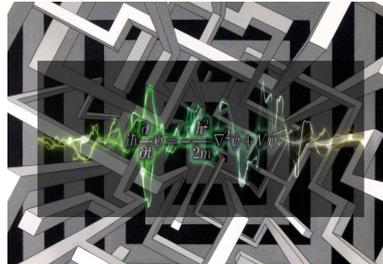


# Una introducción a la Mecánica Cuántica para “no iniciados”



**Renato Álvarez Nodarse**

Departamento de Análisis Matemático,  
Facultad de Matemáticas, Universidad de Sevilla

30 de junio de 2015



# Índice general

<b>1. Breve introducción a la mecánica clásica</b>	<b>1</b>
1.1. Mecánica Hamiltoniana . . . . .	3
1.2. Dos ejemplos representativos . . . . .	4
1.2.1. El oscilador armónico unidimensional . . . . .	4
1.2.2. Movimiento en un campo central de fuerzas . . . . .	6
1.3. Problemas . . . . .	11
<b>2. ¿Cómo se gestó la Mecánica cuántica?</b>	<b>13</b>
2.1. La radiación del cuerpo negro . . . . .	14
2.2. Einstein y el efecto fotoeléctrico . . . . .	18
2.3. Bohr y el modelo atómico . . . . .	19
2.4. El nacimiento de la Mecánica Cuántica . . . . .	21
2.4.1. La dualidad onda-partícula . . . . .	21
2.4.2. La Mecánica matricial de Heisenberg . . . . .	22
2.4.3. La Mecánica ondulatoria de Schrödinger . . . . .	23
2.4.4. Una “deducción” de la ecuación de Schrödinger . . . . .	25
2.5. La interpretación de la Mecánica cuántica . . . . .	25
2.5.1. El gato de Schrödinger . . . . .	27
2.5.2. Los universos paralelos de Everett . . . . .	29
2.6. El principio de incertidumbre de Heisenberg . . . . .	30
2.6.1. El experimento de difracción y el principio de incertidumbre . . . . .	30
2.7. Las matemáticas de la Mecánica Cuántica . . . . .	32
2.8. Sobre la bibliografía . . . . .	33

<b>3. Mecánica Cuántica I: “Movimiento” de una partícula material</b>	<b>35</b>
3.1. Los postulados de la Mecánica Cuántica . . . . .	35
3.2. El principio de incertidumbre . . . . .	42
3.2.1. Los estados estacionarios de la ecuación de Schrödinger . . . . .	43
3.3. Ejemplos . . . . .	43
3.3.1. Una partícula en un pozo de potencial . . . . .	44
3.3.2. El efecto túnel . . . . .	46
3.3.3. Problemas . . . . .	48
<b>4. Mecánica Cuántica II: Espacios de Hilbert</b>	<b>49</b>
4.1. Espacios euclídeos y espacios normados . . . . .	49
4.2. Operadores en $\mathbb{H}$ . . . . .	53
4.3. Los axiomas de la Mecánica Cuántica . . . . .	59
4.4. Discusión e implicaciones de los postulados . . . . .	61
4.4.1. Los proyectores ortogonales y la teoría de mediciones . . . . .	62
4.5. Representación de los operadores $\hat{x}_i$ y $\hat{p}_i$ . . . . .	63
4.6. Las ecuaciones de Heisenberg y de Schrödinger . . . . .	64
4.6.1. Equivalencia de las representaciones de Heisenberg y de Schrödinger . . . . .	66
4.6.2. Integrales de movimiento . . . . .	67
4.6.3. Los estados estacionarios del sistema . . . . .	68
4.6.4. Los operadores unitarios y la evolución temporal . . . . .	68
4.7. El principio de incertidumbre . . . . .	68
4.8. La mecánica matricial . . . . .	69
4.9. La ecuación de Schrödinger y el postulado 4.3.5 . . . . .	70
4.10. Problemas . . . . .	70
<b>5. Resolviendo la ecuación de Schrödinger</b>	<b>73</b>
5.1. El método de Nikiforov-Uvarov . . . . .	73
5.1.1. La ecuación hipergeométrica generalizada . . . . .	73
5.1.2. La ecuación diferencial hipergeométrica . . . . .	74
5.1.3. Los polinomios de Hermite, Laguerre y Jacobi . . . . .	82
5.2. Resolución de la ecuación de Schrödinger . . . . .	85
5.2.1. El oscilador armónico cuántico unidimensional . . . . .	85
5.2.2. La ecuación de Schrödinger en un potencial central . . . . .	87
5.2.3. Los armónicos esféricos . . . . .	88
5.2.4. Resolviendo la parte radial de la ecuación de Schrödinger . . . . .	90

5.2.5. El oscilador armónico tridimensional . . . . . 93

5.3. El método de factorización de Schrödinger . . . . . 94

5.3.1. Introducción . . . . . 94

5.3.2. El oscilador armónico . . . . . 95

5.3.3. El método de factorización . . . . . 98

5.3.4. Ejemplos . . . . . 102

5.4. Factorización de la EDO hipergeométrica . . . . . 109

5.4.1. El *hamiltoniano* y los operadores *escalera* . . . . . 109

5.4.2. Factorización de  $\mathcal{H}(x, n)$  . . . . . 110

5.4.3. Ejemplos . . . . . 112

5.5. Problemas . . . . . 113

**Bibliografía** **115**

**Anexo A: Breve introducción al análisis funcional** **119**

A.1. Introducción: Estacios métricos y espacios normados . . . . . 119

A.2. Espacios de Hilbert separables . . . . . 126

A.2.1. Operadores en espacios de Hilbert . . . . . 131

A.2.2. Teoría Espectral de operadores compactos autoadjuntos . . . . . 135

**Bibliografía** **139**



# Prefacio

Estas notas contienen el contenido de un curso de introducción a la Mecánica cuántica impartido por el autor en la Universidad de Zaragoza en septiembre de 2005 y en Coimbra en febrero de 2006. Las mismas están divididas en 6 capítulos que contienen tanto los conceptos teóricos como distintos métodos de resolución de la ecuación de Schrödinger. Quiero agradecer a todos los que de una forma u otra me han ayudado a que estas notas sean posibles. En primer lugar a mi familia, a los que les he robado un tiempo precioso. Además agradezco a Manuel Alfaro (Universidad de Zaragoza), José Luis Cardoso (Universidade de Trás-os-Montes e Alto Douro), Mario Pérez (Universidad de Zaragoza), Francisco J. (Pacho) Ruiz (Universidad de Zaragoza), José Carlos Petronilho (Universidade de Coimbra) y Juan L. Varona (Universidad de La Rioja) sus comentarios y correcciones que me han permitido mejorar notablemente la exposición. También debo agradecer a Luis Velázquez de la Universidad de Zaragoza y Alberto Grunbaum de la Universidad de California (Berkeley) por las interesantes discusiones que tuvieron lugar en la semana más calurosa de mayo de 2015 en Sevilla que me permitieron corregir e incluir más material.

Renato Álvarez Nodarse  
Sevilla 29 de junio de 2015



# Introducción

Estas notas corresponden a un curso de introducción a la Mecánica cuántica. En él se pretende dar una idea general de cómo surgió la Mecánica cuántica a principios del siglo pasado (XX) y exponer algunos de sus principales principios desde un punto de vista “matemático”. El trabajo estará dividido en dos partes.

En la primera se describirá la evolución de la teoría cuántica desde finales del siglo XIX hasta la aparición de la Mecánica cuántica de Heisenberg y Schrödinger en 1926–1927, es decir, a partir de la idea revolucionaria de Planck sobre los “quanta”, la explicación del efecto fotoeléctrico por Einstein y la descripción de la dualidad onda-partícula de De Broglie, hasta la aparición de las dos principales teorías matemáticas: la mecánica matricial y la mecánica ondulatoria.

En la segunda parte introduciremos la Mecánica Cuántica de una partícula así como sus bases axiomáticas en un espacio de Hilbert separable. Finalmente, veremos algunos de los métodos usados para resolver la ecuación de Schrödinger.



# Capítulo 1

## Breve introducción a la mecánica clásica

La Física se basa en medidas y observaciones experimentales de la realidad que nos rodea, es decir, en cuantificar o caracterizar los distintos fenómenos naturales mediante expresiones cuantitativas o números.

Estas cantidades medibles u observables se denominan *cantidades físicas* (e.g. longitud, velocidad, energía, ...). El objeto o conjunto de objetos a estudiar se denomina *sistema físico* (e.g. una partícula, un átomo, un coche, ...). Cuando conocemos distintas medidas de un sistema que lo caracterizan por completo en un momento de tiempo determinado (e.g. la posición y la velocidad de una partícula de masa  $m$ ) decimos que el sistema se encuentra en un cierto *estado* dado.

El objetivo de toda teoría física es, por tanto:

1. Describir el estado del sistema físico, es decir, dar una representación cuantitativa (matemática) del estado que lo defina biunívocamente.
2. Conocer la *dinámica* del sistema, es decir dado un estado inicial en el momento  $t_0$  conocer su evolución temporal para  $t > t_0$ .
3. Predecir los resultados de las mediciones de las cantidades físicas del sistema.

La teoría física en sí misma está en general constituida, desde el punto de vista abstracto, por tres apartados:

1. El formalismo: Conjunto de símbolos y reglas de deducción a partir de los cuales se pueden deducir proposiciones y enunciados. En general toda teoría comienza fijando un cierto número de *axiomas* comúnmente denominados *postulados*.
2. Ley dinámica: Cierta relación (o relaciones) entre algunos de los principales objetos del formalismo que permitan predecir acontecimientos futuros.
3. Reglas de correspondencia o interpretación física: Conjunto de reglas que permiten asignar valores experimentales a algunos de los símbolos del formalismo.

Como ejemplo ilustrativo vamos a describir la mecánica newtoniana.

En la mecánica newtoniana el estado de un sistema viene dado por el conjunto de trayectorias de todas las partículas que constituyen el sistema. Por ejemplo, para una partícula, el estado estará dado por la función vectorial  $\vec{r}(t) \in \mathbb{R}^3$  que denota la posición en cada instante de tiempo  $t$ . Los observables son las cantidades medibles como la posición  $\vec{r}(t)$ , la velocidad  $\vec{v}(t) = d/dt[\vec{r}(t)]$ , la energía cinética  $T = mv^2(t)$ , etc.

La ley dinámica en este caso es la segunda ley de Newton:

$$m\vec{a}(t) = \vec{F}(t), \quad \vec{a}(t) = \frac{d^2\vec{r}(t)}{dt^2},$$

donde  $\vec{F}$  es la *fuerza* resultante que actúa sobre el sistema, i.e., es una ecuación diferencial de orden 2. Finalmente, las reglas de correspondencia son “evidentes” y consisten en los valores numéricos de las proyecciones de los vectores  $\vec{r}$ ,  $\vec{v}$ , etc. sobre los ejes del correspondiente sistema de coordenadas escogido.

Veamos un ejemplo de sistema físico y algunas de sus propiedades.

Supongamos que tenemos una partícula que se mueve en  $\mathbb{R}^3$  bajo la acción de una fuerza  $\vec{F}(x, y, z)$  que sólo depende de las coordenadas (posición). Supongamos además que existe una función escalar  $V(x, y, z)$  tal que

$$\vec{F}(x, y, z) = -\nabla V(x, y, z) = -\frac{\partial V}{\partial x}\vec{i} - \frac{\partial V}{\partial y}\vec{j} - \frac{\partial V}{\partial z}\vec{k},$$

donde  $\vec{i}$ ,  $\vec{j}$  y  $\vec{k}$ , son los vectores unitarios correspondientes a los ejes  $x$ ,  $y$  y  $z$ , respectivamente. Entonces, usando el producto escalar estándar de los vectores tenemos

$$\vec{F}(x, y, z)d\vec{r} = -\left(\frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz\right) = -dV(x, y, z).$$

Luego, el trabajo de la fuerza  $\vec{F}$  para mover nuestra partícula a lo largo de cierta curva  $\Gamma \in \mathbb{R}^3$  se expresa mediante la integral

$$\int_{\Gamma} \vec{F}(x, y, z)d\vec{r} = -V(x, y, z) \Big|_{\vec{a}}^{\vec{b}},$$

donde  $\vec{a}$  y  $\vec{b}$  son los extremos de dicha curva. En particular, para cualquier curva cerrada,

$$\int_{\Gamma} \vec{F}(x, y, z)d\vec{r} = \oint \vec{F}(x, y, z)d\vec{r} = 0.$$

Las fuerzas con estas características se denominan conservativas y los correspondientes sistemas: sistemas *conservativos*.

La razón de esta denominación se explica por lo siguiente: Si usamos la segunda ley de Newton

$$\vec{F}(x, y, z)d\vec{r} = \frac{d}{dt}(m\vec{v})d\vec{r} = m\vec{v}d\vec{v},$$

luego

$$\int_{\Gamma} \vec{F}(x, y, z) d\vec{r} = \int_{\Gamma} m\vec{v}d\vec{v} = \frac{1}{2}mv^2 \Big|_{\vec{a}}^{\vec{b}}.$$

Juntando esta expresión con la anterior tenemos

$$\left( \frac{1}{2}mv^2 + V(x, y, z) \right) \Big|_{\vec{a}}^{\vec{b}} = 0, \quad v^2 = \vec{v} \cdot \vec{v}.$$

Es decir, la cantidad

$$E = T(\vec{v}) + V(\vec{r})$$

vale lo mismo en los extremos de la curva  $\Gamma$ . Como  $\Gamma$  es arbitraria deducimos que  $E$  es una cantidad invariante en el tiempo. Esta cantidad se denomina energía mecánica del sistema.  $T = \frac{1}{2}mv^2 = \frac{p^2}{2m}$  se denomina energía cinética,  $V$ , energía potencial y  $p = m\vec{v}$ , impulso.

En la mecánica cuántica el formalismo es muy distinto y es el objetivo de este curso. Antes de pasar a discutirlo veamos brevemente otra forma de describir los sistemas mecánicos *clásicos*: El formalismo canónico o hamiltoniano.

## 1.1. Mecánica Hamiltoniana

Por sencillez seguiremos considerando el movimiento de una única partícula.

Vamos a suponer que el espacio físico es un espacio de fases  $(\vec{r}, \vec{p})$ , donde  $\vec{r} = (x, y, z) \in \mathbb{R}^3$  y  $\vec{p} = (p_x, p_y, p_z)$  denotan las componentes del vector posición y momento, respectivamente. Definamos una función  $H$  dependiente de la posición  $\vec{r}$  y el impulso  $\vec{p}$

$$H(\vec{r}, \vec{p}) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z),$$

que denominaremos hamiltoniano del sistema. Entonces, las ecuaciones dinámicas del sistema vienen dadas por las expresiones

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H}{\partial p_x}, & \frac{dp_x}{dt} &= -\frac{\partial H}{\partial x}, \\ \frac{dy}{dt} &= \frac{\partial H}{\partial p_y}, & \frac{dp_y}{dt} &= -\frac{\partial H}{\partial y}, \\ \frac{dz}{dt} &= \frac{\partial H}{\partial p_z}, & \frac{dp_z}{dt} &= -\frac{\partial H}{\partial z}. \end{aligned} \tag{1.1.1}$$

Está claro cómo se generaliza el problema. Supongamos que el hamiltoniano depende de las coordenadas *canónicas*  $q_1, \dots, q_N$  y sus correspondientes momentos  $p_1, \dots, p_N$ . Entonces las ecuaciones dinámicas son

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad i = 1, 2, \dots, N. \tag{1.1.2}$$

De esta forma la dinámica queda determinada por  $2N$  ecuaciones con  $2N$  incógnitas. Finalmente debemos destacar que, en general, las ecuaciones anteriores son equivalentes a las que se obtienen usando la segunda ley de Newton. Así, si

$$H(\vec{r}, \vec{p}) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z),$$

las ecuaciones (1.1.1) nos dan (sólo incluiremos las ecuaciones para la coordenada  $x$ )

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H}{\partial p_x} \implies v_x = \frac{p_x}{m}, \quad p_x = mv_x, \\ \frac{dp_x}{dt} &= -\frac{\partial H}{\partial x} \implies m \frac{dv_x}{dt} = -\frac{\partial V}{\partial x} = F_x \implies m \frac{d^2x}{dt^2} = F_x, \end{aligned}$$

es decir, recuperamos las ecuaciones de Newton de la mecánica clásica.

Dentro del formalismo canónico hamiltoniano hay una operación de especial importancia para entender el paso de la Mecánica clásica a la cuántica: las *llaves de Poisson*.

Se definen las llaves de Poisson de dos cantidades físicas  $A := A(q_1, \dots, q_N, p_1, \dots, p_N)$  y  $B := B(q_1, \dots, q_N, p_1, \dots, p_N)$ , funciones de las variables canónicas  $q_i$  y  $p_i$ ,  $i = 1, 2, \dots, N$ , a la cantidad

$$\{A, B\} := \sum_{k=1}^N \left( \frac{\partial A}{\partial q_k} \frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k} \frac{\partial B}{\partial q_k} \right). \quad (1.1.3)$$

Nótese que las ecuaciones de Hamilton (1.1.2) se pueden escribir como

$$\frac{dq_i}{dt} = \{q_i, H\}, \quad \frac{dp_i}{dt} = \{p_i, H\}, \quad i = 1, 2, \dots, N.$$

Nótese además que para las coordenadas canónicas se tiene

$$\{q_i, q_j\} = 0 = \{p_i, p_j\}, \quad \{q_i, p_j\} = \delta_{i,j}, \quad i, j = 1, 2, \dots, N. \quad (1.1.4)$$

En general se puede probar que la ecuación de evolución para cualquier cantidad física  $A$  es

$$\frac{dA}{dt} = \{A, H\}. \quad (1.1.5)$$

## 1.2. Dos ejemplos representativos

### 1.2.1. El oscilador armónico unidimensional

Comencemos con un sistema clásico de gran importancia: el oscilador armónico. Asumiremos que el eje de coordenadas está situado justo en la posición de equilibrio del oscilador, luego por  $x$  representaremos la desviación del sistema del punto de equilibrio. En este caso

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2}kx^2, \quad (1.2.1)$$

luego

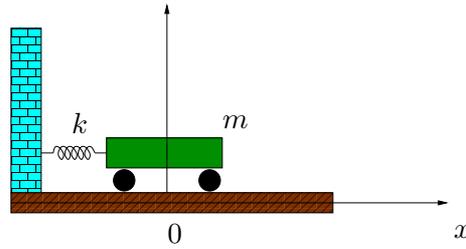


Figura 1.1: El oscilador armónico

$$\frac{dx}{dt} = \frac{\partial H(x, p)}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H(x, p)}{\partial x} \implies$$

$$\frac{dx}{dt} = \frac{p}{m}, \quad \frac{dp}{dt} = -kx \implies mx''(t) + kx(t) = 0.$$

Sus soluciones son:

$$x(t) = A \cos(\omega t + \delta), \quad \omega = \sqrt{\frac{k}{m}},$$

donde  $A$  y  $\delta$  dependerán de las condiciones iniciales  $x_0 = x(0)$ ,  $v_0 = v(0)$  y están dadas por  $v_0 = -\omega x_0 \tan \delta$  y  $x_0 = A \cos \delta$ .

En la figura 1.2 representamos dos soluciones correspondientes a fases  $\delta$  iguales y amplitudes distintas. Nótese que de la solución no se deducen ningún tipo de restricciones para los valores de  $A$  y  $\delta$ .

Si calculamos la energía:

$$E = T + V = \frac{1}{2}m[x(t)']^2 + \frac{1}{2}kx^2 = \frac{1}{2}kA^2 = \text{const.}$$

De lo anterior se deduce que la energía toma los valores reales  $E = \frac{1}{2}kA^2 \geq 0$  y es una cantidad continua.

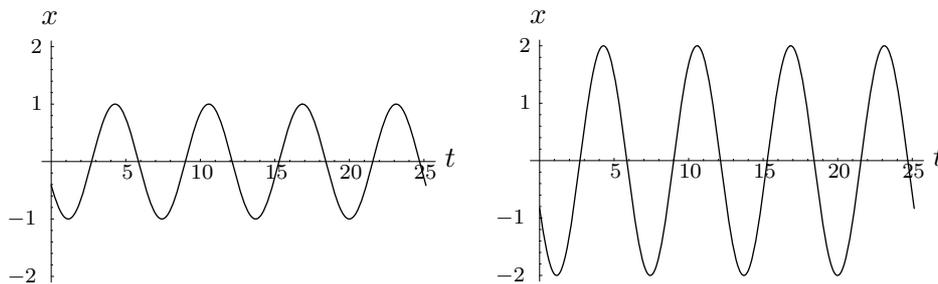


Figura 1.2: El oscilador armónico: soluciones

Obviamente este sistema es demasiado sencillo. Un caso más realista es el oscilador amortiguado, es decir, cuando hay rozamiento. En este caso la ecuación diferencial que se obtiene es  $mx''(t) + \alpha x'(t) + kx(t) = 0$ , donde  $\alpha > 0$  es el coeficiente de viscosidad del medio. Dejamos, al lector que resuelva y analice la ecuación como ejercicio.

### 1.2.2. Movimiento en un campo central de fuerzas

Veamos el caso correspondiente al potencial de una *fuerza central*, es decir,  $V(\vec{r}) = -\alpha/r$ . Ejemplos típicos de dicha fuerza son la fuerza gravitatoria y la electrostática.

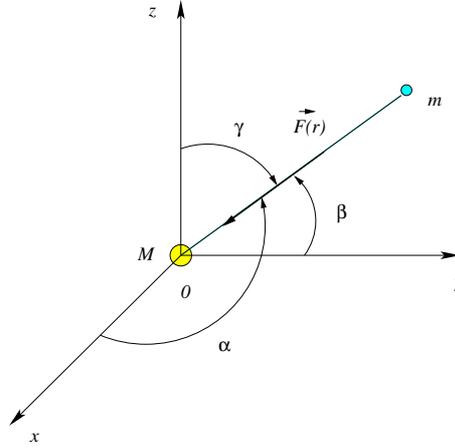


Figura 1.3: Fuerza central de interacción entre dos partículas.

En este caso tenemos

$$H(\vec{r}, \vec{p}) = \frac{p^2}{2m} - \frac{\alpha}{r}, \quad r = \sqrt{x^2 + y^2 + z^2}. \quad (1.2.2)$$

Entonces, las ecuaciones (1.1.1) nos dan

$$mx''(t) = -\alpha \frac{x}{r^3}, \quad my''(t) = -\alpha \frac{y}{r^3}, \quad mz''(t) = -\alpha \frac{z}{r^3}.$$

Vamos a considerar el movimiento de una partícula material de masa  $m$  en un campo de fuerzas centrales (ver figura 1.3). Sea  $\vec{F}(r)$  la fuerza dirigida al origen de coordenadas y que sólo depende de la distancia  $r$  al origen de coordenadas. Usando la Ley de Newton tenemos las siguientes ecuaciones para cada coordenada  $x$ ,  $y$  y  $z$ :

$$\begin{aligned} mx''(t) &= F_x(r), & my''(t) &= F_y(r), \\ mz''(t) &= F_z(r), & r &= \sqrt{x^2 + y^2 + z^2}. \end{aligned}$$

Pero

$$\begin{aligned} F_x(r) &= F(r) \cos \alpha = F(r) \frac{x}{r}, \\ F_y(r) &= F(r) \cos \beta = F(r) \frac{y}{r}, \\ F_z(r) &= F(r) \cos \gamma = F(r) \frac{z}{r}, \end{aligned}$$

luego las ecuaciones del movimiento de nuestra partícula son

$$mx''(t) = F(r) \frac{x}{r}, \quad my''(t) = F(r) \frac{y}{r}, \quad mz''(t) = F(r) \frac{z}{r}. \quad (1.2.3)$$

Vamos a probar que el movimiento de la partícula es plano. Para ello multiplicamos la primera ecuación en (1.2.3) por  $-y$ , la segunda por  $x$  y las sumamos. Ello nos da

$$m(xy'' - yx'') = 0 \implies xy'' - yx'' = 0.$$

Si ahora multiplicamos la segunda por  $z$ , la tercera por  $-y$  y sumamos tenemos

$$m(yz'' - zy'') = 0 \implies yz'' - zy'' = 0.$$

Finalmente, multiplicando la primera por  $z$ , la tercera por  $-x$  y sumando, obtenemos

$$m(zx'' - xz'') = 0 \implies zx'' - xz'' = 0.$$

Ahora bien, integrando por partes en la primera de las tres últimas ecuaciones vemos

$$c_1 = \int (xy'' - yx'') dt = xy' - \int x'y' dt - yx' + \int y'x' dt = xy' - yx'.$$

Análogamente tenemos, para las otras dos,

$$yz' - zy' = c_2, \quad zx' - xz' = c_3.$$

Si multiplicamos la primera de las tres últimas ecuaciones por  $z$ , la segunda por  $x$  y la tercera por  $y$  y las sumamos obtenemos

$$c_1z + c_2x + c_3y = 0,$$

que es precisamente la ecuación de un plano que pasa por el origen. Así pues, tenemos la siguiente propiedad:

**Propiedad 0.** *El movimiento de una partícula sometida a una fuerza central es plano, o sea, su trayectoria está contenida en un plano que pasa por el origen (hacia donde apunta dicha fuerza central).*

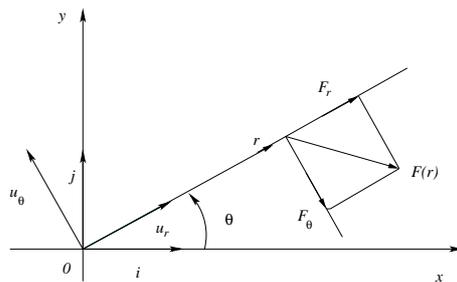


Figura 1.4: Componentes radial y angular de un vector.

Dado un vector cualquiera en el plano, siempre podemos descomponerlo en sus respectivas componentes. Todo vector se puede escribir en función de los vectores unitarios  $\vec{i}$  y  $\vec{j}$  que definen a los ejes  $x$  e  $y$ , respectivamente. En particular, los vectores unitarios de las direcciones radial  $\vec{u}_r$  y angular  $\vec{u}_\theta$  (ver figura 1.4)

$$\vec{u}_r = \vec{i} \cos \theta + \vec{j} \sin \theta, \quad \vec{u}_\theta = -\vec{i} \sin \theta + \vec{j} \cos \theta.$$

Nótese que entonces

$$\frac{d\vec{u}_r}{d\theta} = \vec{u}_\theta, \quad \frac{d\vec{u}_\theta}{d\theta} = -\vec{u}_r. \quad (1.2.4)$$

Además,  $\vec{F} = F_r\vec{u}_r + F_\theta\vec{u}_\theta$ , pero si la fuerza es radial entonces  $F_\theta = 0$ . Escribamos ahora las leyes de Newton para nuestra partícula en coordenadas polares. Comenzamos calculando el vector velocidad

$$\vec{v} = \frac{d(r\vec{u}_r)}{dt} = r\frac{d\vec{u}_r}{dt} + \frac{dr}{dt}\vec{u}_r = r\frac{d\vec{u}_r}{d\theta}\frac{d\theta}{dt} + \frac{dr}{dt}\vec{u}_r = r\frac{d\theta}{dt}\vec{u}_\theta + \frac{dr}{dt}\vec{u}_r, \quad (1.2.5)$$

donde hemos usado (1.2.4). Análogamente

$$\begin{aligned} \vec{a} &= \frac{d\vec{v}}{dt} = \frac{dr}{dt}\frac{d\theta}{dt}\vec{u}_\theta + r\frac{d^2\theta}{dt^2}\vec{u}_\theta + r\frac{d\theta}{dt}\frac{d\vec{u}_\theta}{dt} + \frac{d^2r}{dt^2}\vec{u}_r + \frac{dr}{dt}\frac{d\vec{u}_r}{dt} \\ &= \left[ r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} \right] \vec{u}_\theta + \left[ \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 \right] \vec{u}_r. \end{aligned}$$

Luego, usando la Ley de Newton, tenemos

$$\begin{aligned} 0 = F_\theta = ma_\theta &= m \left[ r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} \right] \implies \\ r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} &= \frac{d}{dt} \left[ r^2\frac{d\theta}{dt} \right] = 0. \end{aligned}$$

Es decir,  $r^2\theta'(t) = c$ , con  $c$  cierta constante. La propiedad anterior se conoce como Ley de las áreas pues el área que barre el radio vector  $\vec{r}$  es tal que  $dA(t) = r^2/2d\theta = r^2/2\theta'(t)dt$ , por tanto  $A'(t) = r^2/2\theta'(t) = c/2$ . Así, hemos probado la siguiente propiedad<sup>1</sup>

**Propiedad 1.** *El movimiento de una partícula sometida a una fuerza central es tal que el radio vector  $\vec{r}$  recorre áreas iguales en intervalos de tiempo iguales, i.e.,  $dA(t)/dt = r^2\theta'(t) = c$ .*

Supongamos ahora que la fuerza es de la forma

$$\vec{F}(r) = -\frac{\alpha m}{r^2}\vec{u}_r, \quad (1.2.6)$$

es decir, una ley del *inverso del cuadrado de la distancia*.

Nuevamente usando la ley de Newton, para la componente radial obtenemos

$$F_r(r) = ma_r \implies \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{\alpha}{r^2}. \quad (1.2.7)$$

Vamos a intentar resolver esta ecuación diferencial. Obviamente es una ecuación diferencial no lineal así que intentaremos convertirla en una ecuación lineal. Para ello haremos el cambio de variable  $r = 1/u(\theta)$  y pasaremos de la variable  $t$  a  $\theta$ , es decir, intentaremos dar con la ecuación de la trayectoria de nuestra partícula. Nótese que en las nuevas variables, la propiedad 1 de las áreas se escribe como  $\theta'(t) = cu^2$ . Tenemos

$$\frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt} = -\frac{u'(\theta)}{u^2(\theta)}cu^2(\theta) = -cu'(\theta),$$

<sup>1</sup>Esta propiedad se conoce como segunda Ley de Kepler, en honor al matemático y astrónomo J. Kepler quien la descubrió en la segunda mitad del siglo XVII cuando estudiaba la órbita del planeta Marte.

$$\frac{d^2r}{dt^2} = -c \frac{du'(\theta)}{dt} = -cu''(\theta)\theta'(t) = -c^2u^2(\theta)u''(\theta),$$

luego (1.2.7) se transforma en

$$u''(\theta) + u(\theta) = \frac{\alpha}{c^2},$$

que es lineal. La solución de la ecuación anterior la escribiremos en la forma

$$u(\theta) = c_1 \cos(\theta) + c_2 \sen(\theta) + \frac{\alpha}{c^2} = \frac{\alpha}{c^2} [1 + e \cos(\theta - \delta)],$$

luego, si escogemos los ejes coordenados de forma que  $\delta = 0$  –o lo que es lo mismo, que  $r$  sea mínimo cuando  $\theta = 0$ , i.e.,  $c_1 > 0$  y  $c_2 = 0$ – obtenemos la ecuación para la trayectoria

$$r(\theta) = \frac{c^2/\alpha}{1 + e \cos(\theta)}.$$

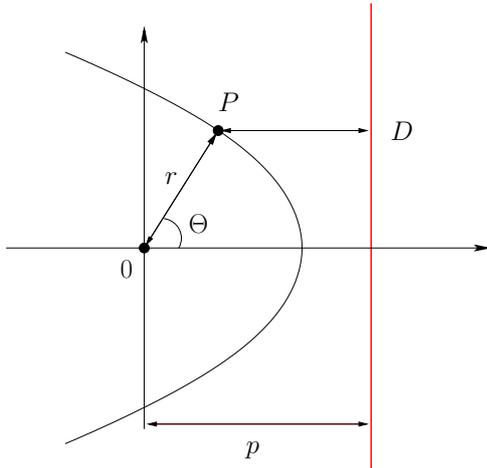


Figura 1.5: Propiedad de las secciones cónicas.

