\forall \alpha \forall \beta \in K, \forall x \in X ((\alpha + \beta)x = \alpha x + \beta x) \quad (\text{distrib. adición escalar}) \quad (\text{EV-8})$$

$$\forall \alpha \forall \beta \in K, \forall x \in X (\alpha(\beta x) = (\alpha\beta)x) \quad (\text{asociativa } \cdot \text{ escalar}) \quad (\text{EV-9})$$

$$\forall x \in X (1x = x) \quad (\text{neutro } \cdot) \quad (\text{EV-10})$$

Algunos autores ([12], pág. 320, [16], pág. 50) definen que X es el espacio vectorial sobre el campo K . Cuando se utilice esta forma de referencia, es decir cuando se hable del espacio vectorial X , debe entenderse que se hace referencia implícita a la estructura algebraica $\mathcal{E}\mathcal{L} = \langle X, K, +, \cdot, 0, 1 \rangle$.

Definición 2.35. *Sea $\mathcal{E}\mathcal{L} = \langle X, K, +, \cdot, 0, 1 \rangle$ un espacio vectorial. Si $K = \mathbb{R}$, entonces $\mathcal{E}\mathcal{L}$ se denomina **espacio vectorial real**. Si $K = \mathbb{C}$, entonces $\mathcal{E}\mathcal{L}$ se denomina **espacio vectorial complejo**.*

Ejemplo 2.6. *Presentamos y a continuación demostramos un ejemplo de un espacio vectorial. La estructura seleccionada como ejemplo de espacio vectorial está definida por*

$$\mathcal{E}\mathcal{L} = \langle C(\mathbb{R}, \mathbb{C}), \mathbb{C}, +, \cdot, \psi_0(x), 1 \rangle \text{ donde,}$$

⁶Sea $x \in X$ y $\alpha \in K$ entonces, la operación $\alpha \cdot x$ la representaremos por αx .

$C(\mathbb{R}, \mathbb{C})$: Es el conjunto compuesto por las funciones de valor complejo de una variable real continuas sobre el intervalo $(-\infty, +\infty)$. Estas funciones son los vectores de \mathcal{EL} y tienen la forma:

$$\psi(x) = u(x) + iv(x), \quad (2.37)$$

donde $u(x)$ y $v(x)$ son funciones reales de la variable real x e i representa la unidad imaginaria.

\mathbb{C} : Es el conjunto de los números complejos. Estos son los escalares de \mathcal{EL} .

$+$: Es la adición vectorial en \mathcal{EL} definida por:

Sean $\psi_1(x) = u_1(x) + iv_1(x)$ y $\psi_2(x) = u_2(x) + iv_2(x)$, entonces,

$$\psi_1(x) + \psi_2(x) = (u_1(x) + u_2(x)) + i(v_1(x) + v_2(x)) \quad (2.38)$$

\cdot : Es la multiplicación por un escalar en \mathcal{EL} definida por:

Sean $\psi(x) \in C$ y $c \in \mathbb{C}$ donde $\psi(x) = u(x) + iv(x)$ y $c = a + ib$, entonces,

$$c\psi(x) = (au(x) - bv(x)) + i(av(x) + bu(x)) \quad (2.39)$$

$\psi_0(x)$: Es el elemento neutro de $C(\mathbb{R}, \mathbb{C})$ definido por: $f_0(x) = 0$.

1: Es el elemento unidad de \mathbb{C} .

Antes de presentar el siguiente teorema, presentamos dos teoremas que nos servirán para su demostración.

Teorema 2.7. Si $f(x), g(x)$ son funciones de valor real de una variable real continuas en un intervalo abierto (a, b) , la suma $f(x) + g(x)$ es continua en el intervalo (a, b) ([1], pág. 68).

Teorema 2.8. Sea $\psi(x) = u(x) + iv(x)$ una función de valor complejo de una variable real continua en un intervalo abierto (a, b) entonces, las funciones de valor real de una variable real $u(x), v(x)$ son continuas en el intervalo abierto (a, b) .

Teorema 2.9. $\mathcal{EL} = \langle C(\mathbb{R}, \mathbb{C}), \mathbb{C}, +, \cdot, \psi_0(x), 1 \rangle$ es un espacio vectorial, es decir, \mathcal{EL} satisface los axiomas con respecto a la adición vectorial (axiomas EV-1, EV-2, EV-3, EV-4 y EV-5) y satisface los axiomas con respecto a la multiplicación por un escalar (axiomas EV-6, EV-7, EV-8, EV-9 y EV-10).

Lema 2.3. \mathcal{EL} satisface el axioma EV-1, es decir:

$$\mathcal{EL} \models \psi_1(x) + \psi_2(x) \in C(\mathbb{R}, \mathbb{C}), \text{ para todo } \psi_1(x), \psi_2(x) \in C(\mathbb{R}, \mathbb{C}).$$

Demostración. Sea $\psi_1(x) = u_1(x) + iv_1(x)$ y $\psi_2(x) = u_2(x) + iv_2(x)$:

$$\begin{aligned} \psi_1(x) + \psi_2(x) &= (u_1(x) + iv_1(x)) + (u_2(x) + iv_2(x)) \\ &= (u_1(x) + u_2(x)) + i(v_1(x) + v_2(x)) \end{aligned}$$

Como $u_1(x), u_2(x), v_1(x)$ y $v_2(x)$ satisfacen los requerimientos de los teoremas 2.7 y 2.8, entonces $\psi_1(x) + \psi_2(x) \in C(\mathbb{R}, \mathbb{C})$ \square

Lema 2.4. \mathcal{EL} satisface el axioma *EV-2*, es decir:

$$\mathcal{EL} \models \psi_1(x) + \psi_2(x) = \psi_2(x) + \psi_1(x), \text{ para todo } \psi_1(x), \psi_2(x) \in C(\mathbb{R}, \mathbb{C}).$$

Demostración. Sea $\psi_1(x) = u_1(x) + iv_1(x)$ y $\psi_2(x) = u_2(x) + iv_2(x)$ entonces:

$$\begin{aligned} \psi_1(x) + \psi_2(x) &= (u_1(x) + iv_1(x)) + (u_2(x) + iv_2(x)) \\ &= (u_1(x) + u_2(x)) + i(v_1(x) + v_2(x)) \\ &= (u_2(x) + u_1(x)) + i(v_2(x) + v_1(x)) \\ &= (u_2(x) + iv_2(x)) + (u_1(x) + iv_1(x)) \\ &= \psi_2(x) + \psi_1(x) \end{aligned}$$

□

Lema 2.5. \mathcal{EL} satisface el axioma *EV-3*, es decir:

$$\mathcal{EL} \models \psi_1(x) + (\psi_2(x) + \psi_3(x)) = (\psi_1(x) + \psi_2(x)) + \psi_3(x), \text{ para todo } \psi_1(x), \psi_2(x), \psi_3(x) \in C(\mathbb{R}, \mathbb{C}).$$

Demostración. Sea $\psi_1(x) = u_1(x) + iv_1(x)$, $\psi_2(x) = u_2(x) + iv_2(x)$ y $\psi_3(x) = u_3(x) + iv_3(x)$ entonces:

$$\begin{aligned} \psi_1(x) + (\psi_2(x) + \psi_3(x)) &= (u_1(x) + iv_1(x)) + ((u_2(x) + iv_2(x)) + (u_3(x) + iv_3(x))) \\ &= (u_1(x) + iv_1(x)) + ((u_2(x) + u_3(x)) + i(v_2(x) + v_3(x))) \\ &= (u_1(x) + (u_2(x) + u_3(x))) + i(v_1(x) + (v_2(x) + v_3(x))) \\ &= ((u_1(x) + u_2(x)) + u_3(x)) + i((v_1(x) + (v_2(x) + v_3(x)))) \\ &= ((u_1(x) + iv_1(x)) + (u_2(x) + iv_2(x))) + (u_3(x) + iv_3(x)) \\ &= (\psi_1(x) + \psi_2(x)) + \psi_3(x) \end{aligned}$$

□

Lema 2.6. \mathcal{EL} satisface el axioma *EV-4*, es decir:

$$\mathcal{EL} \models \psi_1(x) + \psi_0(x) = \psi_0(x) + \psi_1(x) = \psi_1(x), \text{ para todo } \psi_1(x) \in C(\mathbb{R}, \mathbb{C}).$$

Demostración. Sea $\psi(x) = u(x) + iv(x)$ y $\psi_0(x) = 0$ entonces:

$$\begin{aligned} \psi(x) + \psi_0(x) &= (u(x) + iv(x)) + 0 \\ &= (u(x) + 0) + i(v(x) + 0) \\ &= u(x) + iv(x) \\ &= \psi(x) \end{aligned}$$

□

Lema 2.7. \mathcal{EL} satisface el axioma *EV-5*, es decir:

$$\mathcal{EL} \models \forall \psi(x) \in C(\mathbb{R}, \mathbb{C}) - \{\psi_0(x)\}, \exists \psi'(x) \in C(\mathbb{R}, \mathbb{C}) \quad (\psi(x) + \psi'(x) = \psi_0(x)).$$

Demostración. Sea $\psi(x) = u(x) + iv(x)$ entonces se define $\psi'(x)$ por: $\psi'(x) = -u(x) + i(-v(x))$, entonces:

$$\begin{aligned}\psi(x) + \psi'(x) &= (u(x) + iv(x)) + (-u(x) + i(-v(x))) \\ &= (u(x) - u(x)) + i(v(x) - v(x)) \\ &= 0 \\ &= \psi_0(x)\end{aligned}$$

□

No demostraremos los axiomas con respecto a la multiplicación por un escalar.

Presentamos ahora algunos elementos de los espacios vectoriales que nos servirán en nuestra presentación de los espacios de las próximas secciones.

Definición 2.36. Una **combinación lineal** de vectores x_1, \dots, x_m de un espacio vectorial X es una expresión de la forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m,$$

donde los coeficientes $\alpha_1, \alpha_2, \dots, \alpha_m$ son escalares.

Definición 2.37. Sea M un subconjunto ($\neq \emptyset$) de un espacio vectorial X . El conjunto de todas las combinaciones lineales de M es llamado el **span** de M .

Definición 2.38. Lineal independencia y lineal dependencia de un conjunto M de vectores x_1, \dots, x_r ($r \geq 1$) de un espacio vectorial X son definidas por medio de la ecuación

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_r x_r = 0, \tag{2.40}$$

donde $\alpha_1, \alpha_2, \dots, \alpha_r$ son escalares. La ecuación (2.40) es verdadera para $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$. Si esta es la única r -tupla de escalares para la que la ecuación (2.40) es verdadera, el conjunto M se dice que es **linealmente independiente**. M se dice que es **linealmente dependiente** si M no es linealmente independiente, es decir si (2.40) es verdadera para alguna r -tupla de escalares, no todos ceros.

Definición 2.39. Un espacio vectorial X se dice **finito dimensional** si existe un entero positivo n tal que X contiene un conjunto linealmente independiente de n vectores, donde cualquier conjunto de $n + 1$ o más vectores de X es linealmente dependiente. n es llamada la **dimensión** de X y es denotada por $\dim X = n$. Si X no es finito dimensional, entonces se dice que X es **infinito dimensional**.

Definición 2.40. Si $\dim X = n$, un conjunto linealmente independiente de n vectores de X es llamado una **base** para X . Si $B = \{e_1, \dots, e_n\}$ es una base para X , cada $e_i \in B$ es llamado un **vector base**. Cada $x \in X$ tiene una única representación como combinación lineal de los vectores base $x = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$.

En general, si X es un espacio vectorial (finito o infinito dimensional) y B es un subconjunto linealmente independiente de X tal que $\text{span } B = X$, entonces B es llamada una **base** o **base de Hamel** para X .

Finalizamos nuestra presentación de los espacios vectoriales, definiendo el concepto de álgebra.

Definición 2.41. Un **álgebra** \mathcal{A} sobre un campo K es un espacio vectorial \mathcal{A} sobre K que tal que para cada par ordenado de elementos $x, y \in \mathcal{A}$ hay un único producto $xy \in \mathcal{A}$ que satisface los siguientes axiomas, para

todo $x, y, z \in \mathcal{A}$ y para todo $\alpha \in K$:

$$(xy)z = x(yz) \quad (\text{A-1})$$

$$x(y + z) = xy + xz \quad (\text{A-2})$$

$$(x + y)z = xz + yz \quad (\text{A-3})$$

$$\alpha(xy) = (\alpha x)y = x(\alpha y) \quad (\text{A-4})$$

2.3.3. Espacios normados y espacios de Banach

En muchos casos un espacio vectorial \mathcal{X} puede ser considerado al mismo tiempo un espacio métrico \mathcal{X} debido a una métrica d definida sobre \mathcal{X} . Sin embargo, no hay relación entre la estructura algebraica de espacio vectorial y la métrica. Para establecer alguna relación entre las propiedades algebraicas y métricas de \mathcal{X} , se construye un concepto llamado norma a partir de las propiedades algebraicas de \mathcal{X} . A partir de la norma se obtiene una métrica para \mathcal{X} . Esto nos permite hablar entonces de espacios normados. Presentamos a continuación dichos conceptos.

Definición 2.42. Sea X un espacio vectorial sobre un campo escalar K . Una **norma** sobre X es una función $\|\cdot\| : X \rightarrow \mathbb{R}$ que satisface los siguientes axiomas, para todo $x, y \in X$ y para todo $\alpha \in K$:

$$\|x\| \geq 0 \quad (\text{N-1})$$

$$\|x\| = 0 \leftrightarrow x = 0 \quad (\text{N-2})$$

$$\|\alpha x\| = |\alpha| \|x\| \quad (\text{N-3})$$

$$\|x + y\| \leq \|x\| + \|y\| \quad (\text{desigualdad triangular}) \quad (\text{N-4})$$

Definición 2.43. Sea X un espacio vectorial con una norma $\|x\|$. La norma define una métrica d sobre X , dada por $d(x, y) = \|x - y\|$. La métrica d es llamada la **métrica inducida por la norma**.

Definición 2.44. Un **espacio normado** $\langle X, \|\cdot\| \rangle$ es una dupla donde X es un espacio vectorial y $\|\cdot\|$ es una norma sobre X .

Un espacio normado $\langle X, \|\cdot\| \rangle$ puede ser también pensado como un espacio métrico $\langle X, \|\cdot\|_d \rangle$ donde X es un espacio vectorial y $\|\cdot\|_d$ es la métrica inducida por la norma $\|\cdot\|$.

Ejemplo 2.7. Sea $X = \langle C(\mathbb{R}, \mathbb{C}), \mathbb{C}, +, \cdot, \psi_0(x), 1 \rangle$ el espacio vectorial presentado en el ejemplo 2.6. Es posible demostrar que una norma sobre X es:

$$\begin{aligned} \|\psi(x)\| &= \left(\int_{-\infty}^{\infty} \psi(x) \bar{\psi} dx \right)^{\frac{1}{2}} \\ &= \left(\int_{-\infty}^{\infty} |\psi(x)|^2 dx \right)^{\frac{1}{2}} \end{aligned} \quad (2.41)$$

Entonces $\langle X, \|\cdot\| \rangle$ donde $\|\cdot\|$ es la norma definida anteriormente, es un espacio normado. La métrica inducida por esta norma viene dada por:

$$\begin{aligned} d(\psi_1(x), \psi_2(x)) &= \|\psi_1(x) - \psi_2(x)\| \\ &= \left(\int_{-\infty}^{\infty} |\psi_1(x) - \psi_2(x)|^2 dx \right)^{\frac{1}{2}}, \end{aligned}$$

donde el espacio métrico definido por $\langle X, d(\psi_1(x), \psi_2(x)) \rangle$ corresponde al espacio métrico presentado en el ejemplo 2.5.

En los espacios normados es también posible de hablar de espacios normados incompletos y de espacios normados completos, los primeros reciben el nombre de espacios normados y los últimos reciben el nombre de espacios de Banach. Además existe un teorema similar al teorema de completitud de los espacios métricos. Presentamos estos dos elementos a continuación.

Definición 2.45. Un **espacio de Banach** $\langle X, \|\cdot\| \rangle$ es un espacio normado completo. Es decir el espacio métrico formado por X y la métrica inducida por la norma $\|\cdot\|$ es un espacio métrico completo.

Teorema 2.10. Sea $\langle X, \|\cdot\| \rangle$ un espacio normado incompleto. Entonces existe un espacio de Banach \tilde{X} y una isometría T de X a un subespacio W de \tilde{X} el cual es denso en \tilde{X} . El espacio \tilde{X} es único, excepto para isometrías.

Ejemplo 2.8. Es posible demostrar que el espacio presentado en el ejemplo 2.7 es un espacio normado incompleto. La completación de este espacio (métrico) es un espacio de Banach, llamado el espacio complejo $\mathbb{L}^2[-\infty, +\infty]$.

2.3.4. Espacios con producto interior y espacios de Hilbert

En un espacio normado podemos adicionar vectores y multiplicar vectores por escalares. Además, la norma de un vector generaliza el concepto de longitud de un vector. Nos falta conocer si es posible construir un producto punto o producto interior general entre vectores de un espacio normado, que nos permita obtener una generalización del concepto de ortogonalidad (perpendicularidad) entre los mismos. Además, el producto interior permite definir una norma la cual a su vez permite definir una métrica. Los espacios normados (incompletos o completos) que permiten la definición de un producto interior son llamados espacios con producto interior y son el tema que presentamos a continuación.

Definición 2.46. Sea X un espacio vectorial sobre un campo escalar K . Un **producto interior** o **producto escalar** o **producto punto** sobre X es una función $\langle \cdot, \cdot \rangle : X \times X \rightarrow K$ que satisface los siguientes axiomas, para todo $x, y, z \in X$ y para todo $\alpha \in K$:

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad (\text{PI-1})$$

$$\alpha \langle x, y \rangle = \langle \alpha x, y \rangle \quad (\text{PI-2})$$

$$\langle x, x \rangle \geq 0 \text{ y } \langle x, x \rangle = 0 \leftrightarrow x = 0 \quad (\text{PI-3})$$

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \quad (\text{PI-4})$$

Definición 2.47. Sea X en espacio vectorial con un producto interior $\langle x, y \rangle$. El producto interior define una norma $\|x\|$ sobre X , dada por $\|x\| = \sqrt{\langle x, x \rangle}$. La norma $\|x\|$ es llamada la **norma inducida por el producto interior**.

Definición 2.48. Sea X en espacio vectorial con un producto interior $\langle x, y \rangle$. El producto interior define una métrica d sobre X , dada por $d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}$. La métrica d es llamada la **métrica inducida por el producto interior**.

Definición 2.49. Un **espacio con producto interior** o **espacio pre-Hilbert** es una dupla $\langle X, \langle x, y \rangle \rangle$ donde X es espacio vectorial y $\langle x, y \rangle$ es un producto interior definido sobre X .

Un espacio con producto interior $\langle X, \langle x, y \rangle \rangle$ puede ser también pensado como un espacio normado $\langle X, \langle x, y \rangle_n \rangle$ donde X es un espacio vectorial y $\langle x, y \rangle_n$ es la norma inducida por el producto interior $\langle x, y \rangle$.

Adicionalmente un espacio con producto interior $\langle X, \langle x, y \rangle \rangle$ puede ser también pensado como un espacio métrico $\langle X, \langle x, y \rangle_d \rangle$ donde X es un espacio vectorial y $\langle x, y \rangle_d$ es la métrica inducida por el producto interior $\langle x, y \rangle$.

Ejemplo 2.9. Sea $X = \langle C(\mathbb{R}, \mathbb{C}), \mathbb{C}, +, \cdot, \psi_0(x), 1 \rangle$ el espacio vectorial presentado en el ejemplo 2.6. Vamos a demostrar que una producto interior sobre X es:

$$\langle \psi_1, \psi_2 \rangle = \int_{-\infty}^{\infty} \psi_1 \overline{\psi_2} dx \quad (2.42)$$

Entonces $\langle X, \langle \psi_1, \psi_2 \rangle \rangle$ es un espacio con producto interior. La norma inducida por este producto interior está dada por:

$$\begin{aligned} \|\psi\| &= (\langle \psi, \psi \rangle)^{\frac{1}{2}} \\ &= \left(\int_{-\infty}^{\infty} |\psi(x)|^2 dx \right)^{\frac{1}{2}}, \end{aligned}$$

donde el espacio normado definido por $\langle X, \|\psi\| \rangle$ corresponde al espacio normado presentado en el ejemplo 2.7.

Necesitamos demostrar que $\langle \psi_1, \psi_2 \rangle$ satisface los axiomas **PI-1**, **PI-2**, **PI-3** y **PI-4**.

