
DERIVACIÓN DEL OPERADOR DE
DIRAC EN LA ESFERA
POR EL MÉTODO DE CARTAN

TESIS QUE PRESENTA

SERGIO GUTIÉRREZ SOLÍS

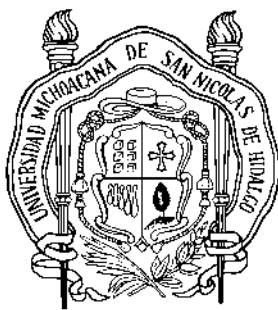
PARA OBTENER EL TÍTULO DE

LICENCIADO EN CIENCIAS
FÍSICO-MATEMÁTICAS

ASESORES

DRA. TATJANA VUKAŠINAC

DR. IDRISH HUET HERNÁNDEZ



UNIVERSIDAD MICHOACANA
DE SAN NICOLÁS DE HIDALGO
FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS

MORELIA, JUNIO DE 2014

A mis abuelos

*La pobreza y la ignorancia
son nuestros peores enemigos, y a nosotros
nos toca resolver el problema de la ignorancia.*

José Vasconcelos
UNAM 1920

Índice general

Resumen	III
Abstract	V
Introducción	VII
1. Mecánica Cuántica	1
1.1. Breve introducción histórica	1
1.2. Espín	5
1.3. Ecuación de Dirac	8
2. Álgebra	19
2.1. Teoría de Grupos	19
2.1.1. Conceptos Básicos	19
2.1.2. Representaciones	22
2.2. Representaciones de $SU(2)$	26
2.2.1. $SU(2)$	26
2.2.2. Operadores de escalera y eigenvalores	31
2.2.3. Representación $\psi_{(1)} \otimes \psi_{(1/2)}$	42
3. Operador de Dirac sobre S^2	45
3.1. Conexión	45
3.1.1. Conceptos básicos	45
3.1.2. Levi-Civita	57
3.2. Conexión de espín en la esfera	58
3.2.1. $SU(2)$ y el momento angular	59
3.2.2. Conexión de espín en la esfera	62
3.3. Operador de Dirac en la esfera	65

3.3.1. Método de Cartan	65
4. Conclusiones	73
A. Grupos de Lie	75
B. Espectro del operador $\mathcal{D}$	79
Bibliografía	81
Índice alfabético	82

Resumen

En el primer capítulo se repasan algunos conceptos fundamentales de la mecánica cuántica y se expone brevemente una construcción de la ecuación de Dirac en $(3 + 1)$ dimensiones. En el segundo capítulo se introducen elementos de la teoría de grupos, se define el grupo $SU(2)$ y se construyen sus representaciones. En el siguiente capítulo se definen conceptos claves de geometría diferencial, como la derivada covariante; se construye la conexión de espín en la esfera y se expone el método de Cartan, usándolo para obtener la ecuación de Dirac en la esfera. En las conclusiones se habla de las ventajas de este método y de las aplicaciones de la ecuación de Dirac en la esfera. Al final se agregan dos apéndices: uno acerca de los grupos de Lie y otro donde se obtiene el espectro del operador de Dirac en la esfera.

Palabras clave: Representaciones, Ecuación de Dirac, Espín.

Abstract

In the first chapter some fundamental concepts of quantum mechanics are reviewed and briefly describes a construction of the Dirac equation in $(3 + 1)$ dimensions. In the second chapter elements of group theory are introduced, the group $SU(2)$ is defined and construct their representations. In the next chapter key concepts of differential geometry, the covariant derivative as defined; the spin connection on the sphere is built and the method of Cartan is exposed, using it to obtain the Dirac equation on the sphere. The conclusions talks about the advantages of this method and applications of the Dirac equation on the sphere. At the end two appendices are added: one on Lie groups and other where the spectrum of the Dirac operator on the sphere is obtained.

Key words: Representations, Dirac's equation, Spin.

Introducción

El presente trabajo consta de tres partes que fueron seleccionadas de acuerdo al tipo de contenido. La primera parte hace referencia a temas de mecánica cuántica, donde se abordan conceptos básicos, situando al lector en contexto. También se introducen algunos elementos más formales con la intención de presentar la ecuación de Dirac (tradicional $(3 + 1)$ dimensiones) y se habla de uno de los conceptos principales de esta tesis, el espín. La segunda parte está dedicada al desarrollo de toda una serie de herramientas y definiciones de teoría de grupos y teoría de representaciones; temas recurrentes cuando se hace un breve estudio de las representaciones del grupo $SU(2)$. La tercera parte se sirve principalmente de la segunda, ya que es en este espacio donde se hace uso de toda la teoría expuesta; se describe el método de Cartan, en cuyo centro se encuentra la teoría de representaciones. También se puntualizan algunas definiciones fundamentales de la geometría diferencial, como apoyo a la exposición de un concepto bastante relevante, la derivada convariante. Al final del núcleo de la tesis, se encuentra la ecuación de Dirac sobre la esfera.

Se añaden dos apéndices, uno que trata de los grupos de Lie y otro donde se calcula el espectro del operador de Dirac sobre la esfera. En las conclusiones se repasa brevemente el método de Cartan y se habla las ventajas que tiene sobre otros enfoques.

La exposición de los temas se ha hecho de tal manera que un lector con cierto conocimiento básico, pueda seguir la línea trazada a través de los puntos claves. Se ha dotado de cierta didáctica y empeño al desarrollo de los detalles, sin entrar de lleno en otros temas, de los cuales se ha procurado incluir una referencia accesible.

Esperamos que los lectores se lleven algo de lo que aquí está escrito.

Capítulo 1

Mecánica Cuántica

Como preámbulo hablaremos muy brevemente del origen y desarrollo de algunos aspectos de la mecánica cuántica, para fijar un contexto y familiarizar al lector. Luego dedicamos una sección a uno de los conceptos centrales de este trabajo, que es el espín; hablamos un poco del experimento de Stern-Gerlach. El capítulo termina con una exposición de la ecuación de Dirac en $3 + 1$ dimensiones.

1.1. Breve introducción histórica

La teoría clásica tanto mecánica como electromagnética formalizadas por Newton, Maxwell y muchos otros más, resultaron insuficientes cuando el hombre comenzó a explorar los territorios expuestos por toda una serie de nuevas observaciones. Una de las ideas más antiguas planteada por pensadores como Demócrito¹, no encajaba en los marcos clásicos de la física.

La idea planteada era realmente sencilla, Demócrito afirmaba que la composición de la materia se basa en pequeñas piezas cuya propiedad principal es ser indivisibles y fundamentales como lo son los ladrillos en un edificio; cualquier cosa en el universo estaría conformada por estos corpúsculos a los que les dió el nombre de *átomos*, que precisamente significa “indivisible”.

A principios del siglo XIX esta idea fue revivida por los químicos princi-

¹Filósofo y matemático griego del siglo V-IV a. C., entre sus máximas se encuentra: “Los principios de todas las cosas son los átomos y el vacío; todo lo demás es dudoso y opinable.”.

palmente. Por ejemplo, el científico inglés Prout² planteó la hipótesis sobre la existencia de pequeñas partículas, que intervienen en las más diversas reacciones químicas sin que pierdan su integridad [1].

En 1904 J. J. Thomson ideó el primer modelo del átomo, el cual es visualizado como una nube de carga positiva, en la que flotan los electrones de carga negativa (que neutraliza la carga del átomo). Este modelo puede explicar hasta cierto punto, la emisión de luz de los objetos al ser calentados, porque el aumento de la temperatura implica un incremento de la energía cinética de los electrones, al encontrarse estos dentro del átomo, son frenados por la atracción de las nubes de cargas positivas, emitiendo así radiación electromagnética.

El modelo de Thomson fue descartado en menos de una década, en 1911, a raíz de un experimento realizado por Rutherford. En dicho experimento se bombardeaban los átomos de una lámina de oro con rayos alfa (recién descubiertos) que eran emitidos por una sustancia radioactiva y están constituidos por partículas de carga positiva. Al estudiar los patrones de dispersión se dió cuenta que, los rayos alfa parecían ser repelidos no por una nube de carga positiva, sino por una carga positiva concentrada en un solo punto, al que llamó *núcleo*. De este modo se reinterpretó la estructura del átomo en analogía con el sistema solar, donde el Sol representa el núcleo (bien ubicado) y los planetas a los electrones (girando alrededor). Pero ¿cuál es el problema con este modelo clásico?

Al estudiar el modelo desde el punto de vista de la electrodinámica clásica se ve que, al existir una fuerza sobre los electrones (constantemente cambian su vector de velocidad), se genera radiación electromagnética de forma continua. Lo anterior se traduce en una pérdida de energía, cuya consecuencia final es la caída del electrón hacia el núcleo, como un satélite al perder velocidad y caer hacia el mar. Esto indica que el átomo tendría un intervalo de vida muy corto, antes de que todos sus electrones se colapsen, convirtiendo al átomo de Rutherford en una estructura muy inestable. Esto entra en contradicción con lo que observamos diariamente a nuestro alrededor: materia estable.

Aquí es donde entra el nombre de Max Planck, quien planteó cierta hipótesis en 1900 para explicar el comportamiento del espectro de emisión electromagnética del *cuerpo negro*. Planck reconoció que esta hipótesis no se susten-

²William Prout, químico y médico del siglo XVIII-XIX, es más recordado por su hipótesis donde dice que, el peso atómico de los elementos es un múltiplo entero del hidrógeno, sugiriéndolo como la única partícula elemental.

taba en nada, lo cual pondría nervioso a cualquiera, sin embargo encajaba perfectamente en los datos experimentales, cosa que nadie había logrado. Su idea se conoce como la *teoría de los cuantos*, y establece que cualquier intercambio o emisión de energía, se lleva a cabo en pequeños bloques de energía. Esta restricción no se encontraba contemplada dentro de las teorías conocidas en aquel entonces, cuando básicamente concebían la transmisión de energía como un continuo. Esta hipótesis también fue requerida por Albert Einstein para su teoría del *efecto fotoeléctrico*, en donde se refería a los cuantos de energía de ondas electromagnéticas, como *fotones*; y por Niels Bohr para su modelo del átomo de hidrógeno. Esto y muchas otras cosas más, nos pueden dar idea de la trascendencia del trabajo de Planck.

El modelo de Bohr está sustentado en una hipótesis adicional, si bien los electrones se mueven alrededor del núcleo, estos lo hacen a través de caminos llamados *órbitas* o *estados estacionarios*. Las órbitas tienen la especial característica de que el electrón no emite luz mientras se encuentra en una de ellas, esto permite que el átomo sea estable. Un átomo puede tener varias órbitas, que se diferencian entre sí por los niveles de energía que requiere un electrón para mantenerse en ella. Entonces un electrón salta de su órbita a una de menor energía, emitiendo un cuanto en forma de luz (un fotón) y continúa moviéndose dentro de su nueva órbita sin radiar.

Este modelo brinda una interpretación sencilla de los espectros observados del hidrógeno, puesto que se pueden medir las distintas frecuencias de la luz emitida y determinar la diferencia de energía entre las órbitas por medio de la relación de Planck

$$E = \omega \hbar = h\nu \quad (1.1)$$

donde E es la energía del fotón, ω la frecuencia angular y $\hbar$ es la constante de Planck³ dividida por 2π . Su nuevo modelo también nos brinda una explicación para los fenómenos de radiación térmica. Un electrón puede adquirir energía por medio de vibraciones y trasladarse a una órbita más lejana (más energética), sucede entonces que el electrón no podrá durar mucho tiempo en este nuevo estado excitado, debido a que el núcleo mantendrá su influencia sobre él, por lo cual el electrón tiende a regresar a su estado base que corresponde a una órbita más baja, provocando la emisión de un fotón en el proceso. Por lo tanto mientras mayor sea la energía absorbida por el electrón, mayor será la energía del fotón emitido; al aumentar la temperatura no solo

³La constante de Planck tiene el valor $h = 6,626069 \times 10^{-34}$ Js, para darnos una idea de las escalas en este contexto

se observarán fotones más energéticos sino en mayor cantidad. Con estos esquemas se logran satisfacer las leyes de Stefan-Boltzmann⁴ y Wien-Golitzin⁵, por ejemplo.

Estos logros pusieron de manifiesto la naturaleza cuántica de la luz y el carácter cuántico de los procesos atómicos, cubriendo así las deficiencias de las teorías clásicas, estimulando además la aceptación de la teoría cuántica entre la comunidad científica. Hasta el momento solamente hemos hablado de un aspecto de la mecánica cuántica que es la novedad de los cuantos de energía, en seguida hablaremos un poco de la naturaleza ondulatoria de la materia.

En 1924 De Broglie propone la existencia de lo que él llamaba *ondas de materia*, que eran producidas por el movimiento de los objetos; presentó una expresión de las ondas de materia dada por

$$\lambda = \frac{2\pi\hbar}{p} \quad (1.2)$$

donde λ es la longitud de la onda y p es el momento del objeto. Esta relación da como resultado que, la magnitud de la longitud de onda sea muy pequeña en el caso de objetos macroscópicos, como podría ser una piedra, un coche o un planeta, que se mueven a velocidades de varios cientos de metros por segundo. En cambio para los objetos microscópicos (de masa pequeña) como podría ser un electrón con una energía cinética de 1eV, le correspondería una longitud de onda del orden de 10^{-9}m , que se encuentra en el rango de los rayos X, lo cual es accesible para una posible medición. En el experimento de G. P. Thomson, por medio de la difracción (que es un fenómeno puramente ondulatorio), se reveló un patrón de anillos oscuros y brillantes en una placa fotográfica sobre la cual incidía un rayo de electrones que era rebotado en un cristal que difractaba los rayos X. En este contexto, es esencial que la rendija u obstáculo usado para provocar la difracción, sea de dimensiones comparables con la longitud de onda, no podría ser mayor porque no se produciría tal efecto de interferencia⁶, es por eso que se usó la estructura cristalina formada por los átomos como un enrejado para producir la difracción.

⁴En un cuerpo negro, la potencia de la radiación térmica sobre unidad de área es proporcional a la cuarta potencia de su temperatura.

⁵Básicamente establece que, la longitud de onda asociada al máximo del espectro emitido de un cuerpo negro, es cada vez menor conforme aumenta la temperatura.

⁶Razón por la cual no se observa el efecto de interferencia de los objetos macroscópicos.

La interpretación que intentaba darle De Broglie a las ondas de materia fracasó, ya que intentaba fusionar mecánicamente ambos conceptos, imaginaba al electrón montado sobre una onda, semejante a un surfista sobre una ola [1]. Con este nuevo concepto la hipótesis de Bohr se aclara un poco más, estipula que sólo las órbitas con radio r tal que $2\pi r = n\lambda$, $n \in \mathbb{N}$ son admitidas en el átomo de hidrógeno, siendo λ la longitud de onda de De Broglie del electrón en la n -ésima órbita. Fue Born quién dió la clave para la correcta interpretación de este fenómeno ondulatorio, grosso modo, los patrones de anillos observados en la difracción del haz de electrones, no era más que la distribución de impactos en la placa fotográfica, de modo que sugiere un trato especial de carácter estadístico. Cuando se lanzaban pocos electrones, no se recolectaba el resultado que era esperado por la mecánica clásica, sino que de alguna manera parecían obedecer una ley que les asignaba un rango de posibles resultados, algunos serían menos probables que otros dentro de un proceso azaroso. Así, al lanzar millones de electrones esencialmente bajo las mismas condiciones en aquel experimento, se les somete a todos bajo la misma ley, que estadísticamente es revelada en la placa fotográfica con regiones más impactadas que otras. Esta distribución de probabilidades tiene el mismo comportamiento que se esperaría de una onda, que es lo que De Broglie llamaba onda de materia.

1.2. Espín

Como hemos visto, los primeros acercamientos de la mecánica cuántica a la física en general, han estado constantemente rodeados de correcciones y el modelo de Bohr no ha sido la excepción. Debido a ciertos problemas en el estudio del espectro de emisión de algunos átomos, conocido como el *efecto Zeeman*⁷, fue introducido en 1916 por parte de Sommerfeld lo que llaman *cuantización espacial*, que no es mas que la cuantización de la proyección (en la dirección z por ejemplo) del momento angular $\mathbf{L}$.⁸ Esta corrección al mo-

⁷Las líneas espectrales sufren una perturbación debido a la presencia de un campo magnético externo, se descomponen en dobletes, tripletes, etc.. Una buena exposición del tema se puede encontrar en [2].

⁸A los vectores los denotamos en negritas. Abusaremos un poco de esta notación y la usaremos con objetos que no necesariamente son vectores, como matrices u operadores. Por ejemplo, un conjunto de matrices $A_1, A_2, \dots, A_n$, expresamos su suma $A_1 + A_2 + \dots + A_n \equiv \mathbf{A}$. También pensamos en $\mathbf{A}$ como un objeto que consta de varias componentes $\mathbf{A} = (A_1, A_2, \dots, A_n)$.

delo de Bohr tiene como resultado una separación en capas de los electrones dentro de una misma órbita, tales separaciones están caracterizadas por los valores permitidos de la componente L_z . Con la intención de verificar estas novedades, en 1922 se llevó a cabo el experimento de Stern-Gerlach que se encuentra entre los experimentos más elegantes dada su relevancia en general y su simplicidad conceptual. El experimento consiste en un horno del cual es emitido a través de un orificio, un haz de átomos de plata en su estado normal (no excitado), el haz atravesará luego un colimador, dirigiendo los átomos hacia un campo magnético externo $\mathbf{B}$ no homogéneo (por simplicidad digamos que apunta en dirección z). El propósito de este campo es producir una desviación del haz, ya que el momento magnético $\boldsymbol{\mu}$ (relacionado con el momento angular de los electrones), permite que se produzca una fuerza sobre los átomos de plata en la dirección z

$$F_z = \frac{\partial}{\partial z}(\boldsymbol{\mu} \cdot \mathbf{B}) = \mu_z \frac{\partial B_z}{\partial z}.$$

Dada la aleatoriedad con que son emitidos del horno los átomos, su momento angular (y por lo tanto el momento magnético) puede resultar en cualquier dirección; desde el punto de vista clásico se esperaría observar en la placa donde se impactaban los átomos, una distribución uniforme y simétrica con respecto al punto $z = 0$ (correspondiente a desviación nula), sin embargo lo que se observó⁹ fue que el haz de átomos de plata se dividió exactamente en dos partes, lo cual reafirmaba la teoría del momento angular cuantizado.

Sin embargo este resultado fue una afortunada casualidad, el experimento no pudo ser interpretado correctamente hasta años más tarde. Un intento por explicar este resultado fue la nueva teoría de Schrödinger, que como veremos más adelante, establece que si el electrón tuviese un momento angular caracterizado por el número cuántico ℓ ($\ell \in \mathbb{N}$) obtendríamos que el haz dejaría de tener un espectro continuo y se dividiría en $2\ell + 1$ partes, por ejemplo si $\ell = 0$ no habría desviación, pero en el caso $\ell = 1$ obtendríamos un espectro de tres componentes. Aun así esta teoría no está conciliada con el experimento, ya que se requiere un valor de $\ell = 1/2$ para obtener el resultado de Stern-Gerlach. En 1925 Pauli propuso añadir un número cuántico s para explicar las configuraciones electrónicas de los átomos¹⁰, restringiendo además que

⁹Según se relata en [3], el resultado no fue percibido a simple vista hasta que por una casualidad, la mala calidad del humo del puro que fumaba Stern produjo una reacción química con los átomos de plata, develando así el resultado del experimento.

¹⁰Y también el llamado *efecto Zeeman anómalo*[2].

cada estado electrónico solo puede ser ocupado por un electrón (*Principio de exclusión de Pauli*). Ese mismo año Goudsmith y Uhlenbeck postularon el concepto de *espín* como ese nuevo número cuántico, que es además una propiedad intrínseca (como la masa y la carga) de las partículas, de modo que no existe una expresión del espín en términos de coordenadas espaciales o del momento. Para ilustrar el concepto se hace referencia al movimiento de La Tierra en el espacio, si bien el planeta gira alrededor del Sol (momento angular orbital), también gira sobre su propio eje (momento angular espino-rial); sin embargo esa visión debe ser abandonada, ya que en este entorno una partícula como el electrón a veces es visualizada como un punto y la noción de un punto girando sobre un eje propio es bastante ambigua.

En 1927 Fraser [4] determina de manera experimental que el átomo de hidrógeno tiene momento angular orbital cero y concluye que esto también sucede con los átomos de sodio y plata, que son hidrogenoides. Ese mismo año Phipps y Taylor [5] reprodujeron el efecto del experimento de Stern-Gerlach con átomos de hidrógeno y, dado que el momento angular orbital del hidrógeno es cero, se concluye que el efecto observado en su experimento se debe a una contribución puramente de espín. Entonces al modelo que se tenía del momento angular total del electrón, se le añadió el término que corresponde al espín, que es el causal de las observaciones de Stern-Gerlach.

El número de proyecciones del espín determinado por el número cuántico s , en la forma $2s+1$ completa la analogía con el momento angular; nos damos cuenta que cuando $2s+1 = 2$, s toma un valor fijo para el electrón y solo pueden ser $s = 1/2$ (en unidades de $\hbar$). Resulta entonces que algunas partículas tienen espín $s = 1/2$ como los protones y neutrones; mientras otras más como el fotón, pueden tener $s = 1$. Esto ofrece una manera de clasificar las partículas como, las que tienen espín con valor semientero $s \in \{1/2, 3/2, 5/2, \dots\}$ y las que toman valores de espín entero $s \in \{0, 1, 2, \dots\}$, las primeras son llamadas *fermiones* porque obedecen la estadística de Fermi-Dirac (y el principio de exclusión de Pauli) y las segundas son *bosones* porque siguen la estadística de Bose-Einstein. Años más tarde (1927) Pauli desarrolla matemáticamente la teoría necesaria para incluir el espín dentro de la mecánica cuántica.

Consideramos que hemos logrado trazar un pequeño bosquejo, el suficiente para que un lector aunque no haya estado familiarizado con el tema, ahora tenga por lo menos una idea del origen y semántica de la mecánica cuántica. Hemos visto cómo, una idea por varios siglos olvidada aparece ante nosotros en el momento menos esperado, motivando a los científicos a explorar toda una diversidad de líneas de investigación que terminan por reanimar

la física. La historia de la ciencia es mucho más rica e interesante de lo que parece en un primer acercamiento, es por eso que motivamos al lector a darle seguimiento al trabajo de estos personajes que de alguna manera (a veces sin darnos cuenta), han influido en nuestra percepción de la realidad. Actualmente existe una cultura de divulgación científica que es perfectamente aprovechable desde cualquier punto de acceso al internet y la cantidad de literatura al respecto es enorme, en lo particular recomendamos a cualquier persona la lectura de [1] y para los estudiantes que deseen un relato en un tono un poco más formal [2, 6].

Estamos preparados para introducir la llamada ecuación de Schrödinger y la *función de onda*. A partir de este momento nos parece conveniente un abordaje del tema centrado más en la estructura y herramientas matemáticas usuales en la mecánica cuántica.

1.3. Ecuación de Dirac

La función de onda (denotada por Ψ) define el estado físico de un sistema (por ejemplo un electrón) y contiene toda la información física asequible del sistema (posición, momento, energía, etc.); a diferencia de la representación clásica de una partícula, cuya trayectoria puede ser descrita por un solo parámetro (el tiempo), una onda requiere dos para ser definida (la amplitud y la fase). De modo que una onda plana tridimensional que describe una partícula libre (que satisface (1.7) con $V = 0$), puede ser expresada como¹¹

$$\Psi(\mathbf{x}, t) = Ce^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$$

donde C es la amplitud de la onda y $\mathbf{k}$ es el llamado *vector de onda* dado por

$$\mathbf{k} = \frac{\mathbf{p}}{\hbar}. \quad (1.3)$$

Usando la expresión anterior (1.3) y la relación de Planck (1.1), se expresa la función de onda como

$$\Psi(\mathbf{x}, t) = Ce^{i(\mathbf{p} \cdot \mathbf{x} - Et)/\hbar}.$$

¹¹El producto interno entre dos vectores lo denotamos $\mathbf{x} \cdot \mathbf{y}$. Frecuentemente usaremos esta notación para expresar (según la convención de Einstein para índices repetidos), sumas de la forma $x_a y_a$, aunque x_a y y_a , no sean las componentes de un vector.

Las ecuaciones (1.2) y (1.3) son conocidas como las *relaciones de De Broglie* las cuales muestran la conexión clave de su hipótesis, por un lado el momento de la partícula y por otro, la longitud y vector de la onda asociada.