¿Qué figura geométrica define la ecuación anterior? Para ello recordemos que el lugar geométrico de los puntos  $P(x, y)$  tales que la distancia  $\overline{OP}$  de  $P$  a un punto fijo  $O$  (foco) es igual a la distancia  $\overline{PD}$  de  $p$  a una recta  $r$  dada (directriz) viene dado por la fórmula  $r = pe/(1 + e \cos \theta)$  (ver figura 1.5) siendo  $e$  cierta constante (excentricidad). Es conocido que si  $e < 1$  la curva es una elipse, si  $e = 1$  una parábola, y si  $e > 1$  una hipérbola. Así pues, hemos demostrado la siguiente propiedad:

**Propiedad 2.** *La trayectoria de una partícula sometida a una fuerza central es una sección cónica, es decir, una elipse, una parábola o una hipérbola.*

Es fácil comprobar que si una fuerza es central entonces existe una función  $U(x, y, z)$  tal que

$$F_x = -\frac{\partial U}{\partial x}, \quad F_y = -\frac{\partial U}{\partial y}, \quad F_z = -\frac{\partial U}{\partial z}.$$

En nuestro caso, además,  $U(x, y, z) = U(r) = -\alpha/r$ . Sumando las tres ecuaciones anteriores y usando la ley de Newton tenemos

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz = -m(x''dx + y''dy + z''dz) \implies$$

$$d\left(U + \frac{m}{2}(x'^2 + y'^2 + z'^2)\right) = 0.$$

En otras palabras, la función *energía*

$$E(t) = U + \frac{m}{2}v^2 = -\frac{\alpha}{r} + \frac{m}{2}v^2 \tag{1.2.8}$$

es constante. Esto es, como ya hemos visto, la Ley de Conservación de la Energía. Si aplicamos esta ley a nuestro sistema tenemos, usando (1.2.5),

$$E = \frac{m}{2}v^2 - \frac{\alpha m}{r} = \frac{m}{2}(v_\theta^2 + v_r^2) - \frac{\alpha m}{r} = \frac{m}{2}[(r\theta'(t))^2 + (r'(t))^2] - \frac{\alpha m}{r}.$$

Escojamos el momento de tiempo cuando  $\theta = 0$ . Usando la propiedad 1 tenemos  $(r\theta'(t))^2 = c^2/r^2$ , pero, según nuestra elección,  $r' = 0$  (es mínimo) y  $r = c^2[\alpha(1+e)]^{-1}$ , luego

$$E = \frac{m\alpha^2(e^2 - 1)}{2c^2} \implies e = \sqrt{1 + E \frac{2c^2}{m\alpha^2}}.$$

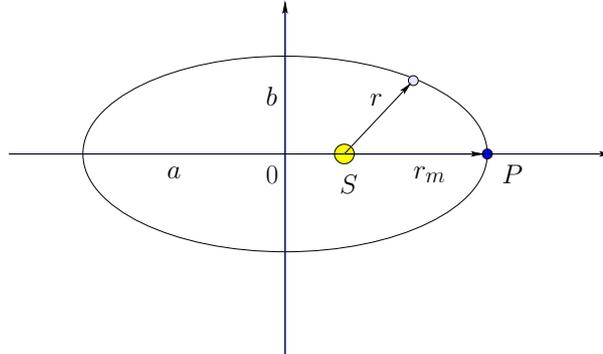


Figura 1.6: Movimiento de un planeta  $P$  alrededor del sol  $S$ .

En el caso de los planetas (ver figura 1.6), al pasar el planeta por el punto más cercano al foco (donde está el sol) se tiene, por la ley de Newton,

$$ma = \frac{\alpha m}{r_m^2} \implies v^2 = \frac{\alpha}{r_m},$$

pero entonces para la energía  $E$  obtenemos el valor

$$E = \frac{m}{2}v^2 - \frac{\alpha}{r} = -\frac{\alpha m}{2r_m} < 0,$$

y por tanto  $e < 1$ , o sea, los planetas se mueven siguiendo órbitas elípticas. Esta es la conocida primera Ley de Kepler.

Finalmente, tenemos que  $r_m = ae$ , y  $(1 - e^2) = b^2/a^2$ , donde  $a$  y  $b$  son los semiejes mayor y menor de la elipse, respectivamente. Entonces, como  $a$  es la semisuma de las distancias máxima y mínima de nuestra partícula (planeta) al foco,

$$a = \frac{1}{2} \left[ \frac{c^2/\alpha}{1+e} + \frac{c^2/\alpha}{1-e \cos} \right] \implies a = \frac{c^2}{\alpha(1-e^2)},$$

de donde deducimos que  $b^2 = c^2 a / \alpha$ . Si ahora usamos que el área de la elipse es  $A = \pi ab$ , y la propiedad 1, obtenemos que  $\pi ab = cT/2$  donde  $T$  es el tiempo que tarda la partícula en dar una vuelta entera sobre la órbita. Sustituyendo el valor de  $b$  en la expresión anterior obtenemos

$$\frac{T^2}{a^3} = \frac{4\pi}{\alpha}, \quad (1.2.9)$$

es decir, el cuadrado del período de revolución de los planetas es proporcional al cubo de sus distancias medias. La propiedad anterior se conoce como tercera Ley de Kepler. Nótese que la constante de proporcionalidad no depende para nada del planeta, sólo de  $\alpha$ , que según la Ley de Gravitación Universal es  $GM_s$ , donde  $G$  es la constante de gravitación universal y  $M_s$  la masa del sol.

## 1.3. Problemas

**Problema 1.3.1** *Estudia como se comporta una partícula material que se mueve en un campo potencial definido por las siguientes funciones potencial:*

$$V(x) = \begin{cases} U_0, & x < 0, \\ 0, & 0 < x < L, \\ U_0, & x > L, \end{cases} \quad V(x) = \begin{cases} 0, & x < 0, \\ U_0, & 0 < x < L, \\ 0, & x > L, \end{cases}$$

*conocidas como pozo potencial y barrera de potencial.*

**Problema 1.3.2** *Discutir lo que ocurre si en vez de dos masas interactuando según la ley de gravitación universal, lo que tenemos es un átomo de hidrógeno. En este caso la fórmula (1.2.6) se convierte en*

$$\vec{F}(r) = -\frac{ke^2}{r^2}\vec{u}_r. \quad (1.3.1)$$

Para terminar con este apartado y pasar a discutir lo que ocurre en el mundo cuántico tenemos que recordar que en el caso del átomo de hidrógeno hay un problema añadido y es que al ser el electrón una partícula cargada en movimiento, está continuamente emitiendo ondas electromagnéticas por lo que su energía va disminuyendo. Esto implica que el electrón va cayendo en espiral al núcleo. Este hecho contradice todos los experimentos conocidos.



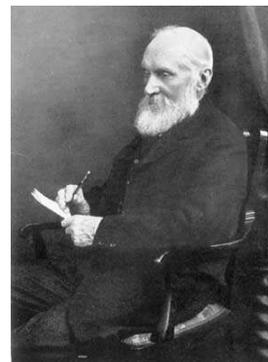
## Capítulo 2

# ¿Cómo se gestó la Mecánica cuántica?

Antes de entrar a discutir el formalismo y las leyes dinámicas de la Mecánica Cuántica *no relativista* vamos a dar un bosquejo histórico de la misma.

William Thomson o, como era conocido, Lord Kelvin, pronunció una conferencia el 27 de abril de 1900 <sup>1</sup> que contenía la siguiente, hoy célebre, frase:

*Hoy día la ciencia física forma, esencialmente, un conjunto perfectamente armonioso, ¡Un conjunto prácticamente acabado! Sólo quedan dos “nubecillas”: la primera, el resultado negativo del experimento de Michelson-Morley. La segunda, las profundas discrepancias de la ley de Rayleigh-Jeans.*



William Thomson

Nadie, en aquel momento, podía imaginar que esas dos nubecillas desembocarían en la *Teoría de la Relatividad* de Einstein y la Mecánica Cuántica, teorías que iban a cambiar radicalmente nuestra concepción de los fenómenos naturales y que representaron una Revolución comparable, en cierta forma, a la Revolución Copernicana y la aparición de los *Principia* de Newton.

En estas notas nos dedicaremos a estudiar las bases de la segunda de estas dos grandes teorías. Antes de entrar es los detalles vamos a dar una breve introducción histórica, necesaria para entender cómo el problema planteado por Lord Kelvin –su segunda “nubecilla”– y su resolución terminó en la moderna teoría cuántica.

---

<sup>1</sup>Esta frase atribuida a Lord Kelvin tiene cierta polémica. Lo que es un hecho es que un año más tarde Lord Kelvin publicó su discurso ampliándolo considerablemente con el título “Nineteenth-Century Clouds over the Dynamical Theory of Heat and Light” que apareció en la *Phil. Mag. S. 6*. Vol. 2. No. 7. July 1901, págs. 1–40. Al inicio de dicho trabajo podemos leer

*The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds. I. The first came into existence with the undulatory theory of light, and was dealt with by Fresnel and Dr. Thomas Young; it involved the question, How could the earth move through an elastic solid, such as essentially is the luminiferous ether? II. The second is the Maxwell-Boltzmann doctrine regarding the partition of energy.*

que esencialmente viene a decir lo mismo que la cita inicial.

## 2.1. La radiación del cuerpo negro

La ley de Rayleigh-Jeans es una fórmula que describe la radiación de un “cuerpo negro”. Todos sabemos que al calentar un cuerpo este cambia de color –quién no ha calentado de pequeño un clavo o un tornillo en el fogón de casa o ha visto el color que toma una parrilla al hacer una parrillada–. Probablemente nuestra curiosidad nos llevó a preguntarnos –al menos todo físico lo debería haber hecho– ¿cómo emiten la luz los cuerpos al calentarse?



Figura 2.1: Metal caliente (izquierda) y modelo de cuerpo negro (derecha).

Para resolver estas dudas, los físicos de finales del siglo XIX idealizaron un cuerpo cualquiera y construyeron un modelo que llamaron *cuerpo negro* –también, “cavidad”, etc–. La idealización consistía en que el cuerpo negro tenía que absorber –y por tanto emitir– ondas electromagnéticas en todo el espectro de frecuencias. De hecho, era un hecho bien conocido a finales del XIX que la luz *visible* estaba constituida por ondas electromagnéticas dentro de un cierto rango de frecuencias.

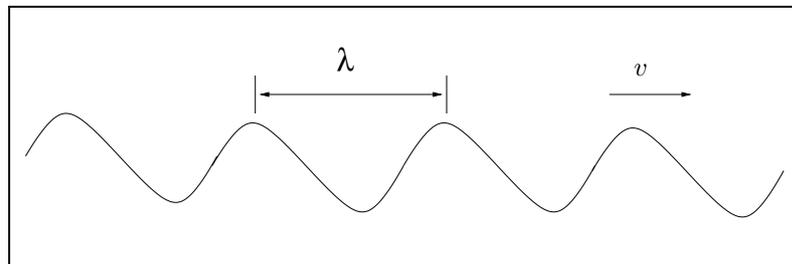


Figura 2.2: Representación esquemática de una onda plana unidimensional: Si  $v$  es la velocidad de la onda,  $T$  el periodo y  $\lambda$  la longitud de onda, entonces  $\lambda = vT$ , la frecuencia angular es  $\omega = 2\pi/T$ , y la frecuencia  $\nu = 1/T$ .

El primero en establecer una ley empírica para la radiación del cuerpo negro fue Wien quien obtuvo la fórmula  $\alpha\omega^3 e^{-\beta\omega/T}$ ,  $\alpha$  y  $\beta$  parámetros experimentales, pero ésta sólo correspondía a la parte ultravioleta –altas frecuencias, o equivalentemente, longitudes de onda pequeñas– del espectro y fallaba en la banda infrarroja –bajas frecuencias, o equivalentemente, longitudes de onda grandes–. Por otro lado, dos físicos ingleses, Rayleigh y Jeans, dedujeron una fórmula para la banda infrarroja pero que no era compatible con la fórmula de Wien.

Definamos la densidad de energía (cantidad de energía por unidad de volumen) mediante  $U(T)$ . Obviamente ésta dependerá de la temperatura  $T$ . Además, cada longitud de onda

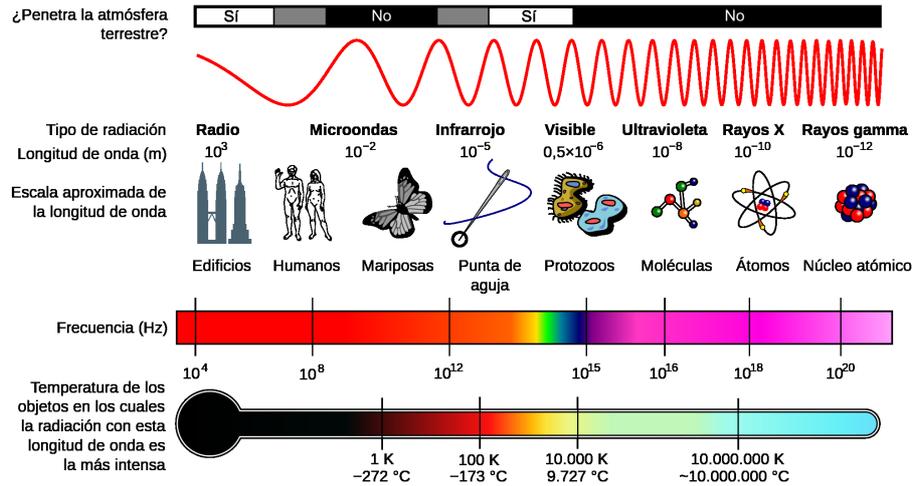


Figura 2.3: Diagrama del espectro electromagnético, mostrando el tipo, longitud de onda con ejemplos, frecuencia y temperatura de emisión de cuerpo negro (tomado de wikipedia). La luz visible está en el rango de 400-790 terahercios ( $10^{12}$  Hz)

–frecuencia– aportará su “granito de arena”, esta densidad por unidad de frecuencia  $\omega$  la denotaremos por  $u(\omega, T)$ , así

$$U(T) = \int_0^\infty u(\omega, T) d\omega. \tag{2.1.1}$$

La fórmula de Rayleigh-Jeans establecía que

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} kT,$$

donde  $c$  es la velocidad de la luz y  $k$  es la constante de Ludwig Boltzmann. En particular, para frecuencias muy altas (zona de frecuencias ultravioletas) es muy grande lo cual está en contradicción con las mediciones experimentales. Esta es la llamada catástrofe ultravioleta y no es más que la segunda nubecilla de Lord Kelvin (ver figura 2.4). Nótese que al sustituir  $u(\omega, T)$  en (2.1.1) obtenemos  $U(T) = \infty$  lo cual no tiene sentido.

La manera de obtener esta fórmula es muy sencilla. Era un hecho establecido en el siglo XIX que el número de ondas estacionarias con frecuencias por unidad de volumen en el interior de un cuerpo en el intervalo  $[\omega, \omega + \Delta\omega]$  era  $dn_\omega = \frac{\omega^2}{\pi^2 c^3}$ . La suposición de Rayleigh y Jeans consistía en que a cada oscilación electromagnética le correspondía, en media, una energía  $\langle \varepsilon \rangle$  igual a  $kT$ . Así

$$u(\omega, T) = \langle \varepsilon \rangle dn_\omega = \frac{\omega^2}{\pi^2 c^3} kT. \tag{2.1.2}$$

El próximo paso lo dio Max Planck. Desde el punto de vista de la física clásica la deducción de Rayleigh y Jeans era impecable, por tanto Planck asumió que debía haber alguna “ley” importante sin descubrir. En octubre de 1900 Planck encontró empíricamente una fórmula que describía perfectamente la ley experimental para la radiación del cuerpo negro. Dicha fórmula

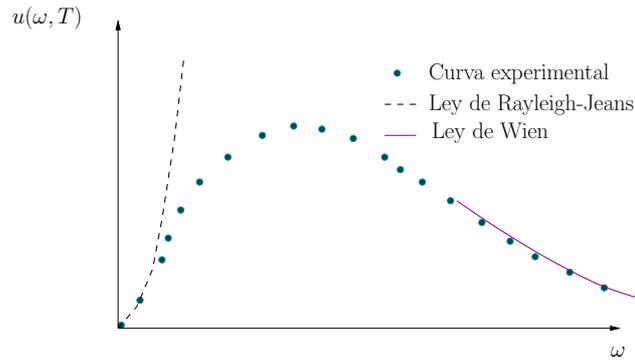


Figura 2.4: Gráfica de la intensidad  $u(\omega, T)$  contra la frecuencia de onda  $\omega$ .

para la energía media de la onda era la siguiente:

$$\langle \varepsilon \rangle = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1},$$

donde  $\hbar$  era cierta constante desconocida. Si  $\hbar\omega/kT \ll 1$ , entonces la fórmula de Planck daba  $\langle \varepsilon \rangle \approx kT$ . Además, si  $\hbar\omega/kT \gg 1$ , Planck recuperaba la fórmula de Wien.



Max Planck

Para explicar su fórmula Planck, rompiendo la concepción clásica, lanza la idea de que los “osciladores” que componen los átomos absorben o emiten luz no de forma continua, como era habitual en la física clásica, sino mediante porciones aisladas proporcionales a la frecuencia, es decir la energía se emitía o absorbía mediante “quantas” de energía  $E = \hbar\omega$ .

Hay varias razones que justifican su audacia. La primera es que Planck consideró su hipótesis como un artificio matemático pues le permitía deducir su fórmula empírica. En efecto, años antes en gran físico austriaco Ludwig Boltzmann había demostrado que la probabilidad de que un sistema en equilibrio tuviese una energía  $E$  era proporcional a la cantidad  $\exp\left(-\frac{E}{kT}\right)$ . Si  $E$  sólo puede tomar valores que sean múltiplos enteros de  $\hbar\omega$ ,  $E_n = n\hbar\omega$ , entonces la probabilidad  $p_n$  de cada uno de estos valores de energía es  $p_n = \exp\left(-\frac{n\hbar\omega}{kT}\right)$  y por tanto la energía media es

$$\langle \varepsilon \rangle = \sum_{n=0}^{\infty} p_n E_n = \frac{\sum_{n=0}^{\infty} n\hbar\omega \exp\left(-\frac{n\hbar\omega}{kT}\right)}{\sum_{n=0}^{\infty} \exp\left(-\frac{n\hbar\omega}{kT}\right)} = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1},$$

que es justamente la fórmula que Planck había encontrado empíricamente. Si ahora multiplica-

mos por  $dn_\omega = \frac{\omega^2}{\pi^2 c^3}$  obtenemos

$$u(\omega, T) = \langle \varepsilon \rangle dn_\omega = \frac{\hbar \omega^3}{4\pi^2 c^2} \frac{1}{\exp\left(\frac{\hbar \omega}{kT}\right) - 1}, \quad (2.1.3)$$

conocida como fórmula de Planck para la densidad de energía de un cuerpo negro.

La segunda razón fue que los cuanta de Planck eran absorbidos o emitidos por la materia que componía al cuerpo negro: es decir sus átomos. Y la estructura atómica era algo desconocido a principios del siglo XX, incluso muchos modelos de la materia consideraban a ésta como un conjunto de osciladores, por lo que no era del todo descabellada la idea de Planck de que estos osciladores sólo pudiesen absorber la energía o emitirla mediante porciones individuales –podría ser una especie de resonancia, o algo parecido–. No obstante para Planck la luz seguía siendo una onda perfectamente continua y perfectamente descrita por las leyes de Maxwell.

Antes de continuar nuestra historia debemos hacer un breve paréntesis para explicar cuáles fueron las principales causas de que la hipótesis de Planck calara tan hondo en la física de principios de siglo.

El primer hecho importante es que, como ya hemos mencionado, la fórmula de Planck se correspondía exactamente con la curva experimental para el cuerpo negro obtenida en los laboratorios. Pero además permitía resolver una de las paradojas de la física clásica: la denominada “catástrofe ultravioleta”, consecuencia de la fórmula de Rayleigh-Jeans. Si la fórmula (2.1.2) era correcta entonces la densidad de equilibrio de la energía  $u(T)$  de la radiación (2.1.1) daba infinito, es decir que nunca se alcanzaría el equilibrio termodinámico entre la materia y la radiación. Si usamos (2.1.3) tenemos

$$U(T) = \int_0^\infty u(\omega, T) d\omega = \frac{\pi^2 k^4}{60 c^2 \hbar^3} T^4 = \sigma T^4,$$

donde  $\sigma$  es una constante conocida también como constante de Boltzmann para la densidad o luminosidad de la energía de radiación. Planck calculó además el valor de la constante  $\hbar$  en su fórmula para que se correspondiera con los valores experimentales obteniendo el valor aproximado  $1,05 \times 10^{-34}$  Jules por segundo<sup>2</sup>. Cuando sustituyó este valor en la fórmula para  $\sigma$  obtuvo exactamente el valor numérico de ésta. Es decir, Planck resolvió de forma brillante la segunda de las nubecillas de Lord Kelvin.



Henri Poincaré

Estos resultados fueron publicados por Planck el 14 de octubre de 1900, día oficial del nacimiento de la teoría cuántica. ¡Quién iba a imaginar que este trabajo iba a revolucionar la física por completo!

El segundo hecho, que además fue crucial en la historia de la mecánica cuántica, fue la prueba de Henri Poincaré (en el otoño de 1911) de que la distribución

$$\langle \varepsilon \rangle = \frac{\hbar \omega}{\exp\left(\frac{\hbar \omega}{kT}\right) - 1},$$

<sup>2</sup>Muchas veces en vez de  $\hbar$  se usa el valor  $h = 2\pi\hbar = 6,62 \times 10^{-34}$  julios×radianes por segundos.

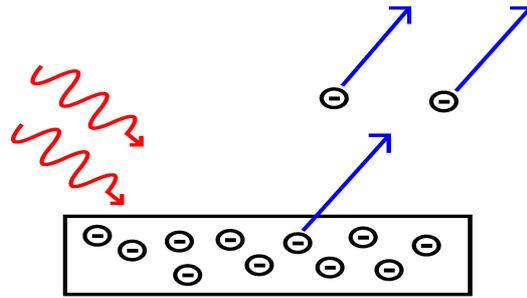


Figura 2.5: Efecto fotoeléctrico

sólo se podía obtener bajo la suposición de que la energía está cuantizada (véase el magnífico artículo de Hendrik Lorentz *Deux Mémoires de Henri Poincaré sur la Physique Mathématique* escrito en 1914 y publicado en *Acta Mathematica* Vol. 38, (1921) págs. 293-308), es decir desde el punto de vista matemático la fórmula de Planck sólo puede ser deducida bajo la suposición de que la energía se emite y absorbe en porciones discretas de energía algo que era totalmente distinto a la concepción clásica del mundo que se tenía hasta la fecha.

## 2.2. Einstein y el efecto fotoeléctrico

El siguiente paso lo dio Einstein en 1905 en un ensayo titulado *Sobre un punto de vista heurístico acerca de la producción y la transformación de la luz* que recibió el mismo Max Planck para su publicación en los *Annalen der Physik*.



Albert Einstein

Einstein, totalmente seducido por los quanta de Planck, lanza la hipótesis de que no sólo los “osciladores” materiales emitían energía cuantificadamente, sino también los “osciladores” lumínicos. Einstein retomando las ideas corpusculares sobre la luz –Newton ya había desarrollado una teoría corpuscular a finales del siglo XVII– considera la luz formada por partículas de masa cero y energía  $\hbar\omega$ : los fotones. Utilizando estas hipótesis dio una explicación sencillísima al *efecto fotoeléctrico* que había descubierto Hertz en 1887 y que seguía sin tener una explicación satisfactoria.

Hertz había descubierto que una placa metálica sometida a una luz ultravioleta –altas frecuencias– emitía electrones. Años más tarde se comprobó que el número de dichos electrones aumentaba proporcionalmente a la intensidad de la radiación pero “incomprensiblemente” la velocidad de éstos no dependía de la intensidad sino de la frecuencia de luz: a mayor frecuencia, mayor velocidad. Además si la frecuencia era lo suficientemente baja –o la longitud de onda muy grande– ya no se emitían electrones independientemente de lo intensa que fuese la luz incidente.

Para resolver el problema Einstein razonó como sigue: Supongamos que usamos una luz monocromática compuesta por quantas de energía luminosa  $\hbar\omega$ , es decir que estamos bombardeando la lámina metálica con partículas luminosas cada una de las cuales tiene una energía

$\hbar\omega$ . Si denotamos por  $W$  la energía necesaria para extraer un electrón del metal, entonces la energía cinética de los electrones,  $E_c$ , se expresará mediante la fórmula

$$E_c = \hbar\omega - W. \quad (2.2.1)$$

¡Todas las observaciones descritas anteriormente son una consecuencia de la fórmula anterior! Años más tarde, el físico estadounidense Robert Millikan comprueba experimentalmente la fórmula de Einstein (2.2.1) y encuentra que  $\hbar$  es la misma  $\hbar$  de Planck. Esta fue la gota que colmó el vaso, pues al parecer la luz, que era aceptada unánimemente por todos como un fenómeno continuo, tenía al parecer cierta naturaleza corpuscular.

## 2.3. Bohr y el modelo atómico

El próximo paso en la historia lo dio el danés Niels Bohr.

Era un hecho aceptado en 1913 que el átomo estaba constituido por un núcleo “pesado” y “denso” que contenía toda la materia del átomo y electrones girando a su alrededor. Este modelo, una especie de sistema solar en miniatura, sencillo y funcional fue propuesto por Rutherford a partir de los resultados de experimentos de dispersión.<sup>3</sup> Sólo tenía un “pequeño” problema: como toda carga en movimiento acelerado emite ondas electromagnéticas, los electrones que giraban alrededor del núcleo debían perder energía y caer al núcleo –además en un tiempo récord:  $10^{-5}$  segundos–. Por tanto, tenía que haber alguna forma de poder “retener” a los electrones en sus órbitas.

Otro hecho curioso tiene que ver con la forma en que radian energía los elementos químicos. Ya sabemos que al calentar un cuerpo este emite energía en todo el espectro electromagnético. Sin embargo, al observar la luz que pasa a través de ciertos elementos en estado gaseoso se veían ciertas franjas negras muy finas (líneas de absorción) en el espectro de la misma, es decir como si faltaran ciertas frecuencias (o longitudes de onda). Algo similar ocurría al excitar ciertos elementos y observar la luz que desprendían: aparecían ciertas líneas brillantes (líneas de emisión) en determinadas longitudes de onda. De hecho para cada elemento había una única secuencia de dichas líneas. Ese es el principio básico de la espectroscopía y es lo que ha permitido, entre otras cosas, saber la composición química de las estrellas, que de otra forma sería imposible de determinar.

En el caso del Hidrógeno era un hecho conocido experimentalmente sus líneas espectrales respondían a una fórmula conocida como fórmula de Balmer (para  $k = 2$  se tiene la conocida serie de Balmer en el espectro visible) y sus generalizaciones que establecía la siguiente expresión para el inverso de la longitud de onda

$$\frac{1}{\lambda} = R_H \left( \frac{1}{k^2} - \frac{1}{n^2} \right), \quad n, k = 1, 2, 3, \dots, n \neq k,$$

siendo  $R_H$  la constante de Rydberg que tomaba el valor  $1,0976776534 \times 10^7$  1/metros. Llamaba tremendamente la atención de que los valores que  $n$  y  $k$  eran números enteros.

---

<sup>3</sup>Rutherford bombardeaba átomos de oro con núcleos de helio (partículas alpha) cuando descubrió que para ángulos muy bajos (es decir cuando se disparaban las partículas alpha frontalmente contra los núcleos de oro) había una gran cantidad de ellas que rebotaban en sentido contrario, como cuando una bola pequeña ligera contra una muy pesada. Eso llevó a Rutherford a deducir que la estructura del átomo era parecida a un sistema solar en miniatura: un núcleo muy denso y pesado en el centro con electrones ligeros girando a su alrededor.

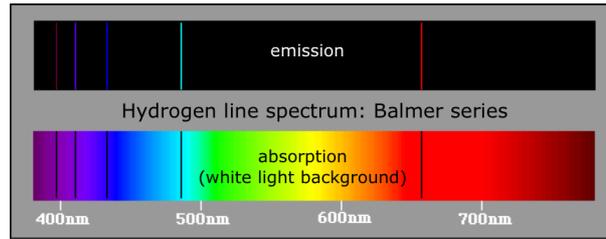


Figura 2.6: La serie de Balmer del átomo de hidrógeno



Niels Bohr

El razonamiento de Bohr fue bastante lógico: obligar al electrón que se mantuviera en ciertas órbitas permitidas (estables) y para pasar de una de estas órbitas a otra este debería “saltar” por encima de todas aquellas no permitidas. Dichas órbitas, que Bohr consideró circulares, debían ser tales que

- De todas las infinitas órbitas posibles sólo son posibles aquellas en la que su momento angular  $L = mvr$ , siendo  $m$  la masa del electrón,  $v$ , su velocidad y  $r$  el radio de la órbita, fuesen múltiplos enteros de  $\hbar$ , i.e.  $mrv = n\hbar$ .
- La energía que absorbe o emite un átomo al saltar un electrón de una órbita permitida a otra es igual a  $\hbar\omega$ , es decir para saltar de una órbita a otra el átomo absorbe o emite un quanta de luz.

Veamos las consecuencias de las hipótesis cuánticas de Bohr.

Si usamos la ley de Newton  $F = ma$ , tenemos

$$m \frac{v^2}{r} = \frac{k_e e^2}{r^2}.$$

Utilicemos que  $L = mvr = n\hbar$  para eliminar la velocidad  $v$ . Ello nos da el valor de los radios  $r_n$  de las órbitas permitidas

$$r_n = \frac{\hbar^2 n^2}{mk_e e^2}.$$

La energía de la órbita es, por tanto<sup>4</sup>,

$$E_n = \frac{1}{2}mv^2 - \frac{k_e e^2}{r} = -\frac{k_e e^2}{2r} = -\frac{mk_e^2 e^4}{2\hbar^2} \frac{1}{n^2}. \quad (2.3.1)$$

Entonces, el salto entre dos órbitas daba para la frecuencia del fotón emitido el valor:

$$\omega = \frac{E_n - E_m}{\hbar} = \frac{mk_e^2 e^4}{2\hbar^3} \left( \frac{1}{k^2} - \frac{1}{n^2} \right), \quad n, k = 1, 2, 3, \dots$$

Cuando Bohr sustituyó en su fórmula el valor de  $m$ ,  $e$  y  $\hbar$ , tomando para esta última el valor de encontró Planck, obtuvo justamente el valor de la constante de Rydberg  $R_H$  (recordemos

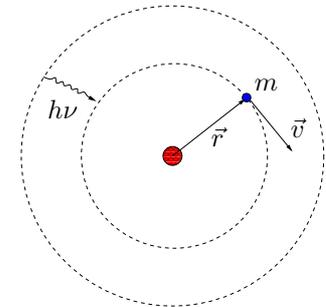


Figura 2.7: El átomo de Bohr

<sup>4</sup>Hemos sustituido el valor de  $v$  que nos da la ley de Newton.

que  $\lambda = (2\pi c/\omega)$  donde  $c$  es la velocidad de la luz. Es decir,  $\hbar$  era más que una simple constante introducida artificialmente por Planck: era una de las constantes más importantes de la naturaleza.

Aunque el modelo de Bohr era muy funcional y explicaba muchos fenómenos, este continuaba siendo muy incompleto y además tenía demasiados interrogantes. ¿De dónde salía ese extraño postulado sobre las órbitas? Este era quizá el punto más “oscuro” de toda la teoría. No obstante, en su artículo Bohr sienta las bases de lo que luego se denominó el *Principio de correspondencia de Bohr*. El dicho trabajo Bohr postula cómo debía ser la teoría cuántica: ésta tenía que ser tal que, para números cuánticos grandes, por ejemplo  $n$  en las fórmulas anteriores, se transformase en la teoría clásica.

Años más tarde, Bohr junto a Sommerfeld mejoran mucho el modelo atómico inicial incluyendo órbitas elípticas, entre otras cosas. Pero no es hasta 1925-1926 que no nace la nueva teoría cuántica a manos de un joven físico alemán: Werner Heisenberg y un físico austriaco: Erwin Schrödinger.

## 2.4. El nacimiento de la Mecánica Cuántica

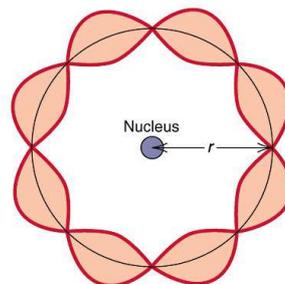
### 2.4.1. La dualidad onda-partícula

Antes de poder describir los fundamentos de la Mecánica Cuántica tenemos que detenernos en un personaje singular: el francés Louis de Broglie. Luis de Broglie era el hermano pequeño del Marqués Maurice De Broglie, un afamado físico experimental francés que dedicaba gran parte de su tiempo y dinero a la investigación experimental –varias veces fue propuesto para el Nobel de Físicas–. Un día de 1923 Louis, influenciado por el trabajo de Einstein sobre el efecto fotoeléctrico, postula que la dualidad onda-partícula que Einstein había proclamado para la luz también había de ser cierta para las partículas materiales, como por ejemplo, el electrón: Sus palabras fueron

*“En la Óptica durante siglos ha sido demasiado despreciado el método corpuscular de estudio en comparación con el ondulatorio. ¿No se habrá cometido el error inverso en la teoría sobre la materia?”.*



Louis de Broglie



Las órbitas del átomo de Bohr según De Broglie

Si Einstein había recuperado la propiedad corpuscular para la luz, De Broglie la postuló la propiedad ondulatoria para la materia. Al enterarse Einstein de las afirmaciones y trabajos de De Broglie afirmó: “*De Broglie ha levantado un extremo del gran velo*”. Aunque Louis no consiguió convencer a ninguno de los físicos del laboratorio de su hermano para que verificase su hipótesis –éstos estaban enfrascados en otros muchos proyectos, como los rayos X–, en 1927 C. Davisson y L. Germer<sup>5</sup> publicaron un artículo donde explicaban el descubrimiento de una figura de difracción al estudiar la dispersión de electrones en un cristal de níquel que luego fue corroborada independientemente por G.P. Thompson y por P. Tartakovsky.