Lema 2.8. $\langle \psi_1, \psi_2 \rangle$ satisface el axioma **PI-1**, es decir:

$$\langle X, \langle \psi_1, \psi_2 \rangle \rangle \models \langle \psi_1 + \psi_2, \psi_3 \rangle = \langle \psi_1, \psi_3 \rangle + \langle \psi_2, \psi_3 \rangle, \text{ para todo } \psi_1, \psi_2, \psi_3 \in X.$$

Demostración. Sea $\psi = u + iv, \psi_1 = u_1 + iv_1$ y $\psi_2 = u_2 + iv_2$ entonces,

$$\begin{aligned} \langle (\psi_1 + \psi_2), \psi \rangle &= \int_{-\infty}^{\infty} (\psi_1 + \psi_2) \overline{\psi} dx \\ &= \int_{-\infty}^{\infty} [\psi_1 \overline{\psi} + \psi_2 \overline{\psi}] dx \\ &= \int_{-\infty}^{\infty} \psi_1 \overline{\psi} dx + \int_{-\infty}^{\infty} \psi_2 \overline{\psi} dx \\ &= \langle \psi_1, \psi \rangle + \langle \psi_2, \psi \rangle \end{aligned}$$

□

Lema 2.9. $\langle \psi_1, \psi_2 \rangle$ satisface el axioma **PI-2**, es decir:

$$\langle X, \langle \psi_1, \psi_2 \rangle \rangle \models \langle \alpha \psi_1, \psi_2 \rangle = \alpha \langle \psi_1, \psi_2 \rangle, \text{ para todo } \psi_1, \psi_2, \psi_3 \in X, \alpha \in K.$$

Demostración.

$$\begin{aligned} \langle \alpha \psi_1, \psi_2 \rangle &= \int_{-\infty}^{\infty} (\alpha \psi_1) \overline{\psi_2} dx \\ &= \int_{-\infty}^{\infty} \alpha (\psi_1 \overline{\psi_2}) dx \\ &= \alpha \int_{-\infty}^{\infty} \psi_1 \overline{\psi_2} dx \\ &= \alpha \langle \psi_1, \psi_2 \rangle \end{aligned}$$

□

Antes de presentar el siguiente lema, es necesario mencionar dos teoremas necesarios para su demostración.

Teorema 2.11. Sea $f(x)$ una función de valor real de una variable real. Si $f(x)$ es continua en el intervalo $(-\infty, +\infty)$, entonces $\int_{-\infty}^{+\infty} f^2(x) dx < \infty$.

Teorema 2.12. Sea $f(x)$ una función de valor real de una variable real. Si $f(x)$ es continua en el intervalo $(-\infty, +\infty)$, entonces $\int_{-\infty}^{+\infty} f^2(x) dx \geq 0$.

Lema 2.10. $\langle \psi, \psi \rangle$ satisface el axioma **PI-3**, es decir:

$$\langle X, \langle \psi_1, \psi_2 \rangle \rangle \models \langle \psi, \psi \rangle \geq 0 \text{ y } \langle \psi, \psi \rangle = 0 \leftrightarrow \psi = 0, \text{ para todo } \psi \in X.$$

Demostración. La demostración se compone de dos partes:

a. $\langle \psi, \psi \rangle \geq 0$.

Sea $\psi(x) = u(x) + iv(x)$ entonces,

$$\begin{aligned} \langle \psi, \psi \rangle &= \int_{-\infty}^{\infty} \psi \bar{\psi} dx \\ &= \int_{-\infty}^{\infty} |\psi|^2 dx \\ &= \int_{-\infty}^{\infty} (u^2(x) + v^2(x)) dx \end{aligned}$$

Como $u(x)$ y $v(x)$ satisfacen los requerimientos de los teoremas 2.8, 2.11 y 2.12, entonces

$$\int_{-\infty}^{\infty} (u^2(x) + v^2(x)) dx \geq 0.$$

b. $\langle \psi, \psi \rangle = 0$ si y sólo si $\psi = 0$.

$$\int_{-\infty}^{\infty} 0\bar{0} dx = \int_{-\infty}^{\infty} 0 dx = 0$$

□

Lema 2.11. $\langle \psi_1, \psi_2 \rangle$ satisface el axioma **PI-4**, es decir:

$$\langle X, \langle \psi_1, \psi_2 \rangle \rangle \models \langle \psi_1, \psi_2 \rangle = \overline{\langle \psi_2, \psi_1 \rangle}, \text{ para todo } \psi_1, \psi_2 \in X$$

Demostración.

$$\begin{aligned} \langle \psi_1, \psi_2 \rangle &= \int_{-\infty}^{\infty} \psi_1 \bar{\psi}_2 dx \\ &= \int_{-\infty}^{\infty} \overline{\overline{\psi_1 \bar{\psi}_2}} dx \\ &= \int_{-\infty}^{\infty} \overline{\overline{\psi_1} \psi_2} dx \\ &= \int_{-\infty}^{\infty} \overline{\psi_2 \overline{\psi_1}} dx \\ &= \overline{\langle \psi_2, \psi_1 \rangle} \end{aligned}$$

□

En los espacios con producto interior es también posible de hablar de espacios con producto interior incompletos y de espacios con producto interior completos, los primeros reciben el nombre de espacios con producto interior y los últimos reciben el nombre de espacios de Hilbert. Además existe un teorema similar al teorema de completitud de los espacios normados. Presentamos estos dos elementos a continuación.

Definición 2.50. Un **espacio de Hilbert** $\langle X, \langle x, y \rangle \rangle$ es un espacio con producto interior completo. Es decir el espacio métrico formado por X y la métrica inducida por el producto interior $\langle x, y \rangle$ es un espacio métrico completo. Dado que un producto interior define una norma, un espacio de Hilbert es un espacio de Banach.

Teorema 2.13. Para cualquier espacio con producto interior X existe un espacio de Hilbert H y un isomorfismo A de X a un subespacio denso $W \subset H$. El espacio H es único excepto para isomorfismos.

Ejemplo 2.10. Es posible demostrar que el espacio presentado en el ejemplo 2.9 es un espacio normado incompleto. La completación de este espacio (métrico) es un espacio de Hilbert, llamado el espacio complejo $\mathbb{L}^2[-\infty, +\infty]$.

2.3.5. Operadores

Podemos afirmar desde una visión muy simplista que la mecánica cuántica oera fundamentalmente con dos elementos matemáticos: espacios de Hilbert y operadores lineales. Presentamos a continuación el segundo de ellos.

Definición 2.51. Un **operador** T es una función entre espacios vectoriales definidos sobre el mismo campo K , es decir, $T : \mathcal{D}(T) \subset X \rightarrow \mathcal{R}(T) \subset Y$ donde X, Y son espacios vectorial sobre algún campo K .

$\mathcal{D}(T)$ denota el dominio del operador T .

$\mathcal{R}(T)$ denota el rango del operador T .

$T : \mathcal{D}(T) \rightarrow \mathcal{R}(T)$ denota que T es un operador desde $\mathcal{D}(T)$ sobre $\mathcal{R}(T)$.

$T : \mathcal{D}(T) \rightarrow Y$ denota que T es un operador desde $\mathcal{D}(T)$ dentro de Y .

Si $\mathcal{D}(T)$ es igual al espacio X entonces escribimos $T : X \rightarrow Y$, para denotar que T es un operador desde X dentro de Y .

Definición 2.52. Un **operador lineal** $T : \mathcal{D}(T) \rightarrow \mathcal{R}(T)$, es un operador que satisface los siguientes axiomas, para todo $x, y \in \mathcal{D}(T)$ y para todo $\alpha \in K$:⁷

$$T(x + y) = Tx + Ty \tag{OL-1}$$

$$T(\alpha x) = \alpha Tx \tag{OL-2}$$

Definición 2.53. El operador lineal $T : \mathcal{D}(T) \rightarrow Y$ se dice **inyectivo** si diferentes puntos en el dominio tiene diferentes imagenes, es decir

$$x_1 \neq x_2 \Rightarrow Tx_1 \neq Tx_2, \text{ para todo } x_1, x_2 \in \mathcal{D}(T). \tag{2.43}$$

En este caso existe el operador $T^{-1} : \mathcal{R}(T) \rightarrow \mathcal{D}(T)$ que envía cada $y \in \mathcal{R}(T)$ sobre el $x \in \mathcal{D}(T)$ para el cual $Tx = y$. El operador T^{-1} es llamado el **operador inverso** de T . Es posible demostrar que si T^{-1} existe, este es un operador lineal.

⁷En funciones se utiliza la notación $f(x)$, pero en el manejo con operadores la notación usual es Tx en lugar de $T(x)$.

Definición 2.54. Sean X y Y espacios normados y sea $T : \mathcal{D}(T) \rightarrow Y$ un operador lineal, donde $\mathcal{D}(T) \subset X$. El operador T se dice que es **acotado** si existe un número real c tal que, para todo $x \in \mathcal{D}(T)$:

$$\|Tx\| \leq c\|x\|, \quad (2.44)$$

donde la norma de la derecha es la de X y la norma de la izquierda es la de Y .

Definición 2.55. Sea $T : H_1 \rightarrow H_2$ un operador acotado lineal, donde H_1 y H_2 son espacios de Hilbert. El **operador adjunto** de T denotado por \bar{T} está definido por $\bar{T} : H_2 \rightarrow H_1$ tal que para todo $x \in H_1$ y $y \in H_2$ se tiene que $\langle Tx, y \rangle = \langle x, \bar{T}y \rangle$ (el producto interior de la izquierda corresponde al espacio de Hilbert H_2 y el producto interior de la derecha corresponde al espacio de Hilbert H_1).

Definición 2.56. Un operador acotado lineal $T : H \rightarrow H$ sobre un espacio de Hilbert H es:

1. **Auto-adjunto o hermítico** si $\bar{T} = T$.
2. **Unitario** si T es biyectivo y $\bar{T} = T^{-1}$.
3. **Normal** si $T\bar{T} = \bar{T}T$.

Teorema 2.14. Sea $T : H \rightarrow H$ un operador lineal acotado sobre un espacio de Hilbert H . Entonces:

1. Si T es auto-adjunto, $\langle Tx, x \rangle$ es real para todo $x \in H$.
2. Si H es complejo y $\langle Tx, x \rangle$ es real para todo $x \in H$, entonces el operador T es auto-adjunto.

Teorema 2.15. El producto de dos operadores lineales acotados auto-adjuntos S y T en un espacio de Hilbert H es un operador auto-adjunto si y solo si los operadores S y T conmutan, es decir $ST = TS$.

Definición 2.57. Un **funcional** es un operador cuyo rango es los números reales \mathbb{R} o los números complejos \mathbb{C} . Desde este punto de vista podemos clasificar los funcionales como un tipo especial de operadores.

Definición 2.58. Un **funcional lineal** es un operador lineal cuyo dominio es un espacio vectorial X y cuyo rango es el campo escalar de X .

Capítulo 3

Algunos elementos de la mecánica cuántica

3.1. Observaciones iniciales

De forma similar al capítulo anterior, se espera que el nombre de este capítulo indique el carácter parcial e incompleto de su contenido. Se puede pensar que hablar de algunos constructos teóricos propuestos por una teoría que tiene aproximadamente 70 años de estar operando es una tarea fácil. Simplemente consistiría en seleccionar el libro de moda usado por algunas instituciones universitarias y verter algunos de sus contenidos. Por fortuna nuestra tarea no ha sido fácil, no ha sido fácil porque este primer contacto para algunos de nosotros o este re-contacto para otros de nosotros, no fue en busca de las aplicaciones y los ejemplos en los que tanto éxito ha tenido la mecánica cuántica; fue por el contrario, una indagación de los por qué; indagación pensarán algunos, innecesaria; sin sentido pensarán otros.

De la bibliografía revisada encontramos textos bastante adecuados para una primera aproximación y textos definitivamente para expertos en el área. Pero de todos los textos revisados y estudiados, uno de ellos destaca por su claridad conceptual, su rigor matemático y su estilo de presentación, nos referimos al texto “Principios de Mecánica Cuántica” de P. M. A. Dirac. La excelente presentación de Dirac nos sorprende aún más si tenemos presente que fue escrito en 1930, es decir, a muy pocos años de los primeros desarrollos de la teoría que nos ocupa. Por supuesto que pueden existir algunos detractores del texto, inclusive el lector puede ser uno de ellos, es cierto que con respecto a las temáticas de mayor actualidad el texto está desactualizado, pero con respecto a los elementos fundamentales de la teoría, el texto tiene la misma vigencia que en la fecha en que fue publicado por primera vez. Además tiene una ventaja adicional con respecto a otros textos, y es que trabaja con la notación de *brackets* o conocida en la actualidad como la notación de Dirac, que corresponde con la notación empleada por los artículos referentes a la máquina cuántica.

Después de intentar varias formas de presentar los elementos expuestos por Dirac en una forma propia y de los fracasos obtenidos debido a su excelente presentación, decidimos la siguiente metodología de exposición: Vamos a presentar textualmente los dos primeros capítulos del texto de Dirac, de estos capítulos se han eliminado algunos párrafos o secciones, pero cuidando de no perder el hilo de la obra. Los capítulos serán intervenidos bajo una figura que hemos denominado genéricamente *Comentario*. Estos comentarios tienen una naturaleza muy variada, podrán ser observaciones aclaratorias con respecto a lo mencionado en los párrafos anteriores o podrán ser vinculaciones con conceptos de capítulos anteriores de este informe o podrán ser

reflexiones con relación a lo expuesto, entre otras cosas. Para diferenciar claramente que parte es de Dirac y que parte es nuestra hemos seleccionado comenzar cada sección de Dirac por “**Dirac:**”, en el momento en que surja un comentario comenzaremos por “**Comentario:**” y cuando continúe de nuevo Dirac otra vez “**Dirac:**” y así, sucesivamente. Además hemos seleccionado este tipo de letra para las secciones escritas por Dirac.

3.2. Comentarios al libro de Dirac: “Principios de M.C.”

Del prólogo de la primera edición

Dirac: Los métodos de trabajo de la física teórica han sufrido un profundo cambio en el presente siglo. La tradición clásica consideraba el mundo como la asociación de objetos observables (partículas, fluidos, campos, etc.) que se mueven según leyes de fuerza fijas, pudiéndose, pues, formar una imagen mental de todo el esquema en el espacio y el tiempo. Ello conducía a una física cuyo método consistía en hacer hipótesis sobre el mecanismo y las fuerzas que relacionan dichos objetos observables, que explicaran su comportamiento de la forma más sencilla posible. En los últimos tiempos se ha ido haciendo cada vez más evidente que la naturaleza actúa en un plano distinto. Sus leyes fundamentales no rigen el mundo tal como éste aparece a nuestra imagen mental, sino que actúan sobre un substrato del que no podemos formarnos ninguna imagen mental sin cometer desatinos. La formulación de dichas leyes exige el empleo de las matemáticas de transformaciones. Los elementos importantes del mundo físico aparecen como los invariantes de dichas transformaciones (o más en general los casi-invariantes, o cantidades que se transforman de un modo sencillo). Las cosas que conocemos de modo inmediato son las relaciones de estos casi-invariantes con un cierto sistema de referencia, que generalmente elegimos de forma que se introduzcan simplificaciones particulares sin importancia desde el punto de vista de la teoría general.

Comentario: Los esquemas mentales que exige la nueva física, no son en nada comparables a los viejos esquemas de la física clásica, donde cada situación es representable por una imagen mental claramente intuible en el terreno de lo causal determinista. Dadas las condiciones iniciales de un sistema es posible saber con toda precisión la evolución del mismo, debido a que su comportamiento es expresable como un sistema de ecuaciones diferenciales lineales, y hasta cierto punto, es posible conocer sus múltiples relaciones con otros objetos. Contrariamente a esa imagen del mundo la física moderna, en particular la mecánica cuántica, no permite esa estructuración de cosas, la intuición causal-determinista es remplazada por la probabilidad de ocurrencia de un fenómeno físico. No podemos entender por esto una imposibilidad de conocer algunos ámbitos de la naturaleza, o un conocimiento parcial de las cosas, sino más bien una nueva forma de ver y entender el mundo. Modernamente el problema de la mecánica cuántica es entonces tratar de encontrar cuáles cantidades permanecen invariantes bajo transformaciones de coordenadas, esos invariantes al ser expresados en un cierto sistema de coordenadas, serán las nuevas cantidades físicamente contrastables con la evidencia experimental. También, concretamente se entiende en mecánica cuántica el proceso de la medida como un proceso de interacción entre un objeto clásico, tal como un aparato de medida y un objeto cuántico que ocurre aparte y con independencia de un observador cualquiera.

I. El principio de superposición

1. Necesidad de una teoría cuántica

Dirac: La mecánica clásica se ha ido desarrollando progresivamente desde los tiempos de Newton y se ha aplicado a un conjunto de sistemas dinámicos cada vez más amplio, del que forma parte el campo electromagnético en interacción con la materia. Las ideas fundamentales y las leyes que rigen su aplicación constituyen un esquema

tan sencillo y elegante, que parece imposible modificarlo seriamente sin destruir todas sus atractivas características. Sin embargo, se ha conseguido construir un nuevo esquema, llamado mecánica cuántica, más adecuado para la descripción de los fenómenos de escala atómica, y que en ciertos aspectos más elegante y satisfactorio que el esquema clásico. Esto ha sido posible gracias a que los cambios introducidos por la nueva teoría son de carácter muy profundo y no simples modificaciones que destruirían las características de la teoría clásica que le confieren su armonía, y así ha resultado que todas estas características han podido ser incorporadas al nuevo esquema.

La necesidad de elegir un camino distinto al de la mecánica clásica viene exigida por los hechos experimentales. . .

Un ejemplo más de la limitación de mecánica clásica lo constituye el comportamiento de la luz. Tenemos, por un lado, los fenómenos de interferencias y de difracción, que sólo pueden ser explicados mediante una teoría ondulatoria; por el otro, fenómenos tales como la emisión fotoeléctrica y la dispersión¹, de electrones libres, que indican que la luz está compuesta por pequeñas partículas. Dichas partículas, llamadas fotones tienen cada una una energía y momento definidos, que dependen de la frecuencia de la luz, y aparecen con una existencia tan real como la de los electrones o cualesquiera otras partículas conocidas en física. No se observa nunca una fracción de fotón.

Los experimentos han demostrado que este extraño comportamiento no es peculiar de la luz, sino que es completamente general. Todas las partículas materiales tienen propiedades ondulatorias, que se ponen de manifiesto en condiciones adecuadas. Éste es un ejemplo muy sorprendente y característico de los fallos de la mecánica clásica, que no radican simplemente en una inexactitud de sus leyes del movimiento, sino en *una insuficiencia de sus conceptos para proporcionarnos una descripción de los fenómenos atómicos*.

La necesidad de apartarse de las ideas clásicas al intentar dar una explicación de la estructura general de la materia se deriva no sólo de los hechos establecidos experimentalmente, sino también de razones filosóficas generales. Es una interpretación clásica de la constitución de la materia, se supondría que ésta está compuesta por un gran número de pequeñas partes, y se postularían leyes de comportamiento de dichas partes, de las cuales se pudieran deducir las leyes de la materia agrupada. Sin embargo, no sería una explicación completa, pues habría quedado sin considerar la estructura y estabilidad de las partes constituyentes. Para responder a esta cuestión se haría necesario postular que cada una de las partes constituyentes está a su vez compuesta de pequeñas partes, en razón de las cuales debe explicarse su comportamiento. Evidentemente no hay límite para este proceso, de tal forma que es imposible llegar por este procedimiento a la estructura general de la materia. Mientras *grande* y *pequeño* sean conceptos meramente relativos, no conduce a nada explicar lo grande en función de lo pequeño. Por tanto, es necesario que modifiquemos las ideas clásicas de forma que el tamaño adquiera un carácter absoluto.

En este punto es importante recordar que a la ciencia solamente la incumben los objetos observables y que únicamente podemos observar un objeto si interacciona con alguna influencia externa. Todo acto de observación va acompañado necesariamente de una alteración del objeto observado. Podemos decir que un objeto es grande cuando la alteración que acompaña a nuestra observación de él puede ser despreciada, y pequeño cuando no pueda serlo. Esta definición está plenamente de acuerdo con el significado corriente de las expresiones grande y pequeño.

Normalmente se supone que, procediendo con cuidado, podemos reducir la alteración que acompaña a nuestra observación a una amplitud tan pequeña como queramos. En este caso, los conceptos de grande y pequeño son meramente relativos y se refieren tanto a la precisión de nuestros medios de observación como al objeto que se describe. Para poder dar un significado absoluto a la magnitud, tal como se requiere en toda teoría de la estructura elemental de la materia, hemos de suponer que existe *un límite de la precisión de nuestro poder de observación y de la magnitud de la alteración que le acompaña -límite que es inherente a la naturaleza de las cosas y que*

¹Para una explicación completa de los fenómenos de interferencia y difracción ver [43] y para los fenómenos de emisión fotoeléctrica y dispersión ver [22] (esta nota ha sido añadida por nosotros).

es imposible superar aunque se perfeccionen las técnicas o se aumente la habilidad práctica del observador. Si el objeto que consideramos es tal que la alteración límite inevitable se puede despreciar, decimos que el objeto es grande en sentido absoluto y podremos aplicarle la mecánica clásica. Si, en cambio, dicha alteración no es despreciable, el objeto es pequeño en sentido absoluto y será necesaria una nueva teoría para cuenta de él.

Como consecuencia de la discusión precedente debemos revisar nuestras ideas sobre la causalidad. La relación causal se aplica exclusivamente a sistemas que no hayan sido alterados. Pero si el sistema es pequeño, nos es imposible observarlo sin producir en él una alteración considerable, y por lo tanto, no podemos esperar encontrar ninguna relación causal entre los resultados de nuestras medidas. Supondremos que sigue existiendo causalidad para los sistemas mientras no se les altere, y las ecuaciones que estableceremos para describir los sistemas no alterados serán ecuaciones diferenciales que expresarán una relación causal entre condiciones en un instante y condiciones en otro instante posterior. Estas ecuaciones estarán en estrecha correspondencia con las ecuaciones de la mecánica clásica, pero sólo estarán indirectamente relacionadas con los resultados de las observaciones. Existe, pues, una indeterminación inevitable en el cálculo de los resultados observables, y la teoría no nos permite calcular, en general, más que la probabilidad de obtener un resultado particular al hacer la observación.