La interpretación probabilística que abordó Born está conectada por medio del cuadrado del módulo de la función de onda¹² $|\Psi(\mathbf{x}, t)|^2$, que describe la densidad de probabilidad. De tal manera que la probabilidad de encontrar a la partícula en el volumen d^3x (localizado entre $\mathbf{x}$ y $\mathbf{x} + d\mathbf{x}$) en algún momento t , es igual a $|\Psi(\mathbf{x}, t)|^2 d^3x$ y la probabilidad de encontrar la partícula en algún lugar del espacio, es

$$\int_{\mathbb{R}^3} |\Psi(\mathbf{x}, t)|^2 d^3x = 1.$$

El estado físico de un sistema lo podemos representar por medio de los *vectores de estado*, los cuales forman parte de un *espacio de Hilbert*¹³, $\mathcal{H}$. La notación de Bra-Ket introducida por Dirac que es común en mecánica cuántica para denotar los vectores estado y sus duales¹⁴, será la que usaremos en los siguientes capítulos. Denotaremos entonces un vector estado con un *ket*, $|\alpha(t)\rangle \in \mathcal{H}$ y su dual por un *bra*, $|\alpha(t)\rangle^\dagger = \langle \alpha(t)| \in \mathcal{H}^*$. Los vectores estado pueden ser expresados en distintos espacios, dependiendo de los aspectos que nos interesen, como podría ser el espacio de momentos $\{|\mathbf{p}\rangle\}$ o el de posiciones $\{|\mathbf{x}\rangle\}$ (cabe mencionar que estos últimos son eigenkets de los operadores definidos más adelante (1.4) y (1.5)). Los detalles y teoría del álgebra de bras y kets, los espacios de posición y momento, pueden ser consultados en [2, 7, 6], a nosotros nos basta señalar las definiciones de normalización, completitud y producto interno en dichos espacios

$$\langle \mathbf{x}' | \mathbf{x} \rangle \equiv \delta^3(\mathbf{x}' - \mathbf{x})$$

$$\int_{\mathbb{R}^3} |\mathbf{x}\rangle \langle \mathbf{x}| d^3x \equiv 1$$

así podemos expresar

$$|\alpha(t)\rangle = \int_{\mathbb{R}^3} |\mathbf{x}\rangle \langle \mathbf{x} | \alpha(t) \rangle d^3x$$

¹²Nótese que solamente depende de la amplitud de la onda, la cual puede ser función de la posición y del tiempo.

¹³En pocas palabras: un espacio vectorial con producto interno definido positivo, dotado con una norma y completo.

¹⁴Hablaremos un poco de espacios vectoriales y sus duales en la sección 3.1

donde $\langle \mathbf{x} | \alpha(t) \rangle \equiv \Psi_{(\alpha)}(\mathbf{x}, t)$ y el producto interno se ve como

$$\langle \beta(t) | \alpha(t) \rangle \equiv \int_{\mathbb{R}^3} \overline{\Psi_{(\beta)}}(\mathbf{x}, t) \Psi_{(\alpha)}(\mathbf{x}, t) d^3x.$$

Normalmente será omitido el subíndice de Ψ cuando sea claro de cuál vector estado es la función de onda.

Las variables físicas medibles tienen asociados operadores llamados *observables*, por ejemplo: el momento lineal, la posición y el momento angular. Al conjunto de valores propios de un observable se le llama *espectro*, término que proviene de la espectroscopía. A cada observable se le puede asociar un operador lineal y hermitiano, cuyos eigenvectores forman una base completa. Ejemplos de operadores son los siguientes

$$\text{Momento lineal: } \mathbf{P}\Psi(\mathbf{x}, t) = -i\hbar\nabla\Psi(\mathbf{x}, t) \quad (1.4)$$

$$\text{Posición: } \mathbf{X}\Psi(\mathbf{x}, t) = \mathbf{x}\Psi(\mathbf{x}, t) \quad (1.5)$$

$$\text{Energía: } E\Psi(\mathbf{x}, t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x}, t) \quad (1.6)$$

donde

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

La evolución temporal de la función de onda $\Psi(\mathbf{x}, t)$ de un sistema (que consta de una partícula), está determinada por la ecuación de Schrödinger

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x}, t) = H\Psi(\mathbf{x}, t) \quad (1.7)$$

donde H es el operador hamiltoniano. En el caso de una partícula está dado por

$$H = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{X}, t) \quad (1.8)$$

$V(\mathbf{X}, t)$ es un potencial y m representa la masa de la partícula, también hemos definido

$$\Delta \equiv \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

La ecuación de Schrödinger tiene la gracia de ser lineal, lo cual implica que satisface el principio de superposición y por lo tanto, las soluciones forman un espacio vectorial. Al ser una ecuación diferencial de primer orden con

respecto al tiempo, es suficiente una sola condición inicial para conocer la función de onda en cualquier momento, siempre y cuando no sea perturbada, por ejemplo, con una medición. El primer caso que trató Schrödinger fue para una partícula relativista y obtuvo la ecuación que se conoce con el nombre de Klein-Gordon, desafortunadamente la descartó porque algunos resultados no los lograba interpretar adecuadamente:¹⁵ la densidad de probabilidad no es definida positiva y el espectro de energía contempla energías negativas. Partiendo de la ecuación de energía de una partícula libre en el caso relativista tenemos¹⁶

$$E^2 = p^2 c^2 + m^2 c^4 \quad (1.9)$$

haciendo uso de los operadores de energía y momento, E y P , obtenemos

$$-\hbar^2 \frac{\partial^2 \Psi(\mathbf{x}, t)}{\partial t^2} = (-\hbar^2 c^2 \Delta + m^2 c^4) \Psi(\mathbf{x}, t) \quad (1.10)$$

y esta última es la ecuación de Klein-Gordon. Definimos el operador diferencial

$$\square \equiv \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \quad (1.11)$$

conocido como el *d'Alambertiano* que también puede ser visto como el operador de Laplace en el espacio de Minkowski con signatura $[1, -1, -1, -1]$. Definiendo los operadores¹⁷

$$\begin{aligned} \partial_\mu &\equiv (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{1}{c} \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \\ \partial^\mu &\equiv \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right) \end{aligned}$$

de modo que (1.11) se escribe de la siguiente forma

$$\square = \partial^\mu \partial_\mu.$$

Dicho operador es invariante ante transformaciones de Lorentz, lo cual es importante en este caso ya que se busca introducir la teoría dentro del marco

¹⁵De esto se habla en [8].

¹⁶En el resto de la sección, m denota la masa en reposo.

¹⁷Como es usual, a veces cambiaremos de notación $(x^1, x^2, x^3) \leftrightarrow (x, y, z)$, según nos convenga.

de la relatividad especial. Usando el operador (1.11) podremos expresar la ecuación (1.10) como

$$\left(\square + \frac{m^2 c^2}{\hbar^2}\right) \Psi = 0. \quad (1.12)$$

La ecuación de Dirac surge como respuesta para resanar los defectos de la ecuación de Klein-Gordon. A continuación presentaremos una construcción de esta ecuación, puede encontrarse otra en [9] cuyo estilo enriquecerá al lector. Buscamos entonces, una ecuación de primer orden en $\partial/\partial t$, la cual bien podría haber sido obtenida tomando la raíz cuadrada de la ecuación (1.9), y sustituir los operadores correspondientes, pero esa misma raíz cuadrada nos dificultaría después la evaluación del operador. A modo de ansatz proponemos que Ψ está formada por N componentes, Ψ^n y satisface la siguiente ecuación

$$\frac{1}{c} \frac{\partial \Psi^n}{\partial t} + \sum_{b=1}^3 \sum_{a=1}^N (A^b)_a^n \frac{\partial \Psi^a}{\partial x^b} + \frac{imc}{\hbar} \sum_{a=1}^N B_a^n \Psi^a = 0. \quad (1.13)$$

Expresando a Ψ como un vector columna, A^b y B resultan ser matrices de $N \times N$ constantes (independientes del tiempo y las coordenadas), usando estos recursos escribimos (1.13) como

$$\frac{1}{c} \frac{\partial \Psi}{\partial t} + \mathbf{A} \cdot \nabla \Psi + \frac{imc}{\hbar} B \Psi = \mathbf{0}. \quad (1.14)$$

Recordando la forma del operador hamiltoniano, en esta representación lo podemos escribir en términos de (1.14) como sigue

$$H \Psi = i\hbar \frac{\partial \Psi}{\partial t} = (-i\hbar c \mathbf{A} \cdot \nabla + mc^2 B) \Psi. \quad (1.15)$$

Dado que el operador H es hermitiano, se deduce que $A^b = (A^b)^\dagger$ y $B = B^\dagger$. Podemos deducir propiedades de las matrices A^b y B multiplicando la ecuación (1.14) por su adjunto¹⁸

$$\left(\frac{1}{c} \frac{\partial}{\partial t} - \mathbf{A} \cdot \nabla - \frac{imc}{\hbar} B\right) \left(\frac{1}{c} \frac{\partial}{\partial t} + \mathbf{A} \cdot \nabla + \frac{imc}{\hbar} B\right) \Psi = \mathbf{0}$$

¹⁸Vale la pena mencionar que un operador T es autoadjunto si $\int_{\mathbb{R}^3} (T\Psi)^\dagger \Psi \, d^3x = \int_{\mathbb{R}^3} \Psi^\dagger T\Psi \, d^3x$, para comprobarlo se debe calcular la primera integral por partes y tomar en cuenta que cuando $\mathbf{x} \rightarrow \pm\infty$, entonces $\Psi(\mathbf{x}) \rightarrow \mathbf{0}$.

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - A^a A^b \nabla_a \nabla_b + \frac{m^2 c^2}{\hbar^2} B^2 - \frac{imc}{\hbar} (BA^a + A^a B) \nabla_a \right) \Psi = \mathbf{0}. \quad (1.16)$$

A continuación comparamos término a término ya que buscamos que esta nueva ecuación coincida con la de Klein-Gordon (1.10); usando que $A^a A^b \nabla_a \nabla_b = \frac{1}{2}(A^a A^b + A^b A^a) \nabla_a \nabla_b$, encontramos las siguientes restricciones

$$A^a A^b + A^b A^a = 2\delta^{ab} \mathbb{I} \quad (1.17)$$

$$BA^a + A^a B = 0 \quad (1.18)$$

$$B^2 = \mathbb{I}. \quad (1.19)$$

De la ecuación (1.18) vemos que

$$BA^a = -A^a B = -\mathbb{I} A^a B$$

y al tomar el determinante en ambos lados

$$\det(B) \det(A^a) = (-1)^N \det(A^a) \det(B)$$

de modo que N debe ser un número par; por otro lado usando la misma ecuación (1.18)

$$(A^a)^{-1} BA^a = -B$$

calculando la traza veremos que

$$\text{Tr}((A^a)^{-1} BA^a) = \text{Tr}(A^a (A^a)^{-1} B) = \text{Tr}(B) = \text{Tr}(-B)$$

por lo tanto $\text{Tr}(B) = 0$, de forma similar obtendremos $\text{Tr}(A^a) = 0$.

Definiremos ahora las siguientes matrices

$$\gamma^0 \equiv B$$

$$\gamma^a \equiv BA^a$$

$$\gamma^\mu = (\gamma^0, \gamma^1, \gamma^2, \gamma^3).$$

Si multiplicamos (1.14) por iB

$$\begin{aligned} \left(iB \frac{1}{c} \frac{\partial}{\partial t} + iB \mathbf{A} \cdot \nabla - \frac{mc}{\hbar} \right) \Psi &= \left(i\gamma^0 \partial_0 + i\gamma^a \partial_a - \frac{mc}{\hbar} \right) \Psi \\ &= \left(i\gamma^\mu \partial_\mu - \frac{mc}{\hbar} \right) \Psi = \mathbf{0} \end{aligned}$$

y definiendo $\not{\partial} \equiv \gamma^\mu \partial_\mu$ escribimos entonces la ecuación de Dirac en una forma muy compacta

$$\left(i\not{\partial} - \frac{mc}{\hbar}\right) \Psi = 0. \quad (1.20)$$

A partir de las propiedades de las matrices A^a y B , se puede mostrar que el total de productos distintos de las matrices γ^μ son 15 en total y que junto a la identidad $(\gamma^\mu)^0 = \mathbb{I}$, generan un espacio lineal de 16 dimensiones; llamemos a estas matrices ζ^a , entonces la ecuación

$$\sum_{a=1}^{16} x_a \zeta^a = 0$$

sólo tiene solución trivial, es decir $x_a = 0$. Esto implica que son requeridas por lo menos 16 ecuaciones para poder resolver el sistema, es decir el menor tamaño que podrán tomar las matrices es de 4×4 .

Notemos que la matriz B puede ser diagonalizada en alguna base, la cual siempre será de nuestra preferencia; recordando que $\text{Tr}(B) = 0$ y la propiedad (1.19), podemos tomar

$$B = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}$$

que junto a la ecuación (1.18) y la condición de hermiticidad de las matrices A^a , encontramos que tienen la siguiente forma

$$A^a = \begin{pmatrix} 0 & \sigma^a \\ (\sigma^a)^\dagger & 0 \end{pmatrix}$$

una elección conveniente para las matrices σ^a son las *matrices de Pauli*

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

y satisfacen las siguientes propiedades

$$(\sigma^a)^2 = \mathbb{I}_2 \quad (1.21)$$

$$(\sigma^a)^\dagger = \sigma^a \quad (1.22)$$

$$\left\{ \frac{\sigma^a}{2}, \frac{\sigma^b}{2} \right\} = \frac{\mathbb{I}_2}{2} \delta^{ab} \quad (1.23)$$

$$\left[\frac{\sigma^a}{2}, \frac{\sigma^b}{2} \right] = i\epsilon^{ab} \frac{\sigma^c}{2} \quad (1.24)$$

donde $\{A, B\} \equiv AB + BA$ (*anticonmutador*) y $[A, B] \equiv AB - BA$ (*conmutador*). Esta representación de las matrices γ^μ , es la llamada representación de Pauli-Dirac

$$\gamma^0 = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}, \quad \gamma^a = \begin{pmatrix} 0 & \sigma_a \\ -\sigma_a & 0 \end{pmatrix}.$$

Consideremos ahora una función de onda Ψ de cuatro componentes Ψ^μ , donde cada una de ellas es solución a la ecuación de Klein-Gordon (1.10)

$$\Psi^\mu = C e^{i(\mathbf{p} \cdot \mathbf{x} - Et)/\hbar}$$

en el caso más sencillo, cuando se trata de una partícula en reposo $\mathbf{p} = \mathbf{0}$, la función de onda no depende de $\mathbf{x}$, así que

$$H\Psi = (-i\hbar c \mathbf{A} \cdot \nabla + mc^2 B)\Psi = mc^2 \gamma^0 \Psi = E\Psi$$

tiene las siguientes soluciones usando la representación de Pauli-Dirac

$$\begin{aligned} \Psi_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad E = mc^2; \quad \Psi_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad E = mc^2 \\ \Psi_3 &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad E = -mc^2; \quad \Psi_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad E = -mc^2. \end{aligned}$$

Al igual como sucedía con la ecuación de Klein-Gordon, obtendremos energía negativa como parte del espectro; esto implicaría que un electrón en un estado de energía positiva puede pasar a un estado de energía negativa y continuar precipitándose hacia $E = -\infty$, lo cual provocaría una emisión infinita de energía. La solución a este problema se encuentra en el principio de exclusión de Pauli. Dirac planteó que los estados de energía negativa se encuentran ocupados por otros electrones, y el principio de exclusión de Pauli prohíbe a los electrones descender infinitamente dentro de este inmenso

“mar” de energías negativas. Este *mar de Dirac*¹⁹ es el vacío, sin embargo bajo este nuevo esquema, el vacío deja de ser la “nada”, convirtiéndose en un reservorio infinito de partículas. Un “hueco” en este mar sería percibido como una *antipartícula*. Los conceptos que transtorna esta visión tan novedosa pueden hacer que a primera vista todo esto parezca un artificio, pero el descubrimiento de antipartículas nos ha dado la confianza que nos podría faltar debido a nuestro escepticismo científico.

Continuamos. Al etiquetar las componentes de Ψ de la siguiente forma

$$\Psi_A = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \quad \Psi_B = \begin{pmatrix} \Psi_3 \\ \Psi_4 \end{pmatrix}, \quad \Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} = \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix}$$

podremos escribir la ecuación de Dirac (1.20) como

$$\begin{pmatrix} i\partial_0 - \frac{mc}{\hbar} & i\sigma^a \nabla_a \\ -i\sigma^a \nabla_a & -i\partial_0 - \frac{mc}{\hbar} \end{pmatrix} \begin{pmatrix} \Psi_A \\ \Psi_B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

de donde obtendremos el siguiente par de ecuaciones

$$i\partial_0 \Psi_A + i\sigma^a \nabla_a \Psi_B = \frac{mc}{\hbar} \Psi_A \quad (1.25)$$

$$i\partial_0 \Psi_B + i\sigma^a \nabla_a \Psi_A = -\frac{mc}{\hbar} \Psi_B. \quad (1.26)$$

Es claro que las ecuaciones anteriores están acopladas, por eso en este punto nos auxiliamos con la representación de Weyl, la cual hace uso del siguiente cambio de variables

$$\Psi_A \equiv \Psi_R + \Psi_L, \quad \Psi_B \equiv \Psi_R - \Psi_L$$

que nos permitirá escribir las ecuaciones (1.25) y (1.26) en una forma diferente

$$i\partial_0 \Psi_R + i\sigma^a \nabla_a \Psi_R = \frac{mc}{\hbar} \Psi_L \quad (1.27)$$

$$i\partial_0 \Psi_L - i\sigma^a \nabla_a \Psi_L = \frac{mc}{\hbar} \Psi_R. \quad (1.28)$$

¹⁹La idea del mar de Dirac se convirtió en un artificio innecesario con el advenimiento de la teoría cuántica de campos, donde la función de onda de Dirac se reinterpreta como un operador.

En el caso de partículas cuya masa en reposo es cero o casi cero como los *neutrinos*, las ecuaciones (1.26) y (1.28) se desacoplan y obtenemos las ecuaciones de Weyl

$$\partial_0 \Psi_R + \sigma^a \nabla_a \Psi_R = 0$$

$$\partial_0 \Psi_L - \sigma^a \nabla_a \Psi_L = 0$$

si observamos cuidadosamente notaremos que podemos expresar las ecuaciones de Weyl en términos del momento de la partícula

$$p^\mu p_\mu = \frac{E^2}{c^2} - p^a p_a = m^2 c^4$$

dado que tratamos el caso de una partícula sin masa, $|\mathbf{p}| = E/c$ y escribimos

$$\sigma^a \frac{P_a}{|\mathbf{p}|} \Psi_R = \Psi_R \quad (1.29)$$

$$\sigma^a \frac{P_a}{|\mathbf{p}|} \Psi_L = -\Psi_L. \quad (1.30)$$

En esta tesis son de gran importancia un par de operadores, el primero es el *operador de momento angular orbital* y que se obtiene de la definición clásica ($\mathbf{L} = \mathbf{x} \times \mathbf{p}$) sustituyendo los operadores (1.14) y (1.15), cuyas componentes podemos expresar como

$$\mathcal{L}_a = -i\hbar \epsilon_{abc} x_b \partial_c \quad (1.31)$$

y el otro es el *operador de espín* $\mathbf{S}$, con componentes

$$S_a = \frac{\hbar}{2} \sigma_a. \quad (1.32)$$

Las componentes de este par de operadores obedecen relaciones idénticas

$$[\mathcal{L}_a, \mathcal{L}_b] = i\epsilon_{abc} \mathcal{L}_c \quad (1.33)$$

$$[S_a, S_b] = i\epsilon_{abc} S_c \quad (1.34)$$

lo cual deja entrever una estructura en común, esto es algo ya muy conocido y es lo que tratamos en los siguientes capítulos. Una de las herramientas que debemos desarrollar es la *teoría de representaciones*, que formularemos dentro de los esquemas de la *teoría de grupos*. La relevancia de los siguientes

temas es central en esta tesis, ya que precisamente el método de Cartan (a través del cual lograremos exponer un caso más de la ecuación de Dirac), está basado en el uso de las representaciones.

Por último, vale la pena mencionar que el operador $\boldsymbol{\sigma} \cdot \mathbf{P}/|\mathbf{p}|$ mide la componente del espín en la dirección del momento lineal, que es conocido como la *helicidad*. Las soluciones a la ecuación de Weyl son *espinores* (en el capítulo 3 veremos como son estos objetos). Así las ecuaciones (1.29) y (1.30) nos dan como solución *eigenespinores* con helicidad positiva o negativa, según sea el caso. El lector debe poner especial atención en la ecuación (1.30) escrita de la siguiente forma

$$\left(\boldsymbol{\sigma} \cdot \frac{\mathbf{P}}{|\mathbf{p}|} + \mathbb{I} \right) \Psi_L = 0 \quad (1.35)$$

ya que resulta ser una curiosa sorpresa en el capítulo 3.

Capítulo 2

Álgebra

En la primera sección de este capítulo, presentamos las definiciones fundamentales de la teoría de grupos, para poder introducir y desarrollar las ideas primordiales de la teoría de representaciones. Una perspectiva amplia de la teoría de grupos puede consultarse en [10]; en cuanto a la parte de representaciones, también se puede encontrar en [11]. En la segunda sección definiremos el grupo $SU(2)$ y mostraremos como construir sus distintas representaciones; en especial, calculamos los generadores de la representación $\psi_{(1)} \otimes \psi_{(1/2)}$. También mostramos como calcular el espectro, en las distintas representaciones, de los generadores K_0 y K^2 .

2.1. Teoría de Grupos

2.1.1. Conceptos Básicos

Definición 2.1 *Un grupo es un conjunto no vacío G junto con una operación binaria que llamaremos producto, denotada por $*$, $*$: $G \times G \rightarrow G$ que satisfice las siguientes propiedades:*

1. *(Existencia del elemento neutro) $\exists e \in G$ tal que $\forall a \in G$, $e * a = a$.*
2. *(Existencia del inverso) $\forall a \in G$, $\exists a' \in G$ tal que $a' * a = e$.*
3. *(Asociatividad) $\forall a, b, c \in G$ se cumple $(a * b) * c = a * (b * c)$.*

Un grupo lo denotamos por $(G, *)$ y si la operación $*$ se sobreentiende lo llamaremos simplemente grupo G , también adoptaremos las siguientes convenciones: al elemento neutro, e , también llamado *identidad*, *uno* y *elemento*

unitario, se denotará por 1 , o 1_G si es necesario especificar de qué grupo estamos hablando; al inverso de algún elemento del grupo digamos a lo denotamos a^{-1} . Es importante mencionar que el producto $*$ en general no es conmutativo, esto significa que el orden en que se opera sí importa. Normalmente omitiremos el símbolo $*$, de modo que la expresión $a * b$ será denotada por ab . Existen algunos resultados inmediatos y de gran utilidad.

Proposición 2.2 *Sea G un grupo, entonces $\forall a \in G$, $a^{-1}a = 1$ implica $aa^{-1} = 1$. Cualquier elemento de G conmuta con su inverso.*

Demostración: Como $a^{-1} \in G$ entonces existe su inverso a' tal que $a'a^{-1} = 1$, calculando $aa^{-1} = 1aa^{-1} = (a'a^{-1})aa^{-1} = a'(a^{-1}a)a^{-1} = a'1a^{-1} = a'a^{-1} = 1$. ■

Proposición 2.3 *Sea G un grupo, entonces $\forall a \in G$, $1a = a$ implica $a1 = a$. Cualquier elemento de G conmuta con el elemento neutro.*

Demostración: $g1 = g(g^{-1}g) = (gg^{-1})g = 1g = g$. ■

Proposición 2.4 *Sea G un grupo, entonces el elemento neutro es único.*

Demostración: Supongamos que 1 y $1'$ son neutros en G , entonces $1 = 1'1 = 1'$. En la primera igualdad se usó que $1'$ es neutro y en la segunda que 1 es neutro. ■

Proposición 2.5 *Sea G un grupo, entonces el inverso de cada elemento del grupo es único.*

Demostración: Supongamos que a^{-1} y a' son inversos de a , entonces $a^{-1} = 1a^{-1} = (a'a)a^{-1} = a'(aa^{-1}) = a'1 = a'$. ■

Definición 2.6 *Un subconjunto S de G es un subgrupo si:*

1. $1 \in S$.
2. Si $x, y \in S$, entonces $xy \in S$.
3. Si $x \in S$, entonces $x^{-1} \in S$.

Si S es subgrupo de G , escribimos $S \leq G$.

Ejemplo 2.7 *Cualquier espacio vectorial con la operación de suma de vectores, forma un grupo; un subespacio vectorial forma un subgrupo del espacio vectorial que lo contiene.*

Definición 2.8 *Sean $(G, *)$ y $(G', \star)$ grupos y una función $f : G \rightarrow G'$ tal que $f(a * b) = f(a) \star f(b)$, $\forall a, b \in G$. En tal caso se dice que f es homomorfismo de grupos.*

Nos referiremos a los homomorfismos de grupos simplemente como homomorfismos, ya que en esta tesis sólo usaremos este concepto al hablar de grupos y se sobreentiende.

Proposición 2.9 *Sea $f : G \rightarrow G'$ un homomorfismo. Entonces f tiene las siguientes propiedades:*

1. $f(1_G) = 1_{G'}$.
2. $\forall x \in G, f(x^{-1}) = f(x)^{-1}$.

Demostración:

1. Sea $x \in G$, entonces $f(x)f(1_G) = f(x1_G) = f(x)$. Multiplicando los extremos de la ecuación por $f(x)^{-1}$ por la izquierda, obtenemos $f(1_G) = 1_{G'}$.
2. Sea $x \in G$, como $x^{-1}x = 1_G$, entonces $f(1_G) = f(x^{-1}x) = f(x^{-1})f(x)$. Multiplicando por $f(x)^{-1}$ por la derecha la ecuación y usando la propiedad que acabamos de demostrar, obtenemos $f(x)^{-1} = f(x^{-1})$. ■

Definición 2.10 *Sea $f : G \rightarrow G'$ un homomorfismo. Definimos el núcleo o kernel de f denotado por*

$$\text{Ker}(f) \equiv \{x \in G : f(x) = 1_{G'}\}$$

y la imagen de f denotada por

$$\text{Im}(f) \equiv \{g' \in G' : g' = f(x) \text{ para algún } x \in G\}.$$

Es evidente que tanto $\text{Ker}(f)$ e $\text{Im}(f)$ son no vacíos cuando f es homomorfismo, esto debido al punto 1 de la proposición 2.9. Si un homomorfismo f es una biyección se le llama *isomorfismo*; un isomorfismo definido de G en G es llamado *automorfismo*. Al conjunto de automorfismos definidos sobre G lo denotamos por $\text{Aut}(G)$.