La *genialidad* de De Broglie fue equiparar un electrón a una onda plana. Por ejemplo, si tenemos un electrón de masa  $m$  y velocidad  $v$ , De Broglie postuló que el momento (impulso)  $p$  del electrón era

$$p = mv = \frac{E}{v} = \frac{\hbar\omega}{v} = \frac{\hbar 2\pi}{vT} = \hbar \frac{2\pi}{\lambda},$$

de esta forma De Broglie daba un significado físico a las órbitas de Bohr: éstas eran justo aquellas órbitas tales que el cociente entre su longitud y la longitud de onda del electrón era un número entero, es decir era una analogía completa a las ondas estacionarias sobre un anillo (círculo).

### 2.4.2. La Mecánica matricial de Heisenberg

La hipótesis de De Broglie fue la clave para una de las dos formulaciones de la Mecánica cuántica: La mecánica ondulatoria de Schrödinger. Pero antes debemos comentar la primera versión de la Mecánica cuántica, la *mecánica matricial*, nacida en 1925 de la mano del joven físico alemán Werner Heisenberg.



Werner Heisenberg

En opinión de Heisenberg, una teoría física correcta ha de hacer uso única y exclusivamente de cantidades o magnitudes observables. Luego haciendo uso del principio de correspondencia de Bohr se lanzó a entender los estados estacionarios del átomo. Su razonamiento era, aproximadamente el siguiente: Una carga en movimiento con una determinada frecuencia debía emitir radiación con dicha frecuencia –como en la teoría clásica–. Este hecho era una consecuencia matemática del análisis de Fourier que Heisenberg aplicaba al mundo cuántico. Como las frecuencias del espectro dependían de dos índices  $\omega_{n,m}$  (véase la fórmula de Balmer), Heisenberg postulaba que debía haber tantos índices como estados estacionarios –no sólo como niveles de energía, pues se sabía que las series espectrales se modificaban al introducir los átomos en fuertes campos magnéticos–. A continuación da un salto cualitativo al afirmar que toda magnitud física clásica  $a(t)$  debe transformarse en el conjunto  $A_{nm}(t)$ . Así, por ejemplo la posición del electrón  $x(t)$  debía ser sustituida por una tabla  $X_{nm}(t)$ . A continuación Heisenberg razona como habría de calcularse  $X_{nm}^2(t)$  hasta obtener la fórmula

$$X_{nm}^2(t) = \sum_k X_{nk}(t)X_{km}(t),$$

<sup>5</sup>Curiosamente Davisson y Germer llevaban haciendo experimentos con electrones desde 1921. En 1925 de forma casual descubrieron al bombardear con electrones cristales de níquel un cuadro de difracción que probaba la dualidad onda partícula.

es decir, las cantidades  $X_{nm}$  eran matrices. Finalmente, deduce, siempre razonando sobre el principio de correspondencia de Bohr que la dinámica que rige las magnitudes cuánticas ha de ser:

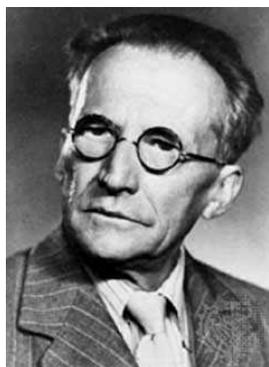
$$\frac{dX}{dt} = \frac{i}{\hbar}(HX - XH) = \frac{i}{\hbar}[H, X], \quad i = \sqrt{-1}, \quad (2.4.1)$$

donde  $H$  representaba la matriz del Hamiltoniano del sistema. Como  $H$ , representa el Hamiltoniano, es decir la energía, y obviamente  $[H, H] = 0$ , entonces, de (2.4.1) se tenía la conservación de la energía. En particular Heisenberg, junto a dos colegas alemanes de Gotinga, M. Born y P. Jordan, desarrollan toda una Mecánica matricial que se ajustaba muy bien a las observaciones –pocas, por cierto– de la época. Aparte de sus “aberrantes” matrices –como les llamaban los físicos de la época, especialmente Schrödinger– Heisenberg descubrió, o más bien postuló, un principio tremendamente polémico: el principio de incertidumbre de Heisenberg. De hecho en sus primeros razonamientos para construir la mecánica matricial Heisenberg descubre la imposibilidad de conocer al mismo tiempo y con una precisión arbitraria la posición y la velocidad del electrón.

Obviamente ese cambio radical no fue bien recibido por la mayoría de los físicos. En primer lugar representaba un cambio drástico de pensamiento –no se podía medir con precisión arbitraria al mismo tiempo ciertas cantidades como la posición y la velocidad de una partícula– en segundo, su aparato matemático –esas “aberrantes” matrices– era lo suficientemente complicado para que no estuviera al alcance de cualquier físico en aquella época. Por eso no es de extrañar que pronto apareciera una formulación alternativa.

### 2.4.3. La Mecánica ondulatoria de Schrödinger

En efecto, en 1926, Erwin Schrödinger andaba buscando una teoría que acabase con esa “aberración” de las matrices que Heisenberg intentaba introducir en la física, cuando, a sugerencia de P. Debye, estudia el trabajo de De Broglie publicado 1924. Según el mismo Schrödinger, un simple vistazo le fue suficiente para dar con la idea: asociar a cada partícula una onda y construir la ecuación diferencial que gobierna dicha onda. Al principio Schrödinger intentó construir una teoría ondulatoria para un electrón atrapado en un átomo. Comienza con un modelo relativista pero no le sale bien y decide estudiar qué ocurre en el caso no relativista.



La idea de Schrödinger era muy simple: Supongamos que tenemos una onda  $\Psi(x, t)$  asociada a un sistema clásico cuya energía viene dada por la función de Hamilton, el Hamiltoniano,  $H(x, p)$ , siendo  $x$  la coordenada y  $p$  el impulso. Entonces, después de un “largo” proceso de prueba y error, y usando el principio de correspondencia de Bohr así como distintos elementos de la mecánica analítica–las ecuaciones de Hamilton-Jacobi, por ejemplo–, Schrödinger **postuló** que la ecuación para una onda  $\Psi(x, t)$  debía ser, en el caso estacionario, es decir cuando no hay dependencia del tiempo,

$$H(x, \hat{p}) \Psi(x, t) = E \Psi(x, t), \quad \hat{p} = -i\hbar \frac{\partial}{\partial x},$$

Erwin Schrödinger donde  $E$  era la energía del sistema asociado a la onda  $\Psi$ . Es decir, en el caso cuando tenemos un Hamiltoniano estándar,

$$H(x, p) = \frac{p^2}{2m} + V(x),$$

siendo  $V(x)$  la función potencial (energía potencial), se tiene la ecuación de diferencial

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi + V(x) \Psi(x, t) = E \Psi(x, t). \quad (2.4.2)$$

Una de las pruebas de fuego de su ecuación fue el caso  $V(x) = 0$ , es decir cuando se tenía el movimiento de una partícula libre. Si consideramos el caso unidimensional, y hacemos  $V(x) = 0$ , la solución debía ser una onda plana del tipo

$$\Psi(x, t) = A \cos(kx + \omega t), \quad k = \frac{2\pi}{\lambda}.$$

Si sustituimos este valor en la ecuación de Schrödinger tenemos el valor

$$E = \frac{\hbar^2 k^2}{2m},$$

que igualado con el valor de la energía cinética –recordemos que  $V(x) = 0$ – nos da

$$p = \hbar k = \frac{2\pi\hbar}{\lambda},$$

que justamente era la fórmula que había postulado De Broglie.

Pero su mayor éxito estaba por llegar. Schrödinger decidió aplicar su ecuación al átomo de hidrógeno. Como en este caso el potencial era

$$V(r) = -\frac{e^2}{r}, \quad r = \sqrt{x^2 + y^2 + z^2},$$

obtuvo la ecuación

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi(x, y, z) - \frac{e^2}{r} \Psi(x, y, z) = E \Psi(x, y, z). \quad (2.4.3)$$

A esta ecuación volveremos más adelante. Lo importante era que Schrödinger sabía como tratar este tipo de ecuaciones y la resolvió. Primero la escribió en coordenadas esféricas y luego aplicó la separación de variables. La parte angular del laplaciano,  $\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ , en coordenadas esféricas era muy sencilla de resolver apareciendo las funciones o armónicos esféricos de Laplace. En particular Schrödinger obtuvo para los valores de la energía en el estado estacionario del átomo de hidrógeno la fórmula

$$E_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2},$$

que era la misma de Bohr (2.3.1).

Finalmente Schrödinger, igual que hizo Heisenberg, “dedujo” una ecuación para la dinámica de un sistema, que en el caso unidimensional tiene la forma

$$H \left( x, -i\hbar \frac{\partial}{\partial x} \right) \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t).$$

Si tenemos un Hamiltoniano estándar, ésta se transforma en

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) + V(x) \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t). \quad (2.4.4)$$

En particular, de (2.4.4) se podía deducir fácilmente la ecuación (2.4.2), para los sistemas estacionarios, introduciendo la factorización

$$\Psi(x, t) = \Psi(x) \exp \left( -\frac{iEt}{\hbar} \right).$$

### 2.4.4. Una “deducción” de la ecuación de Schrödinger

La ecuación de Schrödinger **no se puede deducir**, es sencillamente un postulado impuesto, o mejor, descubierto. No obstante, existe un razonamiento muy sencillo que permite *dar* con la forma de esta ecuación.

Supongamos que tenemos una partícula libre y le asociamos cierta onda plana

$$\Psi(x, t) = Ae^{i(kx - \omega t)}, \quad k = \frac{2\pi}{\lambda}.$$

Según De Broglie  $p = 2\pi\hbar/\lambda$ , así que  $k = p/\hbar$ , y además, usando la fórmula de Planck  $\omega = E/\hbar$ ,

$$\Psi(x, t) = Ae^{\frac{i}{\hbar}(px - Et)}.$$

Ahora hacemos,

$$\frac{\partial^2}{\partial x^2}\Psi(x, t) = -\frac{p^2}{\hbar^2}\Psi(x, t), \quad \text{por tanto} \quad p^2 = -\frac{\hbar^2}{\Psi(x, t)}\frac{\partial^2}{\partial x^2}\Psi(x, t),$$

pero

$$E = \frac{p^2}{2m} \Rightarrow E = -\frac{\hbar^2}{2m\Psi(x, t)}\frac{\partial^2}{\partial x^2}\Psi(x, t), \quad \text{o} \quad -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x, t) = E\Psi(x, t).$$

Si en vez de una partícula libre tenemos una ligada mediante un potencial  $V(x)$ , entonces

$$E = \frac{p^2}{2m} + V(x), \quad E = -\frac{\hbar^2}{2m\Psi(x, t)}\frac{\partial^2}{\partial x^2}\Psi(x, t) + V(x),$$

o, equivalentemente,

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x, t) + V(x)\Psi(x, t) = E\Psi(x, t).$$

Del razonamiento anterior en particular se deduce que al momento  $p$  de una partícula le corresponde el operador

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}.$$

## 2.5. La interpretación de la Mecánica cuántica

La ecuación de Schrödinger trajo cierta paz y tranquilidad a la Física Teórica pues aparentemente dejaba de lado la “infame” mecánica matricial de Heisenberg así que no es de extrañar que la serie de trabajos que publicó Schrödinger en 1926 calaron muy hondo en los físicos de la época. En particular, aparentemente se perdía la cualidad discreta del modelo de Bohr al aparecer nuevamente las “ondas”, en este caso la onda  $\Psi$ . En problema ahora era ¿qué significado físico tenía esta nueva función? El mismo Schrödinger intentó darle un significado a la onda  $\Psi$  de su ecuación usando ciertas analogías con la mecánica de fluidos –ecuación de continuidad para el “fluido electrónico”–, pero no tuvo éxito en su intento.



Max Born

Fue Max Born, colega y amigo de Heisenberg –él mismo ayudó a Heisenberg a construir la Mecánica matricial–, quién rápidamente intuyó una interpretación plausible. Basándose en los resultados experimentales sobre la dispersión de ondas planas –recordemos que el electrón libre se consideraba como tal en la mecánica ondulatoria de Schrödinger– Born aseguró que la función de onda  $\Psi(x)$  daba la probabilidad de que una partícula fuese detectada en la posición  $x$  y que dicha probabilidad era proporcional a  $|\Psi(x)|^2$ , es decir la Mecánica ondulatoria, al igual que la matricial como se vería más tarde, era una teoría estadística incluso para describir una única partícula.

Este trabajo publicado en julio de 1926, apenas un mes después del artículo de Schrödinger sobre la ecuación no estacionaria abrió una de las polémicas más grandes de la historia de la ciencia en los últimos 100 años: La física cuántica es, por principio, no determinista. El mismo Born escribió al final de su artículo:

*Aunque el problema del determinismo ha aparecido [...] yo mismo me inclino a dejar a un lado el determinismo en el mundo de los átomos. Pero esto es una cuestión filosófica para la cual los argumentos físicos no son concluyentes.*

El problema filosófico de Born se agudizó todavía más cuando Dirac por un lado, y el mismo Schrödinger por el otro probaban que las dos formulaciones de la Mecánica cuántica, la matricial y la ondulatoria, eran equivalentes:

*A cada función de la posición y el momento [en la mecánica ondulatoria] se le puede hacer corresponder una matriz de forma que en cada caso dichas matrices satisfacen las reglas formales de cálculo de Born y Heisenberg [...]. La solución del problema de contorno de la ecuación diferencial [en la mecánica ondulatoria] es completamente equivalente a la solución del problema algebraico de Heisenberg*

escribió Schrödinger en 1926.

El problema de la interpretación de la Mecánica Cuántica terminó en una “pelea” abierta entre los que la defendían y la consideraban una teoría completa –Bohr, Heisenberg, Born, Pauli, etc– y la que la consideraban incompleta –Schrödinger, Einstein, etc–. Como ejemplo de esta polémica es representativa la carta que escribe Einstein a Born el 7 de septiembre de 1944:

*Nuestras expectativas científicas nos han conducido a cada uno a las antípodas del otro. Tú crees en un Dios que juega a los dados, y yo en el valor único de las leyes en un universo en el que cada cosa existe objetivamente [...]. El gran éxito de la teoría de los quanta desde sus comienzos no puede hacerme creer en el carácter fundamental de ese juego de dados [...]. Algún día se descubrirá cual de estas dos actitudes instintivas es la buena.*

Parte importante para entender esta polémica viene del hecho de ambas teorías, la mecánica matricial y la ondulatoria, describían rigurosamente muchos de los fenómenos del micromundo, pero ambas tenían un gran problema ¿Cómo definir si una partícula cuántica estaba en un estado determinado o en otro?

$$\begin{array}{cc} \Psi_1, E_1 & \Psi_2, E_2 \\ \Psi = a_1\Psi_1 + a_2\Psi_2, E =? \end{array}$$

Figura 2.8: ¿Cómo sabemos en que estado cuántico se encuentra una partícula?

Un sencillo ejemplo de lo que ocurre es lo siguiente. Imaginemos que definimos los estados mediante la función  $\Psi$  usando la ecuación de Schrödinger y supongamos que nuestro sistema puede encontrarse en dos estados A y B definidos por la función  $\Psi_1$  y  $\Psi_2$ , respectivamente, y que al primero le corresponde una energía  $E_1$  y al segundo  $E_2$ , entonces siempre tenemos que un estado posible es el estado  $\Psi = a_1\Psi_1 + a_2\Psi_2$  (pues la ecuación de Schrödinger es lineal). Resulta entonces que si tenemos un instrumento que nos mide la energía de nuestro sistema éste nos dará unas veces  $E_1$  y otras  $E_2$ , y justo la probabilidad de que nos dé una u otra es proporcional a  $|a_1|^2$  y  $|a_2|^2$ , respectivamente. Esta era la interpretación probabilística que tan poco gustaba a Einstein.

### 2.5.1. El gato de Schrödinger

Para explicar problema descrito al final del apartado anterior (y apoyar la interpretación de Born), Bohr junto a otros propuso lo que hoy día se conoce como la interpretación de Copenhague y que consiste en la suposición de que al hacer la medición la función de onda “colapsa” y el estado queda determinado por la medición. Por ejemplo, cuando medimos la energía en nuestro sistema representado 2.8 la interacción del aparato con el sistema se decanta por una de las dos posibilidades.

Von Neumann intentó dar una explicación lógica a la interpretación de Bohr de la Mecánica Cuántica introduciendo una teoría de mediciones. Como ejemplo, supongamos que tenemos dos posibles estados A y B de un sistema definidos por las funciones de onda  $\Psi_1$  y  $\Psi_2$ , respectivamente. ¿Cómo saber en cual de los estados está el sistema? Este dilema se resolvía al hacer el experimento. De la interacción del aparato de medición con el sistema se concluía en que estado estaba (o más bien, se quedaba) el sistema (ver figura 2.9)<sup>6</sup>.

Esta teoría estaba plagada de efectos “curiosos”. Uno de los más famosos es el hecho de que al medir una cantidad física  $f(a)$  de un sistema físico microscópico (cuántico)  $A$  debemos usar un instrumento que es, en sí mismo otro sistema físico  $M$  y que obviamente es *clásico*, es decir, su comportamiento se puede explicar con las leyes de la física clásica. Pero entonces puede ocurrir que la interacción entre  $A$  y  $M$ , necesaria para poder saber el valor de  $f(a)$  cree una interferencia que se transfiera al mundo macroscópico. Es decir, si tenemos un sistema en el estado

$$\Psi = c_1\Psi_1 + c_2\Psi_2,$$

entonces, al realizar la medición en vez de tener el sistema  $A + M$  en un estado  $\Psi \otimes \Phi$  donde no se mezcle el estado  $\Psi$  de nuestro sistema cuántico con el estado  $\Phi$  de nuestro instrumento

<sup>6</sup>Este efecto se denomina comúnmente el *colapso de la función de onda*.

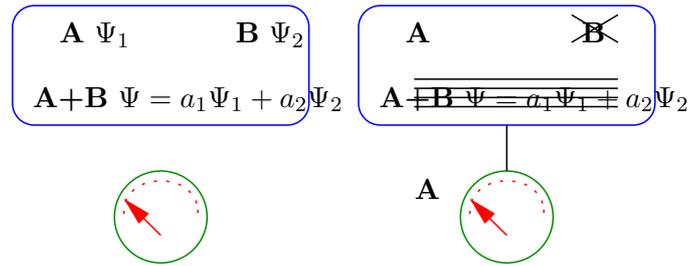


Figura 2.9: Las mediciones selectivas. Un sistema que originalmente puede estar en dos estados  $A$  y  $B$  (izquierda), se decanta por uno de ellos al realizar la medición (derecha). El aparato de medición interfiere en el sistema *acabando* con la incertidumbre del mismo.

clásico, tendremos una superposición más complicada

$$\Psi_{A+M} = c_1\Psi_1 \otimes \Phi_1 + c_2\Psi_2 \otimes \Phi_2,$$

donde se han mezclado estados cuánticos y clásicos. Esto no está de acuerdo con el principio de correspondencia de Bohr, ya que éste insistía que la física cuántica debía estar aparte de la clásica. Obviamente podemos pensar que para resolverlo basta con usar otro aparato  $M'$  que mida lo que mide  $M$ , pero entonces la interferencia pasa a  $M'$  y así sucesivamente terminamos en una cadena infinita. Von Neumann intentó resolver esta paradoja introduciendo en la cadena al observador humano que “*no se deja interferir*” por el sistema cuántico.

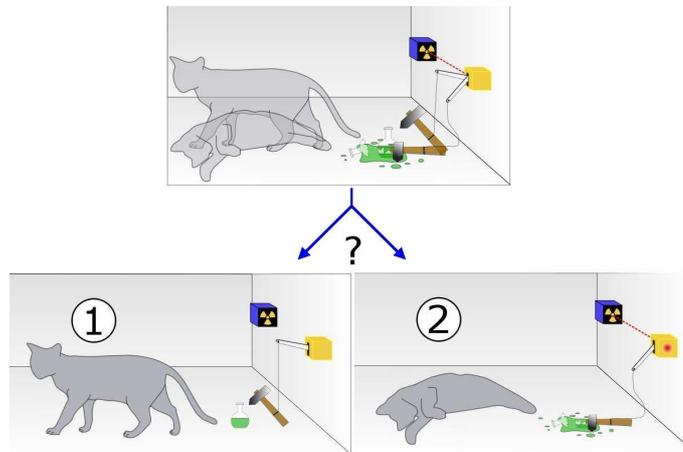


Figura 2.10: La paradoja del Gato de Schrödinger. Arriba se muestra lo que ocurre a puerta cerrada: una dualidad gato vivo-gato muerto que sólo se resuelve al abrir la caja cuando el observador descubre o que el gato está vivo (1), o que está muerto (2).

Esto llevó Schrödinger en 1935 a proponer su famosa paradoja del gato (ver figura 2.10), conocida hoy día como el *Gato de Schrödinger*. En ella Schrödinger usa como fenómeno cuántico la desintegración radioactiva. Las posibilidades son dos: tiene lugar la desintegración del átomo o no tiene lugar. El instrumento de medición es un detector que activa un *diabólico* martillo que, en caso de que ocurra la desintegración, golpea un recipiente de cristal con veneno y lo rompe. Y ahora Schrödinger mete su instrumento dentro de una caja negra sin ventanas ni puertas, junto con ... un pobre gato. La función de onda  $\Psi_{A+M}$  representará entonces una

superposición de los estados gato vivo—gato muerto y sólo un observador “humano” sobre el cual no interferirá el estado cuántico del átomo será capaz de resolver esa paradoja de gato vivo—muerto al mismo tiempo pues al abrir la caja se encontrará con una apacible o cruda realidad —en función del amor que profese a los gatos—. Hemos traspasado una propiedad microscópica (la superposición de estados como desintegrado—no desintegrado) de la desintegración radioactiva, cuántica, al mundo macroscópico, clásico, lo cual está en abierta contradicción con el principio de correspondencia del que ya hemos hablado. Lo que está claro de todo lo anterior es que la interpretación de Copenhagen<sup>7</sup> de la Mecánica Cuántica es algo para meditar y pensar: “*es una cuestión filosófica para la cual los argumentos físicos no son concluyentes*”, como dijo Born en su ya mencionado artículo.

### 2.5.2. Los universos paralelos de Everett

La interpretación de Copenhagen de la Mecánica Cuántica implica el colapso de la función de onda al interactuar con el aparato de medición lo que, como hemos comentado, conduce a no pocas paradojas. Entre los muchos intentos de remediar estos problemas se propusieron una infinidad de teorías como las famosas variables ocultas. No es el objetivo de estas notas tratar aquí de todo ello. El lector interesado puede consultar la bibliografía clásica al respecto, en particular [5].

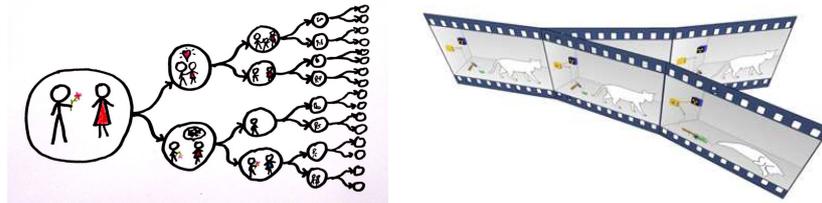


Figura 2.11: Los universos paralelos de Everett. Cada decisión que tomamos o medición que hacemos desdobra nuestro universo en dos o más, de forma que siempre hay alguno donde ocurre cada uno de los sucesos probables.

Por supuesto que existen otras posibles interpretaciones. Una de ellas, que además a dado lugar a muchísimas novelas de ciencia ficción (por ejemplo *La llegada de los gatos cuánticos*, de Frederik Pohl) es la de los universos paralelos propuesta por Hugh Everett en 1957. La idea, *grosso modo*, es la siguiente: Al hacer la medición (o simplemente al haber más de una opción) hay dos posibilidades (aquí estamos considerando de nuevo el ejemplo de dos mediciones de la figura 2.9), entonces en el momento de la medición no hay colapso de la función de onda sino que de forma continua en el tiempo, en ese instante en universo se divide en dos, en uno el resultado ha sido el universo donde ha prevalecido la medición A (como en la figura 2.9), y en el otro, por el contrario, ha sido la medición B la que ha prevalecido. Aunque esto parezca una broma “pesada” no lo es. En [23, §3.7] hay varios ejemplos donde se muestra cómo esta interpretación es plausible y nada descabellada.

<sup>7</sup>Hoy día no hay unanimidad en como interpretar la teoría cuántica. Un magnífico libro sobre el tema es *Lo decible y lo indecible en mecánica cuántica* de John S. Bell (Alianza Universidad, 1990).

## 2.6. El principio de incertidumbre de Heisenberg

Uno de las consecuencias más importantes y controvertidas de la Mecánica Cuántica fue el principio de incertidumbre descubierto por Heisenberg descubre en 1926 (aparentemente descubierto durante una de sus tantas visitas a Borh en Copenhagen cuando paseaba por los bosques de las afueras de la ciudad).

Usando el formalismo ondulatorio, Heisenberg comprobó que debía existir un principio de incertidumbre al medir ciertas cantidades físicas como la posición y el momento. En 1926, Heisenberg consideró una onda “gaussiana” normalizada del tipo

$$\Psi(x, 0) = \|\Psi\|^{-1} e^{-\frac{(x-x_0)^2}{2b^2}} e^{i\hbar p_0 x}, \quad \|\Psi\|^2 = \int_{\mathbb{R}} e^{-\frac{(x-x_0)^2}{b^2}} dx,$$

es decir<sup>8</sup>

$$\int_{\mathbb{R}} \Psi(x, 0) \overline{\Psi(x, 0)} dx = \int_{\mathbb{R}} |\Psi(x, 0)|^2 dx = 1.$$

Si una partícula venía descrita por dicha onda, entonces usando la idea de Born, el valor medio para la posición de la partícula era

$$\langle x \rangle = \int_{\mathbb{R}} \Psi(x, 0) x \overline{\Psi(x, 0)} dx = \int_{\mathbb{R}} x |\Psi(x, 0)|^2 dx = x_0,$$

y para la posición, usando que el operador correspondiente al momento era  $\hat{p} = -i\hbar \frac{\partial}{\partial x}$ , tenemos

$$\langle p \rangle = \int_{\mathbb{R}} \Psi(x, 0) \hat{p} \overline{\Psi(x, 0)} dx = p_0,$$

es decir que nuestra partícula tiene un momento  $p_0$  y está en la posición  $x_0$ . Si ahora intentamos determinar con que precisión estamos calculando los valores de estas dos magnitudes tenemos que calcular las varianzas  $\sigma_x$  y  $\sigma_p$ ,

$$\sigma_x = \int_{\mathbb{R}} \Psi(x, 0) (x - x_0)^2 \overline{\Psi(x, 0)} dx = \frac{b^2}{2}, \quad \sigma_p = \int_{\mathbb{R}} \Psi(x, 0) (\hat{p} - p_0)^2 \overline{\Psi(x, 0)} dx = \frac{\hbar^2}{2b^2},$$

de forma que

$$\sigma_x \sigma_p = \frac{\hbar^2}{4}, \quad \text{o, equivalentemente,} \quad \Delta x \Delta p = \frac{\hbar}{2},$$

con  $\Delta x = \sqrt{\sigma_x}$  y  $\Delta p = \sqrt{\sigma_p}$ . Es decir, no podemos nunca medir con una precisión tan grande como se quiera las dos magnitudes  $\Delta x$  y  $\Delta p$ . De hecho, resulta ser que  $\Delta x \Delta p \geq \frac{\hbar}{2}$ , es decir, Heisenberg descubrió justo la onda que minimizaba el principio que lleva su nombre.

¿De dónde sale esta incertidumbre que no existe en la mecánica clásica?

### 2.6.1. El experimento de difracción y el principio de incertidumbre

Consideremos el siguiente experimento imaginario. Supongamos que tenemos una pared con un agujero en el centro de diámetro  $d$  y lanzamos un electrón cuya trayectoria es perpendicular a la pared y que pasa por dicho orificio.

<sup>8</sup>La operación  $\bar{a}$  denota el complejo conjugado de  $a$ .

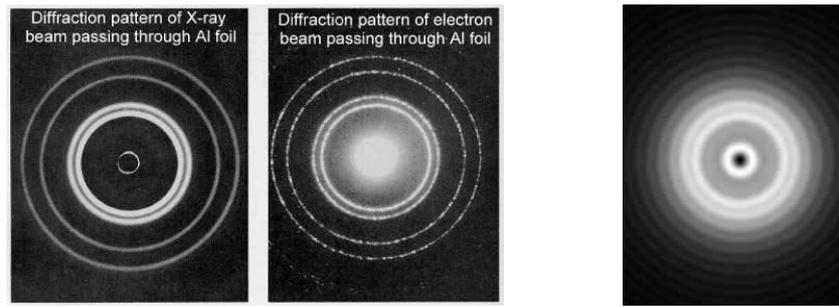


Figura 2.12: Esquema de difracción de rayos X y electrones (izquierda) y de luz monocromática (derecha)

Supongamos que el electrón efectivamente se comporta como una onda monocromática de longitud  $\lambda$ . Entonces al pasar el electrón por el agujero obtendremos la conocida figura de difracción<sup>9</sup> (ver figura 2.12) formada por círculos concéntricos alrededor del punto  $O$ . Nos interesa estimar el radio del primer círculo de dicho patrón.

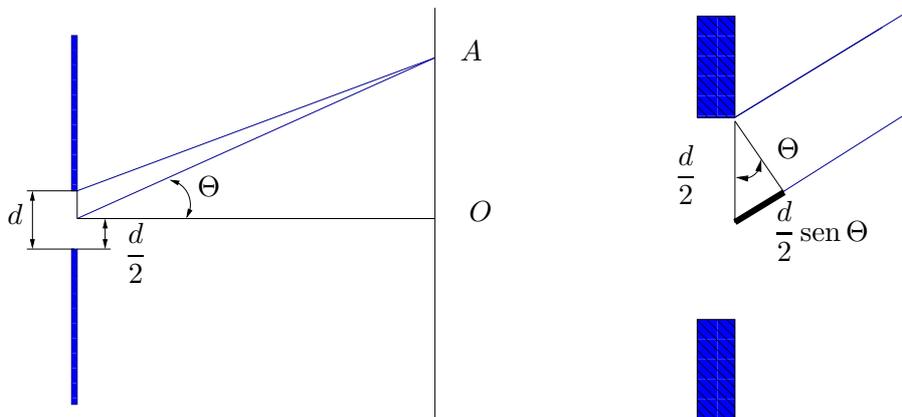


Figura 2.13: Esquema de difracción (izquierda) y detalle del mismo (derecha)

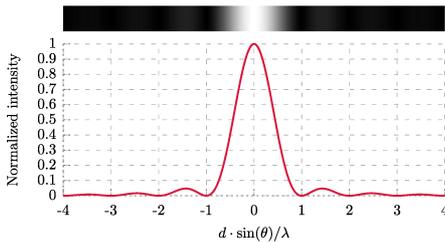
Para ello vamos a utilizar la representación esquemática del experimento que se puede ver en la figura 2.13. Para describir la difracción se suele dividir la rendija en  $2N$  partes iguales. Nos interesa conocer donde aparece el primer mínimo por lo que escogeremos  $N = 1$ , i.e., dividiremos el agujero a la mitad y consideraremos los *rayos* que salen del centro y los extremos, respectivamente, tal y como se representan en la figura 2.13 (cualesquiera otros dos que escojamos nos dan un resultado similar). De la figura 2.13 se deduce que en el punto  $A$  habrá un mínimo si la diferencia del camino recorrido por ambos rayos (el que sale del extremo superior y el del medio del agujero) es un número entero  $m$  de veces la mitad de la longitud de onda de la misma (tendrán una interferencia negativa), luego

$$\frac{d}{2} \sin \Theta = m \frac{\lambda}{2} \quad \implies \quad d \sin \Theta = m \lambda,$$

y por tanto el primer mínimo se obtendrá cuando  $d \sin \Theta = \lambda$ .