4. Superposición e incertidumbre

Dirac: Sea un sistema atómico cualquiera, compuesto de partículas o de cuerpos con propiedades dadas (masa, momento de inercia, etc.) que interactúan de acuerdo con leyes de fuerzas dadas. Existirán diversos movimientos posibles de las partículas o de los cuerpos compatibles con las leyes de la fuerza. Cada uno de tales movimientos recibe el nombre de *estado* del sistema. Según las ideas clásicas, se podría especificar un estado dando el valor numérico de todas las coordenadas y velocidades de los diversos componentes del sistema en un instante de tiempo, queda así, completamente determinado el movimiento de todo el sistema. . . La limitación del poder de observación implica una disminución del número de datos que pueden atribuirse a un sistema. Por tanto, el estado de un sistema atómico tiene que caracterizarse por menos datos o por datos más imprecisos que un conjunto completo de valores numéricos de todas las coordenadas y velocidades en un instante de tiempo particular. . .

Puede definirse el estado de un sistema como un movimiento inalterado restringido por tantas condiciones o datos como sea posible teóricamente sin que se interfieran o se contradigan mutuamente. En la práctica, estas condiciones se pueden imponer mediante una preparación adecuada del sistema, que consista, por ejemplo, en hacerlo pasar a través de una serie de aparatos, tales como rendijas y polarímetros, y dejando al sistema inalterado después de la preparación. La palabra "estado" se emplea indistintamente para designar el estado en un instante particular (después de la preparación) y el estado durante todo el tiempo posterior a la preparación. . .

Comentario: Preparar un sistema cuántico es someter a dicho sistema a una serie de condiciones experimentales con el único fin de posibilitar alguna medida del mismo, al prepararlo, de antemano se prevee cual característica se quiere medir, luego se deja que el sistema evolucione por sí mismo, pudiéndose de este modo caracterizar alguna propiedad del sistema. Hay diferentes maneras de preparar un sistema cuántico, como por ejemplo, hacerlo interactuar con un objeto clásico como una rejilla, un campo magnético o en general cualquier polarímetro. También es posible tener el caso general de una colección de N sistemas idénticos, preparados bajo las mismas condiciones experimentales, al medir sobre todos una misma magnitud física obtendremos una serie de valores de dicha magnitud no necesariamente coincidentes, cuyas frecuencias de repetición en límite al infinito dan origen a una distribución de probabilidad. La colección de tales distribuciones de probabilidad para cada observable, es la encargada de caracterizar el estado del sistema en cuanto a su carácter predictivo. Así, dos estados, aun preparados mediante operaciones distintas, se considerarán iguales cuando conducen a las mismas distribuciones de probabilidad para cada observable (cf. [9]).

Clásicamente es posible, con toda certeza preparar un sistema. Sí, se conocen las condiciones iniciales, podemos tener de alguna manera control sobre su evolución, estando representado el estado, en un instante de tiempo por un punto del espacio fásico. Debido a que la información que se tiene sobre el sistema es máxima se dice entonces que el sistema está en un estado puro, en dicho estado, la medida de cualquier observable sobre estados idénticamente preparados, dará siempre el mismo resultado.

Al tratar de hacer lo mismo con sistemas cuánticos, nos encontramos con muchos tropiezos ya que por ejemplo, la posición y la velocidad no son cantidades que puedan ser conocidas simultáneamente, la medida de la posición indetermina la medida de la velocidad y viceversa, es decir no son cantidades físicamente compatibles, además la interacción entre el aparato de medida y lo que se mide, pueden interactuar de forma incontrolable, es decir, de forma tal que se pierda toda información sobre el mismo. Pese a esas dificultades es posible preparar un sistema con un número maximal de observables compatibles e independientes, o dicho de otra manera, que podamos medir simultáneamente sin modificar el sistema, o mejor aun, sin que las medidas interfieran entre sí. En esas circunstancias tenemos el sistema en un estado puro, es decir, poseemos sobre él la máxima información posible. El caso contrario es cuando no todos los observables son compatibles, de algun modo se pierde información, esos estados se conocen como estados mezclados.

Dirac: El principio general de superposición de la mecánica cuántica se aplica a los estados de todo sistema dinámico con cualquiera de los dos significados precedentes. Según dicho principio existe una relación peculiar entre los estados, de forma que cuando un sistema está en un estado bien definido, a la vez se puede considerar que está parcialmente es cada uno de una serie de estados. El estado original debe considerarse como el resultado de una *superposición* de esta serie de estados, lo que es inconcebible desde el punto de vista clásico. Es evidente que existen infinitas maneras de realizar dicha superposición para un mismo estado dado. Recíprocamente, todo conjunto de dos o más estados puede superponerse para dar lugar a un nuevo estado...

Comentario: En el caso simple de tener un estado que se pueda expresar como una superposición de dos estados independientes, siempre se podrá encontrar dos nuevos estados independientes que superpuestos generen el mismo estado original. Una buena analogía, en el lenguaje matemático, la podemos hacer con los elementos del espacio bidimensional real (plano cartesiano). Consideremos los vectores base ortonormales $\{\hat{e}_1, \hat{e}_2\}$ (base **A**). Cualquier otra base ortonormal (base **B**) se puede expresar como una combinación lineal de la base **A**, así:

$$\begin{aligned}\hat{e}'_1 &= (\cos \theta)\hat{e}_1 + (\sin \theta)\hat{e}_2 \\ \hat{e}'_2 &= (-\sin \theta)\hat{e}_1 + (\cos \theta)\hat{e}_2,\end{aligned}$$

lo que corresponde a una rotación ortonormal en el plano. Podemos observar que

$$\hat{e}'_i \cdot \hat{e}'_j = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j; \\ 1 & \text{si } i = j. \end{cases}$$

Ahora, si un vector \vec{r} viene dado en la base **A** como:

$$\vec{r} = x_1\hat{e}_1 + x_2\hat{e}_2$$

en la base **B** se expresará como:

$$\vec{r} = (x_1 \cos \theta + x_2 \sin \theta)\hat{e}'_1 + (-x_1 \sin \theta + x_2 \cos \theta)\hat{e}'_2$$

Se puede ver entonces de esta última ecuación que existen un número infinito de bases ortonormales en las cuales se puede expresar el vector \vec{r} : el parámetro θ es una variable con rango continuo $[0, 2\pi]$. Es decir, para cada valor de θ en este rango se tiene una nueva base **B**.

Dirac: La relación entre los estados de cualquier sistema que implica el principio de superposición es de tal naturaleza que no es posible expresarla mediante los conceptos físicos habituales. Desde el punto de vista clásico, no es posible imaginar que un sistema esté parcialmente en cada uno de dos estados dados, y que esto se equivalente a estar completamente en otro estado distinto. Ello implica una idea completamente nueva a la que nos debemos ir acostumbrando, y sobre cuya base tenemos que construir una teoría matemática exacta sin disponer de ninguna imagen clásica precisa.

Un estado formado por superposición de otros dos tendrá propiedades que, en cierto sentido, serán intermedias entre las de ambos sistemas de partida, aproximándose en mayor o menor grado a las de cada uno, según se la haya atribuido un "peso" mayor o menor en el proceso de superposición. El nuevo estado estará completamente determinado por los dos estados de partida, cuando se conozcan sus pesos relativos en el proceso de superposición y una diferencia de fase; el significado preciso de pesos y fases en general nos lo proporcionará la teoría matemática. . .

Comentario: El significado de factores de peso y fases puede ser entendido con base en el siguiente ejemplo, aunque se trata de una cuestión puramente clásica, es posible sacar algunas conclusiones desde el punto de vista cuántico.

Cuando se superponen dos ondas polarizadas perpendicularmente como las siguientes:

$$\begin{aligned} E_x &= E \cos(\omega t) \\ E_y &= E \cos(\omega t + \delta) \end{aligned}$$

El resultado de la superposición está determinado por la amplitud resultante, que puede ser identificada en nuestro caso con el factor de peso E y el ángulo δ , en nuestro caso la diferencia de fase. Según la diferencia de fase, si suponemos el mismo factor de peso para la superposición, obtendremos diferentes estados finales, en particular si $\delta = 0$, el estado resultante estará polarizado rectilíneamente según un ángulo de $\pi/4$, si $\delta = \pi/4$ el estado resultante estará polarizado circularmente en todas las direcciones, etc.

Dirac: La naturaleza no clásica del proceso de superposición se pone claramente de manifiesto al considerar dos estados A y B , tales que exista una observación que aplicada al sistema en el estado A dé siempre el resultado a y aplicada al sistema en el estado B dé siempre el mismo resultado particular b distinto de a . ¿Cuál será el resultado de la observación aplicada al sistema en el estado resultante de la superposición? La respuesta es que unas veces obtendremos el resultado a y otras el b , según una ley probabilística que depende únicamente de los pesos relativos de A y B en el proceso de superposición. En ningún caso obtendremos un resultado distinto de a o b . *El carácter intermedio del estado formado por la superposición reside en el hecho de que la probabilidad de obtener un resultado particular en una observación es intermedia entre las probabilidades correspondientes a los estados de partida y no en que el mismo resultado sea intermedio entre los resultados correspondientes a dichos estados.*

Vemos, pues, que un abandono tan radical de las ideas clásicas, como el afirmar la existencia de relaciones de superposición entre los estados, sólo ha sido posible al tener en cuenta explícitamente la importancia de la alteración que acompaña a toda observación y la incertidumbre consiguiente en el resultado. Cuando se lleva a cabo una observación de un sistema atómico, en general no está determinado el resultado, es decir, que si repetimos el experimento varias veces bajo idénticas condiciones, podemos obtener diversos resultados. Sin embargo, es una ley de la naturaleza el hecho que si repetimos el experimento un gran número de veces, cada resultado particular aparece en una fracción bien definida del número total de veces, y por tanto, existe una *probabilidad* determinada de obtenerlo. Esta probabilidad es la que la teoría nos permite calcular. . .

La hipótesis de que existen relaciones de superposición entre los estados nos lleva a establecer una teoría matemática en la cual las ecuaciones que definen el estado son lineales en las incógnitas. . .

Comentario: El requerimiento de la validez del principio de superposición en la mecánica cuántica queda de algún modo justificado gracias a que las ecuaciones diferenciales de evolución con las que se edifica dicha mecánica admiten soluciones para las cuales el principio de superposición es válido. Parece milagroso que las soluciones de las ecuaciones diferenciales de la nueva mecánica resulten ser las mismas funciones conocidas como soluciones de las ecuaciones clásicas: polinomios de Hermite, Legendre, Laguerre, armónicos esféricos, funciones de Bessel, para citar algunas de tales funciones, en todas es válido el principio de superposición e inherente a ellas la estructura de espacio vectorial.

¿Por qué la mecánica cuántica requiere de un principio de superposición un tanto parecido al mismo principio clásico? ¿Será ésto una cuestión puramente física o una limitación de la matemática que estamos acostumbrados a usar? Actualmente no solamente en la mecánica cuántica sino en otras ramas de la física, se usa frecuentemente un artificio, un método de perturbaciones el cual permite de alguna manera extraer los términos lineales en muchos tipos de procesos en la física. ¿Será que los métodos de perturbaciones no son otra cosa que la ignorancia ante otros métodos matemáticos diferentes a los clásicamente conocidos? Podría pensarse y de alguna manera es cierto, que los procesos o fenómenos en la física no son lineales, que ellos están regidos por algunos tipos de ecuaciones diferenciales no lineales, cuya estructura subyacente a sus soluciones son variedades diferenciables, no por esto dejan de satisfacer el principio de superposición sino más bien lo generalizan. Algunos autores en particular han intentado construir una mecánica cuántica no lineal, aunque no han conseguido lograrlo (cf. [42]).

5. Formulación matemática del principio de superposición

Dirac: Durante el presente siglo ha tenido lugar un profundo cambio en las opiniones que los científicos tenían sobre los principios matemáticos de la física. Anteriormente se suponía que la mecánica de Newton constituía la base para la descripción de todos los fenómenos físicos, y que por tanto, todo físico teórico debía contribuir al desarrollo y aplicación de estos principios. El reconocimiento de que no existe razón lógica alguna para que los principios newtonianos y otros principios clásicos tengan que ser válidos fuera del campo en que han sido comprobados experimentalmente, trajo consigo las primeras experiencias que pusieron de manifiesto la absoluta necesidad de apartarse de estos principios. Los intentos de superar las divergencias se concretaron en la introducción de nuevos formalismos matemáticos, de nuevos sistemas de axiomas y reglas en los métodos de la física teórica.

La mecánica cuántica es un buen ejemplo de las nuevas ideas. En ella se exige que los estados de un sistema dinámico estén relacionados con las variables dinámicas de una forma nueva, ininteligible desde el punto de vista clásico. Los estados y las variables dinámicas deben estar representadas por cantidades matemáticas de naturaleza diferente a las que se utilizan corrientemente en física. El nuevo esquema constituirá una teoría física precisa cuando se hayan especificado todos los axiomas y reglas que rigen las cantidades matemáticas se hayan dado ciertas leyes que relacionen los hechos físicos con dichas cantidades, de tal manera que de unas condiciones físicas dadas puedan inferirse ecuaciones matemáticas y recíprocamente.

Comentario: En la actualidad existen varias presentaciones axiomáticas (postulacionales) de la mecánica cuántica (cf. [9], [6] entre otros). Es importante mencionar que *tradicionalmente* una teoría o sistema teórico está axiomatizado, si existe un conjunto de enunciados llamados *los axiomas* o *los postulados* que satisfacen las siguientes condiciones (estas condiciones son tomadas de [29], pág. 69):

El conjunto de axiomas-postulados:

1. No presenta *contradicción* ni permite generarlas.²
2. Es *independiente*, es decir, ningún axioma se puede deducir a partir de los otros.

²Esta condición será ampliada en el próximo comentario.

3. Es *suficiente*, es decir, permite deducir todos los enunciados de la teoría.
4. Es *necesario*, es decir, no deben contener supuestos superfluos.

No es este el lugar ni el momento para realizar una discusión acerca de la axiomatización de teorías en general y en particular de la mecánica cuántica; pero si queremos llamar la atención del lector acerca de la *efectividad* de dicho proceso, así, como de la dificultad de su verificación. No obstante, deseamos mencionar un hecho que nos ha llamado la atención, uno de los principios fundamentales de la mecánica cuántica —el principio de superposición— no aparece como un axioma-postulado de la misma, si no que se deduce como una consecuencia de éstos; de forma similar, el principio del tercio excluido o el principio de no contradicción, no aparecen como axiomas de la lógica formal (de enunciados o de predicados), sino que son consecuencias de sus respectivas axiomáticas. La situación es como si se “seleccionaran” los principios que una teoría debe contener y se buscará un conjunto de axiomas que hicieran “posibles” dichos principios.

Dirac: En la aplicación de la teoría deberíamos disponer de cierta información física, que procederíamos a expresar en forma de ecuaciones entre las cantidades matemáticas. Con ayuda de los axiomas y las reglas de manipulación deduciríamos nuevas ecuaciones y acabaríamos interpretándolas como condiciones físicas. La justificación de todo el esquema depende, aparte de su coherencia intrínseca, de la concordancia de los resultados finales con los experimentos.

Comentario: La última frase del párrafo anterior permite dos comentarios, el primero en relación a la coherencia del sistema y el segundo en relación a la concordancia con los experimentos.

Con respecto a la coherencia del sistema, diremos que ha sido un principio impuesto a los sistemas axiomáticos desde Aristóteles: *El principio de no contradicción*. Popper expone de una manera muy enfática la necesidad de dicho principio: “*El requisito de la compatibilidad o coherencia desempeña un papel especial entre todos los que han de satisfacer los sistemas teóricos, o los sistemas axiomáticos. Puede considerársele la primera condición que ha de cumplir todo sistema teórico, ya sea empírico o no... Caeremos en cuenta de la importancia que tiene el requisito de coherencia si nos percatamos de que los sistemas contradictorios no nos proporcionan ninguna información, pues podemos deducir de ellos las conclusiones que nos plazcan; de modo que no hace discriminación alguna en los enunciados -calificándolos, bien de incompatibles, bien de deductibles-, ya que todos son deductibles. En cambio, un sistema coherente divide el conjunto de todos los enunciados posibles en dos: los que le contradicen y los que son compatibles con él. Es ésta la razón por la que la coherencia constituye el requisito más general que han de cumplir los sistemas, ya sean empíricos o no lo sean, para que puedan tener alguna utilidad.*”([29], pág. 88).

En el caso de los sistemas axiomáticos empíricos, el que acarrea la coherencia del sistema es el sistema lógico subyacente. Es fácil demostrar que en un sistema teórico no coherente (inconsistente) construido sobre la lógica de enunciados es posible demostrar cualquier enunciado (cf. [21], pág. 56).

Desde las primeras formulaciones de la mecánica cuántica se ha sugerido que ésta tiene una lógica subyacente diferente a la tradicional que ha sido llamada por algunos *lógica cuántica* (cf. [13]). Lo importante de este punto es que por lo menos se discute la posibilidad de una lógica diferente para “manipular” los procesos cuánticos. Por otra parte ha diferencia de la época en que Dirac y Popper escribieron su opiniones acerca de la coherencia del sistema, hoy en día existen varias propuestas de lógicas que “violentan” el principio de no contradicción, dichas lógicas reciben en general el nombre de *lógicas paraconsistentes* (cf. [4]). Estas lógicas han servido en la actualidad para empezar a construir propuestas de *sistemas matemáticos inconsistentes* (cf. [25]), lo cual nos lleva a pensar en la posibilidad de sistemas axiomáticos empíricos inconsistentes y en particular la existencia de uno de tales sistemas para la mecánica cuántica.

Con respecto a la concordancia con los experimentos, diremos que podemos dividir las ciencias, en ciencias formales y en ciencias factuales, éstas últimas además de satisfacer los métodos de deducción de

la lógica subyacente sobre la cual operan, necesitan de la observación y de la experimentación para poder “construir” sus respectivos enunciados. No hay una postura epistemológica exacta acerca de que característica deben tener estos enunciados, algunos autores mencionan que dichos enunciados deben ser *verificables en la experiencia* (cf. [5]), otros autores mencionan que dichos enunciados deben ser *falseables* (cf. [29]), etc. Lo que sí es claro, es que los procesos de deducción que prueban la verdad de los teoremas de la ciencias formales (lógica, matemática) son insuficientes para verificar los enunciados de las ciencias factuales.

Dirac: Empezaremos a construir el esquema considerando las relaciones matemáticas entre los estados de un sistema dinámico en un instante de tiempo que se derivan de la formulación matemática del principio de superposición. La superposición es un cierto proceso aditivo, e implica que los estados puedan sumarse de algún modo para dar nuevos estados. Por lo tanto, los estados tienen que estar asociados con cantidades matemáticas que puedan sumarse entre sí, para dar cantidades de la misma clase. Las cantidades matemáticas más sencillas que disfrutan de esta propiedad son los vectores. Los vectores ordinarios definidos en un espacio de un número finito de dimensiones no son suficientemente generales para la mayoría de los sistemas dinámicos de la mecánica cuántica. Nos vemos obligados a generalizar los vectores a un espacio de infinitas dimensiones, con lo que el tratamiento matemático se hace complicado por razones de convergencia. Sin embargo, de momento únicamente vamos a considerar propiedades que puedan ser deducidas sobre la base de un conjunto sencillo de axiomas, y dejaremos a un lado las cuestiones de convergencia y los temas relacionados con ella hasta que nos veamos obligados a tenerlos en cuenta.

Es conveniente designar con un nombre especial a los vectores que se asocian a los estados de un sistema en mecánica cuántica, tanto si forman parte de un espacio de un número finito de dimensiones como de un espacio de infinitas dimensiones. Las denominaremos *vectores ket*, o simplemente *kets*, y los representaremos con el símbolo especial $|\rangle$. Si queremos especificar un ket particular mediante una letra, por ejemplo la A , la colocaremos entre los dos signos así, $|A\rangle$. La conveniencia de esta notación aparecerá clara cuando hayamos desarrollado todo el esquema.

Los vectores ket se pueden multiplicar por números complejos y pueden también sumarse entre ellos para dar nuevos kets; por ejemplo, de los dos kets $|A\rangle$ y $|B\rangle$ podemos formar

$$c_1 |A\rangle + c_2 |A\rangle = |R\rangle \quad (3.1)$$

en donde c_1 y c_2 son números complejos arbitrarios. También es posible efectuar con ellos operaciones lineales más generales, como sumar una serie infinita de kets, o si tenemos un ket $|x\rangle$ que dependa de un parámetro x que puede tomar todos los valores de un cierto intervalo, integrar respecto a x para obtener un nuevo ket

$$\int |x\rangle dx = |Q\rangle$$

Todo vector ket que se pueda expresar linealmente en función de los otros se dice que es dependiente de ellos. Se dice que un conjunto de kets es independiente, si ninguno de ellos se puede expresar linealmente en función de los otros.

Ahora hacemos la hipótesis de que *a cada estado de un sistema dinámico en un instante particular le corresponde un ket, siendo la correspondencia tal que si un estado está definido como superposición de otros dos, su correspondiente ket puede expresarse linealmente en función, de los kets correspondientes a dichos estados y recíprocamente*. Por tanto, el estado R resulta de una superposición de los estados A y B si los correspondientes kets están ligados por (3.1).