Ejemplo 2.11 El conjunto $\text{Aut}(G)$ forma un grupo con la operación $(\circ)$ de composición de funciones. Nos interesa el caso en que G es un espacio vectorial V , como veremos en seguida.

2.1.2. Representaciones

Definición 2.12 Dado un grupo G y un espacio vectorial V , decimos que una función $\psi : G \rightarrow \text{Aut}(V)$, es una representación lineal de G si ψ es homomorfismo.

Los automorfismos pueden ser vistos también como operadores lineales e invertibles, en otras palabras, matrices no singulares las cuales satisfacen cierta estructura algebraica transmitida por medio del homomorfismo dado en la representación. Al espacio vectorial donde están definidos los automorfismos se le llama *espacio de representación*; la *dimensión de una representación* es la dimensión del espacio de representación¹.

Definición 2.13 Dada una representación ψ de G , se dice que es trivial si $\text{Ker}(\psi) = G$. En cambio, si $\text{Ker}(\psi) = 1_G$ se dice que ψ es fiel.

Definición 2.14 Sea G un grupo, V un espacio vectorial y $\psi : G \rightarrow \text{Aut}(V)$ una representación. Un subespacio $U \subset V$ se dice invariante si $\forall g \in G, U$ es invariante bajo $\psi(g)$, esto quiere decir que $\psi(G)U \subset U$, donde $\psi(G)U \equiv \{\mathbf{v} \in V : \mathbf{v} = \psi(g)\mathbf{u} \text{ para algún } g \in G \text{ y algún } \mathbf{u} \in U\}$.

Definición 2.15 Una representación es reducible si existe algún subespacio propio $0 \subsetneq U \subsetneq V$ tal que U es invariante. La representación es irreducible si no es reducible.

Definición 2.16 Sean $\psi_1 : G \rightarrow \text{Aut}(V_1)$ y $\psi_2 : G \rightarrow \text{Aut}(V_2)$ representaciones, definimos la suma directa de representaciones de la siguiente forma: $\psi = \psi_1 \oplus \psi_2 : G \rightarrow \text{Aut}(V_1 \oplus V_2)$ donde para cualesquiera $g \in G, \mathbf{v}_1 \in V_1$ y $\mathbf{v}_2 \in V_2, \psi(g) = (\psi_1 \oplus \psi_2)(g) : \mathbf{v}_1 \oplus \mathbf{v}_2 \mapsto \psi_1(g)\mathbf{v}_1 \oplus \psi_2(g)\mathbf{v}_2$.

Definición 2.17 Una representación es completamente reducible si es reducible y el complemento directo de cada subespacio propio invariante es un subespacio propio invariante.

¹Todos los espacios de representación con los que trataremos serán de dimensión finita.

Definición 2.18 Una matriz A es suma directa de las matrices $A_1, A_2, \dots, A_n$ si existe una base tal que A es diagonal por bloques:

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & A_n \end{pmatrix} = \bigoplus_{i=1}^n A_i.$$

Definición 2.19 Sean $\psi : G \rightarrow \text{Aut}(V)$ y $\psi' : G \rightarrow \text{Aut}(V')$ representaciones de igual dimensión. Si existe un operador lineal e invertible $T : V \rightarrow V'$ tal que $\forall g \in G, \psi'(g) = T\psi(g)T^{-1}$ se dice que las dos representaciones son equivalentes.

En otras palabras, al cambiar la base del espacio de representación se obtienen representaciones equivalentes. Una buena forma de ilustrar algunas consecuencias de la existencia de un subespacio invariante bajo los automorfismos es observando directamente la forma matricial de estos últimos. Por ejemplo, supongamos que $\psi : G \rightarrow \text{Aut}(V)$ es una representación reducible de dimensión n con subespacio propio invariante U de dimensión k ($0 < k < n$). Como U es subespacio invariante, para cualesquiera elementos $g \in G$ y $\mathbf{u} \in U$, $\psi(g)\mathbf{u} = \mathbf{u}' \in U$. Expresamos la ecuación anterior en forma matricial usando alguna base ordenada donde las primeras k entradas corresponden al subespacio U

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} u'_1 \\ \vdots \\ u'_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

La matriz y los vectores los dividiremos por bloques para hacer más sencillos los cálculos

$$\begin{pmatrix} a_{11} & \dots & a_{1k} & a_{1(k+1)} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} & a_{k(k+1)} & \dots & a_{kn} \\ a_{(k+1)1} & \dots & a_{(k+1)k} & a_{(k+1)(k+1)} & \dots & a_{(k+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & a_{n(k+1)} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$$\begin{pmatrix} u_1 \\ \vdots \\ u_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} u'_1 \\ \vdots \\ u'_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{u}' \\ \mathbf{0} \end{pmatrix}.$$

También podemos ver a estos bloques como transformaciones lineales: $T_A : U \rightarrow U$, $T_B : (V - U) \rightarrow U$, $T_C : U \rightarrow (V - U)$ y $T_D : (V - U) \rightarrow (V - U)$. Obtendremos la siguiente expresión

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} A\mathbf{u} \\ C\mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{u}' \\ \mathbf{0} \end{pmatrix}.$$

Dado que $\forall \mathbf{u} \in U$ la ecuación $C\mathbf{u} = \mathbf{0}$ se debe satisfacer, podemos concluir que $C = 0$ (bloque de ceros) de manera que podemos expresar $\psi(g)$ como

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}.$$

Si imponemos ahora la condición de que cada W complemento directo de U , $U \oplus W = V$, sea un subespacio propio invariante, entonces usando la invarianza de W obtenemos que $\forall \mathbf{w} \in W$

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} B\mathbf{w} \\ D\mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{w}' \end{pmatrix}$$

de manera análoga resulta $B = 0$. Entonces $\psi(g)$ se puede expresar en su forma diagonal por bloques

$$\psi(g) = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} = A \oplus D. \quad (2.1)$$

Esto nos muestra claramente que $\psi(g)$ se puede escribir como la suma directa de las matrices A y D , a las cuales podemos asociar un par de representaciones $\psi_1 : G \rightarrow \text{Aut}(U)$ y $\psi_2 : G \rightarrow \text{Aut}(W)$ definidas $\forall g \in G$ como $\psi_1(g)\mathbf{u} = \psi(g)|_U \mathbf{u}$ y $\psi_2(g)\mathbf{w} = \psi(g)|_W \mathbf{w}$, de manera que $\psi = \psi_1 \oplus \psi_2 : G \rightarrow \text{Aut}(U \oplus W)$ y $\forall \mathbf{v} \in V$, se puede escribir $\mathbf{v} = \mathbf{u} + \mathbf{w}$ para algún elemento $\mathbf{u} \in U$ y $\mathbf{w} \in W$

$$\psi(g)\mathbf{v} = (\psi_1 \oplus \psi_2)(g)\mathbf{v} = (\psi_1(g) \oplus \psi_2(g)) \begin{pmatrix} \mathbf{u} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \psi_1(g)\mathbf{u} \\ \psi_2(g)\mathbf{w} \end{pmatrix}.$$

Teorema 2.20 *Sea ψ una representación completamente reducible, entonces ψ se puede expresar como suma de representaciones irreducibles.*

Demostración: Sea V el espacio de representación, como la representación es reducible por lo menos existe un subespacio propio invariante U_1 y como es completamente reducible implica que su complemento directo también será un subespacio propio invariante U_2 , así que $V = U_1 \oplus U_2$. Como hemos visto anteriormente, a cada subespacio propio invariante U_j con $j \in \{1, 2\}$, le podemos asociar una representación definida por medio de la restricción $\psi_j = \psi|_{U_j}$ de modo que $\psi = \psi_1 \oplus \psi_2$. Cada una de estas representaciones ψ_j bien puede o tener un subespacio propio invariante o no tenerlo. Si ψ_j no tiene subespacios propios invariantes, entonces será una representación irreducible; en cambio si tiene algún subespacio propio invariante W este también será subespacio propio invariante de ψ ; sin pérdida de generalidad supongamos que $\{0\} \subsetneq W \subsetneq U_1$, de tal forma que el subespacio $W \oplus U_2$ será subespacio propio invariante de ψ y cada complemento directo, además de ser subespacio propio invariante de ψ también lo será de ψ_1 , entonces ψ_1 será completamente reducible.

Aplicaremos el razonamiento anterior sucesivamente, en cada representación completamente reducible que haya resultando al asociar a los subespacios propios invariantes una representación. Sabemos que este procedimiento tiene fin porque la dimensión de las representaciones va disminuyendo, ya que cada vez se les asocia un subespacio propio invariante de menor dimensión, hasta que, o bien todas han resultado ser irreducibles o se han obtenido representaciones irreducibles y un número finito de representaciones completamente reducibles de dimensión 2. Estas últimas solamente pueden tener dos subespacios propios invariantes a los cuales únicamente se les pueden asociar representaciones irreducibles de dimensión 1. Podemos escribir entonces a ψ como la suma de todas estas representaciones irreducibles. ■

Corolario 2.21 *Sea ψ una representación completamente reducible, entonces existe una base en la cual $\psi(g)$ es diagonal por bloques $\forall g \in G$.*

Demostración: La demostración es inmediata, usando el teorema 2.20 y la forma de ψ , (2.1), cuando es completamente reducible. ■

Definición 2.22 *Sean $\psi_1 : G \rightarrow \text{Aut}(V_1)$ y $\psi_2 : G \rightarrow \text{Aut}(V_2)$ representaciones de dimensiones n_1 y n_2 respectivamente. Definimos el producto tensorial*

de representaciones de la siguiente forma: $\psi = \psi_1 \otimes \psi_2 : G \rightarrow \text{Aut}(V_1 \otimes V_2)$ donde $\forall g \in G$ $\psi(g) = (\psi_1 \otimes \psi_2)(g) = \psi_1(g) \otimes \psi_2(g)$ puede ser expresado como una matriz de $n_1 n_2 \times n_1 n_2$ cuyas entradas están dadas por $[(\psi_1 \otimes \psi_2)(g)]_{ij,kl} = \psi_1(g)_{ik} \psi_2(g)_{jl}$, equivalentemente $[(\psi_1 \otimes \psi_2)(g)](\mathbf{v}_1 \otimes \mathbf{v}_2) = \psi_1(g)\mathbf{v}_1 \otimes \psi_2(g)\mathbf{v}_2$.

Como Weyl menciona en su libro [12] en la parte que llama “Formal Processes”, básicamente da a conocer tres métodos para generar representaciones a partir de otras, nosotros ya hemos hablado de dos de ellos que son precisamente la suma y el producto directo de representaciones, el tercer método a grandes rasgos está basado en la composición de representaciones. Al final de la sección 2.2.2 presentamos el teorema 2.39 de Clebsch-Gordan que establece una relación entre el producto directo y la suma de representaciones irreducibles.

2.2. Representaciones de SU(2)

En esta sección presentamos la teoría necesaria para construir las representaciones irreducibles del grupo SU(2); nos interesan las representaciones irreducibles porque son los elementos a partir de los cuales se pueden construir las representaciones completamente reducibles por medio de la suma.

2.2.1. SU(2)

El grupo SU(2) está formado por matrices cuyas entradas se encuentran en el campo de los números complejos, son unitarias, de tamaño 2×2 y con determinante igual a 1. El producto definido en el grupo es la multiplicación de matrices. En general una matriz de 2×2 con entradas en $\mathbb{C}$ tiene la siguiente forma

$$A = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}$$

esto significa que está determinada por medio de ocho parámetros reales (dos por cada entrada), los cuales se reducen a cuatro luego de aplicar la condición unitaria $A^{-1} = A^\dagger$,

$$A^{-1} = \begin{pmatrix} z_4 & -z_2 \\ -z_3 & z_1 \end{pmatrix} = \begin{pmatrix} \bar{z}_1 & \bar{z}_3 \\ \bar{z}_2 & \bar{z}_4 \end{pmatrix} = A^\dagger$$

donde se ha utilizado la restricción del determinante para encontrar las entradas de A^{-1}

$$\det(A) = z_1 z_4 - z_2 z_3 = 1$$

obteniendo así $z_4 = \bar{z}_1$ y $-z_2 = \bar{z}_3$. Entonces un elemento cualquiera de SU(2) puede ser escrito de la siguiente forma

$$A = \begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix} \quad (2.2)$$

los elementos de SU(2) se parametrizan usando tres variables reales, debido a la condición,²

$$|z_1|^2 + |z_2|^2 = 1. \quad (2.3)$$

Pensemos ahora en una pequeña variación sobre un elemento de SU(2) determinada por la siguiente transformación

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix} \rightarrow \begin{pmatrix} z_1 + dz_1 & z_2 + dz_2 \\ -(\bar{z}_2 + d\bar{z}_2) & \bar{z}_1 + d\bar{z}_1 \end{pmatrix}$$

exigiremos además que al realizar esta variación, el resultado debe pertenecer a SU(2) es decir, su determinante es igual a uno

$$(z_1 + dz_1)(\bar{z}_1 + d\bar{z}_1) + (z_2 + dz_2)(\bar{z}_2 + d\bar{z}_2) = 1$$

la cual se reduce a

$$z_1 d\bar{z}_1 + \bar{z}_1 dz_1 + z_2 d\bar{z}_2 + \bar{z}_2 dz_2 = 0$$

debido a (2.3) y considerando $dz_1 d\bar{z}_1 \approx 0$ y $dz_2 d\bar{z}_2 \approx 0$.

En el caso en que la variación se realiza alrededor del elemento neutro ($z_1 = 1$ y $z_2 = 0$), obtenemos la ecuación $d\bar{z}_1 + dz_1 = 0$, lo que implica $\text{Re}(dz_1) = 0$ de tal manera que dz_1 es un número imaginario puro, lo cual nos permite escribir a un elemento $g \in \text{SU}(2)$ cercano al neutro como

$$g(\delta_1, \delta_2, \delta_3) = \begin{pmatrix} 1 + i\delta_3 & \delta_2 + i\delta_1 \\ -\delta_2 + i\delta_1 & 1 - i\delta_3 \end{pmatrix} \quad (2.4)$$

con parámetros tales que $|\delta_j| \ll 1$ y $\delta_j \in \mathbb{R}$.

²Como no hay más restricciones en las variables z_1 y z_2 , la topología de SU(2) es la misma que la de S^3 .

En general cuando un grupo está parametrizado por n parámetros reales y puede ser expresado por medio de una función

$$g : \mathbb{R}^n \rightarrow G, \quad g : (x_1, x_2, \dots, x_n) \mapsto g(x_1, x_2, \dots, x_n) = g(\mathbf{x}) \in G,$$

tal que g depende suavemente de los parámetros y $g(\mathbf{0}) = 1$, implica que sus representaciones ψ en el espacio vectorial V , del grupo pueden expresarse en forma paramétrica mediante la composición

$$\begin{aligned} \phi &\equiv \psi \circ g : \mathbb{R}^n \rightarrow \text{Aut}(V) \\ &: (\mathbf{x}) \mapsto \phi(\mathbf{x}) = \psi(g(\mathbf{x})). \end{aligned} \quad (2.5)$$

Podemos tomar la serie de Taylor de $\phi(\boldsymbol{\delta})$ alrededor del $\mathbf{0}$ con $\delta \ll 1$ de modo que

$$\phi(\boldsymbol{\delta}) = \mathbb{I}_n + i\delta_a K_a + \dots \quad (2.6)$$

Los operadores K_a son llamados *generadores del grupo* y están definidos de la siguiente forma

$$K_a = -i \frac{\partial}{\partial \delta_a} \phi(\boldsymbol{\delta}) \Big|_{\boldsymbol{\delta}=\mathbf{0}} \quad (2.7)$$

Ejemplo 2.23 La representación de $\text{SU}(2)$ en $\mathbb{C}^2$, $\psi : \text{SU}(2) \rightarrow \text{Aut}(\mathbb{C}^2)$ dada por $\psi(g) = g$ con g dada por (2.4) y ϕ como en (2.5), obtenemos

$$\phi(\boldsymbol{\delta}) = \begin{pmatrix} 1 + i\delta_3 & \delta_2 + i\delta_1 \\ -\delta_2 + i\delta_1 & 1 - i\delta_3 \end{pmatrix}.$$

Usando (2.7) obtenemos los generadores

$$K_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{y} \quad K_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

de modo que

$$\phi(\boldsymbol{\delta}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\delta_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i\delta_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + i\delta_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Los generadores del ejemplo anterior coinciden con las matrices de Pauli de la sección 1.3, usando la misma notación escribimos

$$\phi(\boldsymbol{\delta}) = \mathbb{I}_2 + i\delta_a \frac{\sigma_a}{2}$$

donde hemos escalado δ_a por 2 sin afectar la ecuación.

Dada una representación determinada, un elemento $\phi(\mathbf{x}) \in \text{Aut}(V)$ puede obtenerse a partir de un elemento infinitesimalmente cercano a $\mathbb{I}$. Tomemos $\delta = \frac{\mathbf{x}}{N}$ con $N \gg 1$ entonces

$$\begin{aligned}
 \phi(\mathbf{x}) &= \lim_{N \rightarrow \infty} \left(\mathbb{I} + i \frac{x_a}{N} K_a \right)^N \\
 &= \lim_{N \rightarrow \infty} \left(\binom{N}{0} \mathbb{I} + \frac{1}{N} \binom{N}{1} (ix_a K_a) + \dots + \left(\frac{1}{N} \right)^N \binom{N}{N} (ix_a K_a)^N \right) \\
 &= \mathbb{I} + ix_a K_a + \frac{1}{2!} (ix_a K_a)^2 + \dots \\
 \phi(\mathbf{x}) &= e^{ix_a K_a}
 \end{aligned} \tag{2.8}$$

Esta última expresión es conocida con el nombre de *parametrización exponencial*. En particular, bajo la representación del ejemplo 2.23, podemos encontrar una expresión sencilla para la ecuación (2.8). Sea $\mathbf{x}$ un vector unitario, definimos la matriz $X \equiv \mathbf{x} \cdot \boldsymbol{\sigma} = x_a \sigma_a$, calculamos X^2 con ayuda de las ecuaciones (1.23) y (1.24)

$$\begin{aligned}
 X^2 &= (x_a \sigma_a)^2 = x_a \sigma_a x_b \sigma_b \\
 &= \frac{1}{2} x_a x_b (\{\sigma_a, \sigma_b\} + [\sigma_a, \sigma_b]) \\
 &= x_a x_b (\mathbb{I}_2 \delta_{ab} + i \epsilon_{abc} \sigma_c) = \mathbb{I}_2.
 \end{aligned}$$

Con la ayuda de la última ecuación obtenemos los siguientes resultados

$$X^{2n} = \mathbb{I}_2 \text{ y } X^{2n+1} = X, \quad n \in \mathbb{N}_0$$

los cuales nos permiten agilizar los siguientes cálculos. Sea $\mathbf{x}'$ un vector cualquiera, el cual podemos escribir en términos de un vector unitario $\mathbf{x}$ como $\mathbf{x}' = x' \mathbf{x}$, entonces

$$\begin{aligned}
 g(\mathbf{x}') &= \exp \left(ix'_a \frac{\sigma_a}{2} \right) = \exp \left(ix' \left(x_a \frac{\sigma_a}{2} \right) \right) = \exp \left(ix' \frac{X}{2} \right) \\
 &= \mathbb{I}_2 + i \frac{x'}{2} X - \frac{1}{2!} \left(\frac{x'}{2} \right)^2 X^2 - \frac{i}{3!} \left(\frac{x'}{2} \right)^3 X^3 + \dots
 \end{aligned}$$

$$= \mathbb{I}_2 + i\frac{x'}{2}X - \frac{1}{2!}\left(\frac{x'}{2}\right)^2 \mathbb{I}_2 - \frac{i}{3!}\left(\frac{x'}{2}\right)^3 X + \dots$$

$$g(\mathbf{x}') = \mathbb{I}_2 \cos \frac{x'}{2} + iX \sin \frac{x'}{2} \quad (2.9)$$

$$g(\mathbf{x}') = \begin{pmatrix} \cos \frac{x'}{2} + ix_3 \sin \frac{x'}{2} & (x_2 + ix_1) \sin \frac{x'}{2} \\ (-x_2 + ix_1) \sin \frac{x'}{2} & \cos \frac{x'}{2} - ix_3 \sin \frac{x'}{2} \end{pmatrix}. \quad (2.10)$$

Nótese que la matriz de la ecuación (2.10), efectivamente es de la forma de (2.2). Encontramos también que podemos restringir x' al intervalo $[0, 4\pi]$ dada la estructura periódica de las funciones seno y coseno. Planteamos ahora una pregunta ¿qué expresión usaremos para el producto en el grupo, en esta forma exponencial?

No olvidemos que estamos tratando con operadores que en principio no conmutan, de modo que el producto de dos elementos del grupo, determinados por los vectores $\mathbf{u}$ y $\mathbf{v}$ es tal que

$$g(\mathbf{u})g(\mathbf{v}) = \exp\left(iu_a \frac{\sigma_a}{2}\right) \exp\left(iv_b \frac{\sigma_b}{2}\right) \neq \exp\left(i(u_a + v_a) \frac{\sigma_a}{2}\right).$$

Sin embargo tenemos la certeza de que el producto resulta en otro elemento del grupo, y como ya hemos mencionado, podemos expresarlo mediante la parametrización exponencial

$$\exp\left(iu_a \frac{\sigma_a}{2}\right) \exp\left(iv_b \frac{\sigma_b}{2}\right) = \exp\left(iw_a \frac{\sigma_a}{2}\right). \quad (2.11)$$

El problema surge al buscar una expresión para $\mathbf{w}$ en términos de $\mathbf{u}$ y $\mathbf{v}$, tomemos $A = iu_a \frac{\sigma_a}{2}$ y $B = iv_b \frac{\sigma_b}{2}$, según la fórmula de Baker-Campbell-Hausdorff³ el producto está dado por

$$e^A e^B = \exp\left(A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + \dots\right).$$

Notemos que es posible encontrar una expresión de la forma (2.11) si los generadores del grupo forman un *álgebra*⁴ bajo el conmutador, es decir si satisfacen

$$[K_a, K_b] = if_{abc} K_c \quad (2.12)$$

³Una demostración de esta fórmula se puede encontrar en [13].

⁴Se puede consultar el apéndice A, para tener un poco más de claridad con respecto a los conceptos que se mencionan en esta última parte de la sección.

donde las constantes f_{abc} son conocidas como las *constantes de estructura*, que en el caso de SU(2) son ϵ_{abc} . A este tipo de grupos que pueden ser parametrizados por variables continuas se les conoce como *grupos de Lie* y la relación (2.12) es llamada el *álgebra de Lie* del grupo. En lo siguiente vamos a construir las representaciones hermitianas del álgebra del grupo SU(2).

Se puede consultar [14] y comprobar que las representaciones de SU(2) son equivalentes a representaciones unitarias, lo que implica que los generadores K_a son hermitianos. Si $K_a^\dagger = K_a$ entonces

$$e^{iu_a K_a} e^{(iu_a K_a)^\dagger} = e^{iu_a K_a} e^{-iu_a K_a^\dagger} = \mathbb{I}_2.$$

2.2.2. Operadores de escalera y eigenvalores

En esta sección visualizaremos los generadores de SU(2) (K_1 , K_2 y K_3), expresados en una representación irreducible de dimensión finita (sobre los complejos), como operadores para tratar el problema de eigenvalores. Empezaremos construyendo los siguientes operadores K_+ , K_- , K_0 y K^2 :

$$K_+ = K_1 + iK_2$$

$$K_- = K_1 - iK_2$$

$$K_0 = K_3$$

$$K^2 = K_1^2 + K_2^2 + K_3^2.$$

Proposición 2.24 1. $[K_+, K_-] = 2K_0$

2. $[K_0, K_\pm] = \pm K_\pm$

3. $(K_\pm)^\dagger = K_\mp$

4. $[K_a, K^2] = 0$, $a \in \{1, 2, 3\}$

Demostración:

1. $[K_+, K_-] = [K_1 + iK_2, K_1 - iK_2] = [K_1, K_1 - iK_2] + i[K_2, K_1 - iK_2] = [K_1, K_1] - i[K_1, K_2] + i[K_2, K_1] + [K_2, K_2] = -i(i\epsilon_{123}K_3) + i(i\epsilon_{213}K_3) = K_3 + K_3 = 2K_0.$

2. $[K_0, K_\pm] = [K_3, K_1 \pm iK_2] = [K_3, K_1] \pm i[K_3, K_2] = i\epsilon_{312}K_2 \pm i(i\epsilon_{321}K_1) = \pm K_1 + iK_2 = \pm(K_1 \pm iK_2) = \pm K_\pm.$

3. $(K_{\pm})^{\dagger} = (K_1 \pm iK_2)^{\dagger} = K_1^{\dagger} \mp iK_2^{\dagger} = K_1 \mp iK_2 = K_{\mp}$.
4. $[K_a, K_b K_b] = \{[K_a, K_b], K_b\} = i\epsilon_{abc}\{K_c, K_b\} = 0$. ■

A partir de este momento utilizaremos la notación de Dirac.

Definición 2.25 Sea T un operador lineal sobre un espacio vectorial V sobre un campo $\mathbf{F}$. Un valor propio o eigenvalor de T es un elemento $c \in \mathbf{F}$ tal que existe un vector no nulo $|v\rangle$ en V tal que

$$T |v\rangle = c |v\rangle. \quad (2.13)$$

Cualquier vector $|v\rangle$ de V que cumpla (2.13) se dice que es un vector propio o un eigenvector asociado al eigenvalor c .