<sup>9</sup>Vamos a asumir que tanto la fuente de electrones como el plano donde aparece la figura de difracción están lo suficientemente alejados del agujero y de esta forma usaremos el método de Fraunhofer

El experimento de la difracción de electrones nos permite comprobar la veracidad del principio de incertidumbre de Heisenberg. Ante todo notemos que para nuestro electrón  $\Delta p_y$  es cero ya que el electrón se mueve en el eje de las  $x$ . Ahora bien, cuando pasa por la ranura sabemos que se difracta, es decir que puede cambiar su trayectoria inicial ya que vemos su gráfica de dispersión. Por tanto, la indeterminación  $\Delta y$  del electrón en el momento de pasar por el orificio es del orden  $\Delta y = d$  (pues sabemos que el electrón ha pasado). Ahora bien, como tenemos una gráfica de dispersión, resulta que el electrón tiene que haber ganado cierto impulso  $p_y$  en la dirección perpendicular a su eje, eso nos da una cierta indeterminación del orden  $\Delta p_y$ . Para calcularla usamos la teoría ondulatoria de de Broglie para el electrón,  $\Delta p_y = p_x \sin \theta$ ,  $p_x = 2\pi\hbar/\lambda$ .



Para calcular  $\theta$  asumiremos que el electrón va a parar dentro del círculo que define el primer mínimo en la gráfica de dispersión –los otros círculos los despreciamos al ser la intensidad de éstos bastante más pequeña que la del primero tal y como se ve en la figura de la izquierda–. Ahora bien, como sabemos, el primer mínimo tiene lugar cuando  $d \sin \theta = \lambda$ , de donde se deduce  $\sin \theta = \lambda/d$ , y por tanto

Intensidad normalizada del cuadro de difracción

$$\Delta p_y \Delta y = \left( p_x \frac{\lambda}{d} \right) \Delta y = \frac{2\pi\hbar}{\lambda} \frac{\lambda}{d} d = 2\pi\hbar \geq \frac{\hbar}{2} > 0.$$

El principio de incertidumbre como ya hemos dicho es un pilar de la Mecánica cuántica. Por ejemplo, es el causante de que los electrones no caigan al núcleo pues en ese caso tanto  $p$  como  $x$  valdrían cero. Otra consecuencia del principio de incertidumbre es la desaparición del concepto clásico de trayectoria.

## 2.7. Las matemáticas de la Mecánica Cuántica

Las matemáticas de la Mecánica Cuántica están estrechamente ligadas al problema de la interpretación. La principal razón se debe a que una misma teoría no puede contener dos tipos de postulados, principios o axiomas: los clásicos y los cuánticos. Por tanto los principios de la física clásica deben obtenerse de los axiomas de la Mecánica cuántica al pasar al mundo macroscópico donde la física clásica es aplicable.

La construcción matemática impuso el “orden” en el aparente caos de la interpretación. Los principales axiomas o postulados de la Mecánica Cuántica se pueden resumir en los siguientes:

- I. Cualquier magnitud física se describe a través de un operador lineal “hermítico”  $\hat{A}$  definido sobre un espacio de Hilbert  $\mathcal{H}$ , cuyos vectores  $\Psi$  definen los posibles estados del sistema físico.
- II. Los valores  $f(a')$  que puede tomar una magnitud física son aquellos que corresponden al espectro del operador lineal hermítico  $\hat{A}$  que caracteriza dicha magnitud.

- III. El valor esperado de una magnitud física  $x$  cualquiera de un sistema en el estado  $\Psi$ , es  $\langle \Psi | x \Psi \rangle$ , donde  $\langle \cdot | \cdot \rangle$  representa el producto escalar en el espacio de Hilbert.
- IV. La función de onda  $\Psi$  del sistema está gobernada por la ecuación de Schrödinger  $\hat{\mathcal{H}}\Psi = i\hbar \frac{\partial \Psi}{\partial t}$ , donde  $\hat{\mathcal{H}}$  es el operador de Hamilton del sistema.

Un espacio de Hilbert –que denotaremos aquí por  $\mathcal{H}$ – es un espacio lineal donde está definido un producto escalar  $\langle a | b \rangle$ , cualesquiera sean los vectores  $a, b \in \mathcal{H}$  y que es completo<sup>10</sup> respecto a la norma  $\|x\|$  inducida por el producto escalar<sup>11</sup>  $\|x\| = \sqrt{\langle x | x \rangle}$ . Los operadores lineales definidos sobre  $\mathcal{H}$  se pueden representar mediante matrices –finitas o infinitas– de forma que las matrices de Heisenberg, Born y Jordan se pueden identificar como ciertos operadores lineales sobre  $\mathcal{H}$ . Además, las funciones de onda de Schrödinger pertenecen al espacio de Hilbert  $L^2$ , el espacio de las funciones de cuadrado integrable. Un operador  $\hat{\mathcal{A}}$  se dice hermítico si  $\langle a | \hat{\mathcal{A}} b \rangle = \langle \hat{\mathcal{A}} a | b \rangle$ . Una propiedad fundamental de estos operadores hermíticos era que para ellos existía una base ortonormal de autovectores y además los correspondientes autovalores eran reales. En especial esta última propiedad era decisiva a la hora de identificar los operadores con las magnitudes físicas medibles, que son cantidades reales.

Es decir, la matemática de la Mecánica Cuántica, es la matemática de los operadores en los espacios de Hilbert. Parte de esa teoría era conocida a principios del siglo XX pero gran parte de la misma se desarrolló en Gotinga impulsada por David Hilbert pero fundamentalmente desarrollada por John Von Neumann cuya obra quedaría plasmada en el magnífico libro *Fundamentos Matemáticos de la Mecánica Cuántica* publicado en 1932.

## 2.8. Sobre la bibliografía

Más detalles históricos los puedes encontrar en [3, 8, 12, 15, 18, 24].

Como introducción a nivel elemental de física cuántica se pueden consultar [10, 19, 25]. Textos más avanzados lo constituyen [4, 9, 15, 17, 20, 22, 23]. Una magnífica introducción de lo que pretende la física y de su interacción con las matemáticas se tiene en [11].

---

<sup>10</sup>Es decir que toda sucesión de Cauchy en  $\mathcal{H}$  es convergente.

<sup>11</sup>Para el que usaremos la notación de Dirac: el producto escalar de  $x$  e  $y$  se denotará por  $\langle x | y \rangle$ .



# Capítulo 3

## Mecánica Cuántica I: “Movimiento” de una partícula material

Vamos a comenzar describiendo el sistema cuántico más sencillo: una única partícula sometida a un potencial externo  $V(\vec{r})$ . Comenzaremos describiendo los principales postulados. Una magnífica introducción se puede consultar en [4].

### 3.1. Los postulados de la Mecánica Cuántica

**Postulado 3.1.1** *El estado de una partícula de masa  $m$  en el instante de tiempo  $t$  viene biunívocamente determinado por la función de onda  $\Psi(\vec{r}, t)$ . La densidad de probabilidad de encontrar la partícula en el instante  $t$  en la región del espacio de volumen  $d^3\vec{r}$  alrededor del punto  $\vec{r}$  es  $d^3P(\vec{r}) = |\Psi(\vec{r}, t)|^2 d^3\vec{r}$ .*

Proceden algunos comentarios.

1. Para que el postulado 3.1.1 tenga sentido debe ocurrir que para cada  $t$  fijo

$$\int_{\Omega} |\Psi(\vec{r}, t)|^2 d^3\vec{r} = 1,$$

donde  $\Omega$  es aquella región *accesible* a la partícula, es decir, la región donde con certeza absoluta puede estar la partícula (en general dicha región será  $\mathbb{R}^3$ ). Lo anterior indica que, para cada  $t$ , la función  $\Psi$  es de cuadrado integrable y esta normalizada a la unidad. Al espacio vectorial de las funciones de cuadrado integrable lo denotaremos por  $L^2(\Omega)$  o simplemente  $L^2$ .

Más aún, como veremos más adelante,  $\Psi$  ha de satisfacer una ecuación dinámica, así que es natural que  $\Psi(\vec{r}, t)$  sea una función continua y diferenciable (al menos). Luego,  $\Psi$  ha de ser tal que  $\Psi(\vec{r}, t)$  tiende a cero si  $\vec{r} \rightarrow \infty$  en caso contrario  $\Psi$  no sería de cuadrado integrable en  $\mathbb{R}^3$ . Obviamente si  $\Omega$  es finito, del postulado 3.1.1 se sigue que  $\Psi(\vec{r}, t) = 0$  en la frontera y fuera de  $\Omega$ , ya que en caso contrario existiría una probabilidad finita de que la partícula estuviese en el exterior de  $\Omega$  lo cual sería una contradicción.

2. Obviamente el estado de una partícula en cada instante  $t$  queda determinado a excepción de una *fase*  $e^{i\alpha}$ ,  $\alpha \in \mathbb{R}$ . Es decir, las funciones de onda  $\Psi(\vec{r}, t)$  y  $\Psi(\vec{r}, t)e^{i\alpha}$  describen el mismo estado del sistema. De lo anterior se sigue además que la fase  $\alpha$  no es un observable.

Una cuestión física de vital importancia que surge inmediatamente es: ¿Cómo vamos a medir nuestros observables?

Supongamos que vamos a medir la posición  $\vec{r}$  de la partícula. Para ello asumimos que

1. La medición se hace con un aparato *clásico*, i.e., el aparato no precisa una descripción *cuántica* (por ejemplo una regla o un metro).
2. La precisión del aparato es, en principio, tan alta como se quiera.
3. El resultado de la medición es: *En el instante  $t$  la partícula estaba en la posición  $\vec{r} \pm \delta\vec{r}$  donde  $\delta\vec{r}$  es el error del aparato.*

¿Cómo entender entonces la descripción probabilística del postulado 3.1.1?

Para ello imaginemos que tenemos  $N$  cajas idénticas y  $N$  partículas idénticas de forma que en cada caja hay una única partícula. Asumamos además que todas las partículas están en el mismo estado, i.e.,  $\Psi(\vec{r}, t)$  es la misma para cada una de ellas en el instante de la medición. Si medimos las posiciones de cada partícula independientemente obtendremos los valores  $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N$ , no necesariamente iguales, que estarán distribuidos de acuerdo a la ley de probabilidad  $|\Psi(\vec{r}, t)|^2 d^3\vec{r}$ . El valor esperado de  $r$  es, por tanto

$$\langle \vec{r} \rangle = \int_{\Omega} \vec{r} |\Psi(\vec{r}, t)|^2 d^3\vec{r}, \quad \Leftrightarrow \quad \langle x_i \rangle = \int_{\Omega} x_i |\Psi(\vec{r}, t)|^2 d^3\vec{r},$$

$x_i = x, y, z$  para  $i = 1, 2, 3$ , respectivamente. Su dispersión es

$$\Delta x_i = \sqrt{\langle x_i^2 \rangle - \langle x_i \rangle^2} = \sqrt{\int_{\Omega} (x_i - \langle x_i \rangle)^2 |\Psi(\vec{r}, t)|^2 d^3\vec{r}}.$$

La dispersión  $\Delta x$  tiene un significado muy preciso: si  $\Delta x$  es pequeño, entonces  $|\Psi(\vec{r}, t)|^2$  está muy concentrada alrededor de su valor medio  $\langle x \rangle$ , lo que indica que la partícula se encuentra con probabilidad alta muy cerca de la posición  $\langle x \rangle$ . Por el contrario, si  $\Delta x$  es grande, entonces  $|\Psi(\vec{r}, t)|^2$  está muy dispersa y la partícula tiene una probabilidad baja de estar cerca de  $\langle x \rangle$ .<sup>1</sup>

Para explicar los fenómenos de interferencia y difracción de electrones que mencionamos en el apartado anterior se ha de introducir un segundo postulado fundamental.

**Postulado 3.1.2** *Para cada  $t$  el espacio de las funciones de estado es un subespacio vectorial del espacio  $L^2(\Omega)$ . Es decir, para cada  $t$ , si  $\Psi_1$  y  $\Psi_2$  son funciones de onda, entonces*

$$\forall \alpha_1, \alpha_2 \in \mathbb{C}, \quad \Psi = \alpha_1 \Psi_1 + \alpha_2 \Psi_2,$$

*es una posible función de onda, es decir,  $\Psi$  puede representar un posible estado del sistema.*<sup>2</sup>

<sup>1</sup>Un ejemplo muy ilustrativo es el siguiente. Supongamos que  $\Psi(\vec{r}, t) = \chi_{[a,b]}(x)/(b-a)$ , donde  $\chi_{[a,b]}(x)$  es la función característica del intervalo  $[a, b]$ , i.e., vale 1 si  $x \in [a, b]$  y 0 en otro caso. Entonces,  $\langle x \rangle = \frac{a+b}{2}$  y  $\Delta x = (b-a)^2/12$ . Como vemos si  $a$  y  $b$  están cercanos entonces la dispersión es pequeña y tenemos que la partícula está muy cerca de la posición media. Por el contrario si  $a$  y  $b$  difieren mucho la partícula está equiprobablemente dispersa en el intervalo  $[a, b]$  (muy grande). Este ejemplo se lo debo a de Francisco J. (Pacho) Ruiz Blasco de la Universidad de Zaragoza.

<sup>2</sup>Nótese que si  $\alpha_1 = \alpha_2 = 0$ , se tiene  $\Psi \equiv 0$ , es decir la partícula no está en ningún sitio. Aunque formalmente esta es una posibilidad en la práctica se asume que  $\Psi \neq 0$  en todo  $\Omega$  ( $\Omega$  es la región donde podría estar la partícula, luego en algún punto de  $\Omega$  la función  $\Psi$  ha de ser distinta de cero). Si  $\Psi_1$  y  $\Psi_2$  están normalizadas a la unidad, entonces para que  $\Psi$  lo esté basta que  $|\alpha_1|^2 + |\alpha_2|^2 = 1$ .

Lo anterior indica que, para cada  $t$ , cualquier combinación lineal de funciones es una (posible) función de onda. Ello implica además que la ecuación de evolución que gobierne o determine las funciones de onda ha de ser lineal.

**Postulado 3.1.3 (Ecuación dinámica)** *La función de onda viene determinada por la ecuación de Schrödinger*

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(\vec{r}, t) + V(\vec{r}, t) \Psi(\vec{r}, t), \quad (3.1.1)$$

donde  $\Delta$  es el laplaciano<sup>3</sup> en  $\mathbb{R}^3$ , y  $V$  es el potencial al que esta sometida la partícula.

La ecuación anterior se suele escribir de la forma

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = \hat{\mathcal{H}} \Psi(\vec{r}, t), \quad (3.1.2)$$

donde  $\hat{\mathcal{H}}$  es el operador correspondiente al hamiltoniano del sistema. En nuestro caso está claro que es

$$\hat{\mathcal{H}} := -\frac{\hbar^2}{2m} \Delta + V(\vec{r}, t) \hat{\mathcal{I}}.$$

Antes de continuar debemos destacar que aquí sólo trataremos el caso cuando el potencial  $V$  es independiente del tiempo.

Nuevamente proceden una serie de comentarios.

En primer lugar si  $V := 0$  (partícula libre) una solución de la ecuación viene dada mediante una onda viajera

$$\Psi(\vec{r}, t) = \Psi_0 e^{i(\vec{k}\vec{r} - \omega t)},$$

donde  $\vec{k}$  es el vector de onda. Sustituyendo dicha onda en (3.1.1) y usando que la energía de una onda es, según ya vimos,  $E = \hbar\omega$ , obtenemos

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} \implies E = \frac{\hbar^2 k^2}{2m},$$

de donde, al no haber (energía) potencial  $E = T = p^2/2m$ , por tanto se sigue que  $p = \hbar k$ , que es precisamente la fórmula de De Broglie para expresar la dualidad onda-partícula. Es decir, la ecuación de Schrödinger tiene como solución posible la *onda de materia* de De Broglie.

Un simple vistazo a la expresión para la onda de De Broglie nos revela un *problema*: esta función no es de cuadrado integrable y por tanto no puede describir de acuerdo con el postulado 3.1.1 ningún estado real de una partícula. Para resolver esta aparente contradicción se introducen los *paquetes de ondas*.

Formalmente un paquete de ondas es un conjunto de ondas planas monocromáticas superpuestas. Usando que  $p = \hbar k$  y  $E = \hbar\omega$  tenemos, para cada onda la ecuación

$$\Psi_p(\vec{r}, t) = \Phi(\vec{p}) e^{\frac{i}{\hbar}(\vec{p}\vec{r} - Et)},$$

---

<sup>3</sup>En coordenadas cartesianas es  $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ .

luego la superposición de todas será

$$\Psi(\vec{r}, t) = \int_{\mathbb{R}^3} \Phi(\vec{p}) e^{\frac{i}{\hbar}(\vec{p}\vec{r} - Et)} \frac{d^3\vec{p}}{(2\pi\hbar)^{3/2}} = \int_{\mathbb{R}^3} \tilde{\Psi}(\vec{p}, t) e^{\frac{i}{\hbar}\vec{p}\vec{r}} \frac{d^3\vec{p}}{(2\pi\hbar)^{3/2}}, \quad (3.1.3)$$

donde  $\tilde{\Psi}(\vec{p}, t) = \Phi(\vec{p})e^{-\frac{i}{\hbar}Et}$ . Es decir,  $\tilde{\Psi}$  es la transformada de Fourier de  $\Psi$ . Además, para que lo anterior tenga sentido  $\Phi$  ha de ser tal que  $\Psi$  sea de cuadrado integrable. Lo anterior implica que ambas  $\Psi(\vec{r}, t)$  y  $\tilde{\Psi}(\vec{p}, t)$  son funciones de cuadrado integrable para todo  $t$ .

La relación de  $\Psi(\vec{r}, t)$  y  $\tilde{\Psi}(\vec{p}, t)$  como transformada de Fourier una de la otra tiene una implicación interesante:

$$\tilde{\Psi}(\vec{p}, t) = \int_{\mathbb{R}^3} \Psi(\vec{r}, t) e^{-\frac{i}{\hbar}\vec{p}\vec{r}} \frac{d^3\vec{r}}{(2\pi\hbar)^{3/2}},$$

y

$$\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3\vec{r} = \int_{\mathbb{R}^3} |\tilde{\Psi}(\vec{p}, t)|^2 d^3\vec{p}.$$

Es decir, en principio tanto  $\Psi(\vec{r}, t)$  como  $\tilde{\Psi}(\vec{p}, t)$  valdrían para describir el estado de nuestro sistema.

Además, de la teoría de las transformadas de Fourier tenemos que, si definimos la media y varianza de las componentes del vector de impulso ( $p_i = p_x, p_y, p_z$  para  $i = 1, 2, 3$ , respectivamente)

$$\langle p_i \rangle = \int_{\mathbb{R}^3} p_i |\tilde{\Psi}(\vec{p}, t)|^2 d^3\vec{p}, \quad \Delta p_i = \sqrt{\langle p_i^2 \rangle - \langle p_i \rangle^2}, \quad (3.1.4)$$

entonces  $\Delta x_i \Delta p_i \geq \hbar/2$ , y la igualdad sólo tiene lugar cuando  $\Psi(\vec{r}, t)$  es proporcional a una gaussiana en  $\vec{r}$ , es decir el principio de incertidumbre de Heisenberg.

Para nuestro caso de la partícula libre ello implica que si  $\Psi(\vec{r}, t)$  está muy dispersa en el espacio, entonces el valor del impulso  $\vec{p}$  está muy concentrado y si  $\Psi(\vec{r}, t)$  está muy dispersa entonces  $\vec{p}$  está muy disperso.

En la figura 3.1 se muestra lo que ocurre para el caso unidimensional cuando tomamos en (3.1.3)

$$\Phi(p) = \frac{1}{\sqrt{\hbar\sigma} \pi^{1/4}} e^{-\frac{(p-p_0)^2}{2(\hbar\sigma)^2}}$$

para los valores  $\hbar\sigma = p_0/2$  ( $p$  disperso, figura superior) y  $\hbar\sigma = p_0/10$  ( $p$  concentrado, figura inferior).

Antes de continuar debemos destacar la media y la varianza definidas en (3.1.4) no tienen por que corresponder con la media y varianza del impulso real  $\vec{p}$  de nuestra partícula. Para poder hablar de la cantidad observable  $\vec{p}$  necesitaremos introducir otro postulado, pero antes de ello vamos a probar que los postulados hasta ahora introducidos son consistentes.

### Test de consistencia

Un test de consistencia de la teoría hasta ahora descrita consiste en probar que la *norma* de  $\Psi(\vec{r}, t)$  es independiente del tiempo (en otro caso no podría definir una densidad de probabilidad). Para ello escribimos la ecuación de Schrödinger (3.1.1) para  $\Psi$

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(\vec{r}, t) + V(\vec{r}, t) \Psi(\vec{r}, t)$$

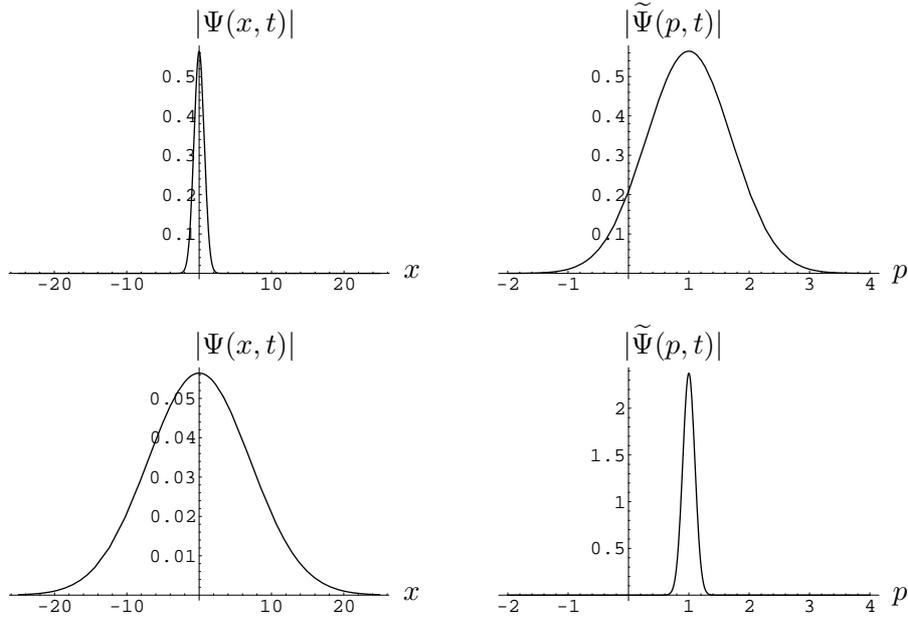


Figura 3.1: Distribución de  $\Psi(x, t)$  y  $\tilde{\Psi}(p, t)$  ( $\hbar\sigma = 1$ )

y su complejo conjugado<sup>4</sup>.  $\bar{\Psi}$  (se entiende que  $V$  es un potencial real)

$$-i\hbar \frac{\partial \bar{\Psi}(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \bar{\Psi}(\vec{r}, t) + V(\vec{r}, t) \bar{\Psi}(\vec{r}, t),$$

multiplicamos la primera por  $\bar{\Psi}$ , la segunda por  $\Psi$ , las restamos y tomamos la integral en toda la región  $\Omega$  (ver postulado 3.1.1)

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \int_{\Omega} |\Psi(\vec{r}, t)|^2 d^3\vec{r} &= \\ &= i\hbar \int_{\Omega} \left( \bar{\Psi}(\vec{r}, t) \frac{\partial \Psi(\vec{r}, t)}{\partial t} + \Psi(\vec{r}, t) \frac{\partial \bar{\Psi}(\vec{r}, t)}{\partial t} \right) d^3\vec{r} \\ &= -\frac{\hbar^2}{2m} \left( \int_{\Omega} \bar{\Psi}(\vec{r}, t) \Delta \Psi(\vec{r}, t) - \Psi(\vec{r}, t) \Delta \bar{\Psi}(\vec{r}, t) \right) d^3\vec{r}. \end{aligned}$$

Si ahora en la última de las integrales usamos la fórmula de Green obtenemos

$$\begin{aligned} &-\frac{\hbar^2}{2m} \int_{\Omega} \left( \bar{\Psi}(\vec{r}, t) \Delta \Psi(\vec{r}, t) - \Psi(\vec{r}, t) \Delta \bar{\Psi}(\vec{r}, t) \right) d^3\vec{r} \\ &= -\frac{\hbar^2}{2m} \int_{\partial\Omega} \left( \bar{\Psi}(\vec{r}, t) \frac{\partial \Psi(\vec{r}, t)}{\partial \vec{n}} - \Psi(\vec{r}, t) \frac{\partial \bar{\Psi}(\vec{r}, t)}{\partial \vec{n}} \right) dS, \end{aligned}$$

donde  $dS$  es el elemento de área de la superficie y  $\frac{\partial f(\vec{r})}{\partial \vec{n}}$  denota la derivada direccional de  $f$  en la dirección del vector normal al elemento de superficie  $dS$ .<sup>5</sup> Si ahora hacemos  $\vec{r} \rightarrow \infty$  la

<sup>4</sup>Como antes denotaremos por  $\bar{z}$  al complejo conjugado del número complejo  $z$

<sup>5</sup>En una dimensión vale simplemente la integración por partes.

última integral se anula pues  $\Psi(\vec{r}, t)$  (y por consiguiente  $\overline{\Psi(\vec{r}, t)}$ ) se anulan cuando  $\vec{r} \rightarrow \infty$  (ver comentario 1 del postulado 3.1.1). Luego

$$\frac{\partial}{\partial t} \int_{\Omega} |\Psi(\vec{r}, t)|^2 d^3\vec{r} = 0.$$

Esto tiene una implicación física inmediata: Dado un estado inicial  $\Psi(\vec{r}, t_0) \in L^2(\Omega)$ , este evoluciona según la ecuación de Schrödinger, i.e., en cada momento de tiempo viene descrito por la función  $\Psi(\vec{r}, t)$  solución de (3.1.1) y además,

$$\int_{\Omega} |\Psi(\vec{r}, t)|^2 d^3\vec{r} = \int_{\Omega} |\Psi(\vec{r}, t_0)|^2 d^3\vec{r} = 1,$$

es decir,  $\Psi(\vec{r}, t)$  es normalizada a la unidad. Luego la ecuación de Schrödinger es consistente con los postulados 3.1.1 y 3.1.2.

Vamos ahora a intentar convencernos de que para los momentos  $\vec{p}$  se tienen las fórmulas (3.1.4). Para ello, como ya comentamos, necesitamos varios postulados que complementen a los anteriores.

**Postulado 3.1.4** *A cualquier cantidad física medible  $A$  le asociaremos un operador hermítico  $\hat{\mathcal{A}}$  que actúa sobre el espacio de las funciones de cuadrado integrable  $L^2(\Omega)$  a las que pertenecen las funciones de estado  $\Psi(\vec{r}, t)$ . i.e.,  $\hat{\mathcal{A}} : L^2(\Omega) \rightarrow L^2(\Omega)$ . Además, dado un estado definido por una función de onda  $\Psi(\vec{r}, t)$  que define cierto estado del sistema, el valor medio de las medidas de  $A$  viene dado por<sup>6</sup>*

$$\langle A \rangle(t) = \int_{\Omega} \overline{\Psi(\vec{r}, t)} \left( \hat{\mathcal{A}}\Psi(\vec{r}, t) \right) d^3\vec{r}$$

Un operador  $\hat{\mathcal{A}} : L^2(\Omega) \rightarrow L^2(\Omega)$  es hermítico<sup>7</sup> si

$$\int_{\Omega} \overline{\Psi_1(\vec{r}, t)} \left( \hat{\mathcal{A}}\Psi_2(\vec{r}, t) \right) d^3\vec{r} = \int_{\Omega} \left( \overline{\hat{\mathcal{A}}\Psi_1(\vec{r}, t)} \right) \Psi_2(\vec{r}, t) d^3\vec{r}.$$

Nótese que si  $\hat{\mathcal{A}}$  es hermítico entonces

$$\begin{aligned} \overline{\int_{\Omega} \overline{\Psi(\vec{r}, t)} \left( \hat{\mathcal{A}}\Psi(\vec{r}, t) \right) d^3\vec{r}} &= \int_{\Omega} \left( \overline{\hat{\mathcal{A}}\Psi(\vec{r}, t)} \right) \Psi(\vec{r}, t) d^3\vec{r} \\ &= \int_{\Omega} \left( \hat{\mathcal{A}}\Psi(\vec{r}, t) \right) \overline{\Psi(\vec{r}, t)} d^3\vec{r}, \end{aligned}$$

i.e.,  $\langle A \rangle(t) \in \mathbb{R}$ .

Vamos a definir el operador  $\hat{r}$  por

$$\hat{r} := \vec{r}\hat{\mathcal{L}}, \quad (3.1.5)$$

<sup>6</sup>Aunque el valor  $\langle A \rangle(t)$  depende de la función de onda  $\Psi$ , no lo vamos a reflejar en la notación para hacer esta lo menos engorrosa posible.

<sup>7</sup>Sobre estos operadores hablaremos con más detalle en el próximo apartado dedicado a la Mecánica cuántica en espacios de Hilbert.

donde  $\widehat{\mathcal{I}}$  es el operador identidad. Entonces

$$\langle \vec{r} \rangle(t) = \int_{\Omega} \overline{\Psi(\vec{r}, t)} \vec{r} \Psi(\vec{r}, t) d^3 \vec{r} = \int_{\Omega} \vec{r} |\Psi(\vec{r}, t)|^2 d^3 \vec{r},$$

que está en concordancia con el postulado 3.1.1.

Definiremos el operador impulso  $\widehat{p}$  por la expresión

$$\widehat{p} := -i\hbar \nabla, \quad (3.1.6)$$

donde  $\nabla$  denota al operador gradiente<sup>8</sup> de  $\mathbb{R}^3$ .

Es fácil comprobar que ambos operadores son hermíticos.

Otros operadores importantes son

1. El operador energía cinética  $\widehat{T} := -\frac{\hbar^2}{2m} \Delta$ , donde  $\Delta$  es, como antes, el laplaciano en  $\mathbb{R}^3$ .
2. El operador energía potencial  $\widehat{V}(\vec{r}) := V(\vec{r})\widehat{\mathcal{I}}$ .
3. El operador energía total  $\widehat{E} := \widehat{T} + \widehat{V}(\vec{r})$ . Este además constituye el operador de Hamilton o hamiltoniano del sistema.
4. El operador momento angular  $\widehat{\mathcal{L}} = \widehat{\vec{r}} \times \widehat{\vec{p}} := -i\hbar \text{rot}$ , donde  $\text{rot}$  denota el vector rotor en  $\mathbb{R}^3$ .

**Postulado 3.1.5** *El resultado de cualquier medición de una cantidad física medible  $A$  debe ser un autovalor del operador asociado  $\widehat{\mathcal{A}}$  y el estado correspondiente a dicha medición estará definido por su correspondiente autofunción.*

Es decir, tenemos la ecuación

$$\widehat{\mathcal{A}}\Psi(\vec{r}, t) = a\Psi(\vec{r}, t).$$

Entonces, por el postulado 3.1.4,

$$\langle A \rangle = \int_{\Omega} \overline{\Psi(\vec{r}, t)} \left( \widehat{\mathcal{A}}\Psi(\vec{r}, t) \right) d^3 \vec{r} = a \int_{\Omega} |\Psi(\vec{r}, t)|^2 d^3 \vec{r} = a,$$

que, como ya hemos visto, es real<sup>9</sup> Además,

$$\langle A^2 \rangle = \int_{\Omega} \overline{\Psi(\vec{r}, t)} \left( \widehat{\mathcal{A}}(\widehat{\mathcal{A}}\Psi(\vec{r}, t)) \right) d^3 \vec{r} = a^2 \implies$$

$\Delta A^2 := \sqrt{\langle A^2 \rangle - \langle A \rangle^2} = 0$ , es decir, el resultado de la medición, en caso de que el sistema se encuentre en el estado  $\Psi_a(\vec{r}, t)$  correspondiente al autovalor  $a$ , es exacto. Este resultado lo discutiremos con más detalle en el marco de los espacios de Hilbert.

<sup>8</sup>En coordenadas cartesianas es el operador  $\nabla f(\vec{r}) = \frac{\partial f(\vec{r})}{\partial x} \vec{i} + \frac{\partial f(\vec{r})}{\partial y} \vec{j} + \frac{\partial f(\vec{r})}{\partial z} \vec{k}$ .