La hipótesis anterior nos lleva a introducir ciertas propiedades del proceso de superposición, que de hecho son necesarias para que sea apropiada la palabra “superposición”. Cuando se superponen dos o más estados, no importa el orden en que entran en el proceso de superposición, y así, dicho proceso es simétrico respecto a los

estados que se superponen. Además, de (3.1) deducimos que (exceptuando el caso en que c_1 o c_2 sean nulos) si el estado R puede formarse por superposición de los estados A y B , el estado A puede formarse por superposición de B y R , y el B por superposición de A y R . Las relaciones de superposición son, pues, simétricas respecto a los tres estados A , B y R .

Cuando un estado está formado por la superposición de otros dos se dice que es dependiente de ellos. Más en general, se dice que un estado es dependiente de un conjunto finito o infinito de otros estados, si su correspondiente ket es dependiente de los kets que corresponden a dichos estados. Se dice que un conjunto de estados es independiente si ninguno de los estados es dependiente de los otros.

Para proseguir con la formulación matemática del principio de superposición tenemos que introducir una nueva hipótesis, y afirmar que por superposición de un estado consigo mismo no podemos construir ningún estado nuevo, sino que siempre obtenemos el mismo estado. Si el estado original correspondía al ket $|A\rangle$, al superponerle consigo mismo el estado resultante corresponde a

$$c_1 |A\rangle + c_2 |A\rangle = (c_1 + c_2) |A\rangle,$$

donde c_1 y c_2 son números. Puede ocurrir que $c_1 + c_2 = 0$, en cuyo caso el resultado de la superposición no representa nada en absoluto, pues las dos componentes se han eliminado mutuamente por un efecto de interferencia. Nuestra nueva hipótesis exige que, salvo en este caso particular, el estado resultante sea el mismo que el original, y por tanto, $(c_1 + c_2) |A\rangle$ tiene que corresponder al mismo estado al que corresponde $|A\rangle$. Pero $c_1 + c_2$ es un número complejo arbitrario, de donde deducimos que *si multiplicamos el ket que corresponde a un estado por un número complejo arbitrario distinto de cero, el ket resultante corresponderá al mismo estado*. Así, pues, un estado viene caracterizado por la dirección de un ket, independientemente de la longitud que le atribuyamos. Los estados de un sistema dinámico están en correspondencia biyectiva con las posibles direcciones de los vectores ket, considerando como una sola las direcciones de $|A\rangle$ y de $-|A\rangle$...

Dados dos estados que correspondan a dos kets $|A\rangle$ y $|B\rangle$, el estado más general que se puede formar por superposición de ambos corresponde a un ket $|R\rangle$ que viene determinado dando dos números complejos: los coeficientes c_1 y c_2 de la ecuación (3.1). Si multiplicamos los dos coeficientes por un mismo factor (también complejo) el ket $|R\rangle$ quedará multiplicado por dicho factor y el estado correspondiente será el mismo de antes. Por lo tanto, únicamente interviene en la determinación del estado R el cociente entre ambos coeficientes. Así, pues, dicho estado queda determinado por un número complejo, o lo que es lo mismo, por dos parámetros reales. Por lo tanto, dados dos estados, por superposición de ambos podemos formar una doble infinidad de estados.

Comentario: Consideremos un espacio vectorial bidimensional (lo que será válido para cualquier dimensión). El campo escalar, puede ser los reales o los números complejos. Ahora, un estado cuántico o ket, se puede expresar (por el solo hecho de ser un vector) como una combinación lineal de dos estados cuánticos independientes, A y B , así:

$$|R\rangle = c_1 |A\rangle + c_2 |B\rangle,$$

donde c_i son escalares.

Supongamos que $c_1 \neq 0$, pudiendo expresar a $|R\rangle$ como:

$$|R\rangle = c_1(|A\rangle + \frac{c_2}{c_1} |B\rangle).$$

Dada la naturaleza de los escalares (reales o complejos) se presentan dos casos: doble infinidad o simple infinidad de estados. Esto aparece cuando los escalares son complejos o reales, respectivamente. Una interpretación de los coeficientes de la combinación lineal en la que se puede expresar el ket en mecánica cuántica es el de una amplitud de probabilidad. En este sentido, el vector $|R\rangle$ está determinado por un sólo parámetro

independiente c_2/c_1 , dado que el ket debe estar normalizado a alguna magnitud real (usualmente a 1) si se le quiere dar el sentido de una probabilidad. Es decir, el coeficiente “global” de la expresión anterior está completamente determinado.

Ahora, si los c_i son reales entonces cualquier ket expresado en esta base se determina por un sólo parámetro (real). En este caso hablamos de *simple infinidad de estados*. De otro lado, si los c_i son complejos, el cociente c_2/c_1 será complejo y por lo tanto se requieren dos parámetros independientes (reales) para determinar cualquier ket en esta base. En este caso hablamos de *doble infinidad de estados*.

6. Vectores bra y ket

Dirac: Siempre que tenemos un conjunto de vectores en una teoría matemática cualquiera podemos construir un segundo conjunto de vectores, que los matemáticos denominan vectores duales. Vamos a describir el método de obtenerlo en el caso de que los vectores de partida sean nuestros kets.

Sea un número ϕ función de un ket $|A\rangle$; es decir, que a cada ket $|A\rangle$ le corresponde un número ϕ , y además exijamos que la función sea lineal, lo que quiere decir que el número que corresponde a $|A\rangle + |A'\rangle$ es igual a la suma de los números que corresponden a $|A\rangle$ y a $|A'\rangle$, y que el número que corresponde a $c|A\rangle$ es igual a c multiplicado por el número que corresponde a $|A\rangle$, siendo c un factor numérico arbitrario. Entonces, el número ϕ que corresponde a $|A\rangle$ puede ser considerado como el producto escalar de dicho $|A\rangle$ con un nuevo vector, existiendo tantos vectores nuevos como funciones lineales de los vectores ket. La justificación de que podamos considerar a ϕ de este modo, reside, como veremos más adelante (véanse las ecuaciones (3.5) y (3.6)), en el hecho de que los nuevos vectores pueden sumarse entre ellos y multiplicarse por números dando nuevos vectores de la misma clase. A pesar de que los nuevos vectores sólo están dados cuando conocemos los productos escalares con los vectores de partida, esto nos basta para construir una teoría matemática con ellos.

Les denominaremos *vectores bra* o simplemente *bras*, y los representaremos con el símbolo $\langle |$, imagen simétrica del símbolo de un ket. Cuando queramos especificar un bra particular mediante una letra B , la escribiremos entre los dos signos así, $\langle B|$. El producto escalar del bra $\langle B|$ y del ket $|A\rangle$ lo escribiremos $\langle B|A\rangle$, es decir, yuxtaponiendo los símbolos del bra y del ket, con el bra a la izquierda, y contrayendo las dos líneas verticales en una sola para abreviar.

Comentario: Es bastante curioso la forma como Dirac introduce el concepto de producto escalar o producto interior entre vectores. Se puede afirmar que es una introducción en *palabras*, a diferencia de la introducción usual que podríamos denominar más formal. De otro modo, la definición de producto escalar ofrecida por Dirac y la definición 2.46 no presentan ninguna diferencia. Continúa Dirac estableciendo las propiedades que debe satisfacer el producto escalar, propiedades similares a los axiomas citados en la definición 2.46.

Dirac: Podemos considerar los símbolos \langle y \rangle como tipos característicos de paréntesis. Un producto escalar $\langle B|A\rangle$ aparece ahora como una expresión entre paréntesis completos, mientras que un bra $\langle B|$ o un ket $|A\rangle$ son expresiones entre paréntesis incompletos. Por lo tanto, son válidas las siguientes reglas: toda expresión entre paréntesis completos expresa un número, y toda expresión entre paréntesis incompletos expresa un vector, que será bra o ket según tenga el primer paréntesis o el segundo.

La condición de que el producto escalar de $\langle B|A\rangle$ sea una función lineal de $|A\rangle$ se puede escribir simbólicamente así:

$$\langle B|\{A + |A'\}\rangle = \langle B|A\rangle + \langle B|A'\rangle \quad (3.2)$$

$$\langle B|\{c|A\rangle\} = c\langle B|A\rangle \quad (3.3)$$

donde c es un número.

Se considera que un bra está completamente definido, cuando se conoce su producto escalar con cualquier ket, de modo que si el producto escalar de un determinado bra con todo ket es nulo, debe considerarse también nulo el bra. En símbolos, sí

$$\langle P | A \rangle = 0, \quad \text{para todo } |A\rangle, \quad \text{entonces } \langle P | = 0 \quad (3.4)$$

Definimos la suma de dos bras $\langle B |$ y $\langle B' |$ por la condición de que su producto escalar con cualquier ket $|A\rangle$ sea igual a la suma de los productos escalares de $\langle B |$ y $\langle B' |$ con $|A\rangle$,

$$\{\langle B | + \langle B' | \} |A\rangle = \langle B | A \rangle + \langle B' | A \rangle, \quad (3.5)$$

y el producto de un bra $\langle B |$ con un número c por la condición de que su producto escalar con cualquier ket $|A\rangle$ sea igual a c multiplicado por el producto escalar de $\langle B |$ y $|A\rangle$,

$$\{c \langle B | \} |A\rangle = c \langle B | A \rangle. \quad (3.6)$$

Comentario: La ecuación (3.5) corresponde al axioma PI-1 y la ecuación (3.6) corresponde al axioma PI-2.

Dirac: Las ecuaciones (3.2) y (3.5) expresan que el producto escalar de un bra y un ket verifica la propiedad distributiva de la multiplicación, y las (3.3) y (3.6) indican que la multiplicación por factores numéricos verifica las propiedades algebraicas ordinarias.

Los vectores bra, tal como los hemos introducido aquí, son de distinta naturaleza que los vectores ket, y hasta aquí, aparte del producto escalar, no existe ninguna otra relación entre bras y kets. Ahora hacemos la hipótesis de que *existe una correspondencia biyectiva entre los bras y los kets, tal que el bra que corresponde a $|A\rangle + |A'\rangle$ es igual a la suma de los bras correspondientes a $|A\rangle$ y $|A'\rangle$ y el bra correspondiente a $c|A\rangle$ es igual a \bar{c} multiplicado por el bra correspondiente a $|A\rangle$, siendo \bar{c} el complejo conjugado de c .* Emplearemos la misma letra para indicar un ket y su bra correspondiente. Así el bra que corresponde a $|A\rangle$ es $\langle A |$.

La relación que existe entre un ket y su correspondiente bra justifica el que llamemos a cada uno el conjugado imaginario del otro. Nuestros bras y kets son cantidades complejas, puesto que se pueden multiplicar por números complejos obteniéndose vectores de la misma naturaleza que antes; pero son cantidades complejas especiales, ya que no pueden separarse en parte real y parte imaginaria pura. El procedimiento habitual de obtener la parte real de una cantidad compleja, que consiste en tomar la semisuma de dicha cantidad y su conjugada, es inaplicable debido a que los bras y los kets son de distinta naturaleza y no se pueden sumar unos con otros. Para llamar la atención sobre esta diferencia, utilizaremos las palabras “complejo conjugado” cuando nos refiramos a números u otras cantidades complejas que puedan separarse en parte real y parte imaginaria pura, y las de “imaginario conjugado” para las que no disfruten de esta propiedad. Para las cantidades del primer tipo, cuando queramos indicar el complejo conjugado de una cantidad dada, emplearemos la notación que consiste en colocar una raya sobre dicha cantidad.

Teniendo en cuenta la correspondencia biyectiva entre bras y kets, *todo estado de nuestro sistema dinámico en un instante particular puede especificarse tanto por la dirección de un bra como por la de un ket.* De hecho toda la teoría será simétrica en lo esencial respecto a bras y kets.

Dados dos kets $|A\rangle$ y $|B\rangle$, podemos formar con ellos un número $\langle B | A \rangle$, producto escalar del primero con el imaginario conjugado del segundo. Tal número depende linealmente de $|A\rangle$ y antilinealmente de $|B\rangle$. Dependencia antilineal significa que el número que se obtiene con $|B\rangle + |B'\rangle$ es igual a la suma de los números que se obtienen con $|B\rangle$ y con $|B'\rangle$, y que el número que se obtiene con $c|B\rangle$ es igual a \bar{c} multiplicado por el número que se obtiene con $|B\rangle$. Hay otro procedimiento de formar un número que dependa linealmente de $|A\rangle$ y antilinealmente

de $|B\rangle$, que consiste en tomar el número complejo conjugado del producto escalar de $|B\rangle$, con el conjugado imaginario de $|A\rangle$. *Supondremos que estos dos números son iguales*, es decir

$$\langle B | A \rangle = \overline{\langle A | B \rangle} \quad (3.7)$$

Comentario: La ecuación (3.7) corresponde al axioma **PI-4**.

Dirac: Poniendo aquí $|B\rangle = |A\rangle$, resulta que el número $\langle A | A \rangle$ es real. Hacemos además la nueva hipótesis de que

$$\langle A | A \rangle > 0, \quad (3.8)$$

salvo si $|A\rangle = 0$.

Comentario: Finalmente la ecuación (3.8) corresponde al axioma **PI-3**. Completa Dirac en esta forma la construcción del producto escalar, presentado implícitamente los axiomas que debe satisfacer. Remarcamos de nuevo, la importancia que concede Dirac a explicar en palabras la justificación de cada axioma.

Dirac: En el espacio ordinario, dados dos vectores, se puede formar con ellos un número—su producto escalar—que es real y simétrico respecto a ellos. En el espacio de los vectores bra o en el de los vectores ket, dados dos vectores cualesquiera también podemos formar un número -el producto escalar de uno de ellos con el imaginario conjugado del otro- que ahora es complejo, y se transforma en su complejo conjugado al cambiar el orden de los vectores. Existe, por tanto, un tipo de perpendicularidad en estos espacios, que es una generalización de la perpendicularidad en el espacio ordinario. Diremos que un bra y un ket son *ortogonales* si su producto escalar es nulo, y que dos bras o dos kets son ortogonales, si el producto escalar de uno de ellos por el imaginario conjugado del otro es cero. Además, diremos que dos estados de nuestro sistema dinámico son ortogonales, si sus vectores correspondientes también lo son.

Definimos la *longitud* de un bra $\langle A |$ o de su ket imaginario conjugado $|A\rangle$ como la raíz cuadrada del número positivo $\langle A | A \rangle$. Si dado un estado queremos caracterizarlo mediante un bra o un ket, únicamente queda determinada la dirección del vector, pero el vector queda indeterminado en un factor numérico arbitrario. A menudo es conveniente elegir dicho factor de forma que el vector tenga longitud unidad. Este procedimiento se llama *normalización* y el vector así elegido se dice que está normalizado. Sin embargo, aún no está completamente determinado el vector, pues se le puede multiplicar por cualquier factor de módulo uno, es decir, cualquier número de la forma $e^{i\gamma}$ con γ real, sin modificar su longitud. A este número le llamaremos *factor de fase*.

Comentario: La expansión de e^x para x real, por series de Taylor bien dada por [2]:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

Si reemplazamos x por iy obtenemos [26]:

$$\begin{aligned} e^{iy} &= 1 + iy - \frac{y^2}{2!} - i\frac{y^3}{3!} + \frac{y^4}{4!} + \dots \\ &= \left(1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \dots\right) + i\left(y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \dots\right) \end{aligned}$$

Como las expansiones en series de Taylor de $\cos y$ y $\sin y$ son [2]:

$$\begin{aligned} \cos y &= 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \dots + (-1)^n \frac{y^{2n}}{(2n)!} + \dots \\ \sin y &= y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \dots + (-1)^{n-1} \frac{y^{2n-1}}{(2n-1)!} + \dots \end{aligned}$$

Entonces podemos interpretar e^{iy} por:

$$e^{iy} = \cos y + i \operatorname{sen} y$$

De donde obtenemos que:

$$\begin{aligned} |e^{iy}|^2 &= |\cos y + i \operatorname{sen} y|^2 \\ &= (\cos y)^2 + (\operatorname{sen} y)^2 \\ &= 1. \end{aligned}$$

La longitud de un ket $|A\rangle$ viene dada por $\sqrt{\langle A|A\rangle}$ y la longitud del ket $|A\rangle$ por un factor de fase $e^{i\gamma}$, es decir la longitud del ket $e^{i\gamma}|A\rangle$ viene dada por:

$$\begin{aligned} \text{longitud de } e^{i\gamma}|A\rangle &= \sqrt{\langle e^{i\gamma}A, e^{i\gamma}A\rangle} \\ &= \sqrt{e^{i\gamma}\langle A, e^{i\gamma}A\rangle} \\ &= \sqrt{e^{i\gamma}e^{-i\gamma}\langle A, A\rangle} \\ &= \sqrt{|e^{i\gamma}|^2\langle A, A\rangle} \\ &= \sqrt{\langle A, A\rangle}. \end{aligned}$$

Y es por esto que un factor de fase $e^{i\gamma}$ no modifica la longitud de un vector.

II. VARIABLES DINÁMICAS Y OBSERVABLES

7. Operadores lineales

Dirac: En la sección anterior hemos considerado funciones lineales numéricas de kets, que nos han llevado al concepto de bra. Ahora consideraremos funciones lineales vectoriales de kets, que nos llevarán al concepto de operador lineal.

Sea un ket $|F\rangle$ función de un ket $|A\rangle$, es decir, que a cada ket $|A\rangle$ le corresponde un ket $|F\rangle$, y supongamos además que dicha función sea lineal. o lo que es lo mismo, que el ket $|F\rangle$ que corresponde a $|A\rangle + |A'\rangle$ es igual a la suma de los kets $|F\rangle$ correspondientes a $|A\rangle$ y a $|A'\rangle$, y que el ket $|F\rangle$ que corresponde a $c|A\rangle$ es igual a c multiplicado por el ket $|F\rangle$ correspondiente a $|A\rangle$, donde c es un factor numérico arbitrario. Bajo estas condiciones, podemos considerar el paso de $|A\rangle$ a $|F\rangle$ como la aplicación de un *operador lineal* a $|A\rangle$. Si introducimos el símbolo α para el operador lineal, se puede escribir

$$|F\rangle = \alpha|A\rangle,$$

donde indicamos la aplicación de α sobre $|A\rangle$ como un producto de α por $|A\rangle$. Establecemos la regla de que en tales productos *el ket debe estar siempre a la derecha del operador lineal*. Podemos expresar ahora las condiciones de linealidad antes mencionadas mediante las ecuaciones

$$\begin{aligned} \alpha\{|A\rangle + |A'\rangle\} &= \alpha|A\rangle + \alpha|A'\rangle \\ \alpha\{c|A\rangle\} &= c\alpha|A\rangle. \end{aligned} \tag{3.9}$$

Diremos que un operador lineal está completamente definido, si conocemos el resultado de su aplicación sobre cada ket. Por tanto, debemos considerar nulo a un operador lineal que aplicado a cualquier ket da cero, y asimismo diremos que dos operadores lineales son iguales, si dan el mismo resultado al aplicarlos a cualquier ket.

Comentario: Introduce ahora Dirac el concepto de operador lineal, observese la similitud con la definición 2.52 y la correspondencia de las ecuaciones representadas por 3.9 con los axiomas OL-1 y OL-2.

Dirac: También podemos sumar operadores, y definimos la suma de dos de ellos como aquel operador lineal que aplicado a cualquier ket da un resultado igual a la suma de los kets que se obtienen al aplicar cada uno de los operadores por separado a dicho, ket. Así pues definimos $\alpha + \beta$ por

$$\{\alpha + \beta\} |A\rangle = \alpha |A\rangle + \beta |A\rangle \quad (3.10)$$

para todo $|A\rangle$. La ecuación (3.10) y la primera de las ecuaciones (3.9) nos indican que el producto de operadores lineales por kets verifica la propiedad distributiva de la multiplicación.

Asimismo podemos multiplicar operadores lineales entre sí, y definimos el producto de dos operadores como aquel operador lineal que aplicado a cualquier ket da el mismo resultado que se obtiene al aplicar sucesivamente los dos operadores dados a dicho ket. Así pues, hemos definido el producto $\alpha\beta$ como el operador lineal que aplicado a cualquier ket $|A\rangle$, lo transforma en el ket que se obtiene aplicando primero β a $|A\rangle$, y después α al resultado. En símbolos

$$\{\alpha\beta\} |A\rangle = \alpha\{\beta |A\rangle\}.$$

Esta definición es equivalente a la propiedad asociativa de la multiplicación del triple producto de α, β y $|A\rangle$, y nos permite por lo tanto escribir dicho producto sin paréntesis $\alpha\beta |A\rangle$. Sin embargo, este triple producto no da en general el mismo resultado que se obtendría aplicando primero α a $|A\rangle$ y después β al resultado, es decir, que en general $\alpha\beta |A\rangle$ es distinto de $\beta\alpha |A\rangle$, y por tanto, en general, $\alpha\beta$ es distinto de $\beta\alpha$. La propiedad conmutativa no es válida para la multiplicación de operadores lineales.