El inciso 4 de la proposición 2.24 es de importancia crucial ya que el hecho de que cada uno de los operadores K_0 , K_1 y K_2 conmuten (uno a uno, ya que K_0 , K_1 y K_2 no son simultáneamente diagonalizables) con K^2 , nos permite garantizar la existencia de una base de eigenvectores que diagonalizan simultáneamente a los operadores, en este caso elegimos a K_0 para buscar tal base. Aunque dos operadores (K_0 y K^2) compartan eigenvectores estos no tendrán asociado el mismo eigenvalor necesariamente, es decir

$$K^2 |\lambda, m\rangle = \lambda |\lambda, m\rangle \quad (2.14)$$

$$K_0 |\lambda, m\rangle = m |\lambda, m\rangle. \quad (2.15)$$

Lema 2.26 Sea $|\lambda, m\rangle$ un eigenvector de los operadores K_0 y K^2 ; si además $K_{\pm} |\lambda, m\rangle$ es no nulo, entonces este último es eigenvector con eigenvalores asociados $m \pm 1$ y λ , de los operadores K_0 y K^2 respectivamente.

Demostración:

$$\begin{aligned} K_0 K_{\pm} |\lambda, m\rangle &= (K_{\pm} K_0 \pm K_{\pm}) |\lambda, m\rangle \\ &= (K_{\pm} K_0) |\lambda, m\rangle \pm K_{\pm} |\lambda, m\rangle = m K_{\pm} |\lambda, m\rangle \pm K_{\pm} |\lambda, m\rangle \\ &= (m \pm 1) K_{\pm} |\lambda, m\rangle. \end{aligned}$$

Luego $K^2 K_{\pm} |\lambda, m\rangle = K_{\pm} K^2 |\lambda, m\rangle = \lambda K_{\pm} |\lambda, m\rangle$. ■

⁵Asumiremos que los eigenvectores $|\lambda, m\rangle$ están normalizados.

Los $K_{\pm}$ a veces son llamados *operadores escalera*, porque al operar sobre un eigenvector $|\lambda, m\rangle$ de K_0 el resultado será nuevamente un eigenvector de K_0 con la diferencia que su eigenvalor habrá aumentado o disminuído en 1 según sea K_+ o K_- el que opera, es de aquí de donde ha resultado también la notación de los subíndices.

Los parámetros λ y m que hemos utilizado en este caso para los eigenvectores $|\lambda, m\rangle$ tienen ciertas restricciones, las cuales salen a la luz al momento de calcular el espectro de los operadores K^2 y K_0 . No sabemos a priori si λ toma valores distintos, sin embargo el lema de Shur excluye esta posibilidad.

Lema 2.27 (Schur) *Si $\psi(g)A = A\psi(g) \forall g \in G$ donde ψ es una representación irreducible de dimensión finita, entonces $A \propto \mathbb{I}$.*

Demostración: Una demostración del lema se puede encontrar en [11]. ■

El operador A es el llamado *operador de Casimir* y para nuestra conveniencia, identificamos $A \equiv K^2$ (recordando el inciso 4 de la proposición 2.24), de modo que $K^2 = \lambda \mathbb{I}$.

Lema 2.28 *Sea λ el eigenvalor asociado al eigenvector $|\lambda, m\rangle$ de K^2 , entonces $\lambda \geq 0$.*

Demostración:

$$\begin{aligned} \langle \lambda, m | K^2 | \lambda, m \rangle &= \lambda \\ \langle \lambda, m | K_0^2 | \lambda, m \rangle + \langle \lambda, m | K_1^2 | \lambda, m \rangle + \langle \lambda, m | K_2^2 | \lambda, m \rangle \\ &= \| K_0 | \lambda, m \rangle \|^2 + \| K_1 | \lambda, m \rangle \|^2 + \| K_2 | \lambda, m \rangle \|^2 \geq 0 \\ \lambda &\geq 0. \blacksquare \end{aligned}$$

A priori, no sabemos si la base ortonormal de eigenvectores comunes de K^2 y K_0 que estamos buscando es degenerada o no, así que por el momento usaremos la notación para la base $\mathbf{B} = \{|\lambda, m; I\rangle\}$ donde el índice I etiqueta la degeneración del eigenvalor m .

Definición 2.29 *Un eigenvector de eigenvalor más alto $|M\rangle$, es tal que si $K_0 |M\rangle = m_0 |M\rangle$ entonces $m \leq m_0$ para cualquier otro eigenvalor m de K_0 . Además $K_+ |M\rangle = \mathbf{0}$.*

Lema 2.30 *Existe por lo menos un eigenvector de eigenvalor más alto $|M\rangle \in \mathbf{B}$.*

Demostración: Como la dimensión del espacio de representación es finita y $\mathbf{B}$ contiene todos los eigenvectores de K_0 , entonces existe al menos un elemento en $\mathbf{B}$ con el mayor eigenvalor m_0 ; y deben cumplir también lo siguiente $K_+ |M\rangle = \mathbf{0}$, porque de otra forma el eigenvector $K_+ |M\rangle$ sería un eigenvector con un eigenvalor más alto que $|M\rangle$ según el lema 2.26, lo que contradice la maximalidad de $|M\rangle$, a menos que $K_+ |M\rangle$ no sea eigenvector. Esto sólo es posible si $K_+ |M\rangle = \mathbf{0}$. ■

De manera inmediata se sigue que, dado un eigenvector de eigenvalor más alto $|M\rangle$, existe un $n \geq 0$ tal que $K_-^{n+1} |M\rangle = \mathbf{0}$, de lo contrario habría una cantidad infinita de eigenvalores negativos.

Lema 2.31 *El eigenvector de eigenvalor más alto es único.*

Demostración: Supongamos que existen dos eigenvectores de eigenvalor más alto distintos entre sí $|M\rangle$ y $|M'\rangle$, por definición deben tener el mismo eigenvalor m_0 . Definimos el subespacio U como

$$U \equiv \text{Span}_{\mathbb{C}}\{(K_-)^k |M\rangle : k \in \{0, 1, 2, \dots, n\}\}.$$

Debido al lema 2.30 el subespacio U tiene por lo menos dimensión 1 sobre $\mathbb{C}$ así que $\{\mathbf{0}\} \subsetneq U$. Se sigue inmediatamente que $|M'\rangle \notin U$ porque el único eigenvector con el mismo eigenvalor en U es $|M\rangle$. Es trivial verificar que el subespacio U es un subespacio invariante usando las relaciones de conmutación para $K_{\pm}$ y K_0 . Dado que $|M'\rangle \notin U$ entonces $U \subsetneq V$, por lo tanto U es subespacio propio invariante de V , lo que contradice la irreducibilidad de la representación. Por lo tanto no puede existir tal $|M'\rangle$ y de hecho $U = V$ a partir del lema 2.30. ■

Lema 2.32 *Los eigenvalores de K_0 son no degenerados.*

Demostración: Supongamos que el lema es falso, consideremos entonces el subespacio U definido en el lema 2.31. Es claro que U contiene un eigenvector de cada eigenvalor pero, como hemos visto, $U = V$ debido a la irreducibilidad de la representación, el resultado se sigue de lo anterior. ■

Se ha mostrado entonces que la base de hecho es $\mathbf{B} = \{|\lambda, m\rangle\}$. Llamaremos j al máximo valor de m , así el eigenvector de peso máximo será $|\lambda, j\rangle$.

Lema 2.33 *El eigenvalor de K^2 es $j(j+1)$, con $j \geq 0$.*

Demostración: Dado el eigenvector $|\lambda, j\rangle$, buscamos una expresión para λ , esto lo haremos resolviendo la ecuación $\|K_+ |\lambda, j\rangle\|^2 = 0$. Entonces⁶

$$\begin{aligned}\|K_+ |\lambda, j\rangle\|^2 &= \langle \lambda, j | K_- K_+ |\lambda, j\rangle \\ &= \langle \lambda, j | K^2 |\lambda, j\rangle - \langle \lambda, j | K_0^2 |\lambda, j\rangle - \langle \lambda, j | K_0 |\lambda, j\rangle \\ &= \lambda - j^2 - j = 0 \\ \lambda &= j(j+1).\end{aligned}$$

Consideremos ahora la traza del conmutador⁷ $[K_+, K_-] = 2K_0$

$$\text{Tr}([K_+, K_-]) = \text{Tr}(K_+ K_-) - \text{Tr}(K_- K_+) = 0$$

de modo que $\text{Tr}(K_0) = 0$. Cuando $j < 0$, dado que es el eigenvalor más alto de K_0 , todos los demás eigenvalores serán negativos, entonces la traza de K_0 sería estrictamente negativa lo cual es una contradicción; como consecuencia $j \geq 0$. ■

Entonces podemos escribir la ecuación (2.14) como

$$K^2 |j, m\rangle = j(j+1) |j, m\rangle. \quad (2.16)$$

Lema 2.34 Si $j(j+1)$ y m son eigenvalores asociados al eigenvector $|j, m\rangle$ de K^2 y K_0 respectivamente, entonces $-j \leq m \leq j$.

Demostración: Sabemos que

$$\|K_+ |j, m\rangle\|^2 = j(j+1) - m(m+1) = (j+m+1)(j-m) \geq 0$$

$$\|K_- |j, m\rangle\|^2 = j(j+1) + m(-m+1) = (j-m+1)(j+m) \geq 0$$

y considerando el lema 2.33 ($j \geq 0$), junto al hecho de que solamente existen tres casos posibles para m , tenemos lo siguiente:

1. $m = 0$ evidentemente cumple $-j \leq m \leq j$.
2. $m > 0$ entonces $j+m+1 > 0$ implica $j-m \geq 0$ y $-j \leq 0 < m \leq j$.
3. $m < 0$ entonces $j-m+1 > 0$ implica $j+m \geq 0$ y $-j \leq m < 0 \leq j$. ■

⁶Es fácil ver que $K_{\mp} K_{\pm} = K^2 - K_0^2 \mp K_0$

⁷La traza de un conmutador siempre es cero cuando se trata de dimensiones finitas.

Lema 2.35 Sea $|j, m\rangle$ eigenvector de K^2 y K_0 con eigenvalores asociados $j(j+1)$ y m respectivamente, entonces

$$m = \pm j \Leftrightarrow K_{\pm} |j, m\rangle = \mathbf{0}.$$

Demostración: $\Rightarrow$) La norma al cuadrado de $K_{\pm} |j, m\rangle$ cuando $m = \pm j$ es

$$\|K_{\pm} |j, \pm j\rangle\|^2 = j(j+1) - j^2 - j = 0.$$

$\Leftarrow$) Si $K_{\pm} |j, m\rangle = \mathbf{0}$, entonces

$$\begin{aligned} \langle j, m | K_{\mp} K_{\pm} |j, m\rangle &= (j(j+1) - m^2 \mp m) \langle j, m | j, m\rangle \\ &= j(j+1) \mp m(\pm m + 1) = (j \mp m)(j \pm m + 1) = 0 \end{aligned}$$

Si $j \pm m + 1 = 0$ entonces $j + 1 = \mp m$ lo que contradice el lema 2.34; resta entonces que $j \mp m = 0$, por lo tanto $m = \pm j$. ■

El siguiente teorema nos indicará como obtener el espectro de K^2 y K_0 .

Teorema 2.36 Sean $j(j+1)$ y m los eigenvalores asociados al eigenvector $|j, m\rangle$ de los operadores K^2 y K_0 respectivamente. Entonces j es cero, semientero o entero. Dado j fijo, m sólo puede tomar $2j+1$ valores: $-j, -j+1, \dots, j-1, j$.

Demostración: A partir del lema 2.34 $-j \leq m \leq j$ es claro que existe un $p \geq 0$ entero tal que

$$j-1 < m+p \leq j. \quad (2.17)$$

Consideremos ahora la sucesión de vectores

$$|j, m\rangle, K_+ |j, m\rangle, (K_+)^2 |j, m\rangle, \dots, (K_+)^p |j, m\rangle \quad (2.18)$$

los cuales según el lema 2.35 serán no nulos y junto con el lema 2.26 resultarán ser eigenvectores $(K_+)^n |j, m\rangle$, con $n \in \{0, 1, 2, \dots, p\}$ de K^2 y K_0 con eigenvalores $j(j+1)$ y $m+n$ respectivamente, lo cual es fácil de ver usando inducción. Supongamos ahora que el eigenvalor $m+p$ asociado al eigenvector $(K_+)^p |j, m\rangle$ es menor que j , es decir $m+p < j$. De acuerdo con el lema 2.26 $K_+(K_+)^p |j, m\rangle$ será eigenvector asociado al eigenvalor $m+p+1$ pero de acuerdo con (2.17) $j < m+p+1$ lo cual contradice el lema 2.34, por lo tanto $m+p = j$, de tal forma que $(K_+)^p |j, m\rangle$ es el eigenvector (de eigenvalor más

alto) de K_0 asociado al eigenvalor j y según el lema 2.35 $K_+(K_+)^p |j, m\rangle$ es nulo. Entonces hemos probado que existe p tal que

$$m + p = j. \quad (2.19)$$

De forma completamente análoga usando los mismos lemas encontraremos $q \geq 0$ tal que

$$m - q = -j \quad (2.20)$$

tomando esta vez $-j \leq m - q < -j + 1$ y la sucesión

$$|j, m\rangle, K_- |j, m\rangle, (K_-)^2 |j, m\rangle, \dots, (K_-)^q |j, m\rangle. \quad (2.21)$$

Usando (2.19) y (2.20) obtenemos una expresión para j

$$p + q = 2j$$

de modo que, o j es igual a un entero positivo dividido entre 2, o es cero. Los vectores de las series (2.18) y (2.21) son eigenvectores de K^2 con eigenvalores $j(j+1)$ y también son eigenvectores de K_0 con eigenvalores

$$-j, -j+1, -j+2, \dots, j-2, j-1, j$$

que son precisamente $2j+1$ los únicos valores que m puede tomar. ■

Así hemos visto que el espectro de estos operadores está determinado por el valor de j , es decir, si j es semientero lo mismo sucederá con los $2j+1$ valores posibles de m ; si j es entero también lo será m . De esta forma solamente es necesario conocer el valor de j para obtener inmediatamente el espectro de K^2 y K_0 .

Ahora nos disponemos a calcular los vectores $K_\pm |j, m\rangle$, donde $|j, m\rangle$ es eigenvector de K^2 y K_0 haciendo uso del lema 2.26:

$$K_\pm |j, m\rangle = a_\pm(j, m) |j, m \pm 1\rangle, \quad a_\pm(j, m) \in \mathbb{C}$$

calcularemos $\|K_\pm |j, m\rangle\|^2$

$$\begin{aligned} \langle j, m | K_\mp K_\pm |j, m\rangle &= \langle j, m \pm 1 | \bar{a}_\pm(j, m) a_\pm(j, m) |j, m \pm 1\rangle \\ &= |a_\pm(j, m)|^2 \langle j, m \pm 1 | j, m \pm 1\rangle = |a_\pm(j, m)|^2 \end{aligned}$$

por otro lado

$$\langle j, m | K_\mp K_\pm |j, m\rangle = \langle j, m | (K^2 - K_0^2 \mp K_0) |j, m\rangle$$

$$= [j(j+1) - m(m \pm 1)] \langle j, m | j, m \rangle = j(j+1) - m(m \pm 1)$$

de modo que

$$|a_{\pm}(j, m)|^2 = j(j+1) - m(m \pm 1). \quad (2.22)$$

Podemos elegir una base de eigenvectores en la cual los $a_{\pm}(j, m)$ son números reales; calculando

$$\langle j, m+1 | K_+ | j, m \rangle = \langle j, m+1 | a_+(j, m) | j, m+1 \rangle = a_+(j, m)$$

por otro lado

$$\langle j, m+1 | K_+ | j, m \rangle = (K_- | j, m+1 \rangle)^{\dagger} | j, m \rangle = \bar{a}_-(j, m+1)$$

así obtenemos que $a_+(j, m) = \bar{a}_-(j, m+1)$ y lo podemos reescribir como

$$|a_+(j, m)|e^{i\alpha_+(j, m)} = |\bar{a}_-(j, m+1)|e^{-i\alpha_-(j, m+1)}$$

donde $\alpha_{\pm}(j, m)$ son las fases (en el intervalo $[0, 2\pi)$) de $a_{\pm}(j, m)$; entonces

$$\alpha_+(j, m) + \alpha_-(j, m+1) = 0. \quad (2.23)$$

Al multiplicar cada uno de los elementos de la base de eigenvectores $|j, m\rangle$ por el factor $e^{-i\omega_{j, m}}$ no se alterarán los eigenvectores mas que en un cambio de fase conservando así las relaciones de ortonormalidad entre ellos

$$|j, m\rangle \mapsto e^{-i\omega_{j, m}} |j, m\rangle \equiv |j, m\rangle'$$

$$\begin{aligned} \langle j, m_1 |' |j, m_2\rangle' &= \langle j, m_1 | e^{i\omega_{j, m_1}} e^{-i\omega_{j, m_2}} |j, m_2\rangle \\ &= e^{i(\omega_{j, m_1} - \omega_{j, m_2})} \langle j, m_1 | j, m_2 \rangle = \delta_{m_1, m_2} \end{aligned}$$

entonces

$$\begin{aligned} K_{\pm} |j, m\rangle' &= e^{-i\omega_{j, m}} a_{\pm}(j, m) |j, m \pm 1\rangle \\ &= e^{-i\omega_{j, m}} |a_{\pm}(j, m)| e^{i\alpha_{\pm}(j, m)} |j, m \pm 1\rangle \\ &= |a_{\pm}(j, m)| e^{i(\alpha_{\pm}(j, m) - \omega_{j, m} + \omega_{j, m \pm 1})} (e^{-i\omega_{j, m \pm 1}} |j, m \pm 1\rangle) \\ &= |a_{\pm}(j, m)| e^{i(\alpha_{\pm}(j, m) - \omega_{j, m} + \omega_{j, m \pm 1})} |j, m \pm 1\rangle'. \end{aligned}$$

Buscamos entonces que el factor $e^{i(\alpha_{\pm}(j, m) - \omega_{j, m} + \omega_{j, m \pm 1})}$ corresponda a un número real, es decir

$$\alpha_{\pm}(j, m) = \omega_{j, m} - \omega_{j, m \pm 1} \quad (2.24)$$

según el teorema 2.36 habrá $2j + 1$ fases $\omega_{j,m}$ por determinar en términos de las $\alpha_{\pm}(j, m)$, sin embargo dada la ecuación (2.23) bastará usar solamente las $\alpha_{+}(j, m)$; notemos que tanto $\alpha_{+}(j, j)$ y $\alpha_{-}(j, -j)$ debido a la ecuación (2.22), pueden tomar cualquier valor de modo que tenemos la libertad de ajustar

$$\alpha_{+}(j, j) = \omega_{j,j}$$

y el resto de las fases $\omega_{j,m}$ ($m < j$) las calculamos usando la ecuación (2.24) recursivamente, de modo que

$$\omega_{j,m} = \sum_{k=m}^j \alpha_{+}(j, k).$$

Como $|a_{\pm}(j, m)| \in \mathbb{R}^{+}$ podemos escribir a partir de (2.22)

$$|a_{\pm}(j, m)| = \sqrt{j(j+1) - m(m \pm 1)}$$

escribiremos $|a_{\pm}(j, m)|$ simplemente como $a_{\pm}(j, m)$ entendiendo que hemos tomado aquella base donde las fases de sus elementos han sido convenientemente ajustadas. Obtendremos así una ecuación que nos describe explícitamente como operan K_{+} y K_{-} sobre los eigenvectores de K^2 y K_0

$$K_{\pm} |j, m\rangle = \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle. \quad (2.25)$$

Ejemplo 2.37 *Calcularemos los operadores $K_1, K_2, K_0, K^2, K_{+}$ y K_{-} en su forma matricial, en el caso en que $j = \frac{1}{2}$. En este ejemplo, los únicos valores permitidos para m resultan ser $\{-\frac{1}{2}, \frac{1}{2}\}$. Tomaremos como base ordenada los eigenvectores $\mathbf{e}_1 = |\frac{1}{2}, \frac{1}{2}\rangle$ y $\mathbf{e}_2 = |\frac{1}{2}, -\frac{1}{2}\rangle$. Entonces*

$$K_0 \mathbf{e}_1 = \frac{1}{2} \mathbf{e}_1$$

$$K_0 \mathbf{e}_2 = -\frac{1}{2} \mathbf{e}_2$$

y resulta

$$K_0 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

$$K_{+} \mathbf{e}_1 = \mathbf{0}$$

$$K_+ \mathbf{e}_2 = \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right) + \frac{1}{2} \left(-\frac{1}{2} + 1 \right)} \mathbf{e}_1 = \mathbf{e}_1$$

entonces

$$K_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

$$K_- \mathbf{e}_1 = \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right) - \frac{1}{2} \left(\frac{1}{2} - 1 \right)} \mathbf{e}_2 = \mathbf{e}_2$$

$$K_- \mathbf{e}_2 = \mathbf{0}$$

entonces

$$K_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

El resto K_1 , K_2 y K^2 los obtenemos despejando de las definiciones.

$$K_1 = \frac{1}{2}(K_+ + K_-) = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$K_2 = \frac{1}{2i}(K_+ - K_-) = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$K^2 = K_1^2 + K_2^2 + K_0^2 = \frac{3}{4} \mathbb{I}_2.$$

En el ejemplo anterior podemos ver que en el caso $j = \frac{1}{2}$, los operadores K_1 , K_2 y K_0 resultan ser una vez más, las matrices de Pauli.

Ejemplo 2.38 *Calcularemos los operadores K_1 , K_2 , K_0 , K^2 , K_+ y K_- en su forma matricial, en el caso en que $j = 1$. En este ejemplo, los únicos valores permitidos para m resultan ser $\{-1, 0, 1\}$. Tomaremos como base ordenada los eigenvectores $\mathbf{e}_1 = |1, 1\rangle$, $\mathbf{e}_2 = |1, 0\rangle$ y $\mathbf{e}_3 = |1, -1\rangle$. Entonces*

$$K_0 \mathbf{e}_1 = \mathbf{e}_1$$

$$K_0 \mathbf{e}_2 = \mathbf{0}$$

$$K_0 \mathbf{e}_3 = -\mathbf{e}_3$$

entonces

$$K_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

$$\begin{aligned} K_+ \mathbf{e}_1 &= \mathbf{0} \\ K_+ \mathbf{e}_2 &= \sqrt{2} \mathbf{e}_1 \\ K_+ \mathbf{e}_3 &= \sqrt{2} \mathbf{e}_2 \end{aligned}$$

entonces

$$K_+ = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Recordemos ahora que $K_- = K_+^\dagger$ y resulta

$$K_- = \sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Nuevamente usamos las definiciones

$$\begin{aligned} K_1 &= \frac{1}{2}(K_+ + K_-) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ K_2 &= \frac{1}{2i}(K_+ - K_-) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \\ K^2 &= K_1^2 + K_2^2 + K_0^2 = 2\mathbb{I}_2. \end{aligned}$$

Notemos que el índice j además de caracterizar el espectro de K_0 , nos dice también cual es la dimensión del espacio de representación de cada representación irreducible. Se denotará entonces a la representación irreducible asociada a j como $\psi_{(j)}$ cuyo espacio de representación tiene dimensión $2j+1$.