<sup>9</sup>Esta propiedad es en realidad una consecuencia de la hermiticidad de  $\widehat{\mathcal{A}}$ , como veremos más adelante.

### 3.2. El principio de incertidumbre

Nótese que los operadores  $\hat{r}$  y  $\hat{p}$  cumplen con las siguientes relaciones:

$$[\hat{x}_i, \hat{x}_j] = 0 = [\hat{p}_i, \hat{p}_j], \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{i,j}\hat{\mathcal{I}},$$

donde  $[\hat{\mathcal{A}}, \hat{\mathcal{B}}]$  es el conmutador de los operadores  $\hat{\mathcal{A}}$  y  $\hat{\mathcal{B}}$  definido por  $[A, B] = AB - BA$ .

Lo curioso es que a partir de esta relación de conmutación se puede probar que si definimos las varianzas de  $\hat{x}_i$  y  $\hat{p}_i$  mediante las expresiones (acordes con los postulados descritos), i.e.,

$$\Delta x_i = \sqrt{\langle (\hat{x}_i - \langle x_i \rangle \hat{\mathcal{I}})^2 \rangle}, \quad \Delta p_i = \sqrt{\langle (\hat{p}_i - \langle p_i \rangle \hat{\mathcal{I}})^2 \rangle},$$

entonces

$$\Delta x_i \Delta p_i \geq \frac{\hbar}{2},$$

es decir, obtenemos el principio de incertidumbre de Heisenberg que ya vimos antes.

Demostremos la fórmula anterior. Por simplicidad escogeremos el caso unidimensional (o equivalentemente nos restringiremos a probar el principio para las componentes en el eje  $x$ , por ejemplo).

Definamos los operadores

$$\Delta \hat{x} := \hat{x} - \langle x \rangle \hat{\mathcal{I}} \text{ y } \Delta \hat{p} := \hat{p} - \langle p \rangle \hat{\mathcal{I}}.$$

Dichos operadores cumplen con la propiedad

$$[\Delta \hat{x}, \Delta \hat{p}] = [\hat{x}, \hat{p}] = i\hbar \hat{\mathcal{I}}.$$

Obviamente  $\Delta x = \sqrt{\langle (\Delta \hat{x})^2 \rangle}$  y  $\Delta p = \sqrt{\langle (\Delta \hat{p})^2 \rangle}$ , ahora bien, por la desigualdad de Cauchy-Schwarz

$$\sqrt{\int_{\Omega} |\Delta \hat{x} \Psi(x)|^2 dx} \sqrt{\int_{\Omega} |\Delta \hat{p} \Psi(x)|^2 dx} \geq \left| \int_{\Omega} \overline{\Delta \hat{x} \Psi(x)} \Delta \hat{p} \Psi(x) dx \right|$$

y

$$\left| \int_{\Omega} \overline{\Delta \hat{x} \Psi(x)} \Delta \hat{p} \Psi(x) dx \right| \geq \Im \left( \int_{\Omega} \overline{\Delta \hat{x} \Psi(x)} \Delta \hat{p} \Psi(x) dx \right) = \mathcal{I}_{xp},$$

i.e.,

$$\Delta x_i \Delta p_i \geq \mathcal{I}_{xp}.$$

Pero

$$\begin{aligned} \mathcal{I}_{xp} &= \frac{1}{2i} \left( \int_{\Omega} (\overline{\Delta \hat{x} \Psi(x)}) (\Delta \hat{p} \Psi(x)) dx - \int_{\Omega} (\overline{\Delta \hat{p} \Psi(x)}) (\Delta \hat{x} \Psi(x)) dx \right) \\ &= \frac{1}{2i} \left( \int_{\Omega} (\overline{\Delta \hat{x} \Psi(x)}) (\Delta \hat{p} \Psi(x)) dx - \int_{\Omega} (\Delta \hat{x} \Psi(x)) (\overline{\Delta \hat{p} \Psi(x)}) dx \right) \\ &= \frac{1}{2i} \left( \int_{\Omega} (\overline{\Psi(x)}) (\Delta \hat{x} \Delta \hat{p} \Psi(x)) dx - \int_{\Omega} (\overline{\Psi(x)}) (\Delta \hat{p} \Delta \hat{x} \Psi(x)) dx \right) \\ &= \frac{1}{2i} \int_{\Omega} (\overline{\Psi(x)}) ([\Delta \hat{x}, \Delta \hat{p}] \Psi(x)) dx = \frac{1}{2i} \int_{\Omega} (\overline{\Psi(x)}) (i\hbar \hat{\mathcal{I}} \Psi(x)) dx = \frac{\hbar}{2}. \end{aligned}$$

Luego

$$\Delta x \Delta p \geq \frac{\hbar}{2},$$

que es justo lo que queríamos probar.

### 3.2.1. Los estados estacionarios de la ecuación de Schrödinger

Supongamos que  $V$  no depende del tiempo. Entonces si hacemos el cambio

$$\Psi(\vec{r}, t) = \Psi_\alpha(\vec{r})e^{-\frac{i}{\hbar}E_\alpha t} \quad (3.2.1)$$

lo sustituimos en (3.1.2) obtenemos la ecuación

$$\hat{\mathcal{H}}\Psi_\alpha(\vec{r}) = E_\alpha\Psi_\alpha(\vec{r}). \quad (3.2.2)$$

Es decir,  $\Psi_\alpha$  son las autofunciones de  $\hat{\mathcal{H}}$  correspondientes a los autovalores  $E_\alpha$ , que por el postulado 3.1.5 corresponden a los posibles valores de energía del sistema.

Nótese que los estados definidos por (3.2.1) son estados con una energía constante (conservada) en el tiempo, además, la función  $\Psi(\vec{r}, t)$  es periódica en el tiempo con una frecuencia  $\omega_\alpha = E_\alpha/\hbar$ . Los estados con estas características se denominan *estados estacionarios*.

Los estados estacionarios cumplen dos propiedades muy importantes (que los hacen de especial relevancia):

1. La densidad de probabilidad es independiente del tiempo:  $|\Psi(\vec{r}, t)|^2 = |\Psi_\alpha(\vec{r})|^2$ .
2. Si cierto observable  $A$  no depende el tiempo, entonces su media

$$\langle A \rangle = \int_{\Omega} \overline{\Psi(\vec{r}, t)} (\hat{\mathcal{A}}\Psi(\vec{r}, t)) d^3\vec{r} = \int_{\Omega} \overline{\Psi_\alpha(\vec{r})} (\hat{\mathcal{A}}\Psi_\alpha(\vec{r})) d^3\vec{r},$$

tampoco depende del tiempo.

Es decir, los estados estacionarios tienen la energía prefijada y además no evolucionan en el tiempo. Lo anterior es muy útil para resolver la ecuación de Schrödinger general.

Supongamos que

$$\Psi(\vec{r}, t = 0) = \sum_{\alpha} c_{\alpha}\Psi_{\alpha}(\vec{r}), \quad (3.2.3)$$

entonces, la función de onda (estado) se escribe como

$$\Psi(\vec{r}, t) = \sum_{\alpha} c_{\alpha}e^{-\frac{i}{\hbar}E_{\alpha}t}\Psi_{\alpha}(\vec{r}).$$

La pregunta de cuándo el desarrollo (3.2.3) tiene lugar y como calcular los coeficientes  $c_{\alpha}$  lo vamos a dejar para más adelante. En realidad lo que tiene que ocurrir es que las autofunciones del hamiltoniano  $\hat{\mathcal{H}}$  constituyan un conjunto completo (y a ser posible ortonormal) de  $L^2(\Omega)$ .

## 3.3. Ejemplos

Vamos a estudiar dos ejemplos representativos unidimensionales.

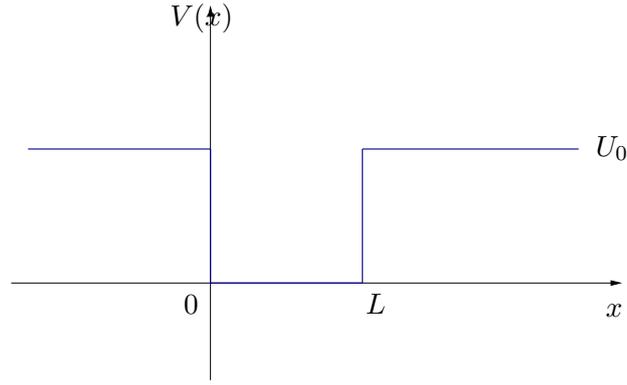


Figura 3.2: Pozo potencial.

### 3.3.1. Una partícula en un pozo de potencial

Sea una partícula que se encuentra en un pozo potencial como el que se muestra en la figura 3.2.

La solución de la ecuación de Schrödinger estacionaria

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x) + V(x)\Psi(x) = E\Psi(x), \quad V(x) = \begin{cases} U_0, & x < 0, \\ 0, & 0 < x < L, \\ U_0, & x > L, \end{cases}$$

para  $E < U_0$  la buscamos de la forma

$$\Psi(x) = \begin{cases} \Psi_1(x) = A_1 e^{kx} + A_2 e^{-kx}, & x < 0, \\ \Psi_2(x) = B_1 e^{iqx} + B_2 e^{-iqx}, & 0 < x < L, \\ \Psi_3(x) = C_1 e^{-kx} + C_2 e^{kx}, & x > L, \end{cases}$$

donde  $k = \frac{1}{\hbar} \sqrt{2m(U_0 - E)}$ ,  $q = \frac{1}{\hbar} \sqrt{2mE}$ .

Como  $\Psi$  ha de ser de cuadrado integrable, entonces  $A_2 = C_2 = 0$ .

Además,  $\Psi$  y  $\Psi'$  han de ser continuas en  $\mathbb{R}$  así que tenemos las siguientes condiciones

$$\Psi_1(0) = \Psi_2(0), \quad \Psi_1'(0) = \Psi_2'(0), \quad \Psi_2(L) = \Psi_3(L), \quad \Psi_2'(L) = \Psi_3'(L),$$

que nos conducen al sistema

$$\begin{cases} A_1 = B_1 + B_2, \\ kA_1 = iq(B_1 - B_2), \\ B_1 e^{iqL} + B_2 e^{-iqL} = C_1 e^{-kL}, \\ iq(B_1 e^{iqL} - B_2 e^{-iqL}) = -kC_1 e^{-kL}. \end{cases}$$

De las dos primeras ecuaciones eliminamos  $A_1$  y de las dos segundas  $C_1$ , así

$$(k - iq)B_1 + (k + iq)B_2 = 0, \quad (k + iq)e^{iqL}B_1 + (k - iq)e^{-iqL}B_2 = 0.$$

Para que este sistema homogéneo tenga solución su determinante ha de anularse

$$\det \begin{vmatrix} k - iq & k + iq \\ (k + iq)e^{iqL} & (k - iq)e^{-iqL} \end{vmatrix} = 0 \implies e^{2iqL} = \left( \frac{k - iq}{k + iq} \right)^2. \quad (3.3.1)$$

Como

$$\left| \frac{k - iq}{k + iq} \right| = 1, \quad e^{i\phi} = \frac{k - iq}{k + iq} \implies$$

$$\cos \phi = \frac{k^2 - q^2}{k^2 + q^2}, \quad \text{sen } \phi = -\frac{2kq}{k^2 + q^2}.$$

Combinando lo anterior con (3.3.1) se tiene  $\phi = qL$ , de donde se sigue que los valores de energía  $E$  permitidos corresponden a la solución de la ecuación trascendental

$$\cos qL = \frac{k^2 - q^2}{k^2 + q^2} \quad \text{o} \quad \text{sen } qL = -\frac{2kq}{k^2 + q^2},$$

o, equivalentemente,

$$\tan qL = -\frac{2kq}{k^2 - q^2}. \quad (3.3.2)$$

Sea  $x = \sqrt{E/U_0}$  y  $\alpha = \sqrt{2mU_0}L/\hbar$ , entonces la ecuación anterior se escribe como

$$\tan \alpha x = -2 \frac{\sqrt{x(1-x)}}{1-2x}. \quad (3.3.3)$$

Es fácil comprobar que para todo  $\alpha$  esta ecuación tiene al menos una raíz real en  $(0, 1)$  por lo que nuestro sistema tiene al menos un estado estacionario. A medida que  $\alpha$  es más grande el número de estados aumenta (véase la figura 3.3).

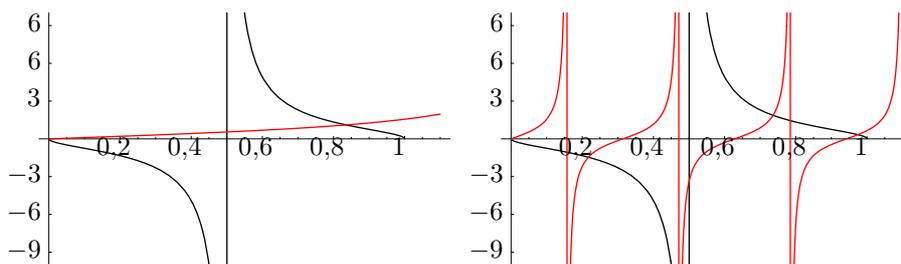


Figura 3.3: Soluciones de la ecuación (3.3.3) para  $\alpha = 1$  (derecha) y  $\alpha = 10$  (izquierda).

Nótese que si tomamos el límite  $U_0 \rightarrow \infty$  en (3.3.2) obtenemos los valores de energía

$$E = \frac{\hbar^2 \pi^2 n^2}{2mL^2}, \quad (3.3.4)$$

que coinciden con los valores de energía para un pozo infinito (ver problema 3.3.1).

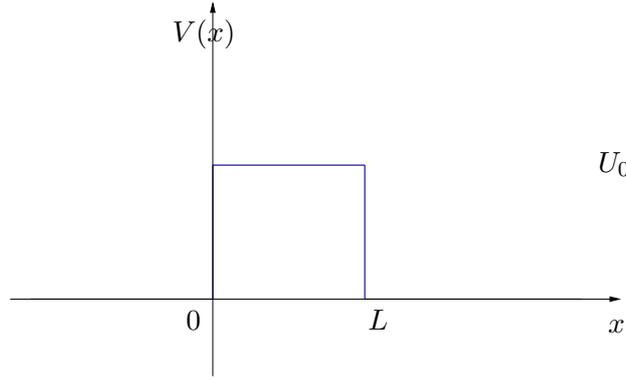


Figura 3.4: Pozo potencial.

### 3.3.2. El efecto túnel

Consideremos ahora el movimiento de una partícula en un potencial del tipo de la figura 3.4

La solución de la ecuación de Schrödinger estacionaria

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x) + V(x)\Psi(x) = E\Psi(x), \quad V(x) = \begin{cases} 0, & x < 0, \\ U_0, & 0 < x < L, \\ 0, & x > L. \end{cases} \quad (3.3.5)$$

Vamos a comenzar considerando el caso  $E > U_0$ . La solución tiene la forma

$$\Psi(x) = \begin{cases} \Psi_1(x) = A_1 e^{ikx} + A_2 e^{-ikx}, & x < 0, \\ \Psi_2(x) = B_1 e^{iqx} + B_2 e^{-iqx}, & 0 < x < L, \\ \Psi_3(x) = C_1 e^{ikx} + C_2 e^{-ikx}, & x > L, \end{cases}$$

donde  $q = \frac{1}{\hbar} \sqrt{2m(E - U_0)}$  y  $k = \frac{1}{\hbar} \sqrt{2mE}$ .

Ante todo notemos que esta solución es de tipo onda viajera, similar a la de una partícula libre, por lo que no es de cuadrado integrable y por tanto no puede describir ningún estado real (véase la discusión del postulado 3.1.3) por lo que habría que definir los correspondientes paquetes de ondas. No obstante podemos considerar nuestra partícula como una onda descrita por la ecuación (3.3.5) y determinar su comportamiento en función de las amplitudes  $A_1, \dots, C_2$  de las mismas. Ante todo notemos que si  $\Psi$  es una solución correspondiente a la energía  $E$ ,  $\bar{\Psi}$  también lo es. Dado un valor de energía podemos considerar dos tipos de soluciones: una  $\Psi_+$  cuando  $C_2 = 0$  (la onda incidente va de izquierda a derecha) y la otra  $\Psi_-$  cuando  $A_1 = 0$  (la onda incidente va de derecha a izquierda). La solución general será una combinación lineal de ambas. Así pues nos centraremos en la solución<sup>10</sup>

$$\Psi_+(x) = \begin{cases} \Psi_1(x) = e^{ikx} + A e^{-ikx}, & x < 0, \\ \Psi_2(x) = B_1 e^{iqx} + B_2 e^{-iqx}, & 0 < x < L, \\ \Psi_3(x) = C e^{ikx}, & x > L. \end{cases}$$

<sup>10</sup>Por la linealidad podemos tomar sin pérdida de generalidad  $A_1 = 1$ .

Los valores  $R(E) = |A|^2$  y  $T(E) = |C|^2$  definen los coeficientes de reflexión y transporte (o paso) de la onda. Usando la continuidad de  $\Psi$  y  $\Psi'$  obtenemos las siguientes condiciones

$$\Psi_1(0) = \Psi_2(0), \quad \Psi'_1(0) = \Psi'_2(0), \quad \Psi_2(L) = \Psi_3(L), \quad \Psi'_2(L) = \Psi'_3(L),$$

que nos conducen al sistema

$$\begin{cases} 1 + A = B_1 + B_2, \\ k(1 - A) = q(B_1 - B_2), \\ B_1 e^{iqL} + B_2 e^{-iqL} = C_1 e^{ikL}, \\ q(B_1 e^{iqL} - B_2 e^{-iqL}) = -kC_1 e^{ikL}. \end{cases} \quad (3.3.6)$$

Multiplicando por  $k$  la tercera y sumándole y restándole la cuarta se obtienen, respectivamente, las ecuaciones

$$B_1 = \frac{(k+q)e^{iL(k-q)}}{2q}C, \quad B_2 = \frac{(q-k)e^{iL(k+q)}}{2q}C.$$

Multiplicando por  $k$  la primera ecuación de (3.3.6) y sumándole y restándole la segunda obtenemos

$$(k+q)B_1 + (k-q)B_2 = 2k, \quad (k-q)B_1 + (k+q)B_2 = 2kA.$$

Si sustituimos en la primera de estas ecuaciones los valores de  $B_1$  y  $B_2$  anteriores obtenemos para  $C$  el valor

$$C = \frac{4kqe^{-ikL}}{(k+q)^2 e^{-iqL} - (k-q)^2 e^{iqL}} = \frac{4kqe^{-ikL}}{(k^2 + q^2) \operatorname{sen}(qL) + 2kqi \cos(qL)}. \quad (3.3.7)$$

Resolviendo respecto a  $A$  obtenemos

$$A = \frac{(k^2 - q^2) \operatorname{sen}(qL)}{(k^2 + q^2) \operatorname{sen}(qL) + 2kqi \cos(qL)}. \quad (3.3.8)$$

De lo anterior se deduce que

$$T(E) = \frac{4k^2 q^2}{(k^2 + q^2)^2 \operatorname{sen}^2(qL) + 4k^2 q^2 \cos^2(qL)} = \frac{1}{1 + \left( \frac{(k^2 - q^2) \operatorname{sen}(qL)}{2kq} \right)^2},$$

y  $R(E) = 1 - T(E)$ .

De la expresión anterior se deduce que  $T(E) \geq 0$ . Además, si  $\operatorname{sen}(qL) = 0$ , entonces  $T(E) = 1$  y  $R(E) = 0$ , es decir, la partícula pasa como si no existiese la barrera: la barrera es transparente. Esto ocurre para los valores de  $qL = n\pi$ , por tanto como  $q = \frac{1}{\hbar} \sqrt{2m(E - U_0)}$   $\implies$

$$E = U_0 + \frac{\hbar^2 \pi^2 n^2}{2mL^2}.$$

Supongamos ahora que  $E < U_0$ . Para obtener el resultado basta hacer el cambio  $q \rightarrow i\kappa$ . Así, obtenemos

$$T(E) = \frac{1}{1 + \left( \frac{(k^2 + \kappa^2) \operatorname{senh}(\kappa L)}{2k\kappa} \right)^2},$$

Nótese que  $T(E) \geq 0$ , es decir, ¡la partícula siempre traspasa la barrera!

Este efecto, imposible en la mecánica clásica, se conoce como efecto túnel. Su gran importancia quedó reflejada en el hecho de que en 1973 L. Esaki, I. Giaever y B.D. Josephson recibieron el premio nobel de física por sus descubrimientos relacionados con este efecto y más tarde en 1986 G. Binnig y H. Rohrer por su diseño de un microscopio electrónico basado en el efecto túnel.

Para terminar este apartado notemos que, para  $E < U_0$ , en el límite  $\kappa L \gg 1$

$$T(E) \sim \frac{16k^2q^2}{(k^2 + q^2)^2} e^{-\kappa L},$$

es decir, el coeficiente de transporte decae exponencialmente a cero con  $L$ . Este límite corresponde al caso  $\sqrt{2m(U_0 - E)}/\hbar \gg 1$ , i.e.,

1.  $m \gg \hbar^2/(2(U_0 - E)L^2)$ , masas muy grandes (límite de la mecánica clásica),
2.  $L^2 \gg \hbar^2/(2(U_0 - E)m)$ , barrera muy ancha,
3.  $U_0 - E \gg \hbar^2/(2mL^2)$ , barrera muy alta.

### 3.3.3. Problemas

**Problema 3.3.1** Encuentra los autoestados de la ecuación de Schrödinger estacionaria para un pozo potencial infinito

$$V(x) = \begin{cases} \infty, & x < 0, \\ 0, & 0 < x < L, \\ \infty, & x > L, \end{cases}$$

y compara el resultado con los valores obtenidos en (3.3.4).

**Problema 3.3.2** Resuelve la ecuación de Schrödinger estacionaria para el potencial escalón

$$V(x) = \begin{cases} 0, & x < 0, \\ U_0, & x > L. \end{cases}$$

Compara el resultado con el del caso de la barrera potencial de anchura infinita.

# Capítulo 4

## Mecánica Cuántica II: Espacios de Hilbert

En este capítulo desarrollaremos la Mecánica Cuántica sobre un espacio de Hilbert general. Aunque el capítulo es autocontenido, es recomendable consultar el Anexo A donde se da una breve introducción a la teoría de espacios funcionales (métricos y normados), así como se desarrolla con más detalles las propiedades de los espacios de Hilbert separables.

### 4.1. Espacios euclídeos y espacios normados

Comenzaremos este apartado recordando algunas definiciones generales.

En adelante asumiremos que  $\mathbb{E}$  es un espacio vectorial complejo<sup>1</sup>.

**Definición 4.1.1** *Se dice que un espacio vectorial  $\mathbb{E}$  es un espacio euclídeo si dados dos elementos cualesquiera  $x, y \in \mathbb{E}$  existe un número denominado producto escalar y que denotaremos por  $\langle x, y \rangle$  tal que*

1. Para todo  $x, y \in \mathbb{E}$ ,  $\langle x, y \rangle = \overline{\langle y, x \rangle}$ .
2. Para todo  $x, y, z \in \mathbb{E}$ ,  $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ .
3. Para todo  $x, y \in \mathbb{E}$  y  $\lambda \in \mathbb{C}$ ,  $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$
4. Para todo  $x \in \mathbb{E}$ ,  $x \neq 0$ ,  $\langle x, x \rangle > 0$  y si  $\langle x, x \rangle = 0$ , entonces  $x = 0$ .

Una consecuencia de la definición anterior es que

1. Para todos  $x, y, z \in \mathbb{E}$ ,  $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ .
2. Para todos  $x, y \in \mathbb{E}$  y  $\lambda \in \mathbb{C}$ ,  $\langle \bar{\lambda}x, y \rangle = \lambda \langle x, y \rangle$

---

<sup>1</sup>En adelante denotaremos por  $\bar{z}$  al complejo conjugado del número complejo  $z$

El ejemplo más sencillo de espacio euclídeo es el espacio  $\mathbb{C}^n$  con el producto escalar estándar. Otro ejemplo sencillo es el espacio  $C[a, b]$  de las funciones continuas en  $[a, b]$  cerrado y acotado con el siguiente producto escalar

$$\langle f, g \rangle = \int_a^b \overline{f(x)}g(x)dx. \quad (4.1.1)$$

Una propiedad importante de los espacios euclídeos es la desigualdad de Cauchy-Schwarz

$$|\langle f, g \rangle|^2 \leq \langle f, f \rangle \langle g, g \rangle. \quad (4.1.2)$$

Para demostrarla basta usar que para todo  $\lambda \in \mathbb{C}$  y  $f, g \in \mathbb{E}$ ,  $\langle \lambda f + g, \lambda f + g \rangle \geq 0$ , o equivalentemente,

$$|\lambda|^2 \langle f, f \rangle + \overline{\lambda} \langle f, g \rangle + \lambda \langle g, f \rangle + \langle g, g \rangle = |\lambda|^2 \langle f, f \rangle + 2\Re(\overline{\lambda} \langle f, g \rangle) + \langle g, g \rangle \geq 0.$$

Escojamos  $\lambda = \alpha \overline{\langle f, g \rangle} / |\langle f, g \rangle|$ ,  $\alpha \in \mathbb{R}$ . Entonces, como

$$0 \leq |\lambda|^2 \langle f, f \rangle + 2\Re(\overline{\lambda} \langle f, g \rangle) + \langle g, g \rangle = |\alpha|^2 \langle f, f \rangle + 2|\alpha| |\langle f, g \rangle| + \langle g, g \rangle,$$

el discriminante de la ecuación cuadrática en  $\alpha$ ,  $|\alpha|^2 \langle f, f \rangle + 2|\alpha| |\langle f, g \rangle| + \langle g, g \rangle = 0$ , ha de ser negativo, luego se tiene  $|\langle f, g \rangle|^2 - \langle f, f \rangle \langle g, g \rangle \leq 0$ , de donde se sigue (4.1.2).

**Definición 4.1.2** *Un espacio vectorial  $\mathbb{X}$  se denomina espacio normado si para todo  $x \in \mathbb{X}$  existe un número real denominado norma, que denotaremos por  $\|x\|$ , que cumple con las condiciones*

1. Para todo  $x \in \mathbb{X}$ ,  $\|x\| \geq 0$  y si  $\|x\| = 0$  entonces  $x = 0$ .
2. Para todo  $x \in \mathbb{X}$  y  $\lambda \in \mathbb{C}$ ,  $\|\lambda x\| = |\lambda| \|x\|$ .
3. Para todo  $x, y \in \mathbb{X}$  se tiene la desigualdad triangular

$$\|x + y\| \leq \|x\| + \|y\|. \quad (4.1.3)$$

**Teorema 4.1.3** *Todo espacio euclídeo  $\mathbb{E}$  es normado si en él definimos la norma mediante la fórmula  $\|x\| = \sqrt{\langle x, x \rangle}$ . Además,  $\|f\| \cdot \|g\| \geq |\langle f, g \rangle|$ .*

La demostración se deja como ejercicio.<sup>2</sup> Por ejemplo, en el espacio  $C[a, b]$  podemos definir la norma por

$$\|f\| = \sqrt{\int_a^b |f(x)|^2 dx}.$$

**Definición 4.1.4** *Un espacio euclídeo  $\mathbb{E}$  completo<sup>3</sup> respecto a la norma inducida por un producto escalar se denomina espacio de Hilbert y lo denotaremos por  $\mathbb{H}$ .*

<sup>2</sup>Es suficiente ver que  $\|f + g\|^2 = \langle f + g, f + g \rangle = \langle f, f \rangle + 2\langle f, g \rangle + \langle g, g \rangle \leq \langle f, f \rangle + 2\sqrt{\langle f, f \rangle} \sqrt{\langle g, g \rangle} + \langle g, g \rangle = (\sqrt{\langle f, f \rangle} + \sqrt{\langle g, g \rangle})^2 = (\|f\| + \|g\|)^2$ . El resto de los axiomas es inmediato.

<sup>3</sup>Un espacio  $\mathbb{E}$  es completo si cualquier sucesión de Cauchy en  $\mathbb{E}$  converge a un vector de  $\mathbb{E}$ .

En adelante nos interesarán los espacios de Hilbert  $\mathbb{H}$  separables, es decir, aquellos espacios de Hilbert que contienen un subconjunto numerable denso.

**Definición 4.1.5** *Sea el sistema de vectores  $\{\phi_n\}_{n=1}^{\infty}$  de  $\mathbb{H}$  linealmente independiente –es decir, que cualquier subsistema finito es linealmente independiente–. Diremos que  $\{\phi_n\}_{n=1}^{\infty}$  es un sistema ortogonal dos a dos si*

$$\langle \phi_n, \phi_m \rangle = \delta_{n,m} \|\phi_n\|^2, \quad \forall n, m \in \mathbb{N}. \quad (4.1.4)$$

Si además  $\|\phi_n\| = 1$  para todo  $n$  diremos que el sistema es ortonormal.

Por ejemplo, el sistema de funciones  $\{1\} \cup \{\sin nx, \cos nx\}_{n=1}^{\infty}$  es un sistema ortogonal dos a dos respecto al producto escalar

$$\langle f, g \rangle = \int_{-\pi}^{\pi} \overline{f(x)}g(x)dx.$$

**Definición 4.1.6** *Dado un vector  $x \in \mathbb{H}$  definiremos la serie de Fourier respecto al sistema ortonormal  $\{\phi_n\}_{n=1}^{\infty}$  a la serie  $s := \sum_{n=1}^{\infty} c_n \phi_n$ , donde los coeficientes vienen dados por las expresiones  $c_n = \langle x, \phi_n \rangle$ , para todo  $n \geq 1$ .*

**Definición 4.1.7** *Dada  $f \in \mathbb{H}$  y una sucesión  $s_n$ , se dice que  $s_n$  converge en norma a  $f$  si*

$$\lim_{n \rightarrow \infty} \|f - s_n\| = 0.$$

Es fácil ver que si los vectores (no nulos)  $\phi_1, \dots, \phi_n$  de un espacio euclídeo son ortogonales, entonces son linealmente independientes.

**Teorema 4.1.8** *En un espacio de Hilbert  $\mathbb{H}$  de cualquier conjunto de vectores linealmente independiente se puede construir un conjunto de vectores ortogonales (ortonormales).*

Demostración: Para probar el teorema tomamos un sistema de vectores linealmente independiente  $(\phi_n)_n$  de  $\mathbb{H}$  cualquiera y definimos un nuevo sistema de vectores  $(\psi_n)_{n=1}^{\infty}$  de la siguiente forma:

1. Tomamos  $\psi_1 = \phi_1 / \|\phi_1\|$ .
2. A continuación escogemos  $\tilde{\psi}_2$  de la forma

$$\tilde{\psi}_2 = \phi_2 + \alpha_{2,1} \psi_1,$$

donde  $\alpha_{2,1}$  es tal que  $\tilde{\psi}_2$  sea ortogonal al vector  $\phi_1$ , i.e.  $\langle \psi_1, \psi_2 \rangle = 0$ , de donde se deduce que  $\alpha_{2,1} = -\langle \psi_1, \phi_2 \rangle$ . Entonces  $\psi_2 = \tilde{\psi}_2 / \|\tilde{\psi}_2\|$  es ortonormal a  $\phi_1$ .

3. Paso  $n$ . Escogemos  $\tilde{\psi}_k$ ,  $k \geq 3$ , de la forma

$$\tilde{\psi}_n = \phi_n + \sum_{k=1}^{n-1} \alpha_{n,k} \psi_k,$$

donde los coeficientes  $\alpha_{n,k}$ ,  $n \in \mathbb{N}$ ,  $k = 1, \dots, n-1$  son tales que  $\tilde{\psi}_n$  sea ortogonal a todos los vectores  $\phi_k$ ,  $k = 1, 2, \dots, n-1$ , anteriores. Usando la ortogonalidad es fácil comprobar que  $\alpha_{n,k} = -\langle \psi_k, \phi_n \rangle$ ,  $k = 1, 2, \dots, n-1$ . Finalmente definimos  $\psi_n = \tilde{\psi}_n / \|\tilde{\psi}_n\|$  que es ortonormal a todos los vectores anteriores  $\phi_k$ ,  $k = 1, 2, \dots, n-1$ . Y así sucesivamente.  $\square$

El proceso anterior se denomina *proceso de ortogonalización de Gram-Schmidt*.