Comentario: Cuando dos operadores conmutan ellos pueden tomar valores simultáneamente determinados, físicamente ello significa que los dos observables asociados a esos operadores pueden ser medidos simultáneamente. Esto de alguna manera implica que la conmutabilidad de operadores es condición necesaria y suficiente para la mensurabilidad simultánea de las magnitudes físicas. En mecánica cuántica representan un papel muy importante los conjuntos de magnitudes físicas que satisfacen la propiedad anterior, puesto que todos ellos pueden medirse simultáneamente. A la totalidad de tales operadores se les llama un conjunto completo. En algunos casos particulares el conjunto completo puede incluso ser de un único observable. Más técnicamente se dice que si existe un conjunto completo de observables que conmutan, dichos observables admiten la misma base propia.

Dirac: Puede ocurrir que en casos especiales dos operadores lineales ξ y η sean tales que $\xi\eta$ y $\eta\xi$ sean iguales. En tal caso diremos que conmuta ξ con η , o que ξ y η conmutan.

Combinando las operaciones dadas de suma y multiplicación de operadores lineales se pueden formar sumas y productos con más de dos operadores, y así podemos construir con ellos un álgebra. En dicho álgebra no será válida la propiedad conmutativa de la multiplicación, y podrá ocurrir también que el producto de dos operadores lineales sea nulo sin que lo sea ninguno de los dos factores. Pero todas las demás propiedades del álgebra ordinaria serán válidas, incluidas las propiedades asociativa y distributiva de la multiplicación, como puede comprobarse fácilmente.

Comentario: Es un resultado conocido del análisis funcional que el conjunto de todos los operadores lineales acotados desde un espacio normado X sobre un espacio normado Y forman un espacio normado

(cf. [16], pág. 118). En nuestro caso es suficiente afirmar que que dicho conjunto de observables forman un espacio vectorial, esto con el fin de satisfacer los requerimientos para un álgebra establecidos por la definición 2.41. Además las propiedades mencionadas por Dirac, con respecto a asociatividad y distributividad con respecto a la multiplicación concuerdan con los axiomas A-1, A-2, A-3 y A-4, exigidos por la definición de álgebra.

Dirac: Si tomamos un número k y lo multiplicamos por vectores ket, aparece como un operador lineal que se aplica a kets, pues verifica las condiciones (3.9) substituyendo k por α . Por tanto, un número es un caso particular de operador lineal. Tiene la propiedad de que conmuta con todo otro operador lineal, lo que le distingue de un operador lineal general.

Hasta aquí únicamente hemos aplicado los operadores lineales a kets. Pero también pueden ser aplicados a bras, con el significado siguiente. Tomemos el producto escalar del bra $\langle B|$ con el ket $\alpha|A\rangle$. Tal producto escalar es un número que depende linealmente de $|A\rangle$ y por lo tanto, dada la definición de bra, puede ser considerado como el producto escalar de $|A\rangle$ por un cierto bra. Dicho bra depende linealmente de $\langle B|$, así que puede considerarse como el resultado de aplicar cierto operador lineal a $\langle B|$. Dicho operador lineal está determinado unívocamente por el operador lineal α de antes y es lógico decir que se trata del mismo operador lineal que actúa ahora sobre un vector bra.

Una notación conveniente para indicar el bra que resulta de aplicar α al bra $\langle B|$ es $\langle B|\alpha$, y en dicha notación, la ecuación que define $\langle B|\alpha$ es

$$\{\langle B|\alpha\}|A\rangle = \langle B|\{\alpha|A\rangle\} \quad (3.11)$$

para cualquier $|A\rangle$, que expresa simplemente la propiedad asociativa de la multiplicación para el triple producto de $\langle B|$, α y $|A\rangle$. Establecemos la regla general de que en todo producto de un bra y un operador lineal, siempre ha de estar el bra a la izquierda. Ahora ya podemos escribir el triple producto de $\langle B|$, α y $|A\rangle$ simplemente $\langle B|\alpha|A\rangle$ sin ningún paréntesis. Puede comprobarse fácilmente que la propiedad distributiva de la multiplicación es válida para los productos de operadores lineales y kets.

En nuestro esquema cabe todavía un nuevo tipo de producto, que es el producto de un ket y un bra con el ket a la izquierda $|A\rangle\langle B|$. Para ver qué significa dicho producto, multipliquémoslo por un ket $|P\rangle$ arbitrario, que colocaremos a la derecha, y supongamos que es válida la propiedad asociativa para esta multiplicación. Dicho producto vale entonces $|A\rangle\langle B|P\rangle$, o sea el ket $|A\rangle$ multiplicado por el número $\langle B|P\rangle$, y es función lineal del ket $|P\rangle$. Por tanto, $|A\rangle\langle B|$ aparece como un operador lineal que puede actuar sobre kets. También puede actuar sobre bras, siendo su producto por un bra $\langle Q|$ a la izquierda $\langle Q|A\rangle\langle B|$, que es igual al número $\langle Q|A\rangle$ multiplicado por el bra $\langle B|$. Debemos distinguir claramente entre los productos $|A\rangle\langle B|$ y $\langle B|A\rangle$ de los mismos factores pero en orden inverso, siendo el último producto un número, como ya sabemos.

Comentario: Cabe mencionar aquí que la operación definida como producto de un ket por un bra es un tipo particular de operador de proyección, el cual al actuar sobre un ket particular, y que puede siempre descomponerse en otros dos ket, uno perteneciente a un subespacio del espacio de Hilbert y otro perteneciente a un subespacio ortogonal al primero, lo proyecta en el subespacio. Por efecto de la misma definición del operador, éste solo posee dos autovalores 0 y 1; el autovalor 0 está asociado a la proyección del ket sobre el subespacio ortogonal, mientras que el autovalor 1 está asociado con la proyección del ket sobre el espacio de Hilbert.

Podemos también mostrar dos casos extremos. Sí el espacio de proyección es el propio espacio de Hilbert, cualquier ket es su propia proyección, en este caso el espacio complemento es nulo. Este es el caso donde el operador de proyección es la identidad.

El otro caso es aquel donde el espacio proyección es idénticamente nulo, el espacio complementario es el propio espacio de Hilbert, la proyección de cualquier ket de este espacio se va al vector nulo. Es el caso donde el operador de proyección es idénticamente cero.

Dirac: Disponemos ahora de un esquema algebraico completo referente a tres clases de cantidades: vectores bra, vectores ket y operadores lineales. Podemos multiplicarlos entre sí de todas las formas indicadas, siendo válidas las propiedades asociativa y distributiva de la multiplicación en todos los casos, pero no la propiedad conmutativa. En el esquema general sigue siendo válida la regla de notación dada en la sección anterior, de que toda expresión entre paréntesis completos, que tenga \langle a la izquierda y \rangle a la derecha, representa un número, mientras que toda expresión entre paréntesis incompletos, que no tenga más que \langle o \rangle , representa un vector.

Respecto al significado físico del esquema, hemos supuesto siempre que los bras y los kets, o mejor dicho las direcciones de dichos vectores, corresponden a los estados de un sistema dinámico en un instante particular. Ahora hacemos la nueva hipótesis de que *los operadores lineales corresponden a las variables dinámicas en dicho instante*. Bajo el nombre de variables dinámicas entendemos cantidades como las coordenadas, las componentes de la velocidad, del momento y del momento angular de las partículas así como funciones de éstas -de hecho las variables que se emplean en mecánica clásica-. La nueva hipótesis exige que dichas cantidades aparezcan también en mecánica cuántica, pero con la notable diferencia de que en ella *están regidas por un álgebra en la cual no es válida la propiedad conmutativa de la multiplicación*.

El hecho de que el álgebra de las variables dinámicas sea distinta en ambas teorías es una, de las diferencias más importantes entre la mecánica cuántica y la clásica. Más adelante veremos, a este respecto, que a pesar de esta diferencia fundamental, las variables dinámicas de la mecánica cuántica siguen gozando de muchas propiedades comunes con sus equivalentes clásicas, y que es posible construir una teoría análoga en gran parte a la clásica que constituye una excelente generalización de ella.

Comentario: Una costumbre particularmente importante y definitivamente fundamental en la mecánica clásica, es poder identificar entre un conjunto de variables dinámicas, cuáles de ellas se conservan en un proceso físico particular. Por ejemplo, es conocido el hecho de que el momento lineal se conserve en una situación física particular, si el espacio en el cual está inmerso dicho proceso es homogéneo, de la misma manera se conserva la energía si el tiempo es homogéneo, o la isotropía del espacio, pone de manifiesto la conservación del momento angular. La mayoría de las situaciones en física permiten como mínimo la definición de alguna de esas cantidades conservadas. Al pie de igualdad la mecánica cuántica, aunque sus variables dinámicas hayan sido definidas, definitivamente de una forma distinta, satisface el mismo principio general, en la práctica también es común en esa teoría tratar de encontrar en sus diferentes situaciones particulares, cuáles cantidades son conservadas en situaciones específicas, esto de algún modo permite simplificar la complejidad de los problemas en la física.

Dirac: Es conveniente utilizar la misma letra para especificar una variable dinámica y su correspondiente operador lineal. De hecho, consideraremos a ambos como una sola cosa, sin que ello dé lugar a confusión.

8. Relaciones conjugadas

Dirac: Nuestros operadores lineales son cantidades complejas, puesto que se pueden multiplicar por números complejos resultando nuevas cantidades de la misma naturaleza. Así pues han de corresponder en general a variables dinámicas complejas, es decir, a funciones complejas de las coordenadas, velocidades, etc. A fin de poder ver qué clase de operador lineal corresponde a una variable dinámica real, precisamos una elaboración más amplia de la teoría.

Consideremos el ket conjugado imaginario de $\langle P | \alpha$. Dicho ket depende antilinealmente de $\langle P |$ y, por tanto, depende linealmente de $|P\rangle$. Por consiguiente, puede ser considerado como el resultado de aplicar un cierto

operador lineal a $|P\rangle$. A este operador lineal se le llama *adjunto* de α y le designaremos por $\bar{\alpha}$. Con esta notación, el conjugado imaginario de $\langle P|\alpha$ es $\bar{\alpha}|P\rangle$.

Comentario: Observe que la definición de operador adjunto de Dirac es equivalente a la definición 2.55.

Dirac: En la fórmula (3.7) del capítulo 1 pongamos $\langle P|\alpha$ en lugar de $\langle A|$, y $\bar{\alpha}|P\rangle$ en lugar de $|A\rangle$. El resultado es

$$\langle B|\bar{\alpha}|P\rangle = \overline{\langle P|\alpha|B\rangle} \quad (3.12)$$

Esta es una fórmula general, válida para todo par de kets $|B\rangle$ y $|P\rangle$ y para todo operador lineal α , que nos expresa una de las propiedades más utilizadas del adjunto.

Substituyendo α por $\bar{\alpha}$ en (3.12), resulta

$$\langle B|\bar{\bar{\alpha}}|P\rangle = \overline{\langle P|\bar{\alpha}|B\rangle} = \langle B|\alpha|P\rangle,$$

después de haber utilizado otra vez (3.12) intercambiando $|P\rangle$ con $|B\rangle$. Esto es válido para todo ket $|P\rangle$, de donde deducimos, con ayuda de (3.4) del capítulo 1,

$$\langle B|\bar{\bar{\alpha}} = \langle B|\alpha,$$

y puesto que esta ecuación es válida para todo bra $\langle B|$, podemos deducir que

$$\bar{\bar{\alpha}} = \alpha.$$

Luego, *el adjunto del adjunto de un operador lineal es igual al operador lineal de partida*. Esta relación entre un operador y su adjunto es análoga a la que existe entre un número y su complejo conjugado, y además podemos comprobar fácilmente que en el caso particular de que el operador lineal sea un número, su operador lineal adjunto es el número complejo conjugado del dado. Por lo tanto, queda justificado suponer que el adjunto de un operador lineal corresponde al complejo conjugado de la variable dinámica. Con este significado físico del adjunto de un operador lineal, podemos llamar al operador adjunto igualmente operador lineal complejo conjugado, lo que está de acuerdo con nuestra notación $\bar{\alpha}$.

Un operador lineal puede ser igual a su adjunto, en cuyo caso se dice que es autoadjunto. Corresponde a una variable dinámica real, y por esto se le denomina también *operador lineal real*.

Comentario: La justificación de que un operador auto-adjunto corresponde a una variable dinámica real, puede ser vista a partir del teorema 2.14 que establece en la notación de Dirac, que si α es un operador auto-adjunto, entonces $\langle A|\alpha|A\rangle$ es un número real. Esto será de bastante importancia cuando se hable posteriormente de los observables de un sistema cuántico.

Dirac: Todo operador lineal puede ser separado en parte real y parte imaginaria pura. Por esta razón para los operadores lineales se aplica la expresión “complejo conjugado” y no la de “imaginario conjugado”.

9. Auto valores y autovectores

Dirac: Es preciso hacer un desarrollo más amplio de la teoría de operadores lineales y estudiar la ecuación

$$\alpha|P\rangle = a|P\rangle, \quad (3.13)$$

en la que α es un operador lineal y a es un número. Esta ecuación se presenta con frecuencia de forma que α es un operador lineal conocido, y el número a y el ket $|P\rangle$ son las incógnitas, que debemos intentar determinar de modo que satisfagan a la ecuación (3.13), sin considerar la solución trivial $|P\rangle = 0$.

La ecuación (3.13) significa que al aplicar el operador α al ket $|P\rangle$, éste queda multiplicado por un factor numérico y no cambia de dirección, o bien que queda multiplicado por el factor cero y deja de tener dirección. Por supuesto, dicho operador α aplicado a otros kets cambiará tanto sus longitudes como sus direcciones. Tengamos presente que en la ecuación (3.13) lo único que importa de $|P\rangle$ es su dirección. El multiplicar $|P\rangle$ por un número distinto de 0, no afecta a la cuestión de si satisface o no a (3.13).

Junto a la ecuación (3.13), consideraremos la ecuación imaginaria conjugada

$$\langle Q|\alpha = b\langle Q|, \quad (3.14)$$

en la que b es un número. Aquí las incógnitas son los números b y los bras $\langle Q|$ que no sean el cero. Dado que las ecuaciones (3.13) y (3.14) son de importancia primordial para la teoría, es conveniente dar un nombre especial al tipo de relación entre las cantidades a que dan lugar. Si una terna α , $|P\rangle$ y a verifican (3.13), diremos que a es un *autovalor* (*eigenvalue*) del operador lineal α o de su correspondiente variable dinámica, y asimismo que $|P\rangle$ es un *autoket* (*eigenket*) del operador lineal α o de la variable dinámica en cuestión. Además diremos que el autoket $|P\rangle$, pertenece al autovalor a . Análogamente, si una terna α , $\langle Q|$ y b , verifican (3.14), diremos que b es un autovalor de α y que $\langle Q|$ es un *autobra* (*eigenbra*) perteneciente a dicho autovalor. Desde luego, las expresiones de autovalor, autoket y autobra únicamente tienen sentido con respecto a un operador lineal o a una variable dinámica.

Con esta terminología, si multiplicamos un autoket de α por un número distinto de cero, el ket resultante es un nuevo autoket perteneciente al mismo autovalor al que pertenecía el ket de partida. Pueden existir dos o más autokets de un operador lineal independientes que pertenezcan a un mismo autovalor; así por ejemplo, la ecuación (3.13) puede tener varias soluciones $|P1\rangle, |P2\rangle, |P3\rangle, \dots$ que pertenezcan a un mismo autovalor a , y que sean independientes. En, tal caso es evidente que cualquier combinación lineal de dichos autokets también es un autoket que pertenece al mismo autovalor del operador lineal, es decir que

$$c_1|P1\rangle + c_2|P2\rangle + c_3|P3\rangle + \dots$$

és también solución de (3.13), siendo c_1, c_2, c_3, \dots números cualesquiera.

Comentario: El conjunto de los autovectores (ket) de un operador lineal relativos a un autovalor dado, forma una estructura de espacio vectorial, generalmente ese espacio vectorial recibe el nombre de subespacio vectorial relativo al valor propio dado, a veces el operador o variable dinámica solo tiene un autovector asociado a un autovalor particular, en ese caso el espacio vectorial formado, solo tiene una dimensión y se dice que el valor propio es no degenerado, puede suceder también que para un mismo autovalor se tengan varios autovectores asociados, en cuyo caso el espacio vectorial formado por estos autovectores tiene varias dimensiones, tantas como autovectores independientes esten asociadas a ese autovalor, en el último caso se dice que el autovalor está degenerado, puede suceder que el orden de degeneración de un autovalor particular sea infinito. Exactamente estas mismas observaciones son aplicables cuando los autovectores de un observable particular son los bra. En el caso general de un operador cualquiera, no existe un conjunto de relaciones sencillas entre los valores y vectores propios asociados a los vectores bra y ket, sin embargo, si el operador es hermítico las cosas son un tanto diferentes, puesto que si es posible desarrollar un conjunto de propiedades comunes a ambos conjuntos de autovalores y autovectores. Podemos mencionar algunas propiedades como por ejemplo: Los espectros de valores propios son los mismos para ambos conjuntos; el conjunto de autovalores es real; el subespacio de los bra propios relativos a un valor propio dado, es el dual del subespacio de los ket propios relativos al mismo valor propio.

Entendemos también que pueden haber vectores propios que no pertenecen al espacio de Hilbert en cuestión, es decir, autovectores de norma infinita, asociados a los valores propios de espectro continuo, este tipo de autovectores satisfacen también las propiedades mencionadas anteriormente, junto con las ya mencionadas características respecto al producto escalar.

Dirac: En el caso particular de que el operador lineal α de las ecuaciones (3.13) y (3.14) sea un número k , evidentemente todo ket $|P\rangle$ y todo bra $\langle Q|$ satisfacen dichas ecuaciones con a y b iguales a k . Por tanto, todo número, considerado como operador lineal, tiene un único autovalor, y todos los kets y todos los bras son autovectores suyos pertenecientes a dicho autovalor.

Comentario: Es consecuente el considerar los números como operadores lineales para el manejo de las constantes en física. Si deseamos medir una constante no esperamos la probabilidad de encontrarla en un rango de valores, por el contrario, esperamos un único valor posible. Entonces al considerar las constantes como operadores lineales es necesario considerar que éstas tienen un único autovalor, el valor de ella misma.

Dirac: La teoría de autovalores y autovectores de un operador lineal α que no sea real, no tiene gran aplicación en mecánica cuántica. Así pues, a estos efectos, nos limitaremos en adelante a considerar operadores lineales reales. Poniendo en lugar de α el operador lineal real ξ , resultan en lugar de las ecuaciones (3.13) y (3.14),

$$\xi |P\rangle = a |P\rangle, \quad (3.15)$$

$$\langle Q| \xi = b \langle Q|. \quad (3.16)$$

Podemos deducir inmediatamente tres importantes resultados.

(i) *Los autovalores son todos reales.*

(ii) *Los autovalores asociados a los autokets son los mismos que los autovalores asociados a los autobras.*

(iii) *El imaginario conjugado de un autoket es un autobra que pertenece al mismo autovalor, y recíprocamente.*

El último resultado nos permite decir con razón que el estado que corresponde a cualquier autoket o a su bra imaginario conjugado es un autoestado de la variable dinámica real ξ .

Los autovalores y los autovectores de algunas variables dinámicas reales se utilizan ampliamente en mecánica cuántica y, por tanto, es conveniente utilizar una notación sistemática para indicarlos. La que vamos a exponer a continuación es adecuada para la mayoría de las aplicaciones. Si ξ es una variable dinámica real, sus autovalores los designaremos por ξ', ξ'', ξ^r , etc. Así una letra sirve para designar una variable dinámica real o un operador lineal real, y la misma letra con primas o índices indica un número, que es un autovalor del operador que simboliza la letra. De este modo, un autovector puede especificarse con el autovalor a que pertenece. Así $|\xi'\rangle$ representa un autoket que pertenece al autovalor ξ' de la variable dinámica ξ .

Teorema: Dos autovectores de una misma variable dinámica real que pertenezcan a distintos autovalores son ortogonales.

10. Observables

Dirac: Hemos hecho un conjunto de hipótesis acerca de cómo se representan los estados y las variables dinámicas en la teoría matemática. Dichas hipótesis por sí solas no constituyen leyes de la naturaleza, pero en cuanto hagamos nuevas hipótesis acerca del significado físico de la teoría, aparecerán como tales. Las nuevas hipótesis tienen que establecer relaciones entre los resultados de observación por un lado, y las ecuaciones del formalismo matemático por el otro.

Toda observación consiste en medir una variable dinámica. Desde el punto de vista físico es evidente que el resultado de dicha medición debe ser siempre un número real, y así supondremos que cualquier variable dinámica que podamos medir debe ser una variable dinámica real. Puede pensarse que sería posible medir una variable

dinámica compleja midiendo su parte real y su parte imaginaria por separado, pero esto llevaría consigo dos medidas o dos observaciones. Si bien en mecánica clásica es posible, en mecánica cuántica no lo es, pues en general las medidas interfieren entre sí -no se puede admitir que dos observaciones se realicen exactamente a la vez, y si se hacen una detrás de otra en rápida sucesión, generalmente la primera altera el estado del sistema e introduce una indeterminación que afecta a la segunda. Por lo tanto, debemos exigir que las variables dinámicas que se pueden medir sean reales, y la condición para ello en mecánica cuántica es la de la sección (8). Sin embargo, no toda variable dinámica real puede ser medida. Como veremos más adelante, es necesaria otra restricción.