Teorema 2.39 (Clebsch-Gordan) Sean $\psi_{(j)}$ y $\psi_{(s)}$ representaciones irreducibles de $SU(2)$, entonces el producto directo $\psi_{(j)} \otimes \psi_{(s)}$ se puede descomponer como suma de representaciones irreducibles de $SU(2)$ de la siguiente manera

$$\psi_{(j)} \otimes \psi_{(s)} = \bigoplus_{l=|s-j|}^{s+j} \psi_{(l)}.$$

Demostración: Una demostración de este teorema se puede consultar en el libro de Sakurai [7]. ■

2.2.3. Representación $\psi_{(1)} \otimes \psi_{(1/2)}$

Construiremos a continuación una representación de $SU(2)$ por medio del producto tensorial entre las representaciones $\psi_{(1)}$ y $\psi_{(1/2)}$, cuyos espacios de representación serán V y W respectivamente de dimensiones 3 y 2 según el teorema 2.36. Tomemos entonces $\psi = \psi_{(1)} \otimes \psi_{(1/2)}$, tal como hicimos en la sección 2.2.1, escribimos la expresión para la serie de Taylor de $\phi_{(1)}(\mathbf{x}) = \psi_{(1)}(g(\mathbf{x}))$ y $\phi_{(2)}(\mathbf{x}) = \psi_{(1/2)}(g(\mathbf{x}))$ alrededor de $\mathbf{0}$

$$\phi_{(1)}(\boldsymbol{\delta}) = \mathbb{I}_3 + i\delta_a \theta_a + \cdots$$

$$\phi_{(2)}(\boldsymbol{\delta}) = \mathbb{I}_2 + i\delta_a \frac{\sigma_a}{2} + \cdots$$

en donde θ_a y σ_a son los generadores de cada representación. Los generadores σ_a ya los conocemos del ejemplo 2.23 y de la sección 1.3, mientras que los generadores θ_a han sido obtenidos por medio de una transformación unitaria de los generadores del ejemplo 2.38, $\theta_a \equiv TK_aT^{-1}$ donde

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 & 1 \\ -i & 0 & -i \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

de tal manera que las representaciones son equivalentes. Hemos elegido estos generadores porque es posible expresarlos componente a componente en términos del símbolo de Levi-Civita

$$(\theta_a)_{bc} = -i\epsilon_{abc} \tag{2.26}$$

cuyas formas explícitas son las siguientes

$$\theta_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$\theta_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$$

$$\theta_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notemos además que las matrices θ_a tienen las siguientes propiedades

$$[\theta_a, \theta_b] = i\epsilon_{abc}\theta_c \quad (2.27)$$

$$\theta_a\theta_a = 2\mathbb{I}_3. \quad (2.28)$$

Al hacer el producto $\phi_{(1)}(\boldsymbol{\delta}) \otimes \phi_{(2)}(\boldsymbol{\delta})$ esperamos encontrar una expresión para los generadores de ψ en términos de los ya conocidos θ_a y σ_a

$$\begin{aligned} \phi_{(1)}(\boldsymbol{\delta}) \otimes \phi_{(2)}(\boldsymbol{\delta}) &= (\mathbb{I}_3 + i\delta_a\theta_a + \cdots) \otimes \left(\mathbb{I}_2 + i\delta_b\frac{\sigma_b}{2} + \cdots \right) \\ &= \mathbb{I}_3 \otimes \mathbb{I}_2 + i\delta_a\theta_a \otimes \mathbb{I}_2 + \mathbb{I}_3 \otimes i\delta_b\frac{\sigma_b}{2} + \cdots \\ &= \mathbb{I}_{V \otimes W} + i\delta_a \left(\theta_a \otimes \mathbb{I}_2 + \mathbb{I}_3 \otimes \frac{\sigma_a}{2} \right) + \cdots \end{aligned}$$

entonces podemos identificar a los generadores de la representación $\psi_{(1)} \otimes \psi_{(1/2)}$ como

$$K_a = \theta_a \otimes \mathbb{I}_2 + \mathbb{I}_3 \otimes \frac{\sigma_a}{2} \quad (2.29)$$

cuyas componentes podemos calcular usando la definición 2.22 del producto tensorial y la ecuación (2.26)

$$\begin{aligned} (K_c)_{a\alpha, b\beta} &= (\theta_c)_{ab}(\mathbb{I}_2)_{\alpha\beta} + \frac{1}{2}(\mathbb{I}_3)_{ab}(\sigma_c)_{\alpha\beta} \\ (K_c)_{a\alpha, b\beta} &= -i\epsilon_{abc}\delta_{\alpha\beta} + \frac{1}{2}\delta_{ab}(\sigma_c)_{\alpha\beta}. \end{aligned} \quad (2.30)$$

Sugerimos la convención en donde las letras griegas toman los valores $\{1, 2\}$ y las latinas $\{1, 2, 3\}$.

Capítulo 3

Operador de Dirac sobre S^2

Una buena parte de las teorías modernas en la física, incluyen elementos de geometría diferencial, los cuales muchas veces un lector novel desconoce o entiende parcialmente; por esta razón se incluye la siguiente sección como preámbulo. Brevemente mostramos como obtenemos la conexión de espín sobre la esfera a partir de la cual se encuentran las ecuaciones que determinan la función de onda y consecuentemente el operador de Dirac sobre una esfera. En adelante usaremos las unidades naturales $\hbar = 1$ y $c = 1$; además la esfera que consideramos es de radio igual a 1.

3.1. Conexión

Introducimos varias definiciones de la geometría diferencial para establecer conceptos como campos vectoriales, derivada covariante, la conexión de Levi-Civita y conexión de espín. Para el desarrollo de las dos primeras partes nos hemos basado en [15], del cual hemos seleccionado cuidadosamente los mínimos conceptos necesarios. Un lector familiarizado con el tema bien puede omitir las secciones 3.1.1 y 3.1.2, y partir de 3.2.2 donde establecemos nuestra conexión de espín.

3.1.1. Conceptos básicos

Aunque nos limitemos a enunciar algunas definiciones, hacemos énfasis en la consulta de libros como [15, 16, 17], tanto para aquellos conceptos que requieren de una reflexión aparte, como de cualquier detalle técnico que haya

sido omitido.

Definición 3.1 Sean $\mathbf{X}_1$ y $\mathbf{X}_2$ espacios topológicos, se dice que una función $f : \mathbf{X}_1 \rightarrow \mathbf{X}_2$ es un homeomorfismo si f es continua y si su inversa $f^{-1} : \mathbf{X}_2 \rightarrow \mathbf{X}_1$ también lo es.

Se dice que $\mathbf{X}_1$ y $\mathbf{X}_2$ son *homeomorfos* si existe un homeomorfismo entre ellos.

Definición 3.2 $\mathcal{M}$ es una variedad diferenciable de dimensión m si satisface las siguientes condiciones:

1. $\mathcal{M}$ es un espacio topológico.
2. $\mathcal{M}$ tiene asociada una familia de pares $\{(U_i, \phi_i) : i \in I\}$ tal que $\{U_i\}$ es una cubierta abierta de $\mathcal{M}$, $\cup_i U_i = \mathcal{M}$; ϕ_i es un homeomorfismo entre U_i y un subconjunto abierto $U'_i \subset \mathbb{R}^m$.
3. Si $U_i \cap U_j \neq \emptyset$, la función $\psi_{ij} = \phi_i \circ \phi_j^{-1}$ que va de $\phi_j(U_i \cap U_j)$ a $\phi_i(U_i \cap U_j)$ es diferenciable.

En el caso en que las funciones ψ_{ij} son de clase C^n , se dice que la variedad es *n-diferenciable* y si son de clase C^∞ entonces se trata de una *variedad suave*. Nosotros estaremos restringidos al caso C^∞ y se debe asumir que nuestras variedades son suaves. Como se verá a lo largo de estas secciones, algunas definiciones están concatenadas y la diferenciabilidad de un objeto depende del orden de diferenciabilidad de la variedad. Varios autores se evitan estos conflictos técnicos exigiendo que las variedades con las que trabajan sean suaves y esta es una medida que adoptamos. La dimensión de una variedad la denotamos con la misma letra que se asigna a la variedad, pero escrita en minúsculas, salvo en casos especiales donde se indicará explícitamente su dimensión.

Podemos imaginar una variedad como un espacio topológico que en cada punto es parecido a $\mathbb{R}^m$ (dentro de una vecindad, aunque globalmente podrían ser muy diferentes). Para darnos una idea, usemos un ejemplo cotidiano: nuestro planeta, que al situarnos sobre un punto en su superficie nos da la impresión de que se trata más de un plano ($\mathbb{R}^2$), que de una esfera (S^2). El par (U_i, ϕ_i) es llamado *carta* o *mapa*; y a la familia $\{(U_i, \phi_i) : i \in I\}$ se le nombra *atlas*.

Ejemplo 3.3 El subconjunto $S^2 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$ de $\mathbb{R}^3$, es un espacio topológico con la topología relativa inducida por la topología usual de $\mathbb{R}^3$. Sea $\tilde{U}_i^+ = \{(x^1, x^2, x^3) \in \mathbb{R}^3 : x^i > 0\}$ y $\tilde{U}_i^- = \{(x^1, x^2, x^3) \in \mathbb{R}^3 : x^i < 0\}$, con $i \in \{1, 2, 3\}$; el conjunto $\{U_i^\pm : U_i^\pm = \tilde{U}_i^\pm \cap S^2\}$ es una cubierta abierta de S^2 . Definimos las funciones $\phi_i^\pm : U_i^\pm \rightarrow \mathbb{R}^2$ como las proyecciones al plano definido por $x^i = 0$, por ejemplo, $\phi_1^+(x^1, x^2, x^3) = (x^2, x^3)$. Las funciones $\phi_i^\pm$ son homeomorfismos entre $U_i^\pm$ y el disco abierto de radio 1 centrado en el origen. Entonces $\{(U_i^\pm, \phi_i^\pm) : i \in \{1, 2, 3\}\}$ es un atlas de S^2 . Siempre que $U_i^\pm \cap U_\pm^j \neq \emptyset$, las funciones formadas como $\phi_i^\pm \circ (\phi_j^\pm)^{-1}$ son de clase C^∞ , e.g. para $x^3 < 0$,

$$\begin{aligned} \phi_3^- \circ (\phi_1^+)^{-1} : (x^2, x^3) &\mapsto \phi_3^-((\phi_1^+)^{-1}(x^2, x^3)) \\ &= \phi_3^-((1 - (x^2)^2 - (x^3)^2)^{1/2}, x^2, x^3) \\ &= ((1 - (x^2)^2 - (x^3)^2)^{1/2}, x^2). \end{aligned}$$

Es por esto que S^2 es una variedad suave.

El uso de cartas y atlas es importante cuando se introducen conceptos en geometría diferencial, por ejemplo para definir la diferenciabilidad de una función f sobre la variedad.

Definición 3.4 Sean $\mathcal{M}$ y $\mathcal{N}$ variedades con cartas (U, ϕ) y (V, ψ) respectivamente. Dada una función $f : U \subset \mathcal{M} \rightarrow V \subset \mathcal{N}$, se dice que f es diferenciable si, al hacer uso de las funciones de coordenadas

$$\phi(p) = (x^1(p), x^2(p), \dots, x^m(p)) \quad (p \in \mathcal{M}),$$

$$\psi(f(p)) = (y^1(f(p)), y^2(f(p)), \dots, y^n(f(p))),$$

la función $\psi \circ f \circ \phi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ es diferenciable con respecto a todas sus entradas x^a . Cuando $\psi \circ f \circ \phi^{-1}$ es de clase C^∞ entonces se dice que f es suave; al conjunto de funciones suaves sobre $\mathcal{M}$ lo denotamos $\mathfrak{F}(\mathcal{M})$.¹

¹Los textos en inglés usan la palabra “map” o “mapping” como sinónimo de “function”; a pesar de la equivalencia algunos autores acostumbran hacer una distinción entre ellas, usando la segunda cuando el rango de la función es un campo, por ejemplo $\mathbb{R}$. Es por eso que el conjunto que denotamos como $\mathfrak{F}(\mathcal{M})$ aparecerá en la literatura refiriéndose a funciones $f : \mathcal{M} \rightarrow \mathbb{R}$ concretamente y esta es la interpretación que adoptamos.

Se debe prestar atención al uso de las funciones de coordenadas como puente para estudiar las propiedades de las funciones sobre variedades, ya que muchas veces se omite la presencia de dichas funciones (incluso se abusa de la notación) y es posible que un lector inexperto llegue a confundirse. En el entorno de las variedades, muchos conceptos se introducen como una extensión o analogía de la geometría y análisis en $\mathbb{R}^n$, e.g., el *vector tangente* a una curva.

Definición 3.5 Sea $\mathcal{M}$ una variedad diferenciable y c una curva diferenciable $c : \mathbb{R} \rightarrow \mathcal{M}$, tal que $c(0) = p$. Definimos entonces al vector tangente $\mathbf{V}$ a la curva c en p como²

$$\mathbf{V} \equiv \left. \frac{d}{dt}(x^a(c)) \right|_{c(t)=p} \left. \frac{\partial}{\partial x^a} \right|_{t=0}.$$

El espacio vectorial formado por los *vectores tangentes* a la variedad $\mathcal{M}$ en el punto p se le llama *espacio tangente* y lo denotamos como $T_p\mathcal{M}$. Los vectores base de dicho espacio son $\mathbf{e}_i \equiv \partial/\partial x^i$, también llamada *base de coordenadas*, de tal modo que $\dim(T_p\mathcal{M}) = \dim(\mathcal{M})$.³ Un vector de dicho espacio $\mathbf{V} = V^a \partial/\partial x^a$, puede operar sobre una función $f \in \mathfrak{F}(\mathcal{M})$, tal operación la denotamos como⁴ $\mathbf{V}[f] \equiv V^a \partial f / \partial x^a$.

Dado un espacio vectorial V se puede definir el llamado *espacio vectorial dual*, el cual denotamos como V^* y cuyos elementos $\mathbf{V}^*$ llamamos *vectores duales* o *covariantes*⁵. El espacio V^* tiene como base los vectores duales $\{\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^n\}$, donde $\mathbf{e}^1 \equiv \mathbf{e}_1^*$; cabe mencionar que en general V y V^* tienen la misma dimensión. Los elementos del espacio dual V^* se pueden interpretar como funciones escalares lineales sobre el espacio V y viceversa, los elementos de V son funciones escalares lineales sobre V^* . La operación entre vectores de dichos espacios se puede definir a través del *producto interno* $\langle \cdot, \cdot \rangle : V^* \times V \rightarrow \mathbb{R}$, de modo que

$$\langle \mathbf{e}^a, \mathbf{e}_b \rangle \equiv \delta^a_b. \quad (3.1)$$

²Nótese las funciones de coordenadas $\phi(c(t)) = (x^1(c(t)), x^2(c(t)), \dots, x^m(c(t)))$.

³Más adelante se puede entrever la razón de estas afirmaciones. En [15] se habla del isomorfismo entre los espacios $\mathbb{R}^m$ y $T_p\mathcal{M}$.

⁴Que en realidad se refiere a $\mathbf{V}[f] = V^a \partial(f \circ \phi^{-1}) / \partial x^a$. Es común omitir el señalamiento de que $\mathbf{V}[f]$ se evalúa en un punto $p \in \mathcal{M}$, como se muestra en la definición 3.5.

⁵En correspondencia, a los elementos de V se les llama vectores *contravariantes*.

En el caso del espacio $T_p\mathcal{M}$, su dual es el llamado *espacio cotangente* $T_p^*\mathcal{M}$; tomando en cuenta que los elementos de nuestra base de coordenadas tienen la forma $\mathbf{e}_a = \partial/\partial x^a$, es conveniente expresar los elementos de la base dual como $\mathbf{e}^a \equiv dx^a$ en concordancia con (3.1). Frecuentemente a los elementos de $T_p^*\mathcal{M}$ se les llama *1-formas*, las cuales se expresan como una combinación lineal de diferenciales totales de funciones, en términos de la base escribimos $\boldsymbol{\omega}^* = \omega_a dx^a$. A veces se omite el asterisco (*) para distinguir a los elementos del dual y en cambio se adopta la convención donde los índices abajo numeran las componentes y los índices arriba los elementos de la base dual; al contrario de los vectores contravariantes donde los índices abajo corresponden a la base y los índices arriba a las componentes. La forma en que $\boldsymbol{\omega} \in T_p^*\mathcal{M}$ y $\mathbf{V} \in T_p\mathcal{M}$ operan entre sí la expresamos explícitamente como

$$\langle \boldsymbol{\omega}, \mathbf{V} \rangle = V^a \omega_b \left\langle dx^b, \frac{\partial}{\partial x^a} \right\rangle = V^a \omega_b \delta^b_a = V^a \omega_a.$$

De manera inmediata definimos un *tensor de tipo* (r, s) , como una función multilinear que toma r elementos de V^* y s elementos de V , y nos regresa un elemento del campo⁶

$$\mathbf{T} : \bigotimes^r V^* \bigotimes^s V \rightarrow \mathbb{R}.$$

Al conjunto de este tipo de tensores lo llamamos *espacio tensorial* de tipo (r, s) y es denotado como T_s^r , de esta manera $T_0^1 = V$ y $T_1^0 = V^*$. En particular cuando el espacio vectorial en cuestión V es $T_p\mathcal{M}$, el espacio tensorial de tipo (r, s) lo denotamos como $T_{s,p}^r\mathcal{M}$, así un elemento de este espacio lo podemos escribir como

$$\mathbf{T} = T^{a_1 \dots a_r}_{b_1 \dots b_s} \frac{\partial}{\partial x^{a_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{a_r}} \otimes dx^{b_1} \otimes \dots \otimes dx^{b_s}.$$

Definición 3.6 *Un campo vectorial $\mathbf{X}$ sobre una variedad $\mathcal{M}$ es una función que asigna a cada punto de la variedad, un vector tangente $\mathbf{X}(p) \in T_p\mathcal{M}$. Si $\forall f \in \mathfrak{F}(\mathcal{M})$, se cumple que $\mathbf{X}[f] \in \mathfrak{F}(\mathcal{M})$, entonces decimos que $\mathbf{X}$ es un campo vectorial suave⁷. Al conjunto de campos vectoriales sobre la variedad lo denotamos como $\mathfrak{X}(\mathcal{M})$.*

⁶El álgebra de tensores y temas afines, se pueden consultar también en el libro de Wald [18].

⁷Si $\mathbf{X}$ es suave, entonces sus componentes son funciones suaves sobre la variedad.

Hablamos entonces de un *campo tensorial suave* de tipo (r, s) como una función suave que asigna a cada punto $p \in \mathcal{M}$, un elemento de $T_{s,p}^r \mathcal{M}$; al conjunto de campos tensoriales tipo (r, s) lo denotamos $\mathfrak{T}_s^r(\mathcal{M})$. En particular cuando se trata de n -formas escribimos $\Omega^n(\mathcal{M})$, que son los campos tensoriales de tipo $(0, n)$, que además tienen la característica que sus componentes son totalmente antisimétricas.

El siguiente paso en este glosario es definir una herramienta que nos permita aplicar las ideas del análisis a variedades diferenciables; la derivación de campos vectoriales es lo primero que tendríamos que considerar. En un espacio como $\mathbb{R}^n$ se define la derivada de un campo vectorial suave $\mathbf{Y}$, en un punto sobre una curva (suave) como la diferencia entre el campo vectorial evaluado en dicho punto y otro sobre la curva, en el límite en que estos puntos se acercan, es decir

$$\left. \frac{d\mathbf{Y}}{dt} \right|_{t_0} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{Y}(t_0 + \Delta t) - \mathbf{Y}(t_0)}{\Delta t}.$$

Es de lo más elemental tomar la diferencia entre dos vectores en $\mathbb{R}^n$, pero cuando intentamos llevar a cabo esta operación sobre una variedad, nos daremos cuenta que los vectores definidos sobre puntos distintos pertenecen a espacios ajenos y por lo tanto una operación entre ellos como la diferencia, no está propiamente definida. El siguiente método nos permitirá hacer a un lado el inconveniente anterior y nos dará una idea gráfica de varios conceptos que introduciremos.

Definición 3.7 Sean $\mathcal{M}$ y $\mathcal{N}$ variedades suaves, la función suave $f : \mathcal{M} \rightarrow \mathcal{N}$ induce en cada punto $p \in \mathcal{M}$, una función $f_* : T_p \mathcal{M} \rightarrow T_{f(p)} \mathcal{N}$ llamada “pushforward”⁸ y definida $\forall \mathbf{V} \in T_p \mathcal{M}$ y $\forall g \in \mathfrak{F}(\mathcal{N})$ por medio de

$$f_*(\mathbf{V})[g] = \mathbf{V}[g \circ f]. \quad (3.2)$$

Sean f y g como en la definición anterior, considere las cartas (U, ϕ) y (V, ψ) en $\mathcal{M}$ y $\mathcal{N}$ respectivamente, de modo que $\mathbf{y} = \psi(f(p))$ y $\mathbf{x} = \phi(p)$. Sean $\mathbf{V} \in T_p \mathcal{M}$ y $f_*(\mathbf{V}) = \mathbf{W} \in T_{f(p)} \mathcal{N}$, entonces haciendo uso de las funciones de coordenadas escribimos (3.2) como

$$f_*(\mathbf{V})[g \circ \psi^{-1}(\mathbf{y})] = \mathbf{V}[g \circ f \circ \phi^{-1}(\mathbf{x})]$$

⁸Este nombre en inglés se usa de manera estándar.

usando que $\mathbf{V} = V^a \partial / \partial x^a$ y $\mathbf{W} = W^\beta \partial / \partial y^\beta$, donde $a \in \{1, 2, \dots, m\}$ y $\beta \in \{1, 2, \dots, n\}$, obtenemos

$$W^\beta \frac{\partial}{\partial y^\beta} (g \circ \psi^{-1}(\mathbf{y})) = V^a \frac{\partial}{\partial x^a} (g \circ f \circ \phi^{-1}(\mathbf{x}))$$

al elegir $g = y^\beta$, tenemos $g \circ f \circ \phi^{-1}(\mathbf{x}) = y^\beta(\mathbf{x})$ y resulta

$$W^\beta = V^a \frac{\partial y^\beta}{\partial x^a}$$

y cuando $\mathbf{V} = \partial / \partial x^a$

$$f_* \left(\frac{\partial}{\partial x^a} \right) = \frac{\partial y^\beta}{\partial x^a} \frac{\partial}{\partial y^\beta}. \quad (3.3)$$

Los coeficientes del lado derecho de la ecuación (3.3) son las entradas de la matriz jacobiana de la función $\psi \circ f \circ \phi^{-1}$.

Definición 3.8 Sea $f : \mathcal{M} \rightarrow \mathcal{N}$ una función suave y $\dim(\mathcal{M}) \leq \dim(\mathcal{N})$. Decimos que f es una inmersión de $\mathcal{M}$ en $\mathcal{N}$, si f_* es inyectiva. Si además f es inyectiva, se dice que f es un encajamiento y la imagen $f(\mathcal{M})$ es una subvariedad de $\mathcal{N}$.

Pensemos en $\mathcal{M}$ como una subvariedad de $\mathbb{R}^n$ y observemos que un subconjunto abierto de una variedad, es una subvariedad con la misma dimensión. Tomemos una carta (U, ϕ) de $\mathcal{M}$, donde⁹ $\phi(U) = U' \subset \mathbb{R}^m$; la función inversa $\phi^{-1} : U' \rightarrow \mathbb{R}^n$ cuya imagen es U , la vemos como un encajamiento de U' en $\mathbb{R}^n$ y lo escribimos en términos de las coordenadas $\mathbf{u} = (u^1, u^2, \dots, u^m) \in U'$ como $\phi^{-1}(\mathbf{u}) = (f^1(\mathbf{u}), f^2(\mathbf{u}), \dots, f^n(\mathbf{u}))$. Sea¹⁰ $\phi_*^{-1} : T_{\phi(p)} \mathbb{R}^m \rightarrow T_p \mathcal{M} \subset T_p \mathbb{R}^n$ el pushforward inducido por ϕ^{-1} , entonces haciendo uso de la ecuación (3.3) aplicada a la base $\{\partial / \partial u^1, \partial / \partial u^2, \dots, \partial / \partial u^m\}$ mediante ϕ_*^{-1} , podemos expresar la base coordenada¹¹ $\mathbf{e}_a(p)$ de $T_p \mathcal{M}$ en términos de la base $\{\partial / \partial x^1, \partial / \partial x^2, \dots, \partial / \partial x^n\}$ de $T_p \mathbb{R}^n$, es decir

$$\mathbf{e}_a(p) = \phi_*^{-1} \left(\frac{\partial}{\partial u^a} \right) = \left(\frac{\partial f^\beta}{\partial u^a} \right) \frac{\partial}{\partial x^\beta}. \quad (3.4)$$

⁹El lector puede convencerse de que ϕ es un encajamiento, y por lo tanto U' es una subvariedad de $\mathbb{R}^m$.

¹⁰Notemos que si $U \subset \mathcal{M}$ y $p \in U$, entonces se identifican $T_p U$ y $T_p \mathcal{M}$, de forma natural. Similarmente sucede con $T_{\phi(p)} U'$ y $T_{\phi(p)} \mathbb{R}^m$.

¹¹Escribimos $\mathbf{e}_a(p)$ para hacer énfasis en que $\mathbf{e}_a$ son campos vectoriales, con los cuales se escriben los elementos de $\mathbf{Y} \in \mathfrak{X}(\mathcal{M})$, como $\mathbf{Y} = Y^a \mathbf{e}_a$.

Como ejemplo podemos considerar el caso que nos concierne, que es el de la esfera S^2 encajada en $\mathbb{R}^3$. Dado un atlas sobre la esfera, se puede construir para cada $p \in S^2$, una base de $T_p S^2$ identificando al espacio tangente como un subespacio en $T_p \mathbb{R}^3$.

Ejemplo 3.9 Para el caso de S^2 , recordemos el ejemplo 3.3 y tomemos la carta (U_1^+, ϕ_1^+) . Sea $\mathbf{u} \in \phi_1^+(U_1^+)$, entonces $(\phi_1^+)^{-1}(u^1, u^2) = ((1 - (u^1)^2 - (u^2)^2)^{1/2}, u^1, u^2)$ e identificamos $f^1(\mathbf{u}) = (1 - (u^1)^2 - (u^2)^2)^{1/2}$, $f^2(\mathbf{u}) = u^1$ y $f^3(\mathbf{u}) = u^2$, de la ecuación (3.4). Obtendremos entonces una base de $T_p S^2$ donde $p = (1/\sqrt{2}, 0, -1/\sqrt{2})$

$$\mathbf{e}_1(p) = \frac{\partial}{\partial x^2}$$

$$\mathbf{e}_2(p) = \frac{\partial}{\partial x^1} + \frac{\partial}{\partial x^3}.$$

Ahora podemos trabajar con los vectores tangentes a la variedad aprovechando la estructura de $\mathbb{R}^n$. Sea $\mathbf{Y}$ un campo vectorial en $\mathbb{R}^n$ y $c(t)$ una curva suave en $\mathcal{M}$, de modo que $\mathbf{Y}(t) = \mathbf{Y}(c(t)) \in \mathbb{R}^n$; entonces calculamos la derivada del campo vectorial a lo largo de la curva en la manera usual

$$\frac{d\mathbf{Y}}{dt} = \frac{dY^\beta}{dt} \frac{\partial}{\partial x^\beta}.$$

Recordemos que nuestro interés es definir una derivada que opere sobre los campos vectoriales de la variedad; por el momento, en el entorno de la variedad sólo son pertinentes los vectores tangentes (definición 3.6). En $\mathbb{R}^n$ los elementos de un campo vectorial (general) sobre $\mathcal{M}$, no son necesariamente vectores tangentes a $\mathcal{M}$. Por esta razón requerimos explícitamente que el campo vectorial $\mathbf{Y}$ sea *tangente* a $\mathcal{M}$, es decir $\mathbf{Y}(c(t)) \in T_{c(t)}\mathcal{M}$. Aunque este requisito no nos garantiza que $\mathbf{Y}'$ sea tangente, podemos hacer uso de un operador lineal π para proyectar $\mathbf{Y}'$ y obtener un campo vectorial tangente a $\mathcal{M}$, es decir

$$\pi \left(\frac{d\mathbf{Y}}{dt} \right) \Big|_{c(t)} \in T_{c(t)}\mathcal{M}. \quad (3.5)$$

Definición 3.10 La proyección $\pi(d\mathbf{Y}/dt)$, será nombrada como la derivada covariante del campo vectorial $\mathbf{Y}$ tangente a $\mathcal{M}$, a lo largo de $c(t)$ y la denotamos como $D\mathbf{Y}/dt$.