Nótese que del proceso anterior se sigue además que, para cada  $n \geq 1$ ,

$$\tilde{\psi}_n = \phi_n + \sum_{k=1}^{n-1} \alpha_{n,k} \tilde{\psi}_k \implies \phi_n = \tilde{\psi}_n + \sum_{k=1}^{n-1} \beta_{n,k} \tilde{\psi}_k.$$

Luego,

$$\langle \tilde{\psi}_k, \tilde{\psi}_n \rangle = 0, \quad \forall k = 0, 1, \dots, n-1 \quad \Leftrightarrow \quad \langle \phi_k, \tilde{\psi}_n \rangle = 0, \quad \forall k = 0, 1, \dots, n-1.$$

Usando lo anterior es fácil ver que  $\tilde{\psi}_n$  admite la siguiente expresión explícita:

$$\tilde{\psi}_n = \begin{vmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_{n-1} \rangle & \phi_1 \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_{n-1} \rangle & \phi_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_{n-1} \rangle & \phi_n \end{vmatrix}. \quad (4.1.5)$$

Para ello basta notar que el producto escalar  $\langle \phi_k, \tilde{\psi}_n \rangle = 0$ ,  $k = 1, 2, \dots, n-1$ , ya que el determinante resultante tiene dos columnas iguales. Nótese además que

$$\langle \tilde{\psi}_n, \tilde{\psi}_n \rangle = \Delta_n,$$

donde  $\Delta_n$  son los determinantes de Gram

$$\Delta_n = \begin{vmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_{n-1} \rangle & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_{n-1} \rangle & \langle \phi_2, \phi_n \rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_{n-1} \rangle & \langle \phi_n, \phi_n \rangle \end{vmatrix}.$$

De lo anterior deducimos también que los  $\psi_n$  son un conjunto linealmente independiente de  $\mathbb{H}$  si y sólo si los  $\Delta_n \neq 0$  (en nuestro caso  $\Delta_n > 0$ ), para todo  $n \in \mathbb{N}$ .

Nótese que los subespacios generados por los vectores  $(\phi_n)_n$  y  $(\psi_n)_n$  coinciden.

**Teorema 4.1.9** *Si el espacio euclídeo  $\mathbb{E}$  es separable, entonces cualquier sistema ortogonal (ortonormal) de  $\mathbb{E}$  es numerable.*

Demostración: Asumamos sin pérdida de generalidad que el sistema  $(\psi_n)_n$  es ortonormal. Entonces  $\|\psi_n - \psi_m\| = \sqrt{2}$  si  $n \neq m$ . Sea el conjunto de las bolas de radio  $1/2$  y centro en cada  $\psi_n$ ,  $B(\psi_n, 1/2)$ . Estas bolas no se interceptan, luego en cada bola hay un único vector  $\psi_n$  de nuestro sistema ortonormal. Sea ahora  $(\phi_k)_k$  un conjunto numerable denso en  $\mathbb{E}$  (pues éste es separable). Entonces, en cada bola  $B(\psi_n, 1/2)$  habrá al menos un  $\phi_k$ , luego el número de bolas y por tanto de elementos  $\psi_n$  es numerable.  $\square$

## 4.2. Operadores en $\mathbb{H}$

**Definición 4.2.1** Un operador  $\widehat{\mathcal{L}}$  es una aplicación de  $\mathbb{H}$  en  $\mathbb{H}_1$ , dos espacios de Hilbert,  $\widehat{\mathcal{L}} : \mathbb{H} \mapsto \mathbb{H}_1$ .

**Definición 4.2.2** El operador  $\widehat{0}$  se denomina operador nulo si  $\forall \Psi \in \mathbb{H}, \widehat{0}\Psi = 0$ . El operador  $\widehat{\mathcal{I}}$  se denomina operador identidad si  $\forall \Psi \in \mathbb{H}, \widehat{\mathcal{I}}\Psi = \Psi$ .

En adelante asumiremos que  $\mathbb{H}_1 \subset \mathbb{H}$ .

**Definición 4.2.3** Un operador  $\widehat{\mathcal{L}}$  es lineal si  $\forall \alpha_1, \alpha_2 \in \mathbb{C}$ , y  $\Psi_1, \Psi_2 \in \mathbb{H}$ ,

$$\widehat{\mathcal{L}}(\alpha_1\Psi_1 + \alpha_2\Psi_2) = \alpha_1\widehat{\mathcal{L}}\Psi_1 + \alpha_2\widehat{\mathcal{L}}\Psi_2.$$

**Definición 4.2.4** Definiremos el producto  $\widehat{\mathcal{L}} = \widehat{\mathcal{A}}\widehat{\mathcal{B}}$  de dos operadores  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  al operador  $\widehat{\mathcal{L}}$  que obtiene al actuar consecutivamente los operadores  $\widehat{\mathcal{B}}$  y luego  $\widehat{\mathcal{A}}$ , i.e.,

$$\Phi = \widehat{\mathcal{B}}\Psi, \quad \widehat{\mathcal{L}}\Psi = \widehat{\mathcal{A}}\Phi \quad \implies \quad \widehat{\mathcal{L}}\Psi = \widehat{\mathcal{A}}(\widehat{\mathcal{B}}\Psi).$$

En general  $\widehat{\mathcal{A}}\widehat{\mathcal{B}} \neq \widehat{\mathcal{B}}\widehat{\mathcal{A}}$ , i.e., la multiplicación de operadores no es conmutativa.

**Definición 4.2.5** Llamaremos conmutador de dos operadores  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  al operador

$$[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] := \widehat{\mathcal{A}}\widehat{\mathcal{B}} - \widehat{\mathcal{B}}\widehat{\mathcal{A}}.$$

Así pues, dos operadores conmutan si y sólo si su conmutador es el operador nulo.

**Definición 4.2.6** El operador  $\widehat{\mathcal{L}}^{-1}$  se denomina operador inverso de  $\widehat{\mathcal{L}}$  si

$$\widehat{\mathcal{L}}\widehat{\mathcal{L}}^{-1} = \widehat{\mathcal{L}}^{-1}\widehat{\mathcal{L}} = \widehat{\mathcal{I}}.$$

En adelante vamos a usar la notación de Dirac para los vectores, los operadores y los productos escalares.

Así, un vector de  $\mathbb{H}$  lo denotaremos por  $|\Psi\rangle$  (ket vector) y su correspondiente conjugado  $\langle\Psi|$  (brac vector)<sup>4</sup>. Así, denotaremos el producto escalar  $\langle\Psi, \Phi\rangle$  por  $\langle\Psi|\Phi\rangle$  y además

$$\langle\Psi|\widehat{\mathcal{L}}|\Phi\rangle := \langle\Psi|\widehat{\mathcal{L}}\Phi\rangle.$$

A los productos anteriores les denominaremos *elementos matriciales* del operador  $\widehat{\mathcal{L}}$ . En adelante, a no ser que se especifique, asumiremos que los vectores  $\Psi$  están normalizados a la unidad, i.e.,  $\|\Psi\| = 1$ .

**Definición 4.2.7** El operador  $\widehat{\mathcal{L}}^+$  se denomina conjugado o adjunto de  $\widehat{\mathcal{L}}$  si,

$$\langle\Psi|\widehat{\mathcal{L}}|\Phi\rangle = \overline{\langle\Phi|\widehat{\mathcal{L}}^+|\Psi\rangle}, \quad (4.2.1)$$

i.e.,  $\langle\Psi|\widehat{\mathcal{L}}|\Phi\rangle = \langle\widehat{\mathcal{L}}^+\Psi|\Phi\rangle = \overline{\langle\Phi|\widehat{\mathcal{L}}^+\Psi\rangle}$

<sup>4</sup>Estos nombres vienen de la palabra inglesa *bracket*.

De la definición anterior se deduce fácilmente que

1.  $(\widehat{\mathcal{L}}^+)^+ = \widehat{\mathcal{L}}$ ,
2.  $(\alpha\widehat{\mathcal{L}})^+ = \overline{\alpha}\widehat{\mathcal{L}}^+, \forall \alpha \in \mathbb{C}$ ,
3.  $(\widehat{\mathcal{L}}\widehat{\mathcal{N}})^+ = \widehat{\mathcal{N}}^+\widehat{\mathcal{L}}^+$  y
4.  $\langle \Psi | \widehat{\mathcal{L}}\widehat{\mathcal{N}} | \Phi \rangle = \langle \widehat{\mathcal{L}}^+\Psi | \widehat{\mathcal{N}} | \Phi \rangle, \forall \Psi, \Phi \in \mathbb{H}$ .

Un ejemplo especialmente importante es el caso cuando  $\mathbb{H}$  es de dimensión finita. En este caso si  $(\phi_n)_{n=1}^N$  es una base (en particular, una base ortogonal) de  $\mathbb{H}$ , entonces

$$\widehat{\mathcal{L}}\phi_n = \sum_{k=1}^N L_{n,k}\phi_k,$$

y por tanto a  $\widehat{\mathcal{L}}$  se le puede hacer corresponder una matriz  $(L_{i,j})_{i,j=1}^N$ . Si denotamos por  $L_{i,j}^+$  la matriz asociada al operador  $\widehat{\mathcal{L}}^+$ , entonces  $L_{i,j}^+ = \overline{L_{j,i}}$ .

Otro ejemplo son los  $\widehat{\mathcal{L}}$  que admiten una representación integral. Por ejemplo, supongamos que  $\mathbb{H} = L^2(\mathbb{R})$  y

$$\widehat{\mathcal{L}}\Psi(x) = \int_{\mathbb{R}} L(x,y)\Psi(y)dy,$$

donde  $L(x,y)$  es el *núcleo* del operador, entonces, si denotamos por  $L^+(x,y)$  al núcleo de  $\widehat{\mathcal{L}}^+$ ,  $L^+(x,y) = \overline{L(y,x)}$ .

**Definición 4.2.8** Si  $\widehat{\mathcal{L}} = \widehat{\mathcal{L}}^+$ , se dice que el operador es *hermítico* o *autoadjunto*.

Por ejemplo, si  $\mathbb{H}$  es de dimensión finita,  $\widehat{\mathcal{L}}$  es hermítico si su correspondiente matriz satisface  $L_{i,j} = \overline{L_{j,i}}$ . Si  $\mathbb{H} = L^2(\mathbb{R})$ , los operadores definidos por

$$\widehat{x}\Psi(x) = x\Psi(x), \quad \widehat{p}\Psi(x) = -i\hbar\frac{d\Psi(x)}{dx}, \quad \widehat{P}\Psi(x) = \Psi(-x),$$

son hermíticos. En el caso de operadores con representación integral, éstos serán hermíticos si su núcleo es tal que  $L(x,y) = \overline{L(y,x)}$ .

**Proposición 4.2.9** El producto  $\widehat{\mathcal{L}} = \widehat{\mathcal{A}}\widehat{\mathcal{B}}$  de dos operadores  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  hermíticos es hermítico si y sólo si  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  conmutan, i.e.,  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = 0$ .

Demostración: Sea  $\widehat{\mathcal{L}} = \widehat{\mathcal{A}}\widehat{\mathcal{B}}$ . Entonces  $\widehat{\mathcal{L}}^+ = \widehat{\mathcal{B}}^+\widehat{\mathcal{A}}^+ = \widehat{\mathcal{B}}\widehat{\mathcal{A}}$ , luego  $\widehat{\mathcal{L}} = \widehat{\mathcal{L}}^+$  si y sólo si  $\widehat{\mathcal{A}}\widehat{\mathcal{B}} = \widehat{\mathcal{B}}\widehat{\mathcal{A}}$ , i.e.,  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = 0$ .  $\square$

**Proposición 4.2.10** El conmutador  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}]$  de dos operadores hermíticos  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  es tal que

$$[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = i\widehat{\mathcal{L}},$$

con  $\widehat{\mathcal{L}}$  hermítico.

Demostración: Supongamos que  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = \widehat{\mathcal{N}}$ . Entonces

$$\widehat{\mathcal{N}}^+ = ([\widehat{\mathcal{A}}, \widehat{\mathcal{B}}])^+ = -[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = -\widehat{\mathcal{N}} \implies \widehat{\mathcal{N}} = i\widehat{\mathcal{L}},$$

con  $\widehat{\mathcal{L}}$  hermítico ( $(i\widehat{\mathcal{L}})^+ = -i\widehat{\mathcal{L}}^+ = -i\widehat{\mathcal{L}}$ ).  $\square$

**Definición 4.2.11** Sea  $|\Psi\rangle \in \mathbb{H}$  con  $\| |\Psi\rangle \| \neq 0$ . Si existe  $\lambda \in \mathbb{C}$  tal que

$$\widehat{\mathcal{L}}|\Psi\rangle = \lambda|\Psi\rangle,$$

entonces se dice que  $|\Psi\rangle$  es un autovector de  $\widehat{\mathcal{L}}$  y  $\lambda$  es su correspondiente autovalor.

**Nota:** En ocasiones es cómodo denotar a un autovector asociado a  $\lambda$  por  $|\Psi_\lambda\rangle$  (suponiendo que  $\lambda$  es un autovalor simple). Si además el conjunto de autovalores es numerable entonces se suele simplificar aún más la notación:  $|n\rangle := |\Psi_{\lambda_n}\rangle$ .

**Proposición 4.2.12** Si  $\widehat{\mathcal{L}}$  es hermítico, entonces sus autovalores son reales.

Demostración:

$$\widehat{\mathcal{L}}|\Psi\rangle = \lambda|\Psi\rangle \implies \langle \Psi | \widehat{\mathcal{L}} | \Psi \rangle = \lambda \langle \Psi | \Psi \rangle = \lambda.$$

Por otro lado de (4.2.1) se sigue que

$$\langle \Psi | \widehat{\mathcal{L}}^+ | \Psi \rangle = \overline{\langle \Psi | \widehat{\mathcal{L}} | \Psi \rangle} = \bar{\lambda} \langle \Psi | \Psi \rangle = \bar{\lambda},$$

luego, como  $\widehat{\mathcal{L}}$  es hermítico  $\widehat{\mathcal{L}} = \widehat{\mathcal{L}}^+$  por tanto  $\lambda = \bar{\lambda}$ .  $\square$

**Proposición 4.2.13** Si  $\widehat{\mathcal{L}}$  es hermítico, entonces los autovectores correspondientes a autovalores distintos son ortogonales.

Demostración: Sea  $\widehat{\mathcal{L}}|\Psi_1\rangle = \lambda_1|\Psi_1\rangle$ ,  $\widehat{\mathcal{L}}|\Psi_2\rangle = \lambda_2|\Psi_2\rangle$ , entonces

$$\begin{aligned} \langle \Psi_2 | \widehat{\mathcal{L}} | \Psi_1 \rangle &= \lambda_1 \langle \Psi_2 | \Psi_1 \rangle = \\ \langle \Psi_2 | \widehat{\mathcal{L}}^+ | \Psi_1 \rangle &= \overline{\langle \Psi_1 | \widehat{\mathcal{L}} | \Psi_2 \rangle} = \lambda_2 \overline{\langle \Psi_1 | \Psi_2 \rangle} = \lambda_2 \langle \Psi_2 | \Psi_1 \rangle, \end{aligned}$$

i.e.  $(\lambda_1 - \lambda_2)\langle \Psi_2 | \Psi_1 \rangle = 0$ , luego como  $\lambda_1 \neq \lambda_2$ ,  $\langle \Psi_2 | \Psi_1 \rangle = 0$ .  $\square$

**Definición 4.2.14** Un operador  $\widehat{\mathcal{U}}$  se denomina unitario si

$$\widehat{\mathcal{U}}\widehat{\mathcal{U}}^+ = \widehat{\mathcal{U}}^+\widehat{\mathcal{U}} = \widehat{\mathcal{I}}.$$

**Proposición 4.2.15** Si  $\widehat{\mathcal{U}}$  es unitario, entonces todos sus autovalores son tales que  $|\lambda| = 1$ .

Demostración: Sea  $\widehat{\mathcal{U}}|\Psi\rangle = \lambda|\Psi\rangle$ , entonces

$$1 = \langle \Psi | \widehat{\mathcal{U}}^+ \widehat{\mathcal{U}} | \Psi \rangle = \lambda \langle \Psi | \widehat{\mathcal{U}}^+ | \Psi \rangle = \overline{\lambda \langle \Psi | \widehat{\mathcal{U}} | \Psi \rangle} = \lambda \bar{\lambda} \langle \Psi | \Psi \rangle = |\lambda|^2$$

$\square$

**Definición 4.2.16** Sea  $\widehat{U}$  un operador unitario. La transformación

$$|\Psi\rangle \mapsto |\psi\rangle = \widehat{U}^+|\Psi\rangle, \quad \widehat{\mathcal{L}} \mapsto \widehat{\ell} = \widehat{U}^+\widehat{\mathcal{L}}\widehat{U},$$

la denominaremos transformación unitaria de  $|\Psi\rangle$  y  $\widehat{\mathcal{L}}$  y la denotaremos por  $\{U\}$ .

**Proposición 4.2.17** Las transformaciones unitarias conservan

1. Las relaciones de conmutación de los operadores.
2. La propiedad de hermiticidad de un operador.
3. Los autovalores.
4. Los productos escalares y elementos matriciales de un operador.

Demostración: 1. Sean  $\widehat{\mathcal{A}}$ ,  $\widehat{\mathcal{B}}$  y  $\widehat{\mathcal{L}}$  tres operadores tales que  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = \widehat{\mathcal{L}}$  y sea  $\{U\}$  una transformación unitaria. Denotemos por  $\widehat{a}$ ,  $\widehat{b}$  y  $\widehat{\ell}$  los operadores correspondientes a  $\widehat{\mathcal{A}}$ ,  $\widehat{\mathcal{B}}$  y  $\widehat{\mathcal{L}}$  después de la transformación. Entonces como  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = \widehat{\mathcal{L}} \implies$

$$\widehat{\mathcal{A}}\widehat{\mathcal{B}} - \widehat{\mathcal{B}}\widehat{\mathcal{A}} = \widehat{\mathcal{L}} \implies \widehat{U}^+\widehat{\mathcal{A}}\widehat{U}\widehat{U}^+\widehat{\mathcal{B}}\widehat{U} - \widehat{U}^+\widehat{\mathcal{B}}\widehat{U}\widehat{U}^+\widehat{\mathcal{A}}\widehat{U} = \widehat{U}^+\widehat{\mathcal{L}}\widehat{U} = \widehat{\ell},$$

pero

$$(\widehat{U}^+\widehat{\mathcal{A}}\widehat{U})(\widehat{U}^+\widehat{\mathcal{B}}\widehat{U}) - (\widehat{U}^+\widehat{\mathcal{B}}\widehat{U})(\widehat{U}^+\widehat{\mathcal{A}}\widehat{U}) = \widehat{a}\widehat{b} - \widehat{b}\widehat{a} \implies [\widehat{a}, \widehat{b}] = \widehat{\ell}.$$

2. Sea  $\widehat{\mathcal{L}} = \widehat{\mathcal{L}}^+$ . Sea  $\widehat{\ell} = \widehat{U}^+\widehat{\mathcal{L}}\widehat{U}$ , entonces

$$\widehat{\ell}^+ = (\widehat{U}^+\widehat{\mathcal{L}}\widehat{U})^+ = \widehat{U}^+\widehat{\mathcal{L}}^+\widehat{U} = \widehat{U}^+\widehat{\mathcal{L}}\widehat{U} = \widehat{\ell}.$$

3. Sea  $\widehat{\mathcal{L}}|\Psi\rangle = \lambda|\Psi\rangle$ , entonces

$$(\widehat{U}^+\widehat{\mathcal{L}}\widehat{U})(\widehat{U}^+|\Psi\rangle) = \lambda(\widehat{U}^+|\Psi\rangle) \implies \widehat{\ell}|\psi\rangle = \lambda|\psi\rangle.$$

4.  $\langle\Psi_1|\widehat{\mathcal{L}}|\Psi_2\rangle = \langle\Psi_1|\widehat{U}\widehat{U}^+\widehat{\mathcal{L}}\widehat{U}\widehat{U}^+|\Psi_2\rangle = \langle\widehat{U}^+\Psi_1|\widehat{U}^+\widehat{\mathcal{L}}\widehat{U}|\widehat{U}^+\Psi_2\rangle = \langle\psi_1|\widehat{\ell}|\psi_2\rangle.$   $\square$

**Definición 4.2.18** Sea  $F(z)$  una función analítica en un entorno de  $z = 0$  y sea  $F(z) = \sum_{n \geq 0} f_n z^n$  su desarrollo en serie de potencias. Definiremos al operador  $F(\widehat{\mathcal{L}})$  mediante la serie

$$F(\widehat{\mathcal{L}}) = \sum_{n \geq 0} f_n \widehat{\mathcal{L}}^n.$$

**Definición 4.2.19** La derivada operacional  $\partial F(\widehat{\mathcal{L}})/\partial \widehat{\mathcal{L}}$  es el operador que se obtiene mediante la formula

$$\frac{\partial F(\widehat{\mathcal{L}})}{\partial \widehat{\mathcal{L}}} = \lim_{\varepsilon \rightarrow 0} \frac{F(\widehat{\mathcal{L}} + \varepsilon \widehat{\mathcal{I}}) - F(\widehat{\mathcal{L}})}{\varepsilon}. \quad (4.2.2)$$

Por ejemplo

$$\frac{\partial \widehat{\mathcal{L}}^n}{\partial \widehat{\mathcal{L}}} = n \widehat{\mathcal{L}}^{n-1}.$$

Sean  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{A}}^+$  tales que  $[\widehat{\mathcal{A}}, \widehat{\mathcal{A}}^+] = \widehat{\mathcal{I}}$ . Entonces es fácil comprobar que

$$[\widehat{\mathcal{A}}, (\widehat{\mathcal{A}}^+)^k] = k(\widehat{\mathcal{A}}^+)^{k-1}, \quad k \geq 1. \quad (4.2.3)$$

**Proposición 4.2.20** *Si  $F(z)$  es una función analítica en un entorno de  $z = 0$  y sean  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{A}}^+$  tales que  $[\widehat{\mathcal{A}}, \widehat{\mathcal{A}}^+] = \widehat{\mathcal{I}}$ . Entonces*

$$[\widehat{\mathcal{A}}, F(\widehat{\mathcal{A}}^+)] = \frac{dF(\widehat{\mathcal{A}}^+)}{d\widehat{\mathcal{A}}^+}.$$

Demostración: Basta escribir la serie de potencias de  $F$  y usar la propiedad (4.2.3).  $\square$

**Nota:** Si  $F$  admite un desarrollo en serie de Laurent el resultado también es válido.

Antes de continuar con nuestro análisis procede hacer un breve comentario sobre un tipo especial de operadores unitarios: los operadores unitarios cercanos a la unidad. Sea  $\epsilon > 0$  tan pequeño como se quiera, y supongamos que  $\widehat{\mathcal{U}}_\epsilon$  admite el desarrollo

$$\widehat{\mathcal{U}}_\epsilon = \widehat{\mathcal{I}} + i\epsilon \widehat{\mathcal{A}} + O(\epsilon^2).$$

Entonces su conjugado es, a primer orden,

$$\widehat{\mathcal{U}}_\epsilon^+ = \widehat{\mathcal{I}} - i\epsilon \widehat{\mathcal{A}}^+ + O(\epsilon^2).$$

Como  $\widehat{\mathcal{I}} = \widehat{\mathcal{U}}^+ \widehat{\mathcal{U}} = \widehat{\mathcal{I}} + i\epsilon(\widehat{\mathcal{A}} - \widehat{\mathcal{A}}^+) + O(\epsilon^2)$ , deducimos que  $\widehat{\mathcal{A}} = \widehat{\mathcal{A}}^+$ , es decir  $\widehat{\mathcal{A}}$  ha de ser un operador hermítico. Si elegimos  $\epsilon = \alpha/N$ , con  $N \in \mathbb{N}$  entonces se puede comprobar que

$$\widehat{\mathcal{U}}_\epsilon^N = \left( \widehat{\mathcal{I}} + \frac{i\alpha \widehat{\mathcal{A}}}{N} \right)^N \rightarrow e^{i\alpha \widehat{\mathcal{A}}} = \widehat{\mathcal{U}}(\alpha).$$

Se dice entonces que  $\widehat{\mathcal{A}}$  es un generador de la transformación  $\widehat{\mathcal{U}}$ .

Los operadores que nos van a interesar son aquellos operadores hermíticos cuyo conjunto de autovectores constituyan un sistema completo de  $\mathbb{H}$ , es decir que todo vector  $|\Psi\rangle \in \mathbb{H}$  se puede *expresar* biunívocamente en función de dicho sistema. Este problema (de encontrar dichos operadores y el conjunto de sus autovalores y autovectores y comprobar que constituyen un sistema completo) es muy complicado y requiere de la potente maquinaria de la teoría de operadores en espacio de Hilbert, el teorema espectral de Riesz, etc. Basta mencionar que si  $\mathbb{H}$  es separable, entonces en  $\mathbb{H}$  existe una base ortonormal completa y si  $\widehat{\mathcal{L}}$  es un operador autoadjunto (hermítico) y compacto, entonces sus autovectores constituyen un sistema ortogonal completo. Para más detalles véase el Apéndice A. Conviene no obstante destacar que muchos operadores importantes de la mecánica cuántica no son siquiera acotados.

En adelante, por sencillez, nos vamos a restringir a considerar aquellos operadores que tengan asociados un conjunto numerable de autovectores y que dicho conjunto sea un sistema completo.

Supongamos que  $\widehat{\mathcal{L}}$  es uno de tales operadores y denotemos por  $(|\Psi_n\rangle)_n$  su conjunto completo de autovectores. Si todos los autovalores son simples entonces, como ya hemos visto, los correspondientes autovectores son ortogonales. En el caso de que tengamos autovalores múltiples sus correspondientes autovectores se pueden ortogonalizar usando el método de Gram-Schmidt que describimos antes. Así pues asumiremos que  $(|\Psi_n\rangle)_n$  es un sistema ortonormal (ortogonal con  $\| |\Psi_n\rangle \| = 1$ ).

Sea  $|\Phi\rangle$  un vector cualquiera de  $\mathbb{H}$ , entonces  $|\Phi\rangle$  se puede desarrollar en serie de Fourier respecto  $(|\Psi_n\rangle)_n$

$$|\Phi\rangle = \sum_n f_n |\Psi_n\rangle, \quad f_n = \langle \Psi_n | \Phi \rangle.$$

En otras palabras,  $(|\Psi_n\rangle)_n$  es una *base ortonormal completa* de  $\mathbb{H}$ .

Las bases juegan un papel fundamental. En particular, las bases asociadas a operadores hermíticos.

Sea  $(|\Psi_n\rangle)_n$  una base ortonormal completa de  $\mathbb{H}$  y sea  $\widehat{\mathcal{A}}$  un operador lineal, entonces

$$\widehat{\mathcal{A}}|\Psi_n\rangle \in \mathbb{H} \implies \widehat{\mathcal{A}}|\Psi_n\rangle = \sum_m A_{m,n} |\Psi_m\rangle \implies A_{m,n} = \langle \Psi_m | \widehat{\mathcal{A}}|\Psi_n\rangle.$$

A la cantidad  $\langle \Psi_m | \widehat{\mathcal{A}}|\Psi_n\rangle$  la denominaremos elemento matricial del operador  $\widehat{\mathcal{A}}$  en la base  $(|\Psi_n\rangle)_n$ .

Si  $(|\Psi_n\rangle)_n$  es la base asociada a cierto operador hermítico  $\widehat{\mathcal{L}}$  se dice que la matriz  $A = (A_{m,n})$  es la matriz del operador  $\widehat{\mathcal{A}}$  en la  $\widehat{\mathcal{L}}$ -representación. Nótese que la matriz del operador  $\widehat{\mathcal{L}}$  en su propia representación (la  $\widehat{\mathcal{L}}$ -representación) es una matriz diagonal con los autovalores en la diagonal.

Más aún, así como el conjunto de números  $(f_n)_n$  define biunívocamente el vector  $|\Phi\rangle$ , la matriz  $A$  define biunívocamente al operador  $\widehat{\mathcal{A}}$  (en la base correspondiente se sobrentiende). Así pues, el operador  $\widehat{\mathcal{A}}$  será hermítico si su matriz  $A$  es autoconjugada, i.e.,  $A_{m,n} = \overline{A_{n,m}}$ ,  $\widehat{\mathcal{A}}$  será unitario si su matriz  $A$  es unitaria, i.e.,  $\sum_k A_{m,k} \overline{A_{n,k}} = \delta_{m,n}$ , etc.

**Nota:** Nótese que si  $\mathbb{H}$  es de dimensión finita, las correspondientes matrices son matrices cuadradas  $N \times N$  donde  $N$  es la dimensión de  $\mathbb{H}$ , pero si  $\mathbb{H}$  es de dimensión infinita, entonces las correspondientes matrices son infinitas.

**Proposición 4.2.21** *Si dos operadores  $\widehat{\mathcal{L}}$  y  $\widehat{\mathcal{N}}$  tienen un sistema completo de autovectores  $(|\Psi_n\rangle)_n$  común, entonces  $[\widehat{\mathcal{L}}, \widehat{\mathcal{N}}] = 0$ .*

Demostración: Sean  $\lambda_n$  los autovalores de  $\widehat{\mathcal{L}}$  correspondientes a los autovectores  $|\Psi_n\rangle$  y  $\nu_n$  los del operador  $\widehat{\mathcal{N}}$  correspondientes al mismo autovector (son comunes). En las premisas del teorema tenemos

$$\widehat{\mathcal{L}}\widehat{\mathcal{N}}|\Psi_n\rangle = \nu_n \widehat{\mathcal{L}}|\Psi_n\rangle = \lambda_n \nu_n |\Psi_n\rangle = \lambda_n \widehat{\mathcal{N}}|\Psi_n\rangle = \widehat{\mathcal{N}}(\lambda_n |\Psi_n\rangle) = \widehat{\mathcal{N}}\widehat{\mathcal{L}}|\Psi_n\rangle.$$

Sea  $|\Phi\rangle \in \mathbb{H}$  cualquiera. Entonces

$$\begin{aligned} \widehat{\mathcal{L}}\widehat{\mathcal{N}}|\Phi\rangle &= \widehat{\mathcal{L}}\widehat{\mathcal{N}} \sum_n f_n |\Psi_n\rangle = \sum_n f_n \widehat{\mathcal{L}}\widehat{\mathcal{N}}|\Psi_n\rangle = \sum_n f_n \widehat{\mathcal{N}}\widehat{\mathcal{L}}|\Psi_n\rangle \\ &= \widehat{\mathcal{N}}\widehat{\mathcal{L}} \sum_n f_n |\Psi_n\rangle = \widehat{\mathcal{N}}\widehat{\mathcal{L}}|\Phi\rangle \implies [\widehat{\mathcal{L}}, \widehat{\mathcal{N}}] = 0. \end{aligned}$$

□

El recíproco también es cierto:

**Proposición 4.2.22** *Si dos operadores  $\widehat{\mathcal{L}}$  y  $\widehat{\mathcal{N}}$  con sistemas completos de autovectores conmutan ( $[\widehat{\mathcal{L}}, \widehat{\mathcal{N}}] = 0$ ), entonces tienen un sistema completo de autovectores  $(|\Psi_n\rangle)_n$  común.*

**Demostración:** Para demostrarlo vamos a probar que ambos se pueden diagonalizar al mismo tiempo. Supongamos que conocemos el conjunto  $(|\Psi_n\rangle)_n$  de los autovectores de  $\widehat{\mathcal{L}}$ . En esa base la matriz de  $\widehat{\mathcal{L}}$  es diagonal:  $L_{m,n} = \lambda_n \delta_{m,n}$ . Como  $[\widehat{\mathcal{L}}, \widehat{\mathcal{N}}] = 0$  entonces las matrices correspondientes a  $\widehat{\mathcal{L}}\widehat{\mathcal{N}}$  y  $\widehat{\mathcal{N}}\widehat{\mathcal{L}}$  son iguales

$$(LN)_{m,n} = (NL)_{m,n} \implies \sum_k L_{m,k} N_{k,n} = \sum_k N_{m,k} L_{k,n} \implies$$

$$\sum_k \lambda_m \delta_{m,k} N_{k,n} = \sum_k N_{m,k} \lambda_n \delta_{k,n} \implies N_{m,n} (\lambda_n - \lambda_m) = 0.$$

Si todos los autovalores son distintos, entonces  $N_{m,n} = 0$  si  $m \neq n$ , es decir la matriz de  $\widehat{\mathcal{N}}$  es diagonal. Resta probar qué ocurre si hay autovalores múltiples. Lo que dejaremos como ejercicio al lector<sup>5</sup>. □

En adelante asumiremos que el espacio de Hilbert  $\mathbb{H}$  es separable.