Hagamos ahora algunas hipótesis acerca de la interpretación física de la teoría. Si *el sistema dinámico está en un autoestado de una variable dinámica real ξ perteneciente al autovalor ξ' , entonces estamos totalmente seguros de que el resultado de medir ξ es el número ξ'* . Recíprocamente, *si el sistema está en un estado tal que al medir una variable dinámica real ξ estamos completamente seguros de obtener un resultado particular determinado (y no distintos posibles resultados, según una ley probabilística, como ocurre en general), entonces el estado es un autoestado de ξ , y el resultado de la medida es el autovalor de ξ a que pertenece dicho autoestado*. Dado que los autovalores de los operadores lineales reales son siempre números reales estas hipótesis son viables.

Señalemos algunas consecuencias inmediatas de estas hipótesis. Si tenemos dos o más autoestados de una variable dinámica real ξ pertenecientes a un mismo autovalor ξ' , todo estado formado por superposición de ellos también será un autoestado de ξ perteneciente al autovalor ξ' . De aquí se deduce que si el sistema está en un estado formado por superposición de otros varios para los que estamos completamente seguros de que al medir ξ en cada uno de ellos obtenemos un mismo resultado ξ' , entonces al medir ξ en él obtenemos también con toda seguridad el resultado ξ' . Esto nos da cierta luz acerca del significado físico del principio de superposición de los estados. Otra consecuencia es que dos autoestados de ξ pertenecientes a distintos autovalores son ortogonales. De aquí se deduce que si tenemos dos estados de un sistema y estamos completamente seguros de que al medir ξ en cada uno de ellos obtenemos un único resultado, distinto para uno y otro, entonces los dos estados son ortogonales. Esto nos da, a su vez, alguna luz acerca del significado físico de la ortogonalidad de los estados.

Cuando medimos una variable dinámica real ξ , la alteración que lleva consigo el acto de la medida produce un cambio del estado del sistema dinámico. Si llevamos a cabo una segunda medición de la misma variable dinámica inmediatamente después de la primera, por continuidad física el resultado debe ser el mismo de antes. Por tanto, después de haber realizado la primera medición no hay ninguna indeterminación en el resultado de la segunda, o sea que después de realizar la primera medición el sistema está en un autoestado de la variable dinámica ξ , siendo el autovalor al que pertenece dicho estado el resultado de la primera medida. Esta conclusión tiene que continuar siendo válida aunque no llevemos a cabo la segunda medida. Así vemos que toda medida obliga al sistema a saltar a un autoestado de la variable dinámica medida, que además pertenece a un autovalor igual al resultado de la medida. Podemos concluir que, cualquiera que sea el estado en que se encuentre el sistema, *todo resultado de medir una variable dinámica real es uno de sus autovalores*. Y recíprocamente, *todo autovalor es un posible resultado de medida de la variable dinámica para algún estado del sistema*, ya que con toda probabilidad será el resultado de la medida cuando el estado del sistema sea un autoestado perteneciente a dicho autovalor. Todo esto nos muestra el significado físico de los autovalores.

Comentario: Dirac ha establecido en el párrafo anterior lo que es uno de los postulados o axiomas de la mecánica cuántica, en particular puede ser leído en la frase “*Así vemos que toda medida obliga al sistema a saltar a un autoestado de la variable dinámica medida, que además pertenece a un autovalor igual al resultado de la medida.*” Queda establecida una de las diferencias fundamentales con la mecánica clásica: la medida afecta al sistema y lo modifica. El por qué esto sucede para algunos sistemas, en particulares para los sistemas pequeños en sentido absoluto y se puede despreciar para los sistemas grandes en sentido absoluto, tal como los denomina Dirac, en otras palabras, por qué “necesitamos” una física para el nivel macroscópico y otra física para el nivel microscópico, constituye una de las preguntas, desde nuestro punto de vista más

importantes de la física contemporánea.

Dirac: El conjunto de autovalores de una variable dinámica real constituye precisamente el espectro de los posibles resultados de medida de dicha variable dinámica, y por esta razón el cálculo de los autovalores constituye un problema importante.

Asimismo hacemos la hipótesis siguiente sobre la interpretación física de la teoría: *cuando medimos una cierta variable dinámica real ξ con el sistema en un estado determinado, los estados a los que puede saltar el sistema a causa de la medida son tales que el estado original es dependiente de ellos.* Pero como por otro lado los estados a los que puede saltar el sistema son todos autoestados de ξ , resulta que el estado original es dependiente de autoestados de ξ . Teniendo en cuenta que el estado original puede ser *cualquiera*, podemos concluir que todo estado es dependiente de autoestados de ξ . Si definimos un *conjunto completo* de estados como aquel para el que todo estado es dependiente de los estados del conjunto, nuestra conclusión también se puede formular así: los autoestados de ξ constituyen un conjunto completo.

No toda variable dinámica real tiene suficientes autoestados para que constituyan un conjunto completo. Pero aquellas cuyos autoestados no forman un conjunto completo no representan cantidades que se puedan medir. Así obtenemos una nueva condición que debe verificar toda variable dinámica, además de la de ser real, para que se pueda medir. Toda variable dinámica real cuyos autoestados constituyan un conjunto completo, la denominaremos con el nombre genérico de *observable*. Por tanto, toda cantidad que se pueda medir es un observable.

Comentario: Dirac establece de nuevo en este punto otro de los postulados actuales de la mecánica cuántica, el relacionado con los observables. Por ejemplo, [10] lo establece como el postulado No. 2 de la siguiente forma:

1. *A cada observable físico \mathcal{A} corresponde en el espacio de Hilbert un operador lineal hermítico \hat{A} que posee un conjunto completo ortonormal de vectores propios $\alpha_1(x), \alpha_2(x), \alpha_3(x), \dots$ y un conjunto correspondiente de valores propios A_1, A_2, A_3, \dots*

$$\hat{A}\alpha_i(x) = A_i\alpha_i(x) \quad i = 1, 2, 3, \dots$$

Recíprocamente, a todo observable de ese tipo en el espacio de Hilbert corresponde un cierto observable físico.

2. *Los únicos valores posibles que puede dar toda medida de \mathcal{A} son los valores propios A_1, A_2, A_3, \dots*

Aunque sabemos que la experimentación confirma los postulados de la mecánica cuántica, no deja de sorprendernos, el proceso por el cual son propuestos. Cada uno de ellos establece una relación bidireccional y compleja entre la matemática y la física, relación que no debemos dejar de cuestionar.

Dirac: La cuestión que se plantea ahora es si todo observable puede ser medido. La respuesta teórica es que sí. En la práctica puede resultar muy difícil o incluso puede estar fuera del alcance del ingenio del experimentador, el imaginar un aparato capaz de medir un observable particular, pero según la teoría siempre existe. . .

A menudo resulta muy difícil establecer desde el punto de vista matemático si una variable dinámica real concreta satisface o no la condición para ser un observable, ya que, en general, el problema de hallar los autovalores y los autovectores es muy difícil de resolver. Sin embargo, pueden existir poderosas razones experimentales que nos permitan admitir que dicha variable dinámica puede ser medida, y en tal caso supondremos lógicamente que es un observable pese a carecer de una demostración matemática. Esto lo haremos con frecuencia durante el desarrollo de la teoría, y así supondremos por ejemplo, que la energía de cualquier sistema dinámico es siempre un observable, si bien con los métodos actuales del análisis matemático sólo se ha podido demostrar en casos muy sencillos.

En el caso particular de que la variable dinámica sea un número, todo estado es autoestado de él, y evidentemente dicha variable dinámica es un observable. Al medirlo obtendremos siempre el mismo resultado, o sea que se trata de una constante física, como por ejemplo la carga del electrón. Por tanto, una constante física se puede considerar en mecánica cuántica bien como un observable con un único autovalor o como un simple número que aparece en las ecuaciones, siendo ambos puntos de vista completamente equivalentes. . .

Implícitamente habíamos supuesto hasta ahora que nuestros bras y kets tenían longitud finita y que sus productos escalares eran finitos. Ahora nos vemos obligados a limitar dicha condición cuando tratamos con autovectores de un observable cuyos autovalores forman un dominio de medida no nula. Si no la limitásemos no sería posible considerar casos en que los autovalores constituyeran un dominio de medida no nula, y nuestra teoría resultaría insuficiente para una gran parte de los problemas reales. . .

El espacio de los bras o el de los kets, cuando exigimos que los vectores tengan longitud finita y que los productos escalares sean finitos, constituye lo que los matemáticos denominan un espacio de Hilbert.

Comentario: Se preguntará el lector porqué en nuestro capítulo de preliminares matemáticos se hizo hincapie en la importancia de los espacios de Hilbert para la mecánica cuántica y hasta ahora es la primera vez que son mencionados. Nuestra respuesta parcial consiste en que en el momento en que Dirac escribía su obra (1930) no era tan *evidente* la relación entre los espacios de Hilbert y la mecánica cuántica. Por el contrario, los autores actuales mencionan de inmediato los espacios de Hilbert y su asociación con los estados de un sistema cuántico.

Por ejemplo, [6] habla del espacio $L^2[-\infty, +\infty]$ como el espacio *natural* de la mecánica cuántica, inclusive menciona que este espacio es demasiado amplio y sugiere un subespacio de trabajo $\mathcal{F} \subset L^2[-\infty, +\infty]$ que estaría formado por funciones con ciertas propiedades de regularidad (definidas en cualquier parte, continuas, infinitamente diferenciables, de dominio acotado).

Continuado con la indagación acerca de la naturaleza del espacio donde opera la cuántica, nos encontramos, por ejemplo con el postulado No. 1 que presentan [9]:

“A cada sistema físico que se pretenda describir en el marco de la mecánica cuántica se le hace corresponder un espacio de Hilbert complejo y separable. Un estado puro de este sistema físico en un instante de tiempo t se representa por un rayo unitario $|\Psi(t)\rangle_R$ perteneciente al espacio de Hilbert correspondiente. Un elemento $|\Psi(t)\rangle$ del rayo $|\Psi(t)\rangle_R$ se llamará vector estado o ket.”([9], pág. 38).

Estos mismos autores mencionan que surgen cuestiones inmediatas acerca de, el por qué aparecen los espacios de Hilbert en la mecánica cuántica, por qué éstos espacios han de ser complejos y separables y por qué los estados puros corresponden a rayos unitarios, y sugieren algunas lecturas donde debido a ciertas hipótesis *plausibles* las cosas son así.

La pregunta en este punto es por los fundamentos físicos, matemáticos, epistemológicos y filosóficos de la mecánica cuántica. Por supuesto no es este el lugar (suponiendo que estuvieramos en condiciones) para afrontar semejante tarea. Afirmamos eso sí que tales preguntas son pertinentes en el marco de una investigación científica, y seguiremos de alguna forma las marcas señaladas por ellas.

Dirac: Los bras y los kets que empleamos aquí dan lugar a un espacio más general de lo que corresponde a un espacio de Hilbert. . .

Comentario: El espacio de los vectores (ket) de la mecánica cuántica, está formado por todos aquellos elementos cuya norma sea finita, es decir, vectores de longitud finita. Es natural que el producto escalar de vectores finitos sea finito, lo cual es un requisito indispensable en la mecánica cuántica. Es común en la teoría que los vectores que representan estados dinámicos son efectivamente de norma finita, ahora bien el estudio de los espectros continuos de alguna manera exige la introducción de vectores propios de longitud infinita, como es el caso de los autovectores del operador de posición, o el operador momento lineal, para citar dos

de ese tipo de autovectores. Estos autovectores deben ser incluidos en la clase de vectores que representan verdaderos estados dinámicos y aunque ellos no pertenecen al espacio de Hilbert en cuestión, es en principio posible exigir que el producto escalar de uno de tales vectores, de norma infinita, con otro vector de norma finita sea finito. Por el contrario el producto escalar de dos vectores de norma infinita no necesariamente converge. En síntesis, los vectores de norma infinita no pertenecen al espacio de Hilbert, sus diferenciales sí (cf. [23], pág. 229). Con esto en mente una expresión de la forma

$$\int \lambda(x) |x\rangle dx,$$

en la cual está expresada una cierta combinación lineal de vectores ket de norma infinita, converge a un elemento del espacio de Hilbert.

12. Interpretación física general

Dirac: Las hipótesis introducidas al principio de la sección (10) nos dan una interpretación física de la teoría matemática que sólo se puede aplicar en casos particulares, pues únicamente sirve para autoestados. Necesitamos alguna hipótesis más general que también nos permita obtener información física a partir de la teoría matemática cuando no consideremos autoestados.

En mecánica clásica decimos que un observable “tiene un valor” para cualquier estado del sistema. ¿Qué es lo que corresponde a esto en mecánica cuántica? Si consideramos un observable ξ y dos estados x e y correspondientes a los dos vectores $\langle x|$ e $|y\rangle$, podemos formar el número $\langle x|\xi|y\rangle$. Dicho número no tiene características análogas al valor que “tiene” un observable en la teoría clásica por tres razones principales: (I) hace referencia a *dos* estados, mientras que el valor clásico siempre se refiere a *uno*, (II) en general no es un número real, y (III) no está unívocamente determinado por el observable y los estados, pues los vectores $\langle x|$ e $|y\rangle$ contienen factores numéricos arbitrarios. Incluso si imponemos la condición de que $\langle x|$ e $|y\rangle$ estén normalizados, $\langle x|\xi|y\rangle$ continuará estando indeterminado en un factor de módulo uno. Sin embargo, si tomamos dos estados idénticos y elegimos $|y\rangle$ de modo que sea el vector imaginario conjugado de $\langle x|$, ninguna de las tres razones apuntadas tiene validez. El número que obtenemos en este caso $\langle x|\xi|x\rangle$ es necesariamente real y está unívocamente determinado si $\langle x|$ está normalizado, pues si multiplicamos $\langle x|$ por el factor numérico e^{ic} siendo c un número real, tenemos que multiplicar $|x\rangle$ por e^{-ic} , y $\langle x|\xi|x\rangle$ toma el mismo valor de antes.

Por tanto, podríamos hacer la sugestiva hipótesis de que cuando el sistema está en el estado x , el observable ξ “tiene un valor” $\langle x|\xi|x\rangle$ en un sentido análogo al clásico. Sin embargo, esta hipótesis no sería satisfactoria por la siguiente razón. Consideremos un segundo observable η que según la hipótesis anterior tendría un valor $\langle x|\eta|x\rangle$ cuando el sistema ‘está en el mismo estado de antes. Por la analogía clásica, para este estado la suma de los dos observables debería tener un valor igual a la suma de los valores de los dos observables por separado, y el producto un valor igual al producto de dichos valores. La hipótesis nos llevaría al valor $\langle x|\xi + \eta|x\rangle$ para la suma de los dos observables, que coincide con la suma de $\langle x|\xi|x\rangle$ y $\langle x|\eta|x\rangle$, pero para el producto nos conduciría al valor $\langle x|\xi\eta|x\rangle$ o bien al $\langle x|\eta\xi|x\rangle$, y ninguno de ellos está relacionado de forma simple con $\langle x|\xi|x\rangle$ y $\langle x|\eta|x\rangle$.

Pero como la hipótesis sólo falla para el producto, sería lógico decir que $\langle x|\xi|x\rangle$ es el *valor medio* (que también se llama valor esperado) del observable ξ en el estado x , ya que el promedio de la suma de dos cantidades tiene que ser igual a la suma de promedios de dichas cantidades, pero el promedio del producto no tiene por qué ser igual al producto de promedios. Por tanto, hacemos la hipótesis general de que *si medimos el observable ξ un gran número de veces cuando el sistema está en el estado que corresponde a $|x\rangle$, el valor medio de todos los resultados obtenidos será $\langle x|\xi|x\rangle$ si $|x\rangle$ está normalizado*. Si $|x\rangle$ no está normalizado, como ocurrirá necesariamente en el caso de que el estado x sea un autoestado de cierto observable perteneciente a un autovalor que forma parte de

un dominio de autovalores de medida no nula, la hipótesis se convierte en que el valor medio de las medida de ξ es proporcional a $\langle x | \xi | x \rangle$. Esta hipótesis general nos da la base para la interpretación física general de la teoría.

Comentario: Es frecuente relacionar la mecánica cuántica con incertidumbre, no determinismo, etc. En este punto mencionamos que la mecánica cuántica nos ofrece un mundo determinista y no determinista de acuerdo a ciertas circunstancias, Penrose lo plantea de esta forma:

“La verdad es que las descripciones cuánticas son muy precisas, . . . , aunque radicalmente diferentes de las familiares descripciones clásicas. Además encontraremos que, pese a la opinión común en contra, las probabilidades no surgen en el ínfimo nivel cuántico de las partículas, átomos o moléculas -éstos evolucionan de modo determinista- sino que, aparentemente, lo hacen por vía de cierta misteriosa acción a mayor escala ligada a la emergencia de un mundo clásico que podemos percibir conscientemente” ([27], pág. 287)

Las ciertas circunstancias a las que hacemos referencia corresponden al proceso de la medida. Si no medimos el sistema, es decir si no interactuamos con él, evoluciona de un modo determinista; pero en el momento en que deseamos conocerlo, es decir en el momento en que deseamos que emerge en nuestra realidad, surge la probabilidad, surge la incertidumbre. Esto admite una situación inversa, podemos decir que mientras el sistema no sea medido, esta en una superposición de estados no determinada, pero una vez se realiza la medición, el sistema se colapsa y adopta un estado determinado por el valor obtenido en la medida. No esperamos resolver esta situación en este momento, deseamos simplemente compartir con el lector nuestros interrogantes.

Dirac: Únicamente podemos decir en mecánica cuántica que un observable “tiene un valor” particular en un estado determinado, cuando estemos totalmente seguros de que el resultado de medirlo es siempre dicho valor particular, o sea, cuando el sistema esté en un autoestado del observable. Si convenimos en utilizar la expresión de que un observable “tiene un valor” con este significado restringido, es fácil comprobar con la ayuda del álgebra introducida, que si dos observables tienen un valor para un estado particular, entonces para dicho estado la suma de los dos observables tiene un valor igual a la suma de los valores de los dos observables, y el producto de ellos tiene un valor igual al producto de los valores de los dos observables.

Comentario: Es común en la mecánica clásica asociar a un sistema físico, por ejemplo, una posición y un momento lineal bien definidos y decir que su estado dinámico queda perfectamente definido. Contrariamente en mecánica cuántica no se puede hacer dicha aseveración, puesto que no tenemos valores bien definidos de estas variables dinámicas. Como las funciones de onda tienen una cierta extensión espacial, no es posible asignarle a ella una posición precisa, sino más bien una probabilidad de encontrar la partícula localizada en una determinada región del espacio, cuando se realiza una determinada medida de la posición. Del mismo modo, no podemos atribuirle a un sistema físico particular un momento lineal bien definido, ya que la onda asociada al sistema es en realidad una superposición de ondas planas de vector de onda variable. Se puede solamente saber cual es la probabilidad de encontrar el sistema físico con un momento lineal específico. En mecánica cuántica no tenemos pues certezas acerca de las variables dinámicas, sino probabilidades de ocurrencia de dichas variables.

Dirac: En el caso general no tiene sentido hablar del valor que tiene un observable para un estado particular del sistema, pero sí podemos hablar del valor medio de dicho observable en ese estado. Podemos incluso ir más allá y hablar de la probabilidad de que al medirlo obtengamos dicho valor. Esta probabilidad puede obtenerse según la hipótesis general del siguiente modo.

Sea el observable ξ y consideremos el estado correspondiente al ket normalizado $|x\rangle$. La hipótesis general no sólo nos da el valor medio de ξ que es $\langle x | \xi | x \rangle$, sino también el valor medio de cualquier función de ξ tal como $f(\xi)$, que será $\langle x | f(\xi) | x \rangle$. Tomemos como $f(\xi)$ una función de ξ que valga 1 cuando $\xi = a$, siendo a cualquier

número real, y cero en cualquier otro caso. De acuerdo con nuestra teoría general de funciones de un observable, dicha función de ξ tiene sentido, y podemos designarla en conformidad con la notación general del símbolo δ con dos subíndices definida por³

$$\delta_{rs} = \begin{cases} 0 & \text{si } r \neq s \\ 1 & \text{si } r = s \end{cases}$$

El valor medio de dicha función de ξ es precisamente la probabilidad P_a , de que ξ tenga el valor a . O sea

$$P_a = \langle x | \delta_{\xi a} | x \rangle. \quad (3.17)$$

Si a no es autovalor de ξ , $\delta_{\xi a}$ multiplicado por cualquier autoket de ξ es cero y, por tanto, $\delta_{\xi a} = 0$ y $P_a = 0$. Esto está de acuerdo con la conclusión de la sección (10), de que todo resultado de la medida de un observable tiene que ser uno de sus autovalores. . .

La hipótesis de la sección (10) de que cuando el sistema está en un autoestado de ξ perteneciente al autovalor ξ' toda medida de ξ da con seguridad el resultado ξ' , está de acuerdo con la hipótesis general relativa a la interpretación física, y de hecho es consecuencia de ella. Según la hipótesis general, si $|\xi'\rangle$ es un autoket de ξ perteneciente al autovalor ξ' , entonces en el caso de que los autovalores de ξ constituyan un conjunto discreto, resulta

$$\delta_{\xi a} |\xi'\rangle = 0 \quad \text{salvo para} \quad a = \xi'.$$

En este caso, la probabilidad de que ξ tenga un valor distinto de ξ' cuando el sistema está en el estado correspondiente a $|\xi'\rangle$ es cero.