Proposición 3.11 Sean $\mathbf{Y}_1(t)$ y $\mathbf{Y}_2(t)$ campos vectoriales tangentes y suaves definidos sobre una curva en $\mathcal{M}$ y $\varphi(t) \in \mathfrak{F}(\mathcal{M})$. Entonces la derivada covariante a lo largo de la curva, tiene las siguientes propiedades:

1. $\frac{D}{dt}(\mathbf{Y}_1 + \mathbf{Y}_2) = \frac{D\mathbf{Y}_1}{dt} + \frac{D\mathbf{Y}_2}{dt}.$
2. $\frac{D}{dt}(\varphi\mathbf{Y}_1) = \frac{d\varphi}{dt}\mathbf{Y}_1 + \varphi\frac{D\mathbf{Y}_1}{dt}.$

Demostración:

1. $\begin{aligned} \frac{D}{dt}(\mathbf{Y}_1 + \mathbf{Y}_2) &= \pi\left(\frac{d}{dt}(\mathbf{Y}_1 + \mathbf{Y}_2)\right) = \pi\left(\frac{d\mathbf{Y}_1}{dt} + \frac{d\mathbf{Y}_2}{dt}\right) \\ &= \pi\left(\frac{d\mathbf{Y}_1}{dt}\right) + \pi\left(\frac{d\mathbf{Y}_2}{dt}\right) = \frac{D\mathbf{Y}_1}{dt} + \frac{D\mathbf{Y}_2}{dt}. \end{aligned}$
2. $\begin{aligned} \frac{D}{dt}(\varphi\mathbf{Y}_1) &= \pi\left(\frac{d}{dt}(\varphi\mathbf{Y}_1)\right) = \pi\left(\frac{d\varphi}{dt}\mathbf{Y}_1 + \varphi\frac{d\mathbf{Y}_1}{dt}\right) \\ &= \pi\left(\frac{d\varphi}{dt}\mathbf{Y}_1\right) + \pi\left(\varphi\frac{d\mathbf{Y}_1}{dt}\right) = \frac{d\varphi}{dt}\pi(\mathbf{Y}_1) + \varphi\pi\left(\frac{d\mathbf{Y}_1}{dt}\right) \\ &= \frac{d\varphi}{dt}\mathbf{Y}_1 + \varphi\frac{D\mathbf{Y}_1}{dt}. \blacksquare \end{aligned}$

Una vez que hemos conseguido construir una derivada, podemos hacer un poco más robusta nuestra definición. Consideremos un campo vectorial tangente a lo largo de una curva $c(t)$ sobre $\mathcal{M}$, escribimos entonces $\mathbf{Y}(t)$ en términos de la base de coordenadas $\mathbf{Y}(t) = Y^a(t)\mathbf{e}_a(t)$ y calculamos su derivada covariante según (3.5)

$$\begin{aligned} \frac{D\mathbf{Y}}{dt} &= \pi\left(\frac{d}{dt}(Y^a\mathbf{e}_a)\right) = \pi\left(\frac{dY^a}{dt}\mathbf{e}_a + Y^a\frac{d\mathbf{e}_a}{dt}\right). \\ &= \frac{dY^a}{dt}\mathbf{e}_a + Y^a\frac{D\mathbf{e}_a}{dt} \end{aligned} \tag{3.6}$$

Para calcular la derivada covariante de $\mathbf{e}_a$ hacemos uso de (3.4)¹²

$$\frac{D\mathbf{e}_a}{dt} = \pi\left(\frac{d}{dt}\left(\frac{\partial f^\beta}{\partial u^a}\frac{\partial}{\partial x^\beta}\right)\right) = \frac{d}{dt}\left(\frac{\partial f^\beta}{\partial u^a}\right)\pi\left(\frac{\partial}{\partial x^\beta}\right)$$

¹²Recordemos que f^β son las componentes del encajamiento de $\mathcal{M}$ en $\mathbb{R}^n$.

$$= \frac{du^c}{dt} \frac{\partial^2 f^\beta}{\partial u^c \partial u^a} \pi \left(\frac{\partial}{\partial x^\beta} \right) \quad (3.7)$$

Debe notarse que $\pi(\partial/\partial x^\beta)$ es un campo vectorial suave tangente a $\mathcal{M}$ y por lo tanto, lo podemos expresar como una combinación lineal de la base coordenada

$$\pi \left(\frac{\partial}{\partial x^\beta} \right) = A^d{}_\beta(u) \mathbf{e}_d.$$

Para usar la expresión anterior en (3.7), es conveniente hacer la siguiente definición

$$\Gamma^d{}_{ca} \equiv \frac{\partial^2 f^\beta}{\partial u^c \partial u^a} A^d{}_\beta \quad (3.8)$$

de modo que (3.7) se escribe como sigue

$$\frac{D\mathbf{e}_a}{dt} = \frac{du^c}{dt} \Gamma^d{}_{ca} \mathbf{e}_d$$

con la cual expresamos (3.6) como

$$\frac{D\mathbf{Y}}{dt} = \left(\frac{dY^a}{dt} + Y^d \frac{du^c}{dt} \Gamma^a{}_{cd} \right) \mathbf{e}_a. \quad (3.9)$$

Podemos hacer un poco más flexible la expresión para la derivada covariante (3.9), ya que únicamente está definida a lo largo de una curva $c(t)$ y nos sería de gran ayuda extenderla a toda la variedad. Sea $\mathbf{Y}$ un campo vectorial tangente a la variedad $\mathcal{M}$, sea p un punto de la variedad que identificamos mediante sus coordenadas locales $\phi(p) = (u_0^1, u_0^2, \dots, u_0^m)$. Pensemos en una curva que en $t = t_0$ pasa por p y en ese mismo punto, el vector tangente a la curva es $\mathbf{X}(p) \in T_p\mathcal{M}$, es decir

$$c(t_0) = p \quad \text{y} \quad \left(\frac{dc}{dt} \right) \Big|_{t_0} = \mathbf{X}(p);$$

podemos hacer uso de las coordenadas locales y escribir $c(t)$ como $\mathbf{u}(t) = (u^1(t), u^2(t), \dots, u^m(t))$ con $\mathbf{u}_0 = \mathbf{u}(t_0) = \phi(p)$ y $(du^a/dt)|_{t_0} = X^a$. Procedemos entonces a calcular

$$\left(\frac{dY^a}{dt} \right) \Big|_{t_0} = \left(\frac{du^b}{dt} \right) \Big|_{t_0} \left(\frac{\partial Y^a}{\partial u^b} \right) \Big|_{\mathbf{u}(t_0)} = \mathbf{X}(p)[Y^a]$$

y a reescribir la ecuación (3.9) evaluada en el punto p como

$$\left. \frac{D\mathbf{Y}}{dt} \right|_{t_0} = (\mathbf{X}(p)[Y^a] + Y^d(\mathbf{u}_0)X^c\Gamma_{cd}^a) \mathbf{e}_a(p). \quad (3.10)$$

Es necesario observar que la derivada covariante de $\mathbf{Y}$ ha quedado definida a través de un vector tangente, de modo que para cada vector de $T_p\mathcal{M}$ se puede asignar una derivada de $\mathbf{Y}$ según (3.10), es decir una función que toma $\mathbf{X}(p) \mapsto \nabla_{\mathbf{X}(p)}\mathbf{Y}$, en donde hemos introducido la notación estándar $\nabla_{\mathbf{X}(p)}\mathbf{Y} \equiv (D\mathbf{Y}/dt)|_{t_0}$. La expresión en (3.10) se interpreta como la razón de cambio del campo vectorial $\mathbf{Y}$ en p en la dirección de $\mathbf{X}(p)$, y cumple con la siguiente.

Definición 3.12 *Una conexión afín es una función $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \rightarrow \mathfrak{X}(\mathcal{M})$ tal que $\forall f \in \mathfrak{F}(\mathcal{M}), \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(\mathcal{M})$ y $\omega \in \Omega^1(\mathcal{M})$; $(\mathbf{X}, \mathbf{Y}) \mapsto \nabla_{\mathbf{X}}\mathbf{Y}$ cumple lo siguiente:*

1. $\nabla_{\mathbf{X}}(f\mathbf{Y} + \mathbf{Z}) = \mathbf{X}[f]\mathbf{Y} + f\nabla_{\mathbf{X}}\mathbf{Y} + \nabla_{\mathbf{X}}\mathbf{Z}.$
2. $\nabla_{(f\mathbf{X}+\mathbf{Y})}\mathbf{Z} = f\nabla_{\mathbf{X}}\mathbf{Z} + \nabla_{\mathbf{Y}}\mathbf{Z}.$

Aunque no forma parte de la definición, añadimos una propiedad más a la lista, cuya relevancia se mostrará en breve

3. $\nabla_{\mathbf{X}}\langle\omega, \mathbf{Y}\rangle = \langle\nabla_{\mathbf{X}}\omega, \mathbf{Y}\rangle + \langle\omega, \nabla_{\mathbf{X}}\mathbf{Y}\rangle.$

Veamos ahora como opera la conexión dados $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(\mathcal{M})$, calculamos recurriendo a las propiedades de la definición 3.12

$$\nabla_{\mathbf{X}}\mathbf{Y} = \nabla_{X^a\mathbf{e}_a}(Y^b\mathbf{e}_b) = X^a\mathbf{e}_a[Y^b]\mathbf{e}_b + X^aY^b\nabla_{\mathbf{e}_a}\mathbf{e}_b$$

y usando la ecuación (3.10) podemos escribir la expansión de $\nabla_{\mathbf{e}_a}\mathbf{e}_b$ en términos de la base

$$\nabla_a\mathbf{e}_b \equiv \nabla_{\mathbf{e}_a}\mathbf{e}_b = \Gamma_{ab}^c\mathbf{e}_c \quad (3.11)$$

donde los Γ_{ab}^c (definidos en (3.8)) son los *coeficientes de la conexión*, también llamados *símbolos de Christoffel*. Usando la expansión (3.11) podemos ver que

$$\nabla_{\mathbf{X}}\mathbf{Y} = X^a(\mathbf{e}_a[Y^b] + Y^c\Gamma_{ac}^b)\mathbf{e}_b$$

lo cual concuerda con la ecuación (3.10). La expresión que hemos encontrado es bastante útil, ya que nos permite ver como es que la conexión puede operar sobre todo tipo de tensores.

Ejemplo 3.13 *Calculemos los coeficientes de la conexión en el caso en que operan sobre 1-formas, es decir $\nabla_a dx^b$. Según la tercera propiedad de la conexión*

$$\begin{aligned}\nabla_a \left\langle dx^b, \frac{\partial}{\partial x^c} \right\rangle &= \left\langle \nabla_a dx^b, \frac{\partial}{\partial x^c} \right\rangle + \left\langle dx^b, \nabla_a \frac{\partial}{\partial x^c} \right\rangle \\ &= (\nabla_a dx^b)_d \left\langle dx^d, \frac{\partial}{\partial x^c} \right\rangle + \left(\nabla_a \frac{\partial}{\partial x^c} \right)^e \left\langle dx^b, \frac{\partial}{\partial x^e} \right\rangle \\ &= (\nabla_a dx^b)_d \delta^d_c + \Gamma^e_{ac} \delta^b_e = (\nabla_a dx^b)_c + \Gamma^b_{ac}\end{aligned}$$

por otro lado

$$\nabla_a \left\langle dx^b, \frac{\partial}{\partial x^c} \right\rangle = \frac{\partial}{\partial x^a} \left(\left\langle dx^b, \frac{\partial}{\partial x^c} \right\rangle \right) = \frac{\partial}{\partial x^a} (\delta^b_c) = 0$$

de modo que

$$\begin{aligned}(\nabla_a dx^b)_c + \Gamma^b_{ac} &= 0 \\ \therefore (\nabla_a dx^b)_c &= -\Gamma^b_{ac}.\end{aligned}$$

Del ejemplo anterior es claro que si $\omega \in \Omega^1(\mathcal{M})$ y $X \in \mathfrak{X}(\mathcal{M})$, entonces

$$\nabla_a dx^b = -\Gamma^b_{ac} dx^c \quad (3.12)$$

y

$$\nabla_X \omega = X^b \left(\frac{\partial \omega_a}{\partial x^b} - \Gamma^c_{ba} \omega_c \right) dx^a.$$

Usando las ecuaciones (3.11) y (3.12), se puede generalizar la derivada covariante al caso en que opera sobre campos tensoriales de tipo (r, s) . Para ello se postula como propiedad adicional la regla de Leibniz para el producto tensorial

$$\nabla_X (T_1 \otimes T_2) = \nabla_X (T_1) \otimes T_2 + T_1 \otimes \nabla_X (T_2). \quad (3.13)$$

De forma análoga al ejemplo 3.13 se establece la conexión que opera sobre elementos de $\mathfrak{T}_s^r(\mathcal{M})$. Como ejemplo mostramos el caso de tensores de tipo $(0, 2)$, los cuales son de especial interés.

Ejemplo 3.14 *A continuación calculamos los coeficientes de la conexión para 2-formas, el cual es directo a partir del resultado del ejemplo 3.13. Sea $\omega \in \mathfrak{T}_2^0(\mathcal{M})$, entonces*

$$\nabla_a \omega = \nabla_a (\omega_{bc} dx^b \otimes dx^c)$$

$$\begin{aligned}
&= (\nabla_a \omega_{bc}) dx^b \otimes dx^c + \omega_{bc} (\nabla_a dx^b) \otimes dx^c + \omega_{bc} dx^b \otimes (\nabla_a dx^c) \\
&= \frac{\partial \omega_{bc}}{\partial x^a} dx^b \otimes dx^c + \omega_{bc} (-\Gamma_{ad}^b) dx^d \otimes dx^c + \omega_{bc} dx^b \otimes (-\Gamma_{ae}^c) dx^e \\
&\therefore (\nabla_a \omega)_{bc} = \frac{\partial \omega_{bc}}{\partial x^a} - \omega_{dc} \Gamma_{ab}^d - \omega_{be} \Gamma_{ac}^e.
\end{aligned}$$

3.1.2. Levi-Civita

Definición 3.15 Sea $\mathbf{g} \in \mathfrak{T}_2^0$, se dice que $\mathbf{g}$ es una métrica riemanniana, si en cada punto $p \in \mathcal{M}$ y $\forall \mathbf{V}, \mathbf{V}_1, \mathbf{V}_2 \in T_p \mathcal{M}$, $\mathbf{g}$ cumple con las siguientes propiedades (para cada punto p en la variedad)

1. $\mathbf{g}_p(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{g}_p(\mathbf{V}_2, \mathbf{V}_1)$.
2. $\mathbf{g}_p(\mathbf{V}, \mathbf{V}) \geq 0$, $(\mathbf{g}_p(\mathbf{V}, \mathbf{V}) = 0 \Leftrightarrow \mathbf{V} = \mathbf{0})$.

Cuando una variedad admite una métrica riemanniana, el par $(\mathcal{M}, \mathbf{g})$ es llamado *variedad riemanniana*. Se dice que un campo vectorial $\mathbf{Y}$ se *transporta paralelamente* sobre una curva, si sobre dicha curva su derivada covariante es cero, es decir $\nabla_{\mathbf{X}} \mathbf{Y} = 0$, donde $\mathbf{X}$ es el campo vectorial tangente a la curva. A través de la métrica se induce un producto interno en $T_p \mathcal{M}$ definido para $\mathbf{V}_1, \mathbf{V}_2 \in T_p \mathcal{M}$, como

$$\mathbf{g}_p(\mathbf{V}_1, \mathbf{V}_2) \equiv \mathbf{V}_1 \cdot \mathbf{V}_2.$$

Una métrica es *covariantemente constante* si para cualquier par de vectores $\mathbf{X}$ y $\mathbf{Y}$, su producto interno se mantiene constante bajo transporte paralelo.

Sea $\mathbf{g}$ una métrica covariantemente constante y $\mathbf{V}$ un vector tangente a la curva sobre la que se transportan paralelamente $\mathbf{X}$ y $\mathbf{Y}$. Entonces dado que el producto interno es covariantemente constante, obtenemos $\nabla_{\mathbf{V}} \mathbf{g}(\mathbf{X}, \mathbf{Y}) = 0$ y por otro lado

$$\begin{aligned}
\nabla_{\mathbf{V}} \mathbf{g}(\mathbf{X}, \mathbf{Y}) &= V^a [(\nabla_a \mathbf{g})(\mathbf{X}, \mathbf{Y}) + \mathbf{g}(\nabla_a \mathbf{X}, \mathbf{Y}) + \mathbf{g}(\mathbf{X}, \nabla_a \mathbf{Y})] \\
&= V^a [(\nabla_a \mathbf{g})(\mathbf{X}, \mathbf{Y}) + \mathbf{g}(\mathbf{0}, \mathbf{Y}) + \mathbf{g}(\mathbf{X}, \mathbf{0})] \\
&= V^a (\nabla_a \mathbf{g})(\mathbf{X}, \mathbf{Y}) = V^a X^b Y^c (\nabla_a \mathbf{g})_{bc}
\end{aligned}$$

finalmente resulta que

$$V^a X^b Y^c (\nabla_a \mathbf{g})_{bc} = 0.$$

Al tomar en cuenta que la anterior es válida para todos los vectores $\mathbf{X}$, $\mathbf{Y}$ y $\mathbf{V}$, que cumplen las condiciones, concluimos que

$$(\nabla_a \mathbf{g})_{bc} = 0$$

según el ejemplo 3.14, lo anterior se puede escribir como

$$\frac{\partial g_{bc}}{\partial x^a} - g_{dc} \Gamma_{ab}^d - g_{be} \Gamma_{ac}^e = 0. \quad (3.14)$$

Cuando una métrica satisface la ecuación (3.14) se dice que la conexión es *compatible con la métrica* o se le llama *conexión métrica*. Cuando además el *tensor de torsión* definido como

$$T_{bc}^a = \Gamma_{bc}^a - \Gamma_{cb}^a$$

es nulo, la conexión lleva el nombre de *conexión de Levi-Civita*. Cuando en una conexión el tensor de torsión es nulo, resulta $\Gamma_{bc}^a = \Gamma_{cb}^a$, es por esto que también se le llama *conexión simétrica*.

Teorema 3.16 (fundamental de la geometría riemanniana.) *En una variedad riemanniana $(\mathcal{M}, \mathbf{g})$ existe una conexión única que es simétrica y compatible con la métrica $\mathbf{g}$. Dicha conexión es la de Levi-Civita.*

Demostración: La demostración se puede consultar en el libro de Nakahara [16], que es de donde se ha tomado el enunciado del teorema. ■

3.2. Conexión de espín en la esfera

Básicamente lo que haremos a continuación es buscar un objeto que reproduce la noción de la conexión de Levi-Civita, en el caso de la variedad S^2 . También será necesario tomar en cuenta que los objetos sobre los que opera dicha conexión, están dotados de un carácter geométrico propio de los espinores. También haremos uso de la idea o método, en el que se construye la derivada covariante como una proyección (sobre la variedad) de una derivada total. En parte la intención en esta sección es, simplificar y poner en práctica estos conceptos, que inevitablemente están contacto con el problema central.

3.2.1. SU(2) y el momento angular

En mecánica cuántica se eligen los operadores de momento angular como los *generadores de rotaciones*, es decir producen los operadores (denotados como $D(R)$) que rotan los vectores estado de un sistema. Como Sakurai menciona [7], las rotaciones afectan a los sistemas físicos, y es de esperarse que el ket asociado a un sistema rotado $|\beta\rangle_R$, difiera del ket asociado al mismo sistema sin rotación¹³ $|\beta\rangle$. Los operadores $D(R)$ se caracterizan a través de una rotación general R en $\mathbb{R}^3$, dicha rotación se puede determinar de varias formas, como las *rotaciones de Euler*, que son una sucesión de rotaciones $R(\alpha, \beta, \gamma) = R_z(\gamma)R_y(\beta)R_z(\alpha)$; también podemos denotar una rotación con un vector (unitario) que establece el eje de rotación y el ángulo total del giro, $R(\hat{n}, \alpha)$. Notemos además que las rotaciones en $\mathbb{R}^3$ son elementos del grupo $SO(3)$, el cual puede verse como una representación (no fiel) de $SU(2)$; vistos como grupos de Lie, $SO(3)$ es *localmente isomorfo* a $SU(2)$. Por lo anterior, podemos elegir entre los grupos $SU(2)$ y $SO(3)$ para trabajar con rotaciones; más aún, se establece la equivalencia entre las relaciones de conmutación, tanto de los operadores de momento angular, como de los generadores de $SU(2)$ (compare entre sí las ecuaciones (1.33), (1.34) y (2.12)). Es en este punto donde las representaciones de $SU(2)$ encajan en la mecánica cuántica, ya que la teoría del momento angular se desarrolla a partir de las relaciones fundamentales de conmutación, y todo lo que hemos desarrollado en el capítulo anterior se puede implementar directamente en mecánica cuántica, basta con hacer una interpretación de las ecuaciones para adaptarlas al contexto físico.

Establecemos entonces a los generadores del grupo, como los operadores de momento angular J_a , que satisfacen en general

$$[J_a, J_b] = i\hbar\epsilon_{abc}J_c. \quad (3.15)$$

A través de estos operadores se obtienen las *rotaciones infinitesimales*

$$D(\hat{n}, \delta\alpha) \equiv \mathbb{I} - \frac{i}{\hbar} \mathbf{J} \cdot \hat{n} \delta\alpha \quad (3.16)$$

comparando esta última con la ecuación (2.6), podemos observar que hay una diferencia en el signo del segundo término, esta diferencia se atribuye al sentido en que se realiza la rotación. Como puede apreciarse en las ecuaciones (1.31) y (1.32), el factor $\hbar^{-1}$ se anula con el factor $\hbar$ de los operadores de

¹³El operador $D(R)$ conecta estos kets a través de $|\beta\rangle_R = D(R) |\beta\rangle$.

momento angular¹⁴. También escribimos la ecuación correspondiente a la (2.8) para *operador de rotación*

$$D(\hat{\mathbf{n}}, \alpha) = \exp(-i\mathbf{J} \cdot \hat{\mathbf{n}}\alpha). \quad (3.17)$$

Los demás operadores ($\mathbf{J}^2$, J_+ y J_-) se construyen en completa analogía y todo el desarrollo con respecto a los eigenvectores y su espectro es inmediato. Tendremos entonces que el eigenket común $|j, m\rangle$ de J_0 y $\mathbf{J}^2$, cumple

$$J_0 |j, m\rangle = m |j, m\rangle$$

$$\mathbf{J}^2 |j, m\rangle = j(j+1) |j, m\rangle$$

compárese con las ecuaciones (2.15) y (2.16).

Los operadores J_a pueden representar el momento angular orbital, el momento angular espinorial o una combinación de ambos dada por un producto tensorial. Las distintas contribuciones al momento angular de un sistema, son añadidas a través del producto tensorial, e.g. un sistema que consta de una partícula con espín $1/2$, en el que se toman en cuenta los grados de libertad internos (momento angular espinorial) y los grados de libertad externos (momento angular orbital), tendrá un *momento angular total*

$$\mathbf{J} \equiv \mathcal{L} \otimes \mathbb{I}_S + \mathbb{I}_L \otimes \mathbf{S}. \quad (3.18)$$

que comúnmente se escribe como $\mathbf{J} = \mathcal{L} + \mathbf{S}$.

Las representaciones ψ_j de $SU(2)$ corresponden a sistemas de espín j . Compare la ecuación anterior con la (2.29), que por ejemplo, corresponde al generador del producto de las representaciones con $j = 1$ y $j = 1/2$. Como ya se había comentado en la sección 1.2, una de las diferencias entre el momento angular orbital y el momento angular espinorial, es que para el primero sus representaciones están determinadas por valores enteros de ℓ y en el segundo se permiten además valores semienteros para s . Una distinción más se puede hacer al notar que, los operadores de momento angular orbital se pueden representar en el espacio de posiciones $\{|\mathbf{x}\rangle\}$, cosa que es imposible para el momento angular espinorial; por esta misma razón, los operadores de momento angular orbital pueden tomar la forma de operadores diferenciales, como se puede observar en la ecuación (1.31). En el ejemplo que hemos dado

¹⁴En adelante no aparecerá puesto que tomamos $\hbar = 1$.

en (3.18), podemos tomar como base, el producto tensorial de la base $\{|\mathbf{x}\rangle\}$ y la de espín $\{|+\rangle, |-\rangle\}$, como sigue

$$|\mathbf{x}, \pm\rangle = |\mathbf{x}\rangle \otimes |\pm\rangle. \quad (3.19)$$

En este mismo contexto, dado un estado arbitrario de la partícula $|\beta\rangle$, escribimos su función de onda como

$$\langle \mathbf{x}, \pm | \beta \rangle = \Psi_{\pm}(\mathbf{x}) \quad (3.20)$$

que también escribimos en forma de columna (como acostumbramos de la sección 1.3)

$$\Psi = \begin{pmatrix} \Psi_+(\mathbf{x}) \\ \Psi_-(\mathbf{x}) \end{pmatrix}. \quad (3.21)$$

Decimos que Ψ es un espinor, porque sus componentes Ψ_+ y Ψ_- , se transforman ante rotaciones según la representación $j = 1/2$. Una característica que los distingue de los vectores, es que ante una rotación de 2π , los vectores no cambian, situación que no sucede con los espinores, los cuales invierten su signo. Esto se puede consultar en las ecuaciones (2.9) y (2.10).