### 4.3. Los axiomas de la Mecánica Cuántica

**Postulado 4.3.1** *A cada sistema físico se le hace corresponder un espacio de Hilbert separable  $\mathbb{H}$  apropiado. Además, para cada  $t \in \mathbb{R}$  (parámetro correspondiente al tiempo) el estado queda completamente caracterizado por un vector  $|\Psi\rangle$  normalizado a la unidad de  $\mathbb{H}$ .*

Es decir, para cada  $t$  el estado está determinado por un vector de  $\mathbb{H}$  tal que  $\|\Psi\| = 1$ . De aquí también se sigue que, dados los estados  $|\Psi_1\rangle, \dots, |\Psi_k\rangle$ , la combinación lineal  $|\Phi\rangle = \sum_{k=1}^n \alpha_k |\Psi_k\rangle$  también es un (posible) estado<sup>6</sup>.

**Postulado 4.3.2** *A cada magnitud física medible (observable)  $L$  se le hace corresponder un operador lineal hermítico  $\widehat{\mathcal{L}}$  que actúa en  $\mathbb{H}$ .*

**Postulado 4.3.3** *Sea  $|\Psi\rangle$  el estado del sistema en el momento  $t$  justo antes de la medición de la magnitud (observable)  $L$  (asociada al operador  $\widehat{\mathcal{L}}$ ). Independientemente de cuál sea el estado original  $|\Psi\rangle$ , el resultado de la medición sólo puede ser un autovalor de  $\widehat{\mathcal{L}}$ .*

<sup>5</sup>Supongamos que la multiplicidad de cierto autovalor  $\lambda_k$  es  $g$  y sean  $\Psi_{k,j}$ ,  $j = 1, 2, \dots, g$ , los correspondientes autovectores asociados. Entonces, como mucho hay  $g(g-1)$  elementos matriciales  $N_{m,n}$  no diagonales distintos de cero, los correspondientes a  $\langle \Psi_{k,i} | \widehat{\mathcal{N}} | \Psi_{k,j} \rangle$ , donde  $i \neq j$ ,  $i, j = 1, 2, \dots, g$ . Basta probar que existen ciertas combinaciones lineales  $\Phi_{k,j}$  de los correspondientes autovectores  $|\Psi_{k,j}\rangle$ ,  $j = 1, 2, \dots, g$ , tales que  $\langle \Phi_{k,i} | \widehat{\mathcal{N}} | \Phi_{k,j} \rangle = 0$ .

<sup>6</sup>Para cada  $t \in \mathbb{R}$  el vector  $|\Psi\rangle$  siempre se puede normalizar a la unidad (a no ser  $|\Psi\rangle = 0$ ).

Este postulado requiere una aclaración y es que al hacer una medición de  $\widehat{\mathcal{L}}$  el sistema cambia (las mediciones interfieren en el sistema). Así pues, antes de medir  $L$  el sistema puede estar en *cualquier* estado  $\Psi$ , pero al realizar la medición, ésta cambia al sistema y lo deja en el estado determinado por el vector  $|\Psi_\lambda\rangle$  que pertenece al autoespacio de  $\widehat{\mathcal{L}}$  correspondiente al autovalor  $\lambda$ .<sup>7</sup>

**Postulado 4.3.4** *El valor esperado  $\langle L \rangle$  de una cantidad física  $L$  cuando el sistema se encuentra en el estado  $|\Psi\rangle$  viene dado por el elemento matricial*

$$\langle L \rangle = \langle \Psi | \widehat{\mathcal{L}} | \Psi \rangle.$$

Nótese que, como  $\widehat{\mathcal{L}}$  es hermítico, entonces

$$\langle \Psi | \widehat{\mathcal{L}} | \Psi \rangle = \overline{\langle \Psi | \widehat{\mathcal{L}}^\dagger | \Psi \rangle} = \overline{\langle \Psi | \widehat{\mathcal{L}} | \Psi \rangle} \implies \langle L \rangle \in \mathbb{R}.$$

**Postulado 4.3.5** *Los elementos matriciales de los operadores  $\widehat{x}_i$  de la posición (coordenadas)  $x_i$  y  $\widehat{p}_i$  de los momentos  $p_i$ ,  $i = 1, 2, 3$ , donde los índices  $i = 1, 2, 3$  corresponden a las proyecciones en los ejes  $x$ ,  $y$  y  $z$ , respectivamente, definidos por  $\langle \Phi | \widehat{x}_i | \Psi \rangle$  y  $\langle \Phi | \widehat{p}_i | \Psi \rangle$ , cualquiera sean  $|\Phi\rangle$  y  $|\Psi\rangle$  de  $\mathbb{H}$  satisfacen las ecuaciones de evolución*

$$\frac{d}{dt} \langle \Phi | \widehat{x}_i | \Psi \rangle = \left\langle \Phi \left| \frac{\partial \widehat{\mathcal{H}}}{\partial \widehat{p}_i} \right| \Psi \right\rangle, \quad \frac{d}{dt} \langle \Phi | \widehat{p}_i | \Psi \rangle = - \left\langle \Phi \left| \frac{\partial \widehat{\mathcal{H}}}{\partial \widehat{x}_i} \right| \Psi \right\rangle, \quad (4.3.1)$$

donde  $\widehat{\mathcal{H}}$  es el operador asociado a la función de Hamilton del correspondiente sistema clásico (si es que lo hay).

Este postulado tiene un significado físico evidente pues nos indica que el promedio de las cantidades medibles posición, impulso y energía (hamiltoniano) satisfacen las ecuaciones dinámicas de la mecánica hamiltoniana (1.1.1), i.e, en el límite apropiado ( $\hbar \rightarrow 0$ ) la Mecánica cuántica se transforma en la clásica (principio de correspondencia de Bohr).

Proceden unas aclaraciones. En general el Hamiltoniano  $H$  de un sistema clásico depende de las coordenadas  $x_i$  y los impulsos  $p_i$ ,  $i = 1, 2, 3$ , por lo que el operador  $\widehat{\mathcal{H}}$  se obtiene cambiando las  $x_i$  por los correspondientes operadores  $\widehat{x}_i$  y  $p_i$  por  $\widehat{p}_i$ . Esto, aunque en apariencia es *trivial*, en general no lo es pues  $\widehat{\mathcal{H}}$  debe ser hermítico (ya que corresponde a la magnitud física energía). A esto regresaremos en breve, pero antes introduciremos nuestro último postulado.

**Postulado 4.3.6** *Los operadores posición  $\widehat{x}_i$  e impulso  $\widehat{p}_i$ ,  $i = 1, 2, 3$ , satisfacen las relaciones de conmutación*

$$[\widehat{x}_k, \widehat{x}_j] = 0 = [\widehat{p}_k, \widehat{p}_j], \quad [\widehat{x}_k, \widehat{p}_j] = i\hbar \delta_{k,j} \widehat{\mathcal{I}}, \quad (4.3.2)$$

donde  $\hbar$  es una constante e  $i = \sqrt{-1}$ .

En particular, de lo anterior se sigue que los operadores  $\widehat{x}_k$  y  $\widehat{p}_k$  no pueden tener un conjunto completo de autovectores comunes. Este postulado es el análogo de las relaciones (1.1.4) (llaves de Poisson).

---

<sup>7</sup>Los operadores cuánticos se postulan en la teoría. En el capítulo 3 vimos varios ejemplos de los mismos. Tres operadores esenciales son el operador posición o coordenadas  $\widehat{x}_k$ , impulso  $\widehat{p}_k$ , y el hamiltoniano del sistema  $\widehat{\mathcal{H}}$ .

## 4.4. Discusión e implicaciones de los postulados

1. Supongamos que tenemos una magnitud física clásica  $L$  que depende en general de  $x_i$  y  $p_i$ . Para construir el operador mecano-cuántico correspondiente sólo tenemos que cambiar los  $x_i$  por los correspondientes operadores  $\hat{x}_i$  y  $p_i$  por  $\hat{p}_i$ . Por ejemplo, la energía cinética viene dada por

$$T = \frac{p_1^2 + p_2^2 + p_3^2}{2m} \implies \hat{T} = \frac{\hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2}{2m},$$

y la potencial  $V(x_1, x_2, x_3)$  por  $\hat{V} = V(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ , donde en ambos casos los operadores son hermíticos. Esto no siempre ocurre. Por ejemplo imaginemos que el hamiltoniano clásico contiene el término  $W_i = x_i p_i$ . Entonces, el operador  $\hat{W}_i = \hat{x}_i \hat{p}_i$  no puede representar al correspondiente operador cuántico ya que no es hermítico (ver Proposición 4.2.9) pues  $\hat{x}_i$  y  $\hat{p}_i$  no conmutan. En este caso hay que definir  $\hat{W}_i$  por

$$\hat{W}_i = \frac{1}{2}(\hat{x}_i \hat{p}_i + \hat{p}_i \hat{x}_i).$$

2. Supongamos el sistema físico se encuentra en el estado definido por  $|\Psi_n\rangle$ , autovector correspondiente al autovalor  $\lambda_n$  de cierto operador  $\hat{\mathcal{L}}$  asociado a la magnitud física  $L$ . Entonces<sup>8</sup>

$$\langle \Psi_n | \hat{\mathcal{L}} | \Psi_n \rangle = \lambda_n, \quad \langle \Psi_n | \hat{\mathcal{L}}^k | \Psi_n \rangle = \lambda_n^k.$$

Supongamos ahora que el sistema se encuentra en el estado  $|\Phi\rangle$  que es en una superposición de los estados  $|\Psi_k\rangle$ ,  $k = 1, 2, \dots, N$ , entonces como  $|\Phi\rangle = \sum_k f_k |\Psi_k\rangle$  tenemos

$$\langle \Phi | \hat{\mathcal{L}} | \Phi \rangle = \sum_k |f_k|^2 \lambda_k.$$

Lo anterior indica, en virtud de postulado 4.3.4, que la cantidad  $|f_k|^2$  es la probabilidad con que se observa el valor  $\lambda_k$  al hacer una medición. Lo anterior implica además que tras la medición el sistema va a parar al estado definido por un vector del espacio generado por los autovectores correspondientes a  $\lambda_k$ . Así pues, en el caso de que el autovalor  $\lambda_k$  sea simple la probabilidad de que el sistema estando en un estado original definido por el vector  $|\Phi\rangle$  termine en el estado definido por  $|\Psi_k\rangle$  es

$$\text{Prob}(|\Phi\rangle \mapsto |\Psi_k\rangle) = |f_k|^2 = |\langle \Psi_k | \Phi \rangle|^2.$$

Nótese que esta probabilidad es por tanto invariante ante transformaciones unitarias:  $\Psi_k \mapsto \hat{\mathcal{U}}\Psi_k = \tilde{\Psi}_k$ ,  $|\Phi\rangle \mapsto \hat{\mathcal{U}}|\Phi\rangle = |\tilde{\Phi}\rangle$  pues  $\text{Prob}(|\Phi\rangle \mapsto |\Psi_k\rangle) = \text{Prob}(|\tilde{\Phi}\rangle \mapsto |\tilde{\Psi}_k\rangle)$ . Luego el sistema físico es invariante frente a cualquier transformación unitaria.

3. Dada cualquier cantidad física clásica  $L$  le podemos adicionar la cantidad  $x_i p_j - p_j x_i$  sin cambiarla. Si transformamos  $L$  en su operador  $\hat{\mathcal{L}}$  ya no le podemos adicionar el correspondiente operador  $\hat{x}_i \hat{p}_j - \hat{p}_j \hat{x}_i$  pues éste no es nulo (ver postulado 4.3.6).

Ahora bien, tomando las derivadas funcionales

$$\frac{\partial}{\partial \hat{x}_i} (\hat{x}_i \hat{p}_j - \hat{p}_j \hat{x}_i) = \frac{\partial}{\partial \hat{p}_i} (\hat{x}_i \hat{p}_j - \hat{p}_j \hat{x}_i) = 0,$$

i.e.,  $\hat{x}_i \hat{p}_j - \hat{p}_j \hat{x}_i$  es proporcional al operador identidad  $\hat{x}_i \hat{p}_j - \hat{p}_j \hat{x}_i = \alpha \hat{\mathcal{I}}$ . Si además  $\hat{x}_i$  y  $\hat{p}_j$  son hermíticos entonces, necesariamente,  $\alpha = i\hbar$  (ver la Proposición 4.2.10) donde  $\hbar \in \mathbb{R}$  que no es más que la relación de conmutación del postulado 4.3.6.

<sup>8</sup>Recuérdese que los estados están normalizados a la unidad.

### 4.4.1. Los proyectores ortogonales y la teoría de mediciones

Como hemos visto en el punto 2 del apartado anterior cuando en un sistema en cierto estado  $|\Psi\rangle$  hacemos una medición para saber el valor de cierto observable  $L$  asociado un operador hermítico  $\widehat{\mathcal{L}}$  el resultado es uno de los autovalores  $\lambda_k$  de dicho operador y el estado después de la medición pasa a ser un vector  $|\Psi_k\rangle$  del autoespacio (recordemos que todos los autovectores están normalizados a la unidad) asociado al autovalor  $\lambda_k$ . Además en el caso de que  $\lambda_k$  sea simple sabemos que la probabilidad de que ello ocurra es  $|f_k|^2 = |\langle\Psi_k|\Psi\rangle|^2$ . Todo lo podemos escribir usando los *proyectores ortogonales*.

Para ello comenzaremos asumiendo el caso más simple. Imaginemos que tenemos la magnitud  $L$  y que el resultado de la medición es en valor  $\lambda_k$  que asumiremos simple. Tras la medición el sistema estará en el estado  $|\Psi_k\rangle$ , donde  $|\Psi_k\rangle$  es el autovector asociado a  $\lambda_k$ .

Definamos el operador de proyección  $P_k$  sobre el subespacio generado por  $|\Psi_k\rangle$  de la siguiente forma

$$\widehat{\mathcal{P}}_k : \mathbb{H} \mapsto \mathbb{H}, \quad \widehat{\mathcal{P}}_k|\Psi\rangle = \langle\Psi_k|\Psi\rangle|\Psi_k\rangle,$$

el cual formalmente podemos escribir como  $\widehat{\mathcal{P}}_k = |\Psi_k\rangle\langle\Psi_k|$ .

De la definición anterior es evidente que  $\widehat{\mathcal{P}}_k^2 := \widehat{\mathcal{P}}_k \circ \widehat{\mathcal{P}}_k = \widehat{\mathcal{P}}_k$  y que  $\widehat{\mathcal{P}}_k$  es hermítico, i.e.,  $\widehat{\mathcal{P}}_k^+ = \widehat{\mathcal{P}}_k$ <sup>9</sup>. Además, claramente sus autovalores son o bien 0 o bien 1. En efecto

$$\widehat{\mathcal{P}}_k|\Psi\rangle = \lambda|\Psi\rangle \quad \Rightarrow \quad \widehat{\mathcal{P}}_k^2|\Psi\rangle = \lambda^2|\Psi\rangle \quad \Rightarrow \quad (\widehat{\mathcal{P}}_k^2 - \widehat{\mathcal{P}}_k)|\Psi\rangle = (\lambda^2 - \lambda)|\Psi\rangle = 0,$$

luego  $\lambda^2 - \lambda = 0$  de donde se sigue el resultado. Por otro lado  $\widehat{\mathcal{P}}_k|\Psi\rangle$  es un autovector asociado al autovalor 1 ( $\widehat{\mathcal{P}}_k(\widehat{\mathcal{P}}_k|\Psi\rangle) = 1(\widehat{\mathcal{P}}_k|\Psi\rangle)$ ) mientras que  $(\widehat{\mathcal{I}} - \widehat{\mathcal{P}}_k)|\Psi\rangle$  es el autovector asociado al autovalor 0. Luego cualquiera sea el vector  $|\Psi\rangle \in \mathbb{H}$ , los vectores  $\widehat{\mathcal{P}}_k|\Psi\rangle$  y  $(\widehat{\mathcal{I}} - \widehat{\mathcal{P}}_k)|\Psi\rangle$  son ortogonales (¿por qué?).

¿Qué ocurre si el autovector  $\lambda_k$  es degenerado, i.e., tiene asociado un subespacio de dimensión  $K > 1$ ?

En ese caso el projector  $\widehat{\mathcal{P}}_k$  es la suma de los proyectores asociados a cada uno de los vectores de la base ortonormal  $(\Psi_{k,j})_{j=1}^K$  del autoespacio asociado a  $\lambda_k$ . Es decir,

$$\widehat{\mathcal{P}}_k = \sum_{j=1}^K |\Psi_{k,j}\rangle\langle\Psi_{k,j}|$$

Es fácil comprobar que en este caso se tiene las mismas propiedades que en el caso cuando  $K = 1$ . También es fácil comprobar que en este caso la probabilidad de obtener el valor  $\lambda_k$  vuelve a ser  $|\langle\Psi_k|\Psi\rangle|^2$  donde ahora  $\Psi_k$  es el projector ortogonal sobre todo el subespacio asociado a  $\lambda_k$ .

Supongamos que tenemos dos autovalores distintos  $\lambda_k$  y  $\lambda_{k'}$ . Entonces los autoespacios correspondientes  $\mathbb{L}_k$  y  $\mathbb{L}_{k'}$  son ortogonales, luego

$$\widehat{\mathcal{P}}_k \circ \widehat{\mathcal{P}}_{k'} = 0, \quad \forall k \neq k'.$$

Si además el operador  $\widehat{\mathcal{L}}$  asociado a la magnitud  $L$  tiene un sistema de autovectores que conforman una base de  $\mathbb{H}$  entonces el operador

$$\widehat{\mathcal{P}}_L := \sum_k \widehat{\mathcal{P}}_k = \widehat{\mathcal{I}}.$$

---

<sup>9</sup> $\langle\Psi|\widehat{\mathcal{P}}_k|\Phi\rangle = \langle\Psi|\Psi_k\rangle\langle\Psi_k|\Phi\rangle = \overline{\langle\Psi_k|\Psi\rangle}\langle\Phi|\Psi_k\rangle = \langle\Phi|\widehat{\mathcal{P}}_k|\Psi\rangle$ .

La identidad anterior se denomina *descomposición de la identidad*. Además de dicha identidad se sigue que para todo  $|\Psi\rangle \in \mathbb{H}$  se tiene

$$|\Psi\rangle = \sum_k \hat{\mathcal{P}}_k |\Psi\rangle = \sum_k \langle \Psi_k | \Psi \rangle |\Psi_k\rangle.$$

Por otro lado,  $\hat{\mathcal{L}} \circ \hat{\mathcal{P}}_k = \lambda_k \hat{\mathcal{P}}_k$ , de donde se sigue que  $(\hat{\mathcal{L}} - \lambda_k \hat{\mathcal{L}}) \circ \hat{\mathcal{P}}_k = 0$ . Entonces, usando la descomposición de la identidad, tenemos que

$$\hat{\mathcal{L}} = \sum_k \hat{\mathcal{L}} \circ \hat{\mathcal{P}}_k = \sum_k \lambda_k \hat{\mathcal{P}}_k.$$

Es decir, todo operador hermítico cuyo conjunto de autovectores es completo está completamente determinado por sus autovalores y autovectores. Este es el conocido Teorema espectral de operadores hermíticos en espacios de Hilbert.

Así pues, si tenemos un sistema en cierto estado  $\Psi$  y sobre el medimos el valor de cierta magnitud  $L$  obtendremos como resultado un autovalor del operador  $\hat{\mathcal{L}}$  asociado a la magnitud  $L$  con probabilidad  $\text{Prob}(|\Phi\rangle \mapsto |\Psi_k\rangle) = \|\hat{\mathcal{P}}_k |\Psi\rangle\|^2$  siendo el estado final del sistema el definido por el vector  $\hat{\mathcal{P}}_k |\Psi\rangle$ , donde  $\hat{\mathcal{P}}_k$  es el projector ortogonal al subespacio asociado al autovalor  $\lambda_k$  obtenido.

## 4.5. Representación de los operadores $\hat{x}_i$ y $\hat{p}_i$

Escojamos como espacio de Hilbert de nuestro sistema el conjunto de las funciones de cuadrado integrable  $\mathbb{H} = L^2(\Omega)$ ,  $|\Psi\rangle = \Psi(x)$ . Definiremos el operador  $\hat{x}_i$  en  $L^2(\Omega)$  como el operador  $\hat{x}_i := x_i \hat{\mathcal{I}}$ . Luego  $\hat{x}^k \Psi(x) = x^k \Psi(x)$ .

¿Quién es  $\hat{p}_j$ ? Por simplicidad vamos a trabajar solamente en dimensión 1 (sólo la proyección en el eje de las  $x$ ). Nuestro objetivo es encontrar un operador  $\hat{p}$  tal que cumpla las relaciones de conmutación (4.3.2).

Del postulado 4.3.6 se sigue que

1.  $[\hat{p}, \hat{x}] = \hat{p}\hat{x} - \hat{x}\hat{p} = -i\hbar \hat{\mathcal{I}}$ ,
2.  $[\hat{p}, \hat{x}^2] = \hat{p}\hat{x}^2 - \hat{x}^2\hat{p} = -2i\hbar \hat{x}$ ,
- ⋮
3.  $[\hat{p}, \hat{x}^n] = \hat{p}\hat{x}^n - \hat{x}^n\hat{p} = -ni\hbar \hat{x}^{n-1} = -i\hbar \frac{\partial \hat{x}^n}{\partial \hat{x}}$ .

Luego, para cualquier función analítica  $F(z)$  tenemos

$$[\hat{p}, F(\hat{x})] = -i\hbar \frac{\partial F(\hat{x})}{\partial \hat{x}}, \quad (4.5.1)$$

que en nuestra representación se puede reescribir como

$$\hat{p}F(\hat{x})\Psi(x) - F(\hat{x})\hat{p}\Psi(x) = -i\hbar \frac{\partial F(\hat{x})}{\partial \hat{x}} \Psi(x),$$

o equivalentemente, usando que  $\widehat{x}^k \Psi(x) = x^k \Psi(x)$ ,

$$\widehat{p}[F(x)\Psi(x)] - F(x)\widehat{p}\Psi(x) = -i\hbar \frac{\partial F(x)}{\partial x} \Psi(x),$$

de donde escogiendo<sup>10</sup>  $\Psi(x) = 1$  obtenemos, sustituyendo  $\widehat{p}1 = \phi(x)$  la expresión

$$\widehat{p}F(x) = -i\hbar \frac{\partial F(x)}{\partial x} + F(x)\phi(x).$$

En general

$$\widehat{p}_k F(x_1, x_2, x_3) = -i\hbar \frac{\partial F(x_1, x_2, x_3)}{\partial x_k} + F(x_1, x_2, x_3)\phi_k(x_1, x_2, x_3).$$

Pero  $[\widehat{p}_k, \widehat{p}_j] = 0$ ,  $k = 1, 2, 3$ , luego

$$\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2} = \frac{\partial \phi_3}{\partial x_2} - \frac{\partial \phi_2}{\partial x_3} = \frac{\partial \phi_1}{\partial x_3} - \frac{\partial \phi_3}{\partial x_1} = 0.$$

Las ecuaciones anteriores son ciertas si  $\phi_k = \frac{\partial \Phi}{\partial x_k}$ ,  $k = 1, 2, 3$ , siendo  $\Phi$  una función lo suficientemente buena.

Así pues, los operadores  $\widehat{p}_k$  deben tener la forma

$$\widehat{p}_k = -i\hbar \frac{\partial}{\partial x_k} + \frac{\partial \Phi}{\partial x_k} \widehat{\mathcal{I}}.$$

Hagamos ahora la transformación unitaria

$$\widehat{p}_k \rightarrow \widehat{\mathcal{U}}^+ \widehat{p}_k \widehat{\mathcal{U}}, \quad \widehat{\mathcal{U}} = \exp(-i\Phi(x_1, x_2, x_3)/\hbar) \widehat{\mathcal{I}}, \quad \widehat{\mathcal{U}}^+ = \exp(i\Phi(x_1, x_2, x_3)/\hbar) \widehat{\mathcal{I}}$$

que, como sabemos, no cambia ni los elementos matriciales, ni las relaciones de conmutación, ni los autovalores, ni la hermiticidad de los operadores (i.e., éstos mantendrán el mismo significado físico de antes):

$$\widehat{p}_k = e^{i\Phi/\hbar} \left( -i\hbar \frac{\partial}{\partial x_k} + \frac{\partial \Phi}{\partial x_k} \widehat{\mathcal{I}} \right) e^{-i\Phi/\hbar} \widehat{\mathcal{I}} = -i\hbar \frac{\partial}{\partial x_k}.$$

Como ejercicio al lector dejamos que pruebe la identidad

$$[\widehat{x}, F(\widehat{p})] = i\hbar \frac{\partial F(\widehat{p})}{\partial \widehat{p}}. \quad (4.5.2)$$

## 4.6. Las ecuaciones de Heisenberg y de Schrödinger

En este apartado vamos a discutir las representaciones de Heisenberg y de Schrödinger para las ecuaciones dinámicas de la Mecánica cuántica. Supongamos que tenemos el sistema en cierto estado  $|\Phi\rangle$  y sea  $\widehat{\ell}$  el operador de cierta magnitud física que queremos estudiar. Si,

<sup>10</sup>Por el momento sólo nos interesa encontrar la expresión del operador independientemente de que luego éste vaya actuar en el espacio  $L^2(\Omega)$ .

en general,  $|\Phi\rangle$  es independiente del tiempo y  $\hat{\ell}$  no lo es diremos que estamos trabajando con la representación de Heisenberg. Si por el contrario,  $|\Phi\rangle$  depende de tiempo y  $\hat{\ell}$  no, entonces diremos que estamos considerando la representación Schrödinger de la Mecánica cuántica.

Por sencillez, en adelante asumiremos que el Hamiltoniano del sistema se expresa mediante la fórmula

$$\hat{\mathcal{H}} = \hat{T} + \hat{V}, \quad \hat{T} = \sum_{i=1}^3 \frac{\hat{p}_i^2}{2m},$$

y  $\hat{V} = V(\hat{x}_1, \hat{x}_2, \hat{x}_3) = V(x_1, x_2, x_3)\hat{\mathcal{I}}$  sólo depende de las coordenadas  $x, y, z$ .<sup>11</sup>

Por simplicidad trabajaremos sólo con la proyección en el eje  $OX$ .

Como  $[\hat{p}, \hat{T}] = 0$ , tenemos, usando (4.5.1) que

$$[\hat{p}, \hat{\mathcal{H}}] = [\hat{p}, V(\hat{x})] = -i\hbar \frac{\partial V(\hat{x})}{\partial \hat{x}} = -i\hbar \frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}} \quad (4.6.1)$$

Supongamos ahora que los vectores de estado no dependen del tiempo pero los operadores sí que pueden, en principio, depender del tiempo (es decir consideremos la representación de Heisenberg). Entonces del postulado 4.3.5 se tiene que

$$\frac{d\hat{p}}{dt} = -\frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}},$$

de donde se sigue que

$$\frac{d\hat{p}}{dt} = -\frac{i}{\hbar} [\hat{p}, \hat{\mathcal{H}}]. \quad (4.6.2)$$

De forma análoga, pero usando (4.5.2), se deduce la segunda ecuación de Heisenberg

$$\frac{d\hat{x}}{dt} = -\frac{i}{\hbar} [\hat{x}, \hat{\mathcal{H}}]. \quad (4.6.3)$$

Las ecuaciones anteriores se conocen como *ecuaciones dinámicas* de la Mecánica cuántica en la *representación de Heisenberg*: es decir, cuando las funciones de onda son vectores independientes del tiempo pero los operadores dependen del tiempo.

Obviamente hay otra posibilidad y es que los operadores no dependan del tiempo y las funciones de onda sí. En este caso usando el postulado 4.3.5 y la fórmula (4.6.1) ( $\frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}} = i/\hbar [\hat{p}, \hat{\mathcal{H}}]$ ), obtenemos

$$\frac{d}{dt} \langle \Phi | \hat{p} | \Psi \rangle = - \left\langle \Phi \left| \frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}} \right| \Psi \right\rangle = -\frac{i}{\hbar} \langle \Phi | [\hat{p}, \hat{\mathcal{H}}] | \Psi \rangle.$$

Luego, por un lado,

$$\left\langle \frac{\partial \Phi}{\partial t} \left| \hat{p} \right| \Psi \right\rangle + \left\langle \Phi \left| \hat{p} \left| \frac{\partial \Psi}{\partial t} \right. \right. \right\rangle = \left\langle \frac{\partial \Phi}{\partial t} \left| \hat{p} \Psi \right. \right\rangle + \left\langle \hat{p} \Phi \left| \frac{\partial \Psi}{\partial t} \right. \right\rangle$$

<sup>11</sup>Recordemos que estamos usando indistintamente la notación  $x_1, x_2, x_3$  y  $x, y, z$  para denotar las coordenadas espaciales.

y, por otro,

$$\frac{i}{\hbar} \langle \Phi | [\hat{p}, \hat{\mathcal{H}}] | \Psi \rangle = \frac{i}{\hbar} \left( \langle \Phi | \hat{p} \hat{\mathcal{H}} | \Psi \rangle - \langle \Phi | \hat{\mathcal{H}} \hat{p} | \Psi \rangle \right) = \left\langle \hat{p} \Phi \left| \frac{i}{\hbar} \hat{\mathcal{H}} \Psi \right. \right\rangle + \left\langle \frac{i}{\hbar} \hat{\mathcal{H}} \Phi \left| \hat{p} \Psi \right. \right\rangle,$$

de donde se sigue que

$$\left\langle \hat{p} \Phi \left| \frac{\partial \Psi}{\partial t} + \frac{i}{\hbar} \hat{\mathcal{H}} \Psi \right. \right\rangle + \left\langle \frac{\partial \Phi}{\partial t} + \frac{i}{\hbar} \hat{\mathcal{H}} \Phi \left| \hat{p} \Psi \right. \right\rangle = 0,$$

cualquiera sean los vectores  $\Phi$  y  $\Psi$ . Por tanto, necesariamente tenemos la ecuación

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = \hat{\mathcal{H}} |\Psi\rangle. \quad (4.6.4)$$

La ecuación anterior se denomina ecuación de Schrödinger y es la ecuación de evolución de la Mecánica cuántica.

Veamos ahora cómo se relacionan las representaciones de Heisenberg y de Schrödinger.

#### 4.6.1. Equivalencia de las representaciones de Heisenberg y de Schrödinger

Las ecuaciones dinámicas del postulado 4.3.5 han de cumplirse independientemente de que escojamos la representación de Schrödinger (S) o la de Heisenberg (H) discutidas en el apartado anterior. Además, los observables que medimos deben tener los mismos valores medios en ambas representaciones. Eso implica que ha de existir una transformación unitaria  $\{U\}$  que pase de S a H y viceversa.

Sean  $|\psi\rangle$  y  $\hat{\ell}$  la función de estado y el observable en la representación de Heisenberg y  $|\Psi\rangle$  y  $\hat{\mathcal{L}}$  en la de Schrödinger. Entonces entre ambas existe la relación:

$$|\Psi\rangle = \hat{U}^+ |\psi\rangle, \quad \hat{\mathcal{L}} = \hat{U}^+ \hat{\ell} \hat{U}, \quad \hat{U}^+ = e^{-i\hat{\mathcal{H}}t/\hbar},$$

donde  $\hat{\mathcal{H}}$  es el operador hamiltoniano del sistema que se asume independiente del tiempo.

En efecto, si  $|\psi\rangle$  no depende del tiempo, entonces

$$\frac{\partial |\Psi\rangle}{\partial t} = \frac{\partial \hat{U}^+}{\partial t} |\psi\rangle = -\frac{i}{\hbar} \hat{\mathcal{H}} e^{-i\hat{\mathcal{H}}t/\hbar} |\psi\rangle = -\frac{i}{\hbar} \hat{\mathcal{H}} |\Psi\rangle,$$

i.e., la ecuación de Schrödinger (4.6.4).

Supongamos que ahora  $\hat{\mathcal{L}}$  no depende de  $t$  (estamos en la representación de Schrödinger). Entonces,

$$\frac{\partial \hat{\ell}}{\partial t} = \frac{\partial \hat{U}}{\partial t} \hat{\mathcal{L}} \hat{U}^+ + \hat{U} \underbrace{\frac{\partial \hat{\mathcal{L}}}{\partial t}}_{=0} \hat{U}^+ + \hat{U} \hat{\mathcal{L}} \frac{\partial \hat{U}^+}{\partial t} = \frac{i}{\hbar} (\hat{\mathcal{H}} \hat{\ell} - \hat{\ell} \hat{\mathcal{H}}) = \frac{i}{\hbar} [\hat{\mathcal{H}}, \hat{\ell}]. \quad (4.6.5)$$

Si escogemos  $\hat{\ell}$  como el operador  $\hat{p}$  y  $\hat{x}$  recuperamos las ecuaciones de Heisenberg (4.6.2) y (4.6.3), respectivamente.

### 4.6.2. Integrales de movimiento

De lo anterior se sigue que en la representación de Heisenberg una cantidad física es independiente del tiempo si el operador asociado a dicha magnitud conmuta con el Hamiltoniano. Esta propiedad es además muy significativa desde el punto de vista físico como veremos a continuación.

**Definición 4.6.1** *Se dice que una cantidad observable  $\widehat{\mathcal{A}}$  es una integral de movimiento si*

$$\frac{d}{dt}\langle a \rangle = \frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = 0.$$

Es decir, una magnitud es una integral de movimiento si dicha magnitud se conserva en media.

Calculamos la derivada del elemento matricial

$$\frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = \left\langle \frac{\partial \Psi}{\partial t} \left| \widehat{\mathcal{A}} \right| \Psi \right\rangle + \left\langle \Psi \left| \frac{\partial \widehat{\mathcal{A}}}{\partial t} \right| \Psi \right\rangle + \left\langle \Psi \left| \widehat{\mathcal{A}} \right| \frac{\partial \Psi}{\partial t} \right\rangle.$$

Supongamos ahora que estamos en la representación de Schrödinger, i.e.,  $A$  no depende de  $t$  y  $|\Psi\rangle$  satisface la ecuación de Schrödinger (4.6.4). Entonces, usando (4.6.4) tenemos

$$\frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = \left\langle \Psi \left| \frac{i}{\hbar} [\widehat{\mathcal{H}}, \widehat{\mathcal{A}}] \right| \Psi \right\rangle. \quad (4.6.6)$$

Entonces,  $\frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = 0$  si y sólo si  $[\widehat{\mathcal{A}}, \widehat{\mathcal{H}}] = 0$ .

Si ahora escogemos la representación de Heisenberg entonces ( $|\Psi\rangle$  no depende del tiempo, pero  $\widehat{\mathcal{A}}$  si puede) tenemos

$$\frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = \left\langle \Psi \left| \frac{\partial \widehat{\mathcal{A}}}{\partial t} \right| \Psi \right\rangle = \frac{i}{\hbar}\langle \Psi | [\widehat{\mathcal{H}}, \widehat{\mathcal{A}}] | \Psi \rangle,$$

donde hemos usado la ecuación de Heisenberg (4.6.5). Es decir, también en la representación de Heisenberg  $\frac{d}{dt}\langle \Psi | \widehat{\mathcal{A}} | \Psi \rangle = 0$  si y sólo si  $[\widehat{\mathcal{A}}, \widehat{\mathcal{H}}] = 0$ .

### El test de consistencia

Probemos que también tenemos aquí el test de consistencia  $d/dt(\|\Psi\|^2) = 0$ . Sea la ecuación de Schrödinger (4.6.4) y sea su conjugada

$$-i\hbar \frac{\partial \langle \Psi |}{\partial t} = \langle \Psi | \widehat{\mathcal{H}}^+.$$

Tomando el producto escalar de esta última por  $|\Psi\rangle$  (por la derecha) y de (4.6.4) por  $\langle \Psi |$  (por la izquierda) obtenemos, respectivamente

$$\langle \Psi | \widehat{\mathcal{H}}^+ | \Psi \rangle = -i\hbar \left\langle \frac{\partial \Psi}{\partial t} \left| \Psi \right. \right\rangle, \quad \langle \Psi | \widehat{\mathcal{H}} | \Psi \rangle = i\hbar \left\langle \Psi \left| \frac{\partial \Psi}{\partial t} \right. \right\rangle,$$

de donde, restando ambas y usando la hermiticidad de  $\widehat{\mathcal{H}}$  tenemos

$$0 = \left\langle \frac{\partial \Psi}{\partial t} \left| \Psi \right. \right\rangle + \left\langle \Psi \left| \frac{\partial \Psi}{\partial t} \right. \right\rangle = \frac{\partial}{\partial t} \langle \Psi | \Psi \rangle = \frac{\partial \|\Psi\|^2}{\partial t}.$$

### 4.6.3. Los estados estacionarios del sistema

Toda la discusión anterior nos conduce a que el estado de un sistema viene dado por un vector de estado  $|\Psi\rangle$  que evoluciona según la ecuación de Schrödinger. Vamos a suponer que el operador hamiltoniano (que es hermítico)  $\widehat{\mathcal{H}}$  es independiente del tiempo y tiene un conjunto de autovectores  $|\Phi_n\rangle$  (independientes de  $t$ ) completo en  $\mathbb{H}$ , i.e.,

$$\widehat{\mathcal{H}}|\Phi_n\rangle = E_n|\Phi_n\rangle,$$

donde los autovalores  $E_n$  del hamiltoniano  $\widehat{\mathcal{H}}$  representan los posibles valores de la energía del sistema.

Supongamos ahora que tenemos un estado del sistema correspondiente a la energía  $E_n$ , que tiene la forma  $|\Psi_n\rangle = \alpha(t)|\Phi_n\rangle$ . Nótese que  $\widehat{\mathcal{H}}|\Psi_n\rangle = E_n|\Psi_n\rangle$ . Como los  $|\Psi_n\rangle$  son estados del sistema, entonces han de satisfacer la ecuación de Schrödinger, i.e.,

$$i\hbar\frac{\partial|\Psi_n\rangle}{\partial t} = \widehat{\mathcal{H}}|\Psi_n\rangle = E_n|\Psi_n\rangle.$$

Resolviendo con respecto al tiempo la ecuación anterior tenemos

$$|\Psi_n\rangle = e^{-iE_nt/\hbar}|\Phi_n\rangle,$$

donde  $|\Phi_n\rangle$  no depende explícitamente del tiempo.

Comúnmente a los estados  $|\Psi_n\rangle$  anteriores se les denominan estados estacionarios del sistema. Además, de lo anterior se deduce que la única dependencia del tiempo de los estados estacionarios es el factor  $e^{-iE_nt/\hbar}$ .

### 4.6.4. Los operadores unitarios y la evolución temporal

La discusión del apartado anterior nos da una pista de una transformación unitaria de especial interés. Concretamente la transformación que define el operador  $U(t) = e^{-i\widehat{\mathcal{H}}t/\hbar}$ , donde  $H$  es el Hamiltoniano del sistema.

Como los sistemas son invariantes frente a las transformaciones unitarias ello implica que si definimos el estado  $|\Psi(t)\rangle = U(t)|\Psi(0)\rangle$ , ambos han de describir el mismo estado. De lo anterior se deduce que para todo  $t_0$ ,  $|\Psi(t+t_0)\rangle = U(t)|\Psi(t_0)\rangle$ , es decir, que los estados físicos son invariantes frente a las traslaciones temporales. Como tomado derivadas respecto a  $t$  en  $|\Psi(t)\rangle = U(t)|\Psi(0)\rangle$  nos conduce a la ecuación de Schrödinger podemos deducir que esta es una consecuencia de la invarianza respecto a las traslaciones temporales de los sistemas físicos.

## 4.7. El principio de incertidumbre

Sean dos operadores hermíticos  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$ . Definamos los operadores

$$\Delta\widehat{\mathcal{A}} = \widehat{\mathcal{A}} - \langle A\rangle\widehat{\mathcal{I}}, \quad \Delta\widehat{\mathcal{B}} = \widehat{\mathcal{B}} - \langle B\rangle\widehat{\mathcal{I}},$$

donde  $\langle A\rangle$  y  $\langle B\rangle$  son los valores medios de  $\widehat{\mathcal{A}}$  y  $\widehat{\mathcal{B}}$  en el estado  $|\Psi\rangle$ . Entonces, como  $[\widehat{\mathcal{A}}, \widehat{\mathcal{B}}] = i\widehat{\mathcal{L}}$ , con  $\widehat{\mathcal{L}}$  hermítico, se sigue que  $[\Delta\widehat{\mathcal{A}}, \Delta\widehat{\mathcal{B}}] = i\widehat{\mathcal{L}}$ .

Las dispersiones de las cantidades  $A$  y  $B$  en el estado  $|\Psi\rangle$  vendrán dadas por

$$\Delta A := \sqrt{\langle \Psi | (\Delta \hat{\mathcal{A}})^2 | \Psi \rangle} = \|\Delta \hat{\mathcal{A}}\Psi\|, \quad \Delta B := \sqrt{\langle \Psi | (\Delta \hat{\mathcal{B}})^2 | \Psi \rangle} = \|\Delta \hat{\mathcal{B}}\Psi\|.$$

Si usamos la desigualdad de Cauchy-Schwarz (4.1.2)

$$\|\Delta \hat{\mathcal{A}}\Psi\| \|\Delta \hat{\mathcal{B}}\Psi\| \geq |\langle \Delta \hat{\mathcal{A}}\Psi | \Delta \hat{\mathcal{B}}\Psi \rangle| \geq |\Im \langle \Delta \hat{\mathcal{A}}\Psi | \Delta \hat{\mathcal{B}}\Psi \rangle|$$

Calculemos la parte imaginaria de

$$\langle \Delta \hat{\mathcal{A}}\Psi | \Delta \hat{\mathcal{B}}\Psi \rangle = \langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle$$

–recordemos que  $A$  es hermítico, luego  $\Delta \hat{\mathcal{A}}$  también lo es pues  $\langle A \rangle$  es real–. Obtenemos

$$\begin{aligned} \Im \langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle &= \frac{1}{2i} \left( \langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle - \overline{\langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle} \right) \\ &= \frac{1}{2i} \left( \langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle - \langle \Psi | \Delta \hat{\mathcal{B}}^+ \Delta \hat{\mathcal{A}}^+ | \Psi \rangle \right) \\ &= \frac{1}{2i} \left( \langle \Psi | \Delta \hat{\mathcal{A}} \Delta \hat{\mathcal{B}} | \Psi \rangle - \langle \Psi | \Delta \hat{\mathcal{B}} \Delta \hat{\mathcal{A}} | \Psi \rangle \right) \\ &= \frac{1}{2i} \langle \Psi | [\Delta \hat{\mathcal{A}}, \Delta \hat{\mathcal{B}}] | \Psi \rangle = \frac{1}{2} \langle \Psi | \hat{\mathcal{L}} | \Psi \rangle. \end{aligned}$$

Como  $\hat{\mathcal{L}}$  es hermítico,  $\langle \Psi | \hat{\mathcal{L}} | \Psi \rangle$  es un número real que denotaremos por  $l$ ; así,

$$\Delta A \Delta B \geq \frac{|l|}{2}.$$

Lo anterior aplicado a los operadores  $\hat{p}$  y  $\hat{x}$  (ver postulado 4.3.6) nos conduce al principio de incertidumbre de Heisenberg

$$\Delta x \Delta p \geq \frac{\hbar}{2}.$$

## 4.8. La mecánica matricial

Supongamos que tenemos una base  $(|\Psi_n\rangle)_n$  completa de vectores de  $\mathbb{H}$ . Entonces, todo vector  $|\Psi\rangle$  de  $\mathbb{H}$  lo podemos escribir, como ya hemos visto, de la forma

$$|\Psi\rangle = \sum_n f_n |\Psi_n\rangle, \quad f_n = \langle \Psi_n | \Psi \rangle.$$

Es decir, a cada vector de  $\mathbb{H}$  le podemos hacer corresponder su vector  $\mathbf{f} = (f_1, f_2, \dots)^T$ . Análogamente, a cada operador  $\hat{\mathcal{L}}$  le podemos hacer corresponder una matriz  $\mathbf{L}$  con entradas  $L_{m,n} = \langle \Psi_m | \hat{\mathcal{L}} | \Psi_n \rangle$ . Luego la ecuación (4.6.5) se puede escribir en la forma

$$\frac{\partial \mathbf{L}}{\partial t} = \frac{i}{\hbar} [\mathbf{H}, \mathbf{L}].$$

donde  $\mathbf{H}$  es la matriz correspondiente al hamiltoniano del sistema, i.e., recuperamos la también antes mencionada *mecánica matricial de Heisenberg*.

## 4.9. La ecuación de Schrödinger y el postulado 4.3.5

Supongamos que la cantidad observable  $L$  es independiente del tiempo y  $|\Psi\rangle$  es la solución de la ecuación de Schrödinger (4.6.4), donde  $\hat{\mathcal{H}}$  es el operador hamiltoniano

$$\hat{\mathcal{H}} = \frac{\hat{p}_1^2 + \hat{p}_2^2 + \hat{p}_3^2}{2m} + V(\hat{x}_1, \hat{x}_2, \hat{x}_3) = \hat{T} + \hat{V}.$$

Entonces, la ecuación (4.6.6) nos da

$$\frac{d}{dt}\langle L \rangle = \frac{i}{\hbar}\langle [\hat{\mathcal{H}}, \hat{\mathcal{L}}] \rangle. \quad (4.9.1)$$

Sea  $\hat{\mathcal{L}} = \hat{x}_k$ . Entonces, usando (4.5.2) tenemos

$$[\hat{\mathcal{H}}, \hat{x}_k] = [\hat{T}, \hat{x}_k] = -i\hbar \frac{\partial \hat{\mathcal{H}}}{\partial \hat{p}_k}.$$

Sustituyendo lo anterior en (4.9.1) obtenemos

$$\frac{d}{dt}\langle x_k \rangle = \left\langle \frac{\partial \hat{\mathcal{H}}}{\partial \hat{p}_k} \right\rangle.$$

Sea  $\hat{\mathcal{L}} = \hat{p}_k$ . Entonces, usando (4.5.1) tenemos

$$[\hat{\mathcal{H}}, \hat{p}_k] = [\hat{V}, x_k] = i\hbar \frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}_k},$$

de donde, usando (4.9.1), se sigue que

$$\frac{d}{dt}\langle p_k \rangle = - \left\langle \frac{\partial \hat{\mathcal{H}}}{\partial \hat{x}_k} \right\rangle.$$

Es decir, si la función de estado evoluciona según la ecuación de Schrödinger (4.6.4), entonces las medias de las coordenadas e impulsos se comportan como en la mecánica hamiltoniana clásica.

En las mismas condiciones de antes se puede probar (de forma totalmente análoga) que, partiendo de la cantidad

$$\frac{d}{dt}\langle \Phi | \hat{\mathcal{L}} | \Psi \rangle,$$

se obtienen las fórmulas de evolución del postulado 4.3.5.

## 4.10. Problemas

**Problema 4.10.1** Dado tres operadores  $\hat{p}$ ,  $\hat{q}$  y  $\hat{r}$ , prueba la identidad de Jacobi

$$[[\hat{p}, \hat{q}], \hat{r}] + [[\hat{q}, \hat{r}], \hat{p}] + [[\hat{r}, \hat{p}], \hat{q}] = 0.$$

**Problema 4.10.2** Probar que

$$e^{\widehat{L}}\widehat{a}e^{-\widehat{L}} = \widehat{a} + \frac{1}{1!}[\widehat{L}, \widehat{a}] + \frac{1}{2!}[\widehat{L}, [\widehat{L}, \widehat{a}]] + \dots$$

*Ayuda:* Encuentra la EDO que satisface el operador  $a(t)$  definido por  $\widehat{a}(t) = e^{t\widehat{L}}\widehat{a}e^{-t\widehat{L}}$ , donde  $\widehat{a}$  no depende del tiempo y desarrolla la función  $\widehat{a}(t)$  en potencias de  $t$ .

**Problema 4.10.3** Para el caso unidimensional encuentra los operadores hermíticos conjugados a los siguientes operadores:

$$\frac{d}{dx}, \quad x\frac{d}{dx}, \quad \widehat{p}_x\frac{d}{dx}, \quad x\frac{d}{dx}.$$

**Problema 4.10.4** Prueba que si  $\widehat{L}$  es hermítico, entonces  $e^{i\widehat{L}}$  es unitario.

**Problema 4.10.5** Prueba que el producto de dos operadores hermíticos  $\widehat{L}$  y  $\widehat{M}$  siempre se puede escribir como

$$\widehat{L}\widehat{M} = \widehat{A} + \widehat{B},$$

donde  $\widehat{A}$  es hermítico y  $\widehat{B}$  es antihermítico:  $\widehat{B}^+ = -\widehat{B}$ .



# Capítulo 5

## Resolviendo la ecuación de Schrödinger

### 5.1. El método de Nikiforov-Uvarov

#### 5.1.1. La ecuación hipergeométrica generalizada

La ecuación hipergeométrica generalizada es una ecuación lineal de segundo orden de la forma

$$u''(z) + \frac{\tilde{\tau}(z)}{\sigma(z)}u'(z) + \frac{\tilde{\sigma}(z)}{\sigma^2(z)}u(z) = 0, \quad (5.1.1)$$

siendo  $\tilde{\tau}(z)$  un polinomio de grado a lo más uno y  $\sigma(z)$  y  $\tilde{\sigma}(z)$  polinomios de grado a lo más dos.

Hagamos el cambio  $u(z) = \phi(z)y(z)$ ,

$$y''(z) + \left(2\frac{\phi'(z)}{\phi(z)} + \frac{\tilde{\tau}(z)}{\sigma(z)}\right)y'(z) + \left(\frac{\phi''(z)}{\phi(z)} + \frac{\phi'(z)\tilde{\tau}(z)}{\phi(z)\sigma(z)} + \frac{\tilde{\sigma}(z)}{\sigma^2(z)}\right)y(z) = 0.$$

El objetivo del cambio es convertir la ecuación anterior en una más sencilla –o por lo menos menos complicada– que (5.1.1), así que al menos debemos tener

$$2\frac{\phi'(z)}{\phi(z)} + \frac{\tilde{\tau}(z)}{\sigma(z)} = \frac{\tau(z)}{\sigma(z)}, \quad \text{o} \quad \frac{\phi'(z)}{\phi(z)} = \frac{\tau(z) - \tilde{\tau}(z)}{2\sigma(z)} = \frac{\pi(z)}{\sigma(z)}, \quad (5.1.2)$$

siendo  $\tau$  un polinomio de grado a lo más uno y, por tanto,  $\pi$  polinomio de grado a lo más uno. Lo anterior transforma nuestra ecuación original (5.1.1) en la siguiente

$$y''(z) + \frac{\tau(z)}{\sigma(z)}y'(z) + \frac{\bar{\sigma}(z)}{\sigma^2(z)}y(z) = 0, \quad (5.1.3)$$
$$\tau(z) = \tilde{\tau}(z) + 2\pi(z),$$

$$\bar{\sigma}(z) = \tilde{\sigma}(z) + \pi^2(z) + \pi(z)[\tilde{\tau}(z) - \sigma'(z)] + \pi'(z)\sigma(z).$$

Como  $\bar{\sigma}$  es un polinomio de grado dos a lo sumo, impongamos que sea proporcional al propio  $\sigma$ , es decir que  $\bar{\sigma}(z) = \lambda\sigma(z)$ . Ello es posible pues  $\bar{\sigma}$  tiene dos coeficientes indeterminados –los coeficientes del polinomio  $\pi$ – y  $\lambda$  es una constante a determinar, lo que nos conduce a

tres ecuaciones –al igualar los coeficientes de  $\bar{\sigma}$  y  $\sigma$ – con tres incógnitas. Hecho esto, nuestra ecuación se transforma en la ecuación hipergeométrica

$$\sigma(z)y'' + \tau(z)y' + \lambda y = 0. \quad (5.1.4)$$

Pasemos a calcular  $\pi$  y  $\lambda$ . Como  $\bar{\sigma} = \lambda\sigma(z)$ , entonces

$$\tilde{\sigma}(z) + \pi^2(z) + \pi(z)[\tilde{\tau}(z) - \sigma'(z)] + \pi'(z)\sigma(z) = \lambda\sigma(z),$$

o, equivalentemente,

$$\pi^2(z) + [\tilde{\tau}(z) - \sigma'(z)]\pi(z) + \{\tilde{\sigma}(z) - [\lambda - \pi'(z)]\sigma(z)\} = 0.$$

Supongamos que  $k = \lambda - \pi'(z)$  es conocido, entonces tenemos una ecuación de segundo orden para  $\pi(z)$ , luego

$$\pi(z) = \frac{\sigma'(z) - \tilde{\tau}(z)}{2} \pm \sqrt{\left(\frac{\sigma'(z) - \tilde{\tau}(z)}{2}\right)^2 - \tilde{\sigma}(z) + k\sigma(z)}, \quad (5.1.5)$$

pero  $\pi(z)$  ha de ser un polinomio de grado a lo sumo uno, por tanto el polinomio

$$\Upsilon(z) = \left(\frac{\sigma'(z) - \tilde{\tau}(z)}{2}\right)^2 - \tilde{\sigma}(z) + k\sigma(z) \quad (5.1.6)$$

ha de ser un cuadrado perfecto, es decir su discriminante debe ser cero, lo que nos conduce a una ecuación para encontrar  $k$ . El  $k$  encontrado lo sustituimos en (5.1.5) y obtenemos  $\pi(z)$ , el cual nos conduce directamente a  $\lambda = \pi'(z) + k$ .

Obviamente el método anterior da distintas soluciones en función del  $k$  que escojamos y del convenio de signos en (5.1.5).

### 5.1.2. La ecuación diferencial hipergeométrica

Como hemos visto en el apartado anterior el estudio de la ecuación (5.1.1) se puede reducir al de la ecuación hipergeométrica (5.1.4) por lo que nos centraremos en el estudio de ésta última. Aquí nos restringiremos a estudiar las soluciones polinómicas de (5.1.4). Para el caso general remitimos al lector [16].

#### La propiedad de hipergeometricidad y la fórmula de Rodrigues

Pasemos a continuación a estudiar la ecuación diferencial

$$\sigma(x)y'' + \tau(x)y' + \lambda y = 0, \quad (5.1.7)$$

donde  $\sigma$  y  $\tau$  son polinomios de grados a lo sumo 2 y 1, respectivamente.

La ecuación (5.1.7) usualmente se denomina *ecuación diferencial hipergeométrica*. La razón fundamental de esta denominación está en la denominada *propiedad de hipergeometricidad* que consiste en que las soluciones  $y$  de la ecuación (5.1.7) son tales que sus  $m$ -ésimas derivadas

$y^{(m)} := y_m$  satisfacen una ecuación del mismo tipo. En efecto, si derivamos (5.1.7)  $m$  veces obtenemos que  $y_m$  satisface una ecuación de la forma

$$\begin{aligned}\sigma(x)y_m'' + \tau_m(x)y_m' + \mu_m y_m &= 0, \\ \tau_m(x) &= \tau(x) + m\sigma'(x),\end{aligned}\tag{5.1.8}$$

$$\mu_m = \lambda + \sum_{i=0}^{m-1} \tau_i'(x) = \lambda + m\tau'(x) + m(m-1)\frac{\sigma''(x)}{2}.$$

Es evidente que grado  $\tau_m \leq 1$  y que  $\mu_m$  es una constante. Además, toda solución de (5.1.8) es necesariamente de la forma  $y_m = y^{(m)}$  siendo  $y$  solución de (5.1.7). La demostración es por inducción y la omitiremos.

Vamos a intentar encontrar las soluciones polinómicas de (5.1.7). Para encontrarlas comenzaremos escribiendo (5.1.7) y (5.1.8) en su forma *simétrica o autoconjugada*

$$\begin{aligned}[\sigma(x)\rho(x)y']' + \lambda\rho(x)y &= 0, \\ [\sigma(x)\rho_m(x)y_m']' + \mu_m\rho_m(x)y_m &= 0,\end{aligned}\tag{5.1.9}$$

donde  $\rho$  y  $\rho_m$  son funciones de simetrización que satisfacen las ecuaciones diferenciales de primer orden (conocidas como ecuaciones de Pearson)

$$\begin{aligned}[\sigma(x)\rho(x)]' &= \tau(x)\rho(x), \\ [\sigma(x)\rho_m(x)]' &= \tau_m(x)\rho_m(x).\end{aligned}\tag{5.1.10}$$

Si  $\rho$  es conocida entonces, utilizando las ecuaciones anteriores, obtenemos para  $\rho_m$  la expresión

$$\rho_m(x) = \sigma^m(x)\rho(x).\tag{5.1.11}$$

**Teorema 5.1.1** *Las soluciones polinómicas de la ecuación (5.1.8) se expresan mediante la fórmula de Rodrigues*

$$P_n^{(m)}(x) = \frac{A_{nm}B_n}{\rho_m(x)} \frac{d^{n-m}}{dx^{n-m}}[\rho_n(x)],\tag{5.1.12}$$

donde  $B_n = P_n^{(n)}/A_{nn}$  y

$$A_{nm} = A_m(\lambda) \Big|_{\lambda=\lambda_n} = \frac{n!}{(n-m)!} \prod_{k=0}^{m-1} [\tau' + \frac{1}{2}(n+k-1)\sigma''].\tag{5.1.13}$$

Además, el autovalor  $\mu_m$  de (5.1.8) es

$$\mu_m = \mu_m(\lambda_n) = -(n-m)[\tau' + \frac{1}{2}(n+m-1)\sigma''].\tag{5.1.14}$$

Demostración: Para demostrar el teorema vamos a escribir la ecuación autoconjugada para las derivadas de la siguiente forma

$$\rho_m(x)y_m = -\frac{1}{\mu_m}[\rho_{m+1}(x)y_{m+1}]',$$

luego

$$\rho_m(x)y_m = \frac{A_m}{A_n} \frac{d^{n-m}}{dx^{n-m}} [\rho_n(x)y_n], \quad A_m = (-1)^m \prod_{k=0}^{m-1} \mu_k, \quad A_0 = 1.$$

Como estamos buscando soluciones polinómicas,  $y := P_n$ , tenemos que  $P_n^{(n)}$  es una constante; por tanto, para las derivadas de orden  $m$ ,  $P_n^{(m)}$ , obtenemos la expresión

$$P_n^{(m)}(x) = \frac{A_{nm}B_n}{\rho_m(x)} \frac{d^{n-m}}{dx^{n-m}} [\rho_n(x)],$$

donde  $A_{nm} = A_m(\lambda) \Big|_{\lambda=\lambda_n}$  y  $B_n = P_n^{(n)}/A_{nn}$ . Como  $P_n^{(n)}$  es una constante, de (5.1.8) obtenemos que  $\mu_n = 0$ , luego, usando la expresión (5.1.8)

$$\mu_n = \lambda_n + n\tau'(x) + n(n-1)\sigma''(x)/2 = 0,$$

deducimos que el valor de  $\lambda_n$  en (5.1.7) se expresa mediante la fórmula<sup>1</sup>

$$\lambda := \lambda_n = -n\tau' - \frac{n(n-1)}{2}\sigma''. \quad (5.1.15)$$

Sustituyendo (5.1.15) en (5.1.8) obtenemos el valor de  $\mu_{nm} = \mu_m(\lambda_n)$

$$\mu_{nm} = \mu_m(\lambda_n) = -(n-m)[\tau' + \frac{1}{2}(n+m-1)\sigma''], \quad (5.1.16)$$

de donde, usando que  $A_{nm} = A_m(\lambda_n) = (-1)^m \prod_{k=0}^{m-1} \mu_{nk}$ , deducimos el valor de la constante  $A_{nm}$ .  $\square$

En la prueba hemos asumido que  $\mu_{nk} \neq 0$  para  $k = 0, 1, \dots, n-1$ . De la expresión explícita (5.1.16) deducimos que para que ello ocurra es suficiente que  $\tau' + n\sigma''/2 \neq 0$  para todo  $n = 0, 1, 2, \dots$ . Nótese que esta condición es equivalente a  $\lambda_n \neq 0$  para todo  $n \in \mathbb{N}$ . Además, de ella se deduce que  $\tau'_n \neq 0$  para todo  $n \in \mathbb{N}$ . Esta condición se conoce como condición de regularidad o de admisibilidad. Nótese además que  $\mu_{nk} = \lambda_n - \lambda_k$ , luego  $\mu_{nk} \neq 0$  para  $k = 0, 1, \dots, n-1$  implica que  $\lambda_n \neq \lambda_k$  si  $n \neq k$ .

Cuando  $m = 0$  la fórmula (5.1.12) se convierte en la conocida fórmula de Rodrigues para los polinomios clásicos

$$P_n(x) = \frac{B_n}{\rho(x)} \frac{d^n}{dx^n} [\sigma^n(x)\rho(x)], \quad n = 0, 1, 2, \dots \quad (5.1.17)$$

La fórmula (5.1.15) determina los autovalores  $\lambda_n$  de (5.1.7) y es conocida como condición de *hipergeometricidad*.

## Ortogonalidad y relación de recurrencia

Veamos ahora cómo a partir de las ecuaciones diferenciales simetrizadas (5.1.9) podemos demostrar la ortogonalidad de las soluciones polinómicas respecto a la función peso  $\rho$ .

<sup>1</sup>Usando la expresión (5.1.15) podemos obtener una expresión alternativa  $A_{nm} = (-n)_m \prod_{k=0}^{m-1} \frac{\lambda_{n+k}}{(n+k)}$ .

**Teorema 5.1.2** *Supongamos que*

$$x^k \sigma(x) \rho(x) \Big|_a^b = 0, \quad \text{para todo } k \geq 0. \quad (5.1.18)$$

Entonces las soluciones polinómicas  $P_n$  de la ecuación (5.1.7) constituyen una sucesión de polinomios ortogonales (SPO) respecto a la función peso  $\rho$  definida por la ecuación  $[\sigma(x)\rho(x)]' = \tau(x)\rho(x)$ , o sea, se cumple que

$$\int_a^b P_n(x) P_m(x) \rho(x) dx = \delta_{n,m} d_n^2, \quad (5.1.19)$$

donde  $\delta_{n,m}$  es el símbolo de Kronecker y  $d_n$  denota la norma de los polinomios  $P_n$ .

Demostración: Sean  $P_n$  y  $P_m$  dos de las soluciones polinómicas de (5.1.7). Partiremos de las ecuaciones simetrizadas para  $P_n$  y  $P_m$ ,

$$\begin{aligned} [\sigma(x)\rho(x)P_n'(x)]' + \lambda_n \rho(x) P_n(x) &= 0, \\ [\sigma(x)\rho(x)P_m'(x)]' + \lambda_m \rho(x) P_m(x) &= 0. \end{aligned}$$

Multiplicando la primera por  $P_m$  y la segunda por  $P_n$ , restando ambas e integrando en  $[a, b]$  obtenemos

$$\begin{aligned} (\lambda_n - \lambda_m) \int_a^b P_n(x) P_m(x) \rho(x) dx &= \\ &= \int_a^b \left( [\sigma(x)\rho(x)P_m'(x)]' P_n(x) - [\sigma(x)\rho(x)P_n'(x)]' P_m(x) \right) dx \\ &= \sigma(x)\rho(x) [P_n(x)P_m'(x) - P_n'(x)P_m(x)] \Big|_a^b \\ &= \sigma(x)\rho(x) W[P_n(x), P_m(x)] \Big|_a^b. \end{aligned}$$

Pero el Wronskiano  $W(P_n, P_m)$  es un polinomio en  $x$ ; por tanto, si imponemos la condición (5.1.18) obtendremos ( $\lambda_n \neq \lambda_m$ ) que los polinomios  $P_n$  y  $P_m$  son ortogonales respecto a la función peso  $\rho$ . Usualmente los valores de  $a$  y  $b$  se escogen de forma que  $\rho$  sea positiva en el intervalo  $[a, b]$ . Una elección puede ser tomar  $a$  y  $b$  como las raíces de  $\sigma(x) = 0$ , si éstas existen.  $\square$

De forma análoga, utilizando la ecuación (5.1.9) para las derivadas  $y_k := P_n^{(k)}$ , se puede demostrar que las  $k$ -ésimas derivadas de los polinomios hipergeométricos también son ortogonales, es decir, que

$$\int_a^b P_n^{(k)}(x) P_m^{(k)}(x) \rho_k(x) dx = \delta_{n,m} d_{kn}^2. \quad (5.1.20)$$

Finalmente, para calcular la norma  $d_n$  de los polinomios podemos utilizar la fórmula de Rodrigues. En efecto, sustituyendo (5.1.17) en (5.1.19) tenemos

$$d_n^2 = B_n \int_a^b P_n(x) \frac{d^n}{dx^n} [\sigma^n(x) \rho(x)] dx,$$

de donde integrando por partes y usando que  $P_n^{(n)} = n! a_n$  concluimos que

$$d_n^2 = B_n (-1)^n n! a_n \int_a^b \sigma^n(x) \rho(x) dx. \quad (5.1.21)$$

**Teorema 5.1.3** *Los polinomios ortogonales satisfacen una relación de recurrencia a tres términos de la forma*