En la práctica es imposible obtener un sistema estrictamente en un autoestado de ξ perteneciente a un autovalor ξ' que forme parte de un intervalo de autovalores, ya que para ello sería necesaria una precisión absoluta. Lo máximo que se puede conseguir en la práctica es que ξ tenga un valor perteneciente a un pequeño intervalo alrededor del valor ξ' . El sistema estará entonces en un estado parecido a un autoestado de ξ . Así pues un autoestado perteneciente a un autovalor que forme parte de un intervalo de autovalores es una idealización matemática, y no se puede realizar en la práctica. A pesar de ello dichos autoestados tienen un papel muy importante en la teoría, y sin ellos no podríamos hacer gran cosa. La ciencia contiene muchos ejemplos de conceptos teóricos que son límites de cosas que se dan en la práctica, y que son de gran utilidad para la formulación precisa de las leyes de la naturaleza, pese a que no se pueden llevar a cabo experimentalmente; y éste no es más que uno de ellos. Es posible que el que la longitud de los kets correspondientes a estos estados sea infinita esté relacionado con su irrealizabilidad práctica, y que todos los estados realizables corresponden a kets que puedan ser normalizados y que, en consecuencia, forman parte de un espacio de Hilbert.

Comentario: Finaliza Dirac la presentación del capítulo presentando de nuevo el problema y la relación entre la matemática y la física. No es desconocida la posición de Dirac acerca de la conveniencia de la *belleza matemática* en la física; el apelar a los estados realizables, es decir, aquellos estados sobre los cuales la experimentación es posible ([8], pág. 100) creemos que es su intento en establecer dicha belleza matemática.

No deseamos terminar esta sección acerca de la mecánica cuántica, sin mencionar que aunque hemos presentado una teoría sólida y aparentemente completa, la realidad es similar a la de cualquier campo del conocimiento. La mecánica cuántica actual —de la cual sólo hemos mencionado una pequeña parte—, corresponde al estado actual del conocimiento en dicha área. Tal como está formulada, en particular con sus interpretaciones, existen detractores de la talla de Einstein, Prigogine y el mismo Dirac, con objeciones

³Esta función es conocida con el nombre de δ de Kronecker.

bastante plausibles. Queremos decir con esto que creemos y esperamos en futuros desarrollos teóricos y experimentales que permitan *ampliar* el campo de aplicación, entender mejor los fundamentos y principios, los problemas y sus posibles soluciones y las limitaciones de la mecánica cuántica, pero esto no es más que una pequeña cara de nuestra ambición por comprender y relacionarnos mejor con lo aquello que denominamos *la realidad*.

3.3. Definiciones adicionales

Presentamos a continuación algunas definiciones adicionales que serán necesarias en el desarrollo de la máquina cuántica.

Definición 3.1. Sea $\mathcal{E}\mathcal{L}_1$ y $\mathcal{E}\mathcal{L}_2$ dos espacios vectoriales y sean los kets $|A\rangle \in \mathcal{E}\mathcal{L}_1$ y $|B\rangle \in \mathcal{E}\mathcal{L}_2$. Se puede formar el ket producto $|A\rangle|B\rangle$ y se representa por $|A\rangle;|B\rangle$. Los kets $|A\rangle;|B\rangle$ subtiende un nuevo espacio vectorial, el espacio $\mathcal{E}\mathcal{L}_1 \otimes \mathcal{E}\mathcal{L}_2$ que se denomina **producto tensorial** de los espacios $\mathcal{E}\mathcal{L}_1$ y $\mathcal{E}\mathcal{L}_2$ (cf. [23], pág. 233). Si $\dim \mathcal{E}\mathcal{L}_1 = n$ y $\dim \mathcal{E}\mathcal{L}_2 = m$ entonces $\dim \mathcal{E}\mathcal{L}_1 \otimes \mathcal{E}\mathcal{L}_2 = nm$.

Capítulo 4

Máquina cuántica

Presentamos en este capítulo la noción de máquina cuántica tal como fue propuesta por David Deutsch en su artículo [8]. Es importante aclarar que el objetivo de este proyecto no era aprehender este concepto, sino el adquirir el lenguaje subyacente al mismo. Sin embargo y por supuesto, consideramos necesaria esta aproximación inicial a la máquina cuántica. La metodología de trabajo será la siguiente, Deutsch construye su máquina a partir de ciertos elementos que son presentados en pocas palabras y definidos usualmente por ecuaciones. Presentamos entonces dichas ecuaciones que serán marcadas por los números *Deutsch – xx*, alrededor de cada ecuación existirán unos comentarios que fueron traducidos y adaptados libremente de los comentarios realizados por Deutsch, estos comentarios y la(s) ecuación(es) respectiva serán encerrados en un caja, de la siguiente manera:

Comentario realizado por Deutsch
$\text{ecuación,} \quad (\text{Deutsch-xx})$
comentario realizado por Deutsch.

De cada ecuación realizaremos una explicación que esperamos permita observar la relación entre los conceptos expuestos en los capítulos anteriores y la comprensión de la misma.

4.1. Elementos estáticos

Comenzaremos nuestra presentación de un modelo cuántico de computación que denotaremos MTQ (Máquina de Turing Cuántica) por los elementos *estáticos* de la misma.

4.1.1. Estados

Una MTQ contiene un conjunto *finito* de estados definido por un conjunto *finito* de M observables 2-estado dado por:

$$\{\hat{n}_i\} \quad (i \in \mathbb{Z}_M) \quad \text{donde} \quad \mathbb{Z}_M = \{0, 1, \dots, M-1\}. \quad (\text{Deutsch-1})$$

Es decir, el conjunto de estados de una MTQ es un conjunto de observables $\{\hat{n}_0, \hat{n}_1, \dots, \hat{n}_{M-1}\}$, este conjunto será representado por \hat{n} .

Miremos un poco más de cerca esta definición. Deutsch afirma que cada observable \hat{n} es un observable 2-estado (*2-state* en el original). Podemos pensar esto como si cada observable \hat{n} tuviera asociado dos auto-vectores, denotemos estos auto-vectores por $|A_1\rangle$ y $|A_2\rangle$ respectivamente, entonces tendríamos que:

$$\begin{aligned} n |A_1\rangle &= a_1 |A_1\rangle, \\ n |A_2\rangle &= a_2 |A_2\rangle, \end{aligned}$$

donde a_1 y a_2 son los auto-valores del observable \hat{n} correspondientes a los auto-vectores $|A_1\rangle$ y $|A_2\rangle$ respectivamente. A partir de la propiedad de los observables de ser 2-estado, también es lícito pensar que los operadores operan sobre un espacio finito bidimensional. De acuerdo los resultados del análisis funcional, un operador lineal que opera sobre un espacio finito puede ser representado por una matriz (cf. [16], sección 2.9). Vamos a presentar como podría ser el operador \hat{n} desde la perspectiva de las matrices, nuestra presentación será realizada con elementos numéricos concretos para enfatizar mejor lo que queremos indicar. Necesitamos una matriz bidimensional n que multiplicada por un vector A_1 obtengamos un escalar a_1 por el mismo vector A_1 y que multiplicada por un vector A_2 obtengamos un escalar a_2 por el mismo vector A_2 .

Sean n , A_1 y A_2 definidos por:

$$n = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \quad A_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Al realizar las multiplicaciones, obtenemos:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Es decir, el vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ es un auto-vector de la matriz $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ y le corresponde el auto-valor 1, asimismo, el vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ es otro auto-vector de la matriz y le corresponde el auto-valor -1 . Finalmente, si multiplicamos la matriz n por otro vector diferente a A_1 y A_2 podemos observar que éste no es un auto-vector de ella.

4.1.2. Memoria

La *memoria* de la MTQ que corresponde a la cinta unidimensional bi-infinita en una máquina de Turing, está definida por un conjunto *infinito* de observables 2-estado representados por:

$$\{\hat{m}_i\} \quad (i \in \mathbb{Z}) \quad \text{donde } \mathbb{Z} \text{ es el conjunto de los números enteros.} \\ \text{(Deutsch-2)}$$

Es decir, la memoria (cinta) en una MTQ es un conjunto de observables definido por $\{\dots, \hat{m}_{-1}, \hat{m}_0, \hat{m}_1, \dots\}$, este conjunto será representado por $\hat{\mathbf{m}}$

El conjunto de observables $\hat{\mathbf{m}}$ y puede ser interpretado afirmando que cada observable \hat{m}_i corresponde a una celda en la cinta infinita bidimensional:

cinta
observables	...	\hat{m}_{-2}	\hat{m}_{-1}	\hat{m}_0	\hat{m}_1	\hat{m}_2	...

Con respecto a este conjunto infinito de observables, afirmamos que tiene las mismas características que los observables definidos para los estados, es decir, cada uno de ellos tiene dos auto-vectores y un auto-valor asociado a cada uno de ellos. Si existe un punto que nos crea algo de incomodidad es con respecto a la infinitud del conjunto. Es difícil concebir un sistema cuántico con infinitos observables en operación, aunque, en forma similar, es difícil concebir una cinta unidimensional bi-infinita para una máquina de Turing. Esta situación se puede entender bajo la hipótesis que aunque el conjunto de observables sea potencialmente infinito, sólo una parte de ellos será usada en cada paso de computación.

4.1.3. Posición actual

Completando el conjunto de observables de un MTQ está la posición actual en la memoria (cinta) que se define por un observable \hat{x} de espectro los \mathbb{Z} .

Es natural comprender que el conjunto de posibles auto-valores (el espectro) del observable \hat{x} sea los \mathbb{Z} debido a la que la MTQ puede estar en cualquier posición de memoria (cinta). Aunque es bastante claro que esto debe ser así desde el punto de vista de los auto-vectores es necesario el siguiente comentario, es posible que un auto-valor tenga asociados dos o más auto-vectores (en este caso se habla de auto-valores degenerados), pero la situación que nos interesa en este caso es que un auto-vector tenga asociados dos o más auto-valores y esto no es posible. Debido que el observable \hat{x} tiene un conjunto infinito de auto-valores, entonces tiene un conjunto infinito de auto-vectores y este conjunto infinito forman una base infinita para el espacio sobre el cual opera el operador, entonces este espacio es infinito dimensional.

4.2. Elementos dinámicos

Para hablar de los elementos *dinámicos* de una MTQ es necesario hablar del *estado* de la MTQ. Diremos que el estado de la MTQ se representa por un vector unitario en el espacio de Hilbert *spanned* por los auto-vectores simultáneos correspondientes

$$|x; \mathbf{n}; \mathbf{m}\rangle \equiv |x; n_0, n_1 \dots n_{M-1}; \dots m_{-1}, m_0, m_1 \dots\rangle. \quad (\text{Deutsch-3})$$

a los observables de la posición actual \hat{x} , al conjunto de estados de la máquina $\hat{\mathbf{n}}$ y a la memoria (cinta) de la máquina $\hat{\mathbf{m}}$, marcados con los correspondientes auto-valores x , \mathbf{n} y \mathbf{m} .

Veamos un poco más de cerca esta definición del estado de una MTQ. El vector de estado es un vector unitario, lo cual quiere decir que está normalizado, es decir que su norma o longitud es uno. Este vector pertenece a un espacio de Hilbert. La notación empleada por Deutsch y los comentarios para el caso por ejemplo del observable \hat{x} significan que su auto-vector se representa por $|x\rangle$ y su auto-valor por x . Observe el lector que en la ecuación (Deutsch-3) se emplea un signo de equivalencia (\equiv) que corresponde a escribir el elemento de la izquierda en la forma del elemento de la derecha, porque de acuerdo a la convención utilizada hasta ahora tenemos que $\mathbf{n} \equiv \{n_1, n_2, \dots, n_M\}$ y que $\mathbf{m} \equiv \{\dots, m_1, m_0, m_1, \dots\}$.

Con base en esta equivalencia nos limitaremos a hablar acerca del elemento $|x; \mathbf{n}; \mathbf{m}\rangle$ en la ecuación (Deutsch-3). La notación empleada corresponde a la definición 3.1 de espacio tensorial. El término *spanned* corresponde a la formación de este producto tensorial. Podríamos hacer una analogía diciendo que el estado de la MTQ pertenece a un espacio de tres “dimensiones” y cada uno de los elementos x , \mathbf{n} y \mathbf{m} representa una “coordenada” de dicho espacio; en otras palabras, para representar el estado de una MTQ necesitamos tres “coordenadas” diferentes.

El estado de la MTQ puede ser representado por la función de estado $|\psi(t)\rangle \in \mathcal{H}$. Su evolución temporal está representada por un operador unitario constante \mathbf{U} sobre el espacio de Hilbert \mathcal{H} definida por:

$$|\psi(nT)\rangle = \mathbf{U}^n |\psi(0)\rangle \quad (n \in \mathbb{Z}^+). \quad (\text{Deutsch-4})$$

En este caso diremos que la relación implícita entre un ket $|A\rangle$ y una función de estado $\psi(t)$ está relacionada con algo llamado *la representación* del ket. Por el momento simplemente aceptaremos que $|x; \mathbf{n}; \mathbf{m}\rangle \equiv |\psi(t)\rangle$, para efectos de observar la evolución del estado de la MTQ. El término $|\psi(nT)\rangle = \mathbf{U}^n |\psi(0)\rangle$ en la

ecuación (Deutsch-4) puede ser entendido así:

$$\begin{aligned}
 |\psi(1)\rangle &= \mathbf{U} |\psi(0)\rangle \\
 |\psi(2)\rangle &= \mathbf{U} |\psi(1)\rangle \\
 &= \mathbf{U}(\mathbf{U} |\psi(0)\rangle) \\
 &= \mathbf{U}^2 |\psi(0)\rangle \\
 |\psi(3)\rangle &= \mathbf{U} |\psi(2)\rangle \\
 &= \mathbf{U}(\mathbf{U} |\psi(1)\rangle) \\
 &= \mathbf{U}(\mathbf{U})(\mathbf{U} |\psi(0)\rangle) \\
 &= \mathbf{U}^3 |\psi(0)\rangle \\
 &\dots \\
 |\psi(nT)\rangle &= \mathbf{U}^n |\psi(0)\rangle
 \end{aligned}$$

donde, $\psi(0)$ representa el estado inicial (posición actual, información en la memoria (cinta), estado actual) de la MTQ.

El término ($n \in \mathbb{Z}^+$) en la ecuación (Deutsch-4) significa que la computación evoluciona en el sentido *positivo* del tiempo y en unidades discretas de éste. No tiene sentido pensar en una computación hacia atrás, esto no quiere decir que la computación no pueda ser reversible, porque la reversibilidad significa llegar al mismo estado anterior en un estado posterior de tiempo, es esta situación no estamos devolviendo el tiempo, estamos regresando el sistema a su estado inicial.

Para ejemplificar el comportamiento de un operador U , vamos a simplificar el espacio de operación a un espacio 4-dimensional. Este ejemplo será construido con base en [33].

Sea

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},$$

nuestro operador de evolución.

Sean

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

una base para nuestro espacio 4-dimensional. Supongamos que el sistema se encuentra en un estado inicial definido por $|\psi(0)\rangle = |11\rangle$. En este caso la evolución del sistema viene dada por:

$$\begin{aligned}
 |\psi(1)\rangle &= \mathbf{U} |\psi(0)\rangle, \\
 \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.
 \end{aligned}$$

El nuevo estado $|\psi(1)\rangle$ puede ser expresado como combinación lineal de elementos de la base por $|\psi(1)\rangle = \frac{1}{\sqrt{2}}(|10\rangle - |11\rangle)$.

Supongamos ahora que el sistema se encuentra en un estado inicial definido por $|\psi(0)\rangle = |00\rangle$. Su evolución está dada por:

$$|\psi(1)\rangle = U |\psi(0)\rangle,$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Es este caso observamos que el estado no cambio, esto quiere decir que el estado $|00\rangle$ es un auto-vector del operador U con un auto-valor asociado de 1.

Continuando con la evolución del estado de la MTQ decimos que la computación comienza en $t = 0$. En este momento \hat{x} y \hat{n} son preparados con el valor cero. El estado de un número finito de los \hat{n} es preparado con la entrada a la MTQ. Además

$$|\psi(0)\rangle = \sum_m \lambda_m |0; \mathbf{0}; \mathbf{m}\rangle \quad (\text{Deutsch-6a})$$

$$\sum_m |\lambda_m|^2 = 1 \quad (\text{Deutsch-6b})$$

donde solamente un número finito de los λ_m son no cero.

Hablemos inicialmente de la ecuación (Deutsch-6a), los términos $|0\rangle$ y $|\mathbf{0}\rangle$ corresponden a inicializar en cero los observables \hat{x} y \hat{n} respectivamente, por lo que la parte “interesante” de la ecuación (Deutsch-6a) corresponde a:

$$|\psi(0)\rangle = \sum_m \lambda_m |\mathbf{m}\rangle \quad (4.1)$$

Desafortunadamente, Deutsch utiliza la misma letra m para varias cosas, no obstante escribiremos la ecuación (4.1) por:

$$|\psi(0)\rangle = \sum_m \lambda_m |m_m\rangle \quad (4.2)$$

Observe que hemos representado el conjunto de auto-vectores $|\mathbf{m}\rangle$ por $|m_m\rangle$. ¿Qué quiere decir entonces la ecuación (4.2)? Quiere decir que el estado $|\psi(0)\rangle$ puede ser descrito como una combinación lineal de la base formada por los auto-vectores $|m_m\rangle$ correspondiente al observable \hat{n} . Cada uno de estos auto-vectores $|m_m\rangle$ tiene asociado un auto-valor que denotaremos por m_m . El término $|\lambda_m|^2$ expresa la probabilidad de encontrar el auto-valor m_m cuando se mida el observable \hat{n} si el sistema está en el estado $|\psi(0)\rangle$ (cf. [17], pág. 10). La suma de estas probabilidades debe ser igual a 1 y esto es precisamente lo que expresa la ecuación (Deutsch-6b).

La probabilidad $|\lambda_m|^2$ puede ser también expresada de la siguiente forma:

Sea

$$|A\rangle = \sum_m \lambda_m |B_m\rangle,$$

un vector expresado como combinación lineal de elementos de una base $B = \{B_1, B_2, \dots\}$, correspondiente a un observable \hat{B} . Si multiplicamos a ambos lados por un elemento $|B_n\rangle$ arbitrario de la base, obtenemos:

$$\begin{aligned}\langle B_n | A \rangle &= |B_n\rangle \sum_m \lambda_m |B_m\rangle \\ &= \sum_m \lambda_m \langle B_n | B_m \rangle.\end{aligned}\tag{4.3}$$

Pero los elementos de una base debido a su ortogonalidad y a su ortonormalidad satisfacen la propiedad

$$\langle B_n | B_m \rangle = \delta_{mn},$$

siendo $\delta_{m,n}$ el delta de Kronecker definido por:

$$\delta_{mn} = \begin{cases} 0 & \text{si } m \neq n \\ 1 & \text{si } m = n. \end{cases}$$

Entonces podemos escribir la ecuación (4.3) por:

$$\begin{aligned}\langle B_n | A \rangle &= \sum_m \lambda_m \delta_{mn} \\ &= \lambda_n.\end{aligned}$$

Si obtenemos el módulo al cuadrado a ambos lados, tenemos:

$$|\langle B_n | A \rangle|^2 = |\lambda_n|^2.$$

Es decir la probabilidad de encontrar el auto-valor B_n correspondiente al auto-vector $|B_n\rangle$, cuando se mide el observable \hat{B} , en un sistema que está en un estado $|A\rangle = \sum_m \lambda_m |B_m\rangle$ viene dada por $|\langle B_n | A \rangle|^2$, o lo que es lo mismo, viene dada por $|\lambda_n|^2$.

4.3. Observación final

Como esperamos haya podido observar el lector, la *esencia* de la MTQ corresponde al operador de evolución temporal U . Cada MTQ corresponde a un operador U diferente. Este operador evoluciona el sistema de un estado $|x; \mathbf{n}; \mathbf{m}\rangle$ a un estado $|x'; \mathbf{n}'; \mathbf{m}'\rangle$ y de acuerdo a Deutsch puede ser representado por una matriz con ciertas características. Suspendemos en este punto la presentación de la MTQ. Somos consientes que hemos dejado algunas lagunas importantes en nuestra exposición, pero por la justa razón que al elaborar este informe lo son también para nosotros. Diremos finalmente, que hemos presentado un nuevo proyecto de investigación, continuación inmediata de éste, en cual esperamos poder resolver algunas de nuestras inquietudes y por supuesto generar algunas nuevas.

Capítulo 5

Conclusiones

Inicialmente mencionamos que no nos satisface la palabra “conclusiones” en su lugar sería más conveniente utilizar la palabra “continuaciones”. Es difícil registrar y plasmar los diferentes momentos vividos durante una investigación. Momentos que van desde la angustia por los infinitos caminos de exploración que se abren, de desilusión por los obstáculos encontrados, hasta la satisfacción (ingenua) de haber aprehendido algo nuevo.

Hemos decidido exponer nuestras “continuaciones” ofreciendo al lector algunos apartes del proyecto de investigación que hemos presentado para ser desarrollado en el año de 1999 *¿Máquina de Turing Cuántica Autorreferencial: Una Posibilidad?*. Este proyecto se presentó varias semanas antes de realizar este informe y estamos seguros que el lector observará, que dicha presentación exigió una competencia *a-priori* en algunos conceptos de la mecánica cuántica que son necesarios para manipular el constructo denominado *máquina cuántica*. El desarrollo de dicha competencia, fue la idea generatriz del proyecto de investigación que hemos presentamos en este informe.