Para un estudiante inexperto puede causar cierta confusión mencionar que el operador de momento angular orbital se escribe en términos de la base $\{|\mathbf{x}\rangle\}$ (un espacio de representación de dimensión infinita), cuando ya hemos mencionado que podemos usar cualquier resultado del capítulo 2, que incluye la técnica para encontrar la base de eigenkets $\{|j, m\rangle\}$, donde el espacio de representación es de dimensión finita¹⁵. En mecánica cuántica frecuentemente se recurre a distintas representaciones de un mismo operador, ya que pueden verse reflejados aspectos diferentes en cada una de ellas; siempre se debe poner atención a este tipo de técnicas, porque se convierten en un recurso estándar en muchas áreas de la física teórica. Quien ya esté familiarizado con la literatura se preguntará ¿por qué hemos hecho el abordaje de la teoría del momento angular, de esta forma? En parte nuestra intención es mostrarle al lector, que dentro de una teoría, que acostumbra desprenderse de la intuición, es necesario ver un poco hacia el fondo y estudiar los andamios matemáticos que sustentan aquello, que a primera vista, puede parecer artificial o una casualidad. Además ha sido necesario introducir la teoría de representaciones por su papel central en el método de Cartan.

¹⁵En el apéndice B, se puede ver un ejemplo de esto que mencionamos.

3.2.2. Conexión de espín en la esfera

En la sección 3.1 describimos una forma para obtener la conexión a partir de la proyección de la derivada total

$$\nabla = \mathcal{P}\partial.$$

Nosotros estamos interesados en particular, en la conexión de espín en la esfera; los teoremas de existencia y unicidad nos garantizan que nuestra búsqueda no será en vano.

Recordemos que las derivadas están relacionadas con la noción de cambios infinitesimales, por ejemplo, en un espacio como $\mathbb{R}^3$, estos cambios están dados por traslaciones infinitesimales. Un cambio infinitesimal sobre puntos de una esfera, está dado por rotaciones infinitesimales. Al buscar la derivada de un objeto como la función de onda de la ecuación (3.21), es necesario notar previamente las partes que la componen, ya que cada una de estas partes se transformará (bajo rotaciones), según la representación que le corresponde. Una de estas partes es la variable $\mathbf{x}'$, que está asociada al espacio de posiciones en el que se representa el momento angular orbital; la otra parte $\Psi_{\pm}$ son componentes de la función de onda, correspondientes a la parte del momento angular espinorial.

Ejemplo 3.17 *Veamos como los operadores $\mathcal{L}_a$ producen rotaciones infinitesimales. Según (3.16), tenemos¹⁶*

$$\begin{aligned} D(\hat{n}, \delta\alpha) &= \mathbb{I} - i(\delta\alpha)_a \mathcal{L}_a = \mathbb{I} - i\epsilon_{abc}(\delta\alpha)_a X_b P_c \\ &= \mathbb{I} - i\epsilon_{abc}(\delta\alpha)_a P_c X_b \end{aligned}$$

de modo que¹⁷

$$\begin{aligned} D(\hat{n}, \delta\alpha) |x, y, z\rangle &= (\mathbb{I} - i\epsilon_{abc}(\delta\alpha)_a P_c X_b) |x, y, z\rangle \\ &= (\mathbb{I} - i\epsilon_{abc}(\delta\alpha)_a P_c x_b) |x, y, z\rangle \\ &= |x + \epsilon_{ab1}x_b(\delta\alpha)_a, y + \epsilon_{ab2}x_b(\delta\alpha)_a, z + \epsilon_{ab3}x_b(\delta\alpha)_a\rangle. \end{aligned}$$

¹⁶Los operadores X_a y P_b cumplen con la siguiente $[X_a, P_b] = i\delta_{ab}$

¹⁷Hemos usado que el operador de traslaciones $T(\delta\mathbf{x}) = \mathbb{I} - i\mathbf{P} \cdot \delta\mathbf{x}$. Nótese que las traslaciones son generadas por el operador de momento lineal, $T(\delta\mathbf{x})|\mathbf{x}\rangle = |\mathbf{x} + \delta\mathbf{x}\rangle$.

En particular, una rotación infinitesimal alrededor del eje z ($(\delta\alpha)_1 = (\delta\alpha)_2 = 0$) se ve como

$$D(\hat{k}, \delta\alpha) |x, y, z\rangle = |x - y\delta\alpha, y + x\delta\alpha, z\rangle$$

donde el ket restante es claramente el resultado de la transformación debida a una rotación infinitesimal (positiva, levógira o activa) alrededor del eje z .

Nos interesa ver como se transforma la función de onda, ante rotaciones infinitesimales en su argumento $\Psi(\mathbf{x}) \rightarrow \Psi(\mathbf{x}')$, para averiguarlo observemos lo siguiente, si $\Psi(\mathbf{x}) = \langle x, y, z | \beta \rangle$, entonces

$$\begin{aligned} \Psi(\mathbf{x}') &= \langle x, y, z | (\mathbb{I} - i\epsilon_{abc}(\delta\alpha)_a X_b P_c) | \beta \rangle \\ &= \langle x, y, z | \beta \rangle - i\epsilon_{abc}x_b(\delta\alpha)_a \langle x, y, z | P_c | \beta \rangle \\ &= \langle x, y, z | \beta \rangle - \epsilon_{abc}x_b(\delta\alpha)_a \partial_c \langle x, y, z | \beta \rangle \\ \Psi(\mathbf{x}') &= \Psi(\mathbf{x}) - \epsilon_{abc}x_b(\delta\alpha)_a \partial_c \Psi(\mathbf{x}). \end{aligned} \quad (3.22)$$

La transformación infinitesimal de las componentes de Ψ está dada por la ecuación (3.16) para el caso de espinores, es decir $\mathbf{J} = \mathbf{S}$

$$\Psi'(\mathbf{x}) = \left(\mathbb{I}_2 - \frac{i}{2} \delta\alpha_d \sigma_d \right) \Psi(\mathbf{x}).$$

Podemos ahora calcular el cambio total en el espinor $\Psi(\mathbf{x})$ debido a rotaciones infinitesimales (despreciando los términos de orden cuadrático en $\delta\alpha$)

$$\begin{aligned} \Delta\Psi(\mathbf{x}) &\equiv \Psi'(\mathbf{x}') - \Psi(\mathbf{x}) \\ &= \left(\mathbb{I}_2 - \frac{i}{2} \delta\alpha_d \sigma_d \right) \Psi(\mathbf{x}') - \Psi(\mathbf{x}) \\ &= \left(\mathbb{I}_2 - \frac{i}{2} \delta\alpha_d \sigma_d \right) (\Psi(\mathbf{x}) - \epsilon_{abc}x_b \delta\alpha_a \partial_c \Psi(\mathbf{x})) - \Psi(\mathbf{x}) \\ &= \Psi(\mathbf{x}) - \frac{i}{2} \delta\alpha_d \sigma_d \Psi(\mathbf{x}) - \epsilon_{abc}x_b \delta\alpha_a \partial_c \Psi(\mathbf{x}) - \Psi(\mathbf{x}) \\ &= -\frac{i}{2} \delta\alpha_d \sigma_d \Psi(\mathbf{x}) - \epsilon_{abc}x_b \delta\alpha_a \partial_c \Psi(\mathbf{x}) \\ \Delta\Psi(\mathbf{x}) &= -i\delta\alpha_a \left(-i\epsilon_{abc}x_b \partial_c + \frac{\sigma_a}{2} \right) \Psi(\mathbf{x}). \end{aligned} \quad (3.23)$$

Recordando la forma de los operadores de momento angular orbital (1.31), y el de espín (1.32), podemos comparar con (3.23) y darnos cuenta que efectivamente, los operadores de la ecuación (3.18) son los que generan los cambios de nuestra función de onda, bajo rotaciones infinitesimales. Es decir, ya conocemos la parte que corresponde a la derivada total ($\partial = \mathcal{L} + \boldsymbol{\sigma}/2$), falta ahora trabajar la proyección a la esfera ($\mathcal{P}$).

Consideremos una esfera de radio 1 centrada en el origen, sobre algún punto $\mathbf{x}$ de la esfera tenemos un vector $\mathbf{v}$ apuntando en una dirección arbitraria, queremos encontrar el operador que proyecta $\mathbf{v}$ a su componente tangencial a la esfera $\mathbf{v}_\tau$. La componente tangencial la obtendremos restándole al vector su parte radial $\mathbf{v}_\rho$

$$(\mathbf{v}_\tau)_a = v_a - (\mathbf{v}_\rho)_a \quad (3.24)$$

la parte radial la obtenemos con el producto punto entre $\mathbf{v}$ y $\mathbf{x}$

$$\mathbf{v}_\rho = (\mathbf{x} \cdot \mathbf{v})\mathbf{x}$$

así podremos reescribir (3.24) como

$$(\mathbf{v}_\tau)_a = v_a - (\mathbf{x} \cdot \mathbf{v})x_a = (\delta_{ab} - x_ax_b)v_b = \mathcal{P}_{ab}v_b.$$

Donde $\mathcal{P}_{ab} = \delta_{ab} - x_ax_b$ es el operador de proyección que buscamos; para comprobar que efectivamente $\mathcal{P}$ es una proyección, veamos que satisface $\mathcal{P}^2 = \mathcal{P}$

$$\begin{aligned} \mathcal{P}^2 \mathbf{v} &= \mathcal{P}(\mathcal{P}\mathbf{v}) = \mathcal{P}(\mathbf{v} - (\mathbf{x} \cdot \mathbf{v})\mathbf{x}) \\ &= \mathbf{v} - (\mathbf{x} \cdot \mathbf{v})\mathbf{x} - (\mathbf{x} \cdot \mathbf{v})\mathbf{x} + (\mathbf{x} \cdot \mathbf{v})(\mathbf{x} \cdot \mathbf{x})\mathbf{x} \\ &= \mathbf{v} - (\mathbf{x} \cdot \mathbf{v})\mathbf{x} = \mathcal{P}\mathbf{v}. \end{aligned}$$

La conexión que buscamos, será entonces $\nabla = \mathcal{P}(\mathcal{L} + \boldsymbol{\sigma}/2)$. Aprovechando el siguiente resultado

$$\begin{aligned} \mathcal{P}_{ab}\mathcal{L}_b &= (\delta_{ab} - x_ax_b)\mathcal{L}_b = \mathcal{L}_a - x_ax_b\mathcal{L}_b \\ &= \mathcal{L}_a + i\epsilon_{bcd}x_ax_bx_c\partial_d = \mathcal{L}_a \end{aligned}$$

podemos escribir la conexión de espín como $\nabla = \mathcal{L} + \mathcal{P}\boldsymbol{\sigma}/2$, cuyas componentes están dadas por

$$\nabla_a = \mathcal{L}_a + \frac{1}{2}\mathcal{P}_{ab}\sigma_b. \quad (3.25)$$

3.3. Operador de Dirac en la esfera

3.3.1. Método de Cartan

La forma en que actúa la conexión ∇ sobre la función de onda Ψ , la podemos expresar componente a componente

$$\nabla_a \Psi_\alpha = \mathcal{L}_a \Psi_\alpha + \frac{1}{2} \mathcal{P}_{ab} (\sigma_b)_\alpha^\beta \Psi_\beta \quad (3.26)$$

recordemos que el subíndice $a \in \{1, 2, 3\}$ y $\alpha \in \{+, -\}$ (o $\alpha \in 1, 2$ según convenga), como consecuencia el objeto $\nabla_a \Psi_\alpha$ consta de 6 componentes. La parte ∇_a se transforma de acuerdo a la representación $\psi_{(1)}$, y Ψ_α según $\psi_{1/2}$, por lo tanto $\nabla_a \Psi_\alpha$ transforma bajo las reglas de la representación $\psi_{(1)} \otimes \psi_{(1/2)}$. Del teorema de Clebsch-Gordan (2.39) sabemos que $\psi_{(1)} \otimes \psi_{(1/2)} = \psi_{(1/2)} \oplus \psi_{(3/2)}$, donde es notable la presencia de la representación $\psi_{(1/2)}$, la cual sabemos está asociada a los espinores. Nos interesa extraer estas componentes espinoriales de $\nabla_a \Psi_\alpha$, y para lograrlo será necesario construir un operador de proyección sobre las representaciones, es decir $\Pi : \psi_{(1)} \otimes \psi_{(1/2)} \rightarrow \psi_{(1/2)}$.¹⁸ Precisamente este es el método que usa Cartan en su libro "The theory of spinors" [19], para construir el operador de Dirac de los espacios euclidiano y de Minkowski. Una diferencia entre nuestro trabajo y el de Cartan, es que él plantea los tensores en términos de polinomios. Calcula la derivada de un espinor, que es un polinomio de dos variables, las cuales se transforman de acuerdo a la representación fundamental de $SU(2) \times SU(2)$ (o el grupo de Lorentz) y lo descompone en representaciones irreducibles, finalmente proyecta sobre la parte de espín 1/2 y obtiene la ecuación de Dirac.

Para calcular dicho proyector, será necesario calcular el operador de Casimir (C) de la representación $\psi_{(1)} \otimes \psi_{(1/2)}$ (lema 2.27), usando los generadores descritos en la ecuación (2.29)

$$\begin{aligned} C &= K_a K_a = \left(\theta_a \otimes \mathbb{I}_2 + \mathbb{I}_3 \otimes \frac{\sigma_a}{2} \right)^2 \\ &= \theta_a \theta_a \otimes \mathbb{I}_2 + 2 \left(\theta_a \otimes \frac{\sigma_a}{2} \right) + \mathbb{I}_3 \otimes \frac{\sigma_a \sigma_a}{2} \\ &= 2\mathbb{I}_3 \otimes \mathbb{I}_2 + \theta_a \otimes \sigma_a + \mathbb{I}_3 \otimes \frac{3}{4} \mathbb{I}_2 \end{aligned}$$

¹⁸La parte que estamos eliminando $\psi_{(3/2)}$, es la correspondiente al objeto del tipo Rarita-Schwinger en la esfera.

$$C = \frac{11}{4}(\mathbb{I}_3 \otimes \mathbb{I}_2) + \theta_a \otimes \sigma_a$$

cuyas componentes están dadas por

$$C_{a\alpha;b\beta} = \frac{11}{4}\delta_{ab}\delta_{\alpha\beta} - i\epsilon_{abc}(\sigma_c)_{\alpha\beta} \quad (3.27)$$

De tal modo que C se puede diagonalizar como

$$C = \begin{pmatrix} \frac{15}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{15}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{15}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{15}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{4} \end{pmatrix}.$$

En la forma matricial de C , es claro que los primeros 4 términos en la diagonal, corresponden a la parte de la representación $\psi_{(3/2)}$, que es la parte que deseamos eliminar, pues recordemos que para $j = 3/2$, $j(j+1) = 15/4$. Entonces

$$\begin{aligned} \Pi &= \frac{C - \frac{15}{4}\mathbb{I}_6}{\frac{3}{4} - \frac{15}{4}} = -\frac{1}{3} \left(\left(\frac{11}{4} - \frac{15}{4} \right) \mathbb{I}_6 + \theta_a \otimes \sigma_a \right) \\ &= \frac{1}{3} (\mathbb{I}_6 - \theta_a \otimes \sigma_a) \end{aligned}$$

usando la expresión (3.27), vemos que sus componentes (en la base que diagonaliza C) son

$$\Pi_{a\alpha;b\beta} = \frac{1}{3}(\delta_{ab}\delta_{\alpha\beta} + i\epsilon_{abc}(\sigma_c)_{\alpha\beta}). \quad (3.28)$$

Antes de aplicar el proyector Π a $\nabla\Psi$, introducimos la siguiente notación para auxiliarnos un poco en los cálculos

$$(\eta_a)_{\alpha\beta} \equiv \frac{1}{2}\mathcal{P}_{ab}(\sigma_b)_{\alpha\beta} \quad (3.29)$$

y reescribimos (3.26) como

$$\nabla_a \Psi_\alpha = \mathcal{L}_a \Psi_\alpha + (\eta_a \Psi)_\alpha. \quad (3.30)$$

Imponemos la condición

$$\Pi \nabla \Psi = 0 \quad (3.31)$$

que corresponde físicamente a la ecuación que describe un fermión libre sin masa. En el caso general, la interpretación de la ecuación de Dirac en términos de representaciones, consiste en imponer que la componente espinorial de $\nabla\Psi$ es proporcional a Ψ , donde la constante de proporcionalidad es la masa del fermión.

Hacemos los cálculos de la ecuación (3.31) por medio de sus componentes, para ello haremos uso de las expresiones (3.28) y (3.31)

$$\frac{1}{3}(\delta_{ab}\delta_{\alpha\beta} + i\epsilon_{abc}(\sigma_c)_{\alpha\beta})(\mathcal{L}_b\Psi_\beta + (\eta_b\Psi)_\beta) = 0$$

es decir

$$\mathcal{L}_a\Psi_\alpha + (\eta_a\Psi)_\alpha + i\epsilon_{abc}(\sigma_c)_{\alpha\beta}[\mathcal{L}_b\Psi_\beta + (\eta_b\Psi)_\beta] = 0. \quad (3.32)$$

De la expresión anterior resultarán 6 ecuaciones, calcularemos el caso en que $a = 1$ y $\alpha = 1$

$$\begin{aligned} & \mathcal{L}_1\Psi_1 + (\eta_1\Psi)_1 + i\epsilon_{1bc}(\sigma_c)_{1\beta}[\mathcal{L}_b\Psi_\beta + (\eta_b\Psi)_\beta] \\ &= \mathcal{L}_1\Psi_1 + (\eta_1\Psi)_1 + i\epsilon_{123}(\sigma_3)_{1\beta}[\mathcal{L}_2\Psi_\beta + (\eta_2\Psi)_\beta] + i\epsilon_{132}(\sigma_2)_{1\beta}[\mathcal{L}_3\Psi_\beta + (\eta_3\Psi)_\beta] \\ &= \mathcal{L}_1\Psi_1 + (\eta_1\Psi)_1 + i(\sigma_3)_{11}[\mathcal{L}_2\Psi_1 + (\eta_2\Psi)_1] - i(\sigma_2)_{12}[\mathcal{L}_3\Psi_2 + (\eta_3\Psi)_2] \\ &= \mathcal{L}_1\Psi_1 + (\eta_1\Psi)_1 + i[\mathcal{L}_2\Psi_1 + (\eta_2\Psi)_1] - i(-i)[\mathcal{L}_3\Psi_2 + (\eta_3\Psi)_2] \\ &= (\mathcal{L}_1 + i\mathcal{L}_2)\Psi_1 + [(\eta_1 + i\eta_2)\Psi]_1 - [\mathcal{L}_3\Psi_2 + (\eta_3\Psi)_2] = 0 \end{aligned} \quad (3.33)$$

para los casos con $a = 2$, $\alpha = 1$ y $a = 3$, $\alpha = 2$ obtendremos ecuaciones idénticas a la (3.33); algo parecido sucede con el resto de los casos $a = 1$, $\alpha = 2$; $a = 2$, $\alpha = 2$ y $a = 3$, $\alpha = 1$, producen la siguiente ecuación

$$(\mathcal{L}_1 - i\mathcal{L}_2)\Psi_2 + [(\eta_1 - i\eta_2)\Psi]_2 + [\mathcal{L}_3\Psi_1 + (\eta_3\Psi)_1] = 0 \quad (3.34)$$

de 6 ecuaciones, solo 2 son independientes, las demás resultan ser copias.

Como es costumbre, definimos las siguientes combinaciones de ∇_a

$$\nabla_\pm \equiv \nabla_1 \pm i\nabla_2$$

$$\nabla_0 \equiv \nabla_3$$

para poder reescribir las ecuaciones (3.33) y (3.34); respectivamente obtendremos las siguientes

$$\nabla_+\Psi_1 - \nabla_0\Psi_2 = 0 \quad (3.35)$$

$$\nabla_- \Psi_2 + \nabla_0 \Psi_1 = 0. \quad (3.36)$$

Recordemos que el proyector Π nos dará como resultado un espinor, es por eso que las ecuaciones (3.35) y (3.36), serán las componentes de un espinor. Podemos verificar lo anterior realizando una rotación y comprobar que efectivamente, se transforman como las componentes de espinores; además nos ayudará a deducir cuál ecuación corresponde a cuál componente. Tomemos por simplicidad una rotación infinitesimal por ϵ , alrededor del eje z ; denotamos por Ψ' a la función de onda transformada, entonces

$$\begin{aligned} \Psi' &= \left(\mathbb{I}_2 - i\epsilon \frac{\sigma_0}{2} \right) \Psi \\ &= \begin{pmatrix} \Psi_1 - i\frac{\epsilon}{2}\Psi_1 \\ \Psi_2 + i\frac{\epsilon}{2}\Psi_2 \end{pmatrix} \end{aligned}$$

un vector se transforma como

$$\begin{aligned} \mathbf{x}' &= (\mathbb{I}_3 - i\epsilon\theta_0)\mathbf{x} \\ &= \begin{pmatrix} x_1 - \epsilon x_2 \\ x_2 + \epsilon x_1 \\ x_0 \end{pmatrix}. \end{aligned}$$

Entonces

$$\begin{aligned} \nabla'_1 &= \nabla_1 - \epsilon \nabla_2 \\ \nabla'_2 &= \nabla_2 + \epsilon \nabla_1 \\ \nabla'_0 &= \nabla_0 \end{aligned}$$

así resulta

$$\begin{aligned} \nabla'_+ &= \nabla'_1 + i\nabla'_2 \\ &= \nabla_1 - \epsilon \nabla_2 + i(\nabla_2 + \epsilon \nabla_1) \\ &= \nabla_1 + i\nabla_2 + i\epsilon(\nabla_1 + i\nabla_2) \\ \nabla'_+ &= \nabla_+ + i\epsilon \nabla_+ \end{aligned}$$

análogamente

$$(\nabla_-)' = \nabla_- - i\epsilon \nabla_-.$$

Procedemos a calcular como se transforma la ecuación (3.35)

$$(\nabla_+ \Psi_1 - \nabla_0 \Psi_2)' = (\nabla_+ + i\epsilon \nabla_+) \left(\Psi_1 - i\frac{\epsilon}{2}\Psi_1 \right) - \nabla_0 \left(\Psi_2 + i\frac{\epsilon}{2}\Psi_2 \right)$$

$$\begin{aligned}
&= \nabla_+ \Psi_1 - \nabla_0 \Psi_2 - i\frac{\epsilon}{2} \nabla_+ \Psi_1 + i\epsilon \nabla_+ \Psi_1 - i\frac{\epsilon}{2} \nabla_0 \Psi_2 \\
&= (\nabla_+ \Psi_1 - \nabla_0 \Psi_2) + i\frac{\epsilon}{2} \nabla_+ \Psi_1 - i\frac{\epsilon}{2} \nabla_0 \Psi_2 \\
&= (\nabla_+ \Psi_1 - \nabla_0 \Psi_2) + i\frac{\epsilon}{2} (\nabla_+ \Psi_1 - \nabla_0 \Psi_2)
\end{aligned}$$

el último resultado nos indica que, la expresión (3.35) se transforma como la segunda componente de un espinor, por lo tanto, la expresión (3.36) se transforma como la primera componente de un espinor, es decir podemos hacer la siguiente identificación

$$\begin{aligned}
\nabla_- \Psi_2 + \nabla_0 \Psi_1 &= \Phi_1 \\
\nabla_+ \Psi_1 - \nabla_0 \Psi_2 &= \Phi_2
\end{aligned}$$

donde Φ_1 y Φ_2 , son componentes de un espinor tal que

$$\not{D}\Psi = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}. \quad (3.37)$$

Esta última es la ecuación de Dirac que buscamos, está escrita en términos de un operador $\not{D}$, cuya expresión aún no conocemos y un espinor Φ fijado por consideraciones físicas. Como ya mencionamos, en el caso de fermiones con masa, será $\Phi = m\Psi$ donde m es la masa del fermión.

Antes de empezar a calcular las expresiones del operador de Dirac, de manera auxiliar establecemos las siguientes definiciones

$$\eta_{\pm} \equiv \eta_1 \pm i\eta_2$$

y $\eta_0 \equiv \eta_3$, los cuales calculamos explícitamente

$$\begin{aligned}
\eta_0 &= \frac{1}{2} \mathcal{P}_{3a} \sigma_a = \frac{1}{2} (\delta_{3a} - x_3 x_a) \sigma_a = \frac{1}{2} (\sigma_3 - x_3 x_a \sigma_a) \\
&= \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - x_3 \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} \right] \\
\eta_0 &= \frac{1}{2} \begin{pmatrix} 1 - x_0^2 & -x_0 x_- \\ -x_0 x_+ & -1 + x_0^2 \end{pmatrix}
\end{aligned}$$

donde hemos aplicado la ya usual notación $x_0 = x_3$ y $x_{\pm} = x_1 \pm ix_2$; de forma similar obtenemos η_1 y η_2

$$\eta_1 = \frac{1}{2} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - x_1 \begin{pmatrix} x_0 & x_- \\ x_+ & -x_0 \end{pmatrix} \right]$$

$$\begin{aligned}
&= \frac{1}{2} \begin{pmatrix} -x_1x_0 & 1 - x_1x_- \\ 1 - x_1x_+ & x_1x_0 \end{pmatrix} \\
\eta_2 &= \frac{1}{2} \left[\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - x_2 \begin{pmatrix} x_0 & x_- \\ x_+ & -x_0 \end{pmatrix} \right] \\
&= \frac{1}{2} \begin{pmatrix} -x_2x_0 & -i - x_2x_- \\ i - x_2x_+ & x_2x_0 \end{pmatrix}
\end{aligned}$$

finalmente calculamos η_+ y $\eta_- = \eta_+^\dagger$

$$\begin{aligned}
\eta_+ &= \frac{1}{2} \begin{pmatrix} -x_1x_0 - ix_2x_0 & 1 - x_1x_- + (1 - ix_2x_-) \\ 1 - x_1x_+ + (-1 - ix_2x_+) & x_1x_0 + ix_2x_0 \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} -x_0x_+ & 1 + x_0^2 \\ -x_+^2 & x_0x_+ \end{pmatrix} \\
\eta_+^\dagger \eta_- &= \frac{1}{2} \begin{pmatrix} -x_0x_- & -x_-^2 \\ 1 + x_0^2 & x_0x_- \end{pmatrix}.
\end{aligned}$$

Ahora podemos calcular Φ_1 y Φ_2

$$\begin{aligned}
\Phi_1 &= \mathcal{L}_- \Psi_2 + (\eta_-)_{21} \Psi_1 + (\eta_-)_{22} \Psi_2 + \mathcal{L}_0 \Psi_1 + (\eta_0)_{11} \Psi_1 + (\eta_0)_{12} \Psi_2 \\
&= \mathcal{L}_- \Psi_2 + \mathcal{L}_0 \Psi_1 + \frac{1}{2} [((1 + x_0^2) + (1 - x_0^2)) \Psi_1 + ((x_0x_-) + (-x_0x_-)) \Psi_2] \\
\Phi_1 &= \mathcal{L}_- \Psi_2 + \mathcal{L}_0 \Psi_1 + \Psi_1
\end{aligned}$$

análogamente

$$\begin{aligned}
\Phi_2 &= \mathcal{L}_+ \Psi_1 + (\eta_+)_{11} \Psi_1 + (\eta_+)_{12} \Psi_2 - \mathcal{L}_0 \Psi_2 - (\eta_0)_{21} \Psi_1 - (\eta_0)_{22} \Psi_2 \\
&= \mathcal{L}_+ \Psi_1 - \mathcal{L}_0 \Psi_2 + \frac{1}{2} [((-x_0x_+) - (-x_0x_+)) \Psi_1 + ((1 + x_0^2) - (-1 + x_0^2)) \Psi_2] \\
\Phi_2 &= \mathcal{L}_+ \Psi_1 - \mathcal{L}_0 \Psi_2 + \Psi_2
\end{aligned}$$

Volviendo a la ecuación

$$\begin{aligned}
D\Psi &= \begin{pmatrix} \mathcal{L}_- \Psi_2 + \mathcal{L}_0 \Psi_1 + \Psi_1 \\ \mathcal{L}_+ \Psi_1 - \mathcal{L}_0 \Psi_2 + \Psi_2 \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{L}_- \Psi_2 + \mathcal{L}_0 \Psi_1 \\ \mathcal{L}_+ \Psi_1 - \mathcal{L}_0 \Psi_2 \end{pmatrix} + \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \begin{pmatrix} \mathcal{L}_0 & \mathcal{L}_- \\ \mathcal{L}_+ & -\mathcal{L}_0 \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} + \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}
\end{aligned}$$

$$\not{D}\Psi = \left[\begin{pmatrix} \mathcal{L}_0 & \mathcal{L}_- \\ \mathcal{L}_+ & -\mathcal{L}_0 \end{pmatrix} + \mathbb{I}_2 \right] \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \quad (3.38)$$

dado que

$$\boldsymbol{\sigma} \cdot \mathbf{L} = \begin{pmatrix} \mathcal{L}_0 & \mathcal{L}_- \\ \mathcal{L}_+ & -\mathcal{L}_0 \end{pmatrix}$$

escribimos el operador de Dirac en términos de los operadores de momento angular

$$\not{D} = \boldsymbol{\sigma} \cdot \mathcal{L} + \mathbb{I} \quad (3.39)$$

nótese inmediatamente la semejanza que guarda con la ecuación (1.35), lo cual es una curiosidad. Así la ecuación de Dirac para fermiones con masa queda como

$$(\boldsymbol{\sigma} \cdot \mathcal{L} + \mathbb{I})\Psi = m\Psi.$$

Se puede comprobar consultando en la literatura [20], que nuestro resultado que obtenemos es el correcto.

Capítulo 4

Conclusiones

Hemos logrado hacer una revisión de una buena variedad de conceptos, tanto físicos como matemáticos. Presentamos una construcción estándar de la ecuación de Dirac en (3+1) dimensiones. Repasamos los fundamentos de la teoría de grupos y de la teoría de representaciones, que sirvió de preámbulo para definir el grupo $SU(2)$ y estudiar algunos de sus aspectos. Concretamente las representaciones de $SU(2)$ fue el problema que más trabajamos; calculamos explícitamente las expresiones de los generadores de las representaciones $\psi_{(1)}$ y $\psi_{(1/2)}$. Hablamos de la relación estrecha que existe entre las representaciones de $SU(2)$ y la formulación del momento angular en la mecánica cuántica.

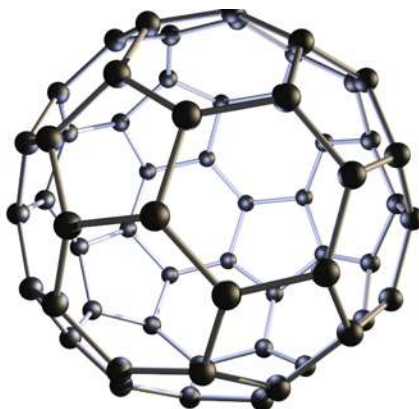
También tuvimos la oportunidad de introducir algunas nociones de geometría diferencial, entre las que destacamos la noción de derivada covariante y sobre todo, el método de obtenerla a partir de la proyección de la derivada total. Construimos de esta forma la conexión de espín sobre la esfera, identificando primero a los operadores de momento angular de nuestro sistema $\mathcal{L} + \mathcal{S}$ y luego construyendo el operador de proyección sobre la esfera $\mathcal{P}$.

Al implementar el método de Cartan, los recursos previos de la teoría de representaciones fueron necesarios para una rápida identificación del operador de proyección Π . Dicha identificación se hizo observando que el objeto $\nabla_a \Psi_\alpha$, se transforma de acuerdo a la representación $\psi_{(1)} \otimes \psi_{(1/2)}$. Recurriendo al teorema de Clebsch-Gordan, descomponemos dicho producto de representaciones en una suma $\psi_{(1/2)} \oplus \psi_{(3/2)}$, que es donde se observa la presencia de la representación de espín $1/2$, que es la única que nos interesa preservar al momento de usar la proyección Π . Usando la descomposición espectral del operador de Casimir de la representación, es que obtuvimos la proyección.

Imponiendo la condición $\Pi \nabla \Psi = 0$, que corresponde al caso de una partícula sin masa, se obtuvieron dos ecuaciones que identificamos con las componentes de un espinor, al observar como se transforman ante rotaciones. Después de un poco de álgebra identificamos el operador de Dirac y recordamos, que en el caso de un fermión con masa, el operador de Dirac cumple $(\boldsymbol{\sigma} \cdot \mathcal{L} + \mathbb{I})\Psi = m\Psi$.

Una de las ventajas del método de Cartan, es que no es necesario introducir un atlas ni coordenadas locales, y los cálculos usando las representaciones son más sencillos. Por ejemplo, se puede consultar el artículo de Abrikosov [21] y comparar ambos métodos. En el mismo artículo se menciona que, entre las aplicaciones de la ecuación de Dirac en la esfera, se encuentran las relacionadas a las sustancias llamadas *fullerenos*. Los fullerenos son moléculas compuestas de átomos de carbono, que están distribuidos en los vértices de poliedros o sobre una estructura esferoidal. Entre los fullerenos más abundantes están los llamados *buckminsterfullerenos* (o buckyball), que son moléculas de C_{60} y su forma se parece a una pelota de fútbol. Los buckminsterfullerenos son de especial interés por sus propiedades de superconductividad. En este caso, la ecuación de Dirac en la esfera sirve como un modelo teórico, con el cual podemos comparar, o aproximar, el comportamiento de los electrones en este tipo de moléculas [22].

También incluimos un apéndice donde calculamos el espectro del operador de Dirac, apoyándonos en la teoría de representaciones; en el artículo de Abrikosov también se encuentra el espectro y los eigenespinores, pero el método que usa, a través de los polinomios de Jacobi, implica más cálculos, he aquí otra ventaja de las representaciones.



Apéndice A

Grupos de Lie

Algunos conceptos a los que nos referimos en este apéndice se han incluido en la sección 3.1. Cualquier tema que pudiésemos presentar con respecto a grupos de Lie o álgebras de Lie, sería muy reducido, dado que el desarrollo teórico que envuelve a dichos grupos es tan extenso como sus aplicaciones. Presentamos las definiciones necesarias del tema y proporcionamos un par de ejemplos, ya que no pretendemos otra cosa más que señalarle al lector la existencia de dichos elementos.

Definición A.1 Sea $\mathcal{G}$ una variedad diferenciable a la que se ha dotado con una estructura de grupo $(\mathcal{G}, *)$. Cuando la operación $*$: $(g_1, g_2) \mapsto g_1 * g_2$ junto con $^{-1} : g \mapsto g^{-1}$ resultan ser diferenciables, entonces se dice que $\mathcal{G}$ es un grupo de Lie.

Ejemplo A.2 Los siguientes grupos son grupos de Lie: el grupo lineal general $\mathrm{GL}(n, \mathbb{R})$, el grupo ortogonal $\mathrm{O}(n)$, el grupo lineal especial $\mathrm{SL}(n, \mathbb{R})$ y el grupo ortogonal especial $\mathrm{SO}(n, \mathbb{R})$.

- $\mathrm{GL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det(A) \neq 0\}$, $M_n(\mathbb{R})$ es el conjunto de matrices reales de $n \times n$.
- $\mathrm{O}(n) = \{A \in \mathrm{GL}(n, \mathbb{R}) : AA^T = \mathbb{I}_n\}$, matrices ortogonales.
- $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) : \det(A) = 1\}$.
- $\mathrm{SO}(n, \mathbb{R}) = \mathrm{SL}(n, \mathbb{R}) \cap \mathrm{O}(n)$.

Cuando se trata con variable compleja, los siguientes son grupos de Lie: el grupo lineal general complejo $GL(n, \mathbb{C})$, el grupo unitario $U(n)$, el grupo lineal especial complejo $SL(n, \mathbb{C})$ y el grupo unitario especial $SU(n)$.

- $GL(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) : \det(A) \neq 0\}$.
- $U(n) = \{A \in GL(n, \mathbb{C}) : AA^\dagger = \mathbb{I}_n\}$.
- $SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) : \det(A) = 1\}$.
- $SU(n) = SL(n, \mathbb{C}) \cap U(n)$.

Definición A.3 Sea $\mathcal{G}$ un grupo de Lie y $g, h \in \mathcal{G}$. Le llamamos desplazamiento por la izquierda¹ por h a la función $L_h : \mathcal{G} \rightarrow \mathcal{G}$, $g \mapsto L_h g = hg$.

Es claro que los desplazamientos por la izquierda L_h , son funciones suaves y por lo tanto inducen el pushforward $L_{h*} : T_g \mathcal{G} \rightarrow T_{hg} \mathcal{G}$. Sea $\mathbf{X} \in \mathfrak{X}(\mathcal{G})$, cuando $L_{h*} \mathbf{X}|_g = \mathbf{X}|_{hg}$ se dice que $\mathbf{X}$ es *izquierdo-invariante*. Al conjunto de campos vectoriales izquierdo invariantes de $\mathcal{G}$ lo denotamos por $\mathfrak{g}$. Es fácil ver que $\mathfrak{g}$ es un espacio vectorial y que además es cerrado bajo el corchete de Lie definido² para campos vectoriales como

$$[\mathbf{X}, \mathbf{Y}]f = \mathbf{X}[\mathbf{Y}[f]] - \mathbf{Y}[\mathbf{X}[f]]$$

donde $f \in \mathfrak{F}(\mathcal{G})$. Consideremos una base de $\mathfrak{g}$, $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$; dado que $[\mathbf{X}_a, \mathbf{X}_b]$ es un elemento de $\mathfrak{g}$, se puede expandir en términos de la base:

$$[\mathbf{X}_a, \mathbf{X}_b] = f_{ab}^c \mathbf{X}_c$$

donde los f_{ab}^c son las *constantes de estructura* del grupo de Lie $\mathcal{G}$. El caso en que el grupo de Lie consta de matrices, con la operación de producto de matrices, entonces el corchete de Lie pasa a ser el conmutador que ya conocemos. Es un buen ejercicio revisar que las constantes de estructura tienen las siguientes propiedades:

1. Antisimetría

$$f_{ab}^c = -f_{ba}^c$$

¹El *desplazamiento por la derecha* se define de manera análoga.

²Dado que $f_*[\mathbf{X}, \mathbf{Y}] = [f_*\mathbf{X}, f_*\mathbf{Y}]$.

2. Identidad de Jacobi

$$f_{ab}{}^c f_{cd}{}^e + f_{da}{}^c f_{cb}{}^e + f_{bd}{}^c f_{ca}{}^e = 0$$

Definición A.4 Se le llama álgebra de Lie (del grupo de Lie $\mathcal{G}$) al conjunto $\mathfrak{g}$, junto con el corchete de Lie $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$.

Según la literatura, para referirse al álgebra de Lie de un grupo, se usan los correspondientes caracteres góticos en minúsculas. Por ejemplo el álgebra de Lie de $SU(2)$ se escribe $\mathfrak{su}(2)$.

Ejemplo A.5 Las siguientes son las álgebras de Lie de los grupos del ejemplo A.2:

- $\mathfrak{gl}(n, \mathbb{R}) = M_n(\mathbb{R})$.
- $\mathfrak{o}(n) = \{A \in M_n(\mathbb{R}) : A^T + A = 0\}$, matrices antisimétricas.
- $\mathfrak{sl}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \text{Tr}(A) = 0\}$.
- $\mathfrak{so}(n) = \mathfrak{o}(n)$.
- $\mathfrak{gl}(n, \mathbb{C}) = M_n(\mathbb{C})$.
- $\mathfrak{u}(n) = \{A \in M_n(\mathbb{C}) : A^\dagger + A = 0\}$, matrices antihermitianas.
- $\mathfrak{sl}(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) : \text{Tr}(A) = 0\}$.
- $\mathfrak{su}(n) = \mathfrak{sl}(n, \mathbb{C}) \cap \mathfrak{u}(n)$.

Apéndice B

Espectro del operador $\mathcal{D}$

La función de onda de la ecuación (3.21), se puede escribir usando la base de eigenkets comunes de $\mathcal{L}^2$, $\mathcal{L}_z$, $\mathbf{S}$ y S_z , $\{|\ell, m\rangle \otimes |\pm\rangle\}$, en lugar de la de posiciones. La representación correspondiente a $\mathcal{L}^2$ es $\oplus_{\ell=0}^{\infty} \psi_{(\ell)}$, de modo que al momento angular total $\mathbf{J}^2$ le corresponde $\oplus_{\ell=0}^{\infty} \psi_{(\ell)} \otimes \psi_{(1/2)}$, la cual podemos descomponer como suma de representaciones irreducibles, según el teorema de Clebsch-Gordan

$$\bigoplus_{\ell=0}^{\infty} \psi_{(\ell)} \otimes \psi_{(\frac{1}{2})} = \bigoplus_{\ell=1}^{\infty} \psi_{(\ell-\frac{1}{2})} \bigoplus_{\ell=0}^{\infty} \psi_{(\ell+\frac{1}{2})}. \quad (\text{B.1})$$

Así mismo, también podemos encontrar la expresión para la base de eigenkets comunes de $\mathbf{J}^2$, J_z , $\mathcal{L}^2$ y $\mathbf{S}^2$, denotados como $|J, M\rangle$; en términos de la base $\{|\ell, m\rangle \otimes |\pm\rangle\}$. Los elementos de la matriz de cambio de base son los coeficientes de Clebsch-Gordan, obtenemos entonces

$$|J, M\rangle = \sum_{m+\alpha=M} C_{\ell, m; \frac{1}{2}\alpha}^{JM} |\ell, m\rangle \otimes |1/2, \alpha\rangle \quad (\text{B.2})$$

donde $\alpha \in \{1/2, -1/2\}$ (que es otra forma de escribir $|\pm\rangle$).

Para calcular el espectro del operador $\mathcal{D}$, será necesario hacer la siguiente descomposición en términos de operadores de Casimir

$$\begin{aligned} \mathcal{D} &= (\boldsymbol{\sigma} \cdot \mathcal{L} + \mathbb{I}) = \left(\mathcal{L} + \frac{\boldsymbol{\sigma}}{2} \right)^2 - \mathcal{L}^2 + \frac{1}{4} \\ &= \mathbf{J}^2 - \mathcal{L}^2 + \frac{1}{4}. \end{aligned}$$

Usando la base $\{|J, M\rangle\}$ podemos calcular el espectro, para ello es necesario observar la descomposición (B.1), y notar que los valores que puede tomar J son $\{\ell + 1/2; \ell \geq 0\}$ y $\{\ell - 1/2; \ell \geq 1\}$. Obtenemos entonces la siguiente ecuación para los eigenvalores

$$\lambda = J(J+1) - \ell(\ell+1) + \frac{1}{4}. \quad (\text{B.3})$$

Sustituyendo los valores de J en la ecuación (B.3) encontraremos el espectro. Para $J = \ell + 1/2$, con $\ell \geq 0$, vemos que $\lambda \in \{1, 2, 3, \dots\}$; mientras que para $J = \ell - 1/2$ con $\ell \geq 1$, resulta $\lambda \in \{-1, -2, -3, \dots\}$. Es decir el espectro del operador de Dirac en la esfera, es $\{\pm\ell; \ell \in \mathbb{N} \setminus \{0\}\}$. La degeneración del eigenvalor λ es $2|\lambda|$.

También podemos encontrar los eigenespinores a partir de la ecuación (B.2), notando primero que $\langle \mathbf{x} | \ell, m \rangle = Y_\ell^m(\mathbf{x})$ son los armónicos esféricos ($\mathbf{x} \cdot \mathbf{x} = 1$); esto se puede consultar en [7], basta observar que el operador $\mathcal{L}^2$ corresponde a la parte angular del operador laplaciano. La función de onda de un eigenket $|J, M\rangle$ se escribe como

$$\Psi_{J,M}^\ell = \langle \mathbf{x} | J, M \rangle = \sum_{m+\alpha=M} C_{\ell,m;\frac{1}{2},\alpha}^{J,M} Y_m^\ell(\mathbf{x}) \otimes |\pm\rangle.$$

Bibliografía

- [1] V. I. Ridnik, *¿Qué es la mecánica cuántica?*, Mir (1977).
- [2] L. de la Peña, *Introducción a la mecánica cuántica*, FCE (2006).
- [3] B. Friedrich, D. Herschbach, *Stern and Gerlach: How a bad cigar helped reorient atomic physics*, Physics Today, **56** December (2003), 53-59.
- [4] R. G. J. Fraser, *The efective cross section of the oriented hydrogen atom*, Proc. R. Soc. A, **114** (1927), 212-221.
- [5] T. E. Phipps, J. B. Taylor, *The magnetic moment of the hydrogen atom*, Phys. Rev., **29** (1927), 309-320.
- [6] N. Zettili, *Quantum mechanics, concepts and applications*, Wiley (2001).
- [7] J. J. Sakurai, *Modern quantum mechanics*, Addison-Wesley (1993).
- [8] W. Greiner, *Relativistic quantum mechanics, wave equations*, Springer (2000).
- [9] L. H. Ryder, *Quantum field theory*, Cambridge University Press (1996).
- [10] J. J. Rotman, *Advanced modern algebra*, Prentice Hall (2003).
- [11] H. Georgi, *Lie algebras in particle physics: From isospin to unified theories*, Westview Press (1999).
- [12] H. Weyl, *The theory of groups and quantum mechanics*, Dover Publications (1950).
- [13] B. C. Hall, *An elementary introduction to groups and representations*, arXiv: math-ph/0005032.

- [14] M. R. Sepanski, *Compact Lie groups*, Springer (2007).
- [15] W. M. Boothby, *An introduction to diferentiable manifolds and Riemannian geometry*, Academic Press Inc. (1986).
- [16] M. Nakahara, *Geometry, topology and physics*, Institute of Physics Publishing (2003).
- [17] M. P. do Carmo, *Riemannian geometry*, Birkhäuser (1992).
- [18] R. M. Wald, *General relativity*, The University of Chicago Press (1984).
- [19] E. J. Cartan, *The theory of spinors*, Dover Publications (1981).
- [20] C. Jayewardena, *Schwinger model on S^2* , helv. Phys. Acta, **61** (1988), 636-711.
- [21] A. A. Abrikosov, *Dirac operator on the Riemann sphere*, arXiv: hep-th/0212134.
- [22] J. González, F. Guinea, M. A. H. Vozmediano, *The electronic spectrum of fullerenes from the Dirac equation*, Nucl. Phys. B, **406** (1993), 771-794.

Índice alfabético

- Álgebra de Lie, 31, 77
- Automorfismo, 21
- Baker-Campbell-Hausdorff, 30
- Campo vectorial, 49
- Clebsch-Gordan, 41, 79
 - coeficientes de, 79
- Conexión
 - afín, 55
 - de espín, 62, 64, 65
 - de Levi-Civita, 58
 - métrica, 58
- Constantes de estructura, 31, 76
- Derivada covariante, 52
- Ecuación de Weyl, 17, 18
- Ecuación de Dirac, 14
 - en la esfera, 71, 74
- Eigenespinor, 74, 80
- Encajamiento, 51
- Espacio
 - cotangente, 49
 - tangente, 48
- Espectro, 10, 37, 74, 80
- Espín, 7, 60, 65
- Espinor, 61, 68
- Experimento de Stern Gerlach, 6, 7
- Función de onda, 8, 10, 61, 63
- Función suave, 47
- Generadores
 - del grupo, 28, 30
 - de rotaciones, 59
- Grupo de Lie, 31, 75
- Homeomorfismo, 46
- Homomorfismo, 21
- Identidad de Jacobi, 77
- Inmersión, 51
- Isomorfismo, 21
- Izquierdo-invariante, 76
- Ecuación de Klein-Gordon, 11, 13, 15
- Longitud de onda de De Broglie, 4
- Matrices de Pauli, 14, 28, 40
- Métrica riemanniana, 57
- Momento angular
 - espinorial, 7
 - orbital, 7
 - total, 60
- Operador
 - de momento angular, 59
 - de Casimir, 33, 65
 - de Dirac, 14
 - en la esfera, 65, 71, 79
 - de energía, 10
 - de espín, 17
 - de helicidad, 18

- de momento lineal, 10
- de momento angular orbital, 17, 60, 62
- de posición, 10
- de rotación, 59
- Operadores escalera, 33
- Parametrización exponencial, 29
- Principio de exclusión, 7, 15
- Producto tensorial de representaciones, 25
- Pushforward, 50
- Relación de Planck, 3
- Representación de Weyl, 16
- Rotaciones infinitesimales, 59
- Símbolos de Christoffel, 55
- Subvariedad, 51
- Suma directa de representaciones, 23
- Tensor tipo (r, s) , 49
- Variedad
 - diferenciable, 46, 75
 - riemanniana, 57
- Vector de estado, 9
- Vector tangente, 48

Agradecimientos

Sobre todo, le agradezco a mi familia por formar a la persona que soy. A mi hermana Melissa que con su alegría y ocurrencias, siempre me saca una sonrisa. A mis padres Lupita y Sergio, que nunca han dudado en apoyarme y a quienes debo todo lo que tengo. Les agradezco a ellos su esfuerzo y reconozco que este logro es de ellos.

También estoy en deuda con todos mis amigos y compañeros, con ellos he crecido y he aprendido; juntos celebramos en la bonanza y juntos batallamos en la desdicha. Le agradezco a Victoria por hacer de mí una persona menos solitaria y por todo el cariño que me ha brindado.

Agradezco en especial a mi amigo Juan Pablo, a los profesores Alberto Mendoza, Joaquín y Malú; que gracias a ellos decidí dedicarme a la ciencia.

A mis asesores por nunca escatimar en soluciones, respuestas y consejos. Han sido ejemplares y he encontrado en ellos una guía en mi carrera. Se les debe reconocer la paciencia que han tenido con este alumno tan terco.

Donde sea que estén, incluyo a Beethoven, Dvořák y a Hendrix; por hacer de mis noches de trabajo una experiencia más llevadera.