Planteamiento del problema: Proyecto ¿Máquina de Turing Cuántica Autorreferencial: Una Posibilidad?

1. Introducción

Con base en la noción de deseo de Lyotard: *“El deseo no pone en relación una causa y un efecto, sean cuales fueren, sino que es el movimiento de algo que va hacia lo otro como hacia lo que le falta a sí mismo. Eso quiere decir que lo otro está presente en quien desea, y lo está en forma de ausencia. Quien desea ya tiene lo que le falta, de otro modo no lo desearía, y no lo tiene, no lo conoce, puesto que de otro modo no lo desearía”* ([20], pàg. 81), afirmamos que es nuestro deseo intentar construir la *máquina cuántica autorreferencial*. Esta construcción nos posibilita aproximarnos a un deseo subyacente que consiste en comprender las limitaciones actuales de la física cuántica para manipular procesos no computables.

Para ello es necesario apropiarnos de la noción de *Máquina de Turing Cuántica (MTQ)* (lo cual se realiza en el actual proyecto de investigación). Una vez adquirida esta noción, examinaremos la MTQ propuesta por Deutsch [8] en relación con la noción de *Máquina de Turing Probabilística (MTP)* y en relación con la *Máquina de Turing (MT)*; como un caso particular de MTQ se estudiará la MTQ universal (MTQU), ya sea la sugerida por Deutsch [8] o la construída por Bernstein y Vazirani [3]. Adicionalmente se establecerá la relación entre la MTQ y la tesis de Church-Turing. Finalmente, pero como *núcleo* de nuestro proyecto, abordaremos el problema de la construcción de la *Máquina de Turing*

Cuántica Autorreferencial (MTQA).

El proyecto cuenta además con otro matiz que puede ser considerado complementario y anexo a los anteriores, este matiz representa la línea aplicada del pro de investigación y consiste en desarrollar la habilidad del planteamiento y la implementación de algoritmos bajo el paradigma de la computación cuántica, más específicamente, estos algoritmos serán desarrollados en un lenguaje de programación cuántico.

A continuación presentamos más explícitamente la descripción de que se va a realizar en cada una de las etapas que constituyen nuestro proyecto de investigación.

2. Máquina de Turing (MT) y Máquina de Turing Probabilística (MTP)

Se define formalmente una *Máquina de Turing Determinística* (MT) por una cuadrúpla¹

$$\mathcal{MT} = \langle Q, \Sigma, M, \delta \rangle \text{ donde}$$

$Q = \{q_0, q_1, q_2, \dots, q_n\}$: Conjunto finito de estados de la máquina ($Q \neq \emptyset$).

$\Sigma = \{s_0, s_1, s_2, \dots, s_m\}$: Alfabeto. Conjunto finito de símbolos de entrada-salida. Adoptamos por convención que $s_0 = \square$ (símbolo vacío) ($\Sigma - s_0 \neq \emptyset$).

$M = \{L, R, N\}$: Conjunto de movimientos (L : izquierda, R : derecha, N : no movimiento).

δ : Es una *función* definida de un subconjunto $K \times \Sigma$ en $\Sigma \times M \times K$. El lector debe notar que la función δ está definida sobre un subconjunto de $K \times \Sigma$ porque no es necesario que exista una instrucción para cada una de las situaciones (estado, símbolo) teórica (el conjunto formado por $K \times \Sigma$) en las que puede estar la máquina. Además, el *concepto de función refleja la propiedad determinística* de las máquinas de Turing determinísticas. δ también puede ser definida como un conjunto *finito*² de instrucciones $\delta = \{i_0, i_1, i_2, \dots, i_p\}$ donde cada i_j es una quintupla de la forma: $q_m s_m s_n m q_n$, donde $q_m, q_n \in K$; $s_m, s_n \in \Sigma$; $m \in M$.

Por otra parte, una *Máquina de Turing No Determinística* (MTND) es similar a una Máquina de Turing Determinística *excepto* por el carácter “funcional” de δ . En una MTND, δ es una *relación* definida de un subconjunto $K \times \Sigma$ en $\Sigma \times M \times K$. Bajo estas circunstancias la relación δ determina *parcialmente* el movimiento de la MTND y se hace necesario contar con un *operador externo* ([39], pàg. 232) que realice una selección del movimiento a realizar. Este operador externo puede ser implementado bajo diferentes modalidades, siendo una de ellas una función de probabilidad; en este caso se habla de máquinas de Turing Probabilística (MTP).

Nuestra primer etapa será entonces establecer la relación entre una MT y una MTP.

3. Máquina de Turing Probabilística (MTP) y Máquina de Turing Cuántica (MTQ)

Presentamos (parcialmente) la descripción de una Máquina de Turing Cuántica (MTQ).

¹Esta definición puede tener varias variantes de acuerdo al “gusto” de cada autor. Una de ellas es definir la máquina de Turing como un conjunto finito de instrucciones, ya que este conjunto da implícitamente el conjunto de símbolos utilizados y los posibles estados de la máquina.

²La finitud de este conjunto está determinada implícitamente por la finitud del conjunto de estados K y la finitud del alfabeto Σ .

Con relación a los elementos *estáticos* de una MTQ tenemos los siguientes:

a) Estados

La MTQ contiene un conjunto *finito* definido por un conjunto *finito* de M observables de espectro $\{0, 1\}$ ³ representados por:

$$\{\hat{n}_i\} \quad (i \in \mathbb{Z}_M) \quad \text{donde } \mathbb{Z}_M = \{0, 1, \dots, M-1\}.$$

Es decir, el conjunto de estados de una MTQ es un conjunto de observables $\{\hat{n}_0, \hat{n}_1, \dots, \hat{n}_{M-1}\}$, este conjunto será representado por $\hat{\mathbf{n}}$.

b) Cinta

La *memoria* de la MTQ que corresponde a la cinta unidimensional bi-infinita en la MT, está definida por un conjunto *infinito* de observables de espectro $\{0, 1\}$ representados por:

$$\{\hat{m}_i\} \quad (i \in \mathbb{Z}) \quad \text{donde } \mathbb{Z} \text{ es el conjunto de los números enteros.}$$

Es decir, la memoria (cinta) en una MTQ es un conjunto de observables $\{\dots, \hat{m}_{-1}, \hat{m}_0, \hat{m}_1, \dots\}$, este conjunto será representado por $\hat{\mathbf{m}}$ y puede ser interpretado afirmando que cada observable \hat{m}_i corresponde a una celda en la cinta infinita bidimensional:

cinta
observables	...	\hat{m}_{-2}	\hat{m}_{-1}	\hat{m}_0	\hat{m}_1	\hat{m}_2	...

c) Posición actual

La posición actual en la memoria (cinta) en una MTQ se define por un observable \hat{x} de espectro los \mathbb{Z} .

Para hablar de los elementos *dinámicos* de una MTQ es necesario hablar del *estado* (función de estado) de la MTQ. El estado de la MTQ se representa por un vector unitario en el espacio de Hilbert *spanned* por los vectores propios simultáneos correspondientes a los observables de la posición actual \hat{x} , al conjunto de estados de la máquina \hat{n}_i y a la memoria (cinta) de la máquina \hat{m}_i ; este estado se representa por:

$$|x; \mathbf{n}; \mathbf{m}\rangle \equiv |x; n_0, n_1 \dots n_{M-1}; \dots m_{-1}, m_0, m_1 \dots\rangle.$$

Este estado de la MTQ puede ser representado por la función de estado $\psi(T)$ cuya evolución temporal está dada por:

$$|\psi(nT)\rangle = \mathbf{U}^n |\psi(0)\rangle \quad (n \in \mathbb{Z}^+).$$

En este caso, $\psi(0)$ representa el estado inicial (posición actual, información en la memoria (cinta), estado actual) de la MTQ, $\psi(nT)$ representa la evolución a un estado siguiente en el tiempo y \mathbf{U} es el operador que *realiza* la evolución temporal de la MTQ.

³Aunque el espectro de los observables en física es comunmente $\{-\frac{1}{2}, +\frac{1}{2}\}$, se selecciona el espectro $\{0, 1\}$ por tener una interpretación natural con un *bit* de memoria ([8], pàg. 103)

La relación entre la MTP y la MTQ constituye la siguiente etapa de nuestro proyecto. Por el momento suponemos que el operador \mathbf{U} de la MTQ “corresponde” a la relación δ de la MTP, con lo cual diferentes operadores \mathbf{U} permitirán construir diferentes MTQ. Lo anterior se puede “justificar” si se observa que los observables definidos en la MTQ son discretos, los cual nos conduce a distribuciones de probabilidad discretas, que podrían ser “implementadas” en una MTP.

4. Máquina de Turing Cuántica (MTQ) y tesis de Church-Turing

La reinterpretación *física* que realiza Deutsch de la tesis de Church-Turing: “*Every 'function which naturally be regarded as computable' can be computed by the universal Turing machine*” ([8], pàg. 99) en lo que él denomina *principio* de Church-Turing: “*Every finitely realizable physical system can be perfectly simulated by a universal model computing machine operation by finite means*” ([8], pàg. 99), permite “interpretar” una posible invalidación de la tesis de Church-Turing (por lo menos desde el punto de vista Deutsch).

El propio Deutsch es bastante enfático en este punto y afirma lo siguiente “... *there is no a priori reason why physical laws should respect the limitacions of the mathematical processes we call “algorithms” ...*”, lamentablemente para nuestro propositos continua con la siguiente frase “*Although I shall not in this paper find it necessary to do so, there is nothing paradoxical or inconsistent in postulating physical systems which compute functions not in $C(F)$* ” ([8], pàg. 101).⁴

Sin embargo, existe un resultado contradictorio (bajo ciertas restricciones) con la idea de Deutsch, este resultado afirma lo siguiente: “*Uno de los argumentos más sólidos en favor de la tesis de Church ha sido propuesto por R. Gandy, quien ha demostrado que los mecanismos de todas las máquinas constructibles por la mecánica newtoniana no pueden calcular sino funciones programables*” ([7], pàg. 81).

Una forma de aceptar la compatibilidad entre la interpretación de Deutsch y el resultado obtenido por Gandy, es suponer que es la teoría cuántica la que permitiría construir sistemas físicos que computaran funciones no computables (clásicamente). Este es el contexto donde deseamos instaurar la siguiente etapa de nuestro proyecto: Establecer la relación entre la MTQ y algunas interpretaciones de la tesis de Church-Turing.

5. Máquina de Turing Cuántica Universal (MTQU)

Con base en lo mencionado en la sección 3 y en particular lo relacionado con la dinámica de la MTQ representada por:

$$|\psi(nT)\rangle = \mathbf{U}^n |\psi(0)\rangle \quad (n \in \mathbb{Z}^+),$$

podemos afirmar que la MTQU corresponde a un operador \mathbf{U} muy especial. Con respecto a la MTQU, Deutsch afirma que es posible su construcción [8], pero no lo hace; por el contrario, Bernstein y Vazirani sí la construyen, aunque mencionan que surgen ciertas complicaciones en la “implementación” de algunas operaciones [3]. Esta etapa de nuestro proyecto tiene como objetivo construir la MTQU a partir de las indicaciones ofrecidas por Deutsch, o estudiar la propuesta de MTQU realizada por Bernstein y Vazirani.

6. Máquina de Turing Cuántica Autorreferencial (MTQA)

Como mencionamos en la sección 1 esta etapa es el *núcleo* de nuestro proyecto de investigación. Una vez realizadas las etapas anteriores y con la competencia adquirida en su desarrollo, emprenderemos

⁴ $C(F)$ es el conjunto de la funciones Turing-computables en la notación de Deutsch.

la construcción de una MTQA.

En un sistema cuántico diferenciamos entre el *estado del sistema (función de estado)* y los *observables (operadores)* del sistema. Como su nombre lo indica, los observables del sistemas son aquellas propiedades que podemos medir en el sistema. Bajo la teoría cuántica existe una relación entre el medir un observable y el estado del sistema, esta relación puede ser descrita en lenguaje técnico afirmando que una vez se obtiene un valor de un observable; es decir, se obtiene un valor propio correspondiente al operador que representa el observable, el estado del sistema colapsa al vector propio asociado al valor propio que se obtuvo en la medida. Lo que afirma esta relación es que el estado de un sistema cuántico se modifica una vez se realice una medida sobre él.

Con base en lo anterior, nuestra hipótesis consiste en construir una MTQ que sea capaz de observarse ella misma, si esto es posible implicaría que una vez que la MTQ se observa, ella modifica su estado y se convertiría en una MTQA. Por supuesto nuestra hipótesis exige construir un contexto adecuado para la noción de *medir*, que posiblemente no sea el contexto tradicionalmente aceptado, además exige realizar las construcciones formales para modelar las auto-modificaciones realizadas por la MTQA.

7. Lenguaje de programación cuántico

Aunque la investigación que proponemos a ustedes es en esencia teórica el tema de la computación cuántica contiene una gran elemento pragmático que no queremos dejar pasar completamente inadvertido. Por una parte, aunque la idea de la máquina cuántica es un constructo teórico, existen investigaciones acerca de la posible construcción de estas máquinas cuánticas [18]. Pero en donde la computación cuántica presenta su mayor actividad y sus mejores resultados, es en el área de la complejidad algorítmica. Se ha demostrado que es posible realizar la factorización de enteros, en complejidad temporal de tipo polinomial utilizando para ello un algoritmo cuántico realizado por Peter Shor [7]; este resultado está muy relacionado con algunos de los sistemas de criptografía actuales que se basan en el hecho de que en una máquina tradicional, la factorización de enteros es un problema con una complejidad temporal de tipo exponencial. Este ejemplo ilustra el área en donde se ha demostrado que la computación cuántica ofrece sus mayores beneficios: La *disminución de la complejidad temporal computacional* de algunos de los problemas computables que caen bajo la categoría de los problemas llamado *intratables* [37, 32].

Con base en lo anterior y en el trabajo realizado por Omer [44] el cual consiste en el diseño y la implementación de un lenguaje de programación cuántico (diponible como *free software*) bajo ambiente Linux, el profesor Sicard y cuatro estudiantes de la carrera de Ingeniería de Sistemas, tiene como objetivo realizar programación bajo el paradigma de la computación cuántica. Para ello, se desarrollaran las etapas de instalación del lenguaje, estudio del lenguaje, estudio de un algoritmo en programación cuántica (el algoritmo de Shor para factorizar enteros) y finalmente la implementación de dos problemas bajo el paradigma de la computación cuántica; estos problemas se esperan que sean de complejidad temporal exponencial, para poder observar los beneficios del paradigma de la computación cuántica. Por supuesto, el desarrollo de las etapas anteriores exige *a-priori* ofrecer a los estudiantes los elementos de física cuántica y en particular de computación cuántica necesarios para poder “acercarse” al lenguaje de programación cuántico.

Bibliografía

- [1] Apostol, Tom M.: *Mathematical Analysis*, volumen 1 de *Series in Mathematics*. Addison-Wesley, 1957.
- [2] Apostol, Tom M.: *Calculus*, volumen 1. John Wiley & Sons, 2ª edición, 1967.
- [3] Bernstein, Ethan y Umesh Vazirani: *Quantum Complexity Theory*. SIAM J. Comput., 26(5):1411–1473, Octubre 1997.
- [4] Bobenrieth, M. Andrés: *¿Inconsistencias, Por Qué No?* Santafé de Bogotá: Tercer Mundo Editores, División Gráfica, 1996.
- [5] Bunge, Mario: *La Ciencia, Su método y Su filosofía*. Panamericana Editorial Ltda., 2ª edición, 1997.
- [6] Cohen-Tannoudji, Claude, Bernard Diu y Franck Laloe: *Quantum Mechanics*, volumen 1. Hermann and John Wiley and Sons, 1997.
- [7] Delahaye, Paul: *Creaciones Informáticas: Un Juego Universal de Herramientas de Cálculo*. Investigación y Ciencia, páginas 80–83, Abril 1996.
- [8] Deutsch, David: *Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer*. Proc. R. Soc. Lond. A, 400:97–117, 1985.
- [9] Galindo, A. y P. Pascual: *Mecánica Cuántica*. Editorial Alhambra, S.A., 1978.
- [10] Gillespie, D. T.: *Introducción a la Mecánica Cuántica*. Editorial Reverté S.A., 1976.
- [11] Herstein, I. N.: *Álgebra Moderna*. Editorial Trillas, 1974.
- [12] Herstein, I. N. y David J. Winter: *Álgebra Lineal y Teoría de Matrices*. Grupo Editorial Iberoamérica, 1989.
- [13] Hughes, R. I. G.: *Quantum Logic*. Scientific American, 245, Octubre 1981.
- [14] IEEE: *Proceedings of the 35th Symposium on Foundations of Computer Science (FOCS '94)*, 1994.
- [15] Kampis, George: *Life-Like Computing Beyond the Machine Metaphor*. (Unpublished), 1992.
- [16] Kreyszig, Erwin: *Introductory Functional Analysis with Applications*. John Wiley & Sons, 1978.
- [17] Landau, L. D. y E. M. Lifshitz: *Mecánica Cuántica No Relativista*, volumen III. Editorial Reverté S. A., 2ª edición, 1972.

- [18] Lloyd, Seth: *A Potentially Realizable Quantum computer*. Science, 261:1569–1571, 1993.
- [19] Lovaglia, Florence M., Merrit A. Elmore y Donald Conway: *Álgebra*. Harla, 1972.
- [20] Lyotard, Jean F.: *¿Por Qué Filosofar?* Colección: Grandes Obras del Pensamiento. Ediciones Altaya S.A., 1964.
- [21] López López, León y Raúl Gómez Marín: *Matemáticas Básicas para la Informática*, volumen II. Medellín: Universidad EAFIT, 1995.
- [22] Marcelo, Alonso. y Edward Finn: *Fundamentos Cuánticos y Estadísticos*, volumen III. Addison-Wesley, 1980.
- [23] Messiah, Albert: *Mecánica Cuántica*, volumen I. Editorial Tecnos, 1973.
- [24] Morin, Edgar: *Introducción al Pensamiento Complejo*. Colección: Ciencias Cognitivas. Editorial GEDISA, 1990.
- [25] Mortensen, Chris: *Inconsistent mathematics*. Stanford Encyclopedia of Philosophy plato.stanford.edu/entries/mathematics-inconsistent/, 1996.
- [26] Palka, Bruce P.: *An Introduction to Complex Function Theory*. Springer-Verlag, 1991.
- [27] Penrose, Roger: *The Emperor's New Mind*. Oxford University Press, 1989.
- [28] Penrose, Roger: *Las Sombras de la Mente*. Colección: Drakontos. Crítica, 1996.
- [29] Popper, Karl R.: *La Lógica De La Investigación Científica*. Colección: Estructura y Función. Editorial Tecnos, S.A., 1994.
- [30] Rocha, Luis M.: *Artificial Semantically Closed Objects*. Communication and Cognition - Artificial Intelligence, 12(1–2):63–89, 1995.
- [31] Rogers, Hartley: *Theory of Recursive Functions and Effective Computability*. MIT Press, third printing edición, 1992.
- [32] Shor, Peter W.: *Algorithms for Quantum Computation: Discrete Log and Factoring*. En *Proceedings of the 35th Symposium on Foundations of Computer Science (FOCS '94)* [14], páginas 124–134.
- [33] Shor, Peter W.: *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. SIAM J. Comput., 26(5):1484–1509, 1997.
- [34] Sicard, Andrés: *¿Más Allá de la Computabilidad? (Algunas Reflexiones acerca de)*. Rev. U. EAFIT, 34(112):91–100, 1998.
- [35] Sicard Ramírez, Andrés: *Máquinas de Turing Dinámicas: Historia y Desarrollo de una Idea*. Tesis de Licenciatura, Departamento de Informática y Sistemas. Universidad EAFIT, 1998.
- [36] Sieg, Wilfred: *Step by Recursive Step: Church's Analysis of Effective Calculability*. The Bulletin of Symbolic Logic, 3(2):154–180, 1997.
- [37] Simon, Daniel R.: *On The Power of Quantum Computation*. En *Proceedings of the 35th Symposium on Foundations of Computer Science (FOCS '94)* [14], páginas 116–123.

- [38] Soare, Robert I.: *Computability and Recursion*. The Bulletin of Symbolic Logic, 2(3):284–321, 1996.
- [39] Turing, Alan M.: *On Computable Numbers, with an Application to the Entscheidungsproblem*. Proc. London Math. Soc., 42:230–265, 1936.
- [40] von Bertalanffy, Ludwing: *Teoría General de los Sistemas*. Fondo de Cultura Económica (FCE), 1994.
- [41] von Foerster, Heinz: *Principios de autoorganización en un contexto socioadministrativo*. En Pakman, Marcelo (editor): *Las Semillas de la Cibernética*, Colección: Terapia Familiar, página ? Gedisa, 1984.
- [42] Weinberg, Steven: *El Sueño De Una Teoría Final*. Crítica. Crítica, 1994.
- [43] Zajac, Hecht: *Óptica*. Addison-Wesley, 1986.
- [44] Ömer, Bernhard: *A Procedural Formalism for Quantum Computing*. Tesis de Licenciatura, Technical University of Viene, Department of Theoretical Physics, 1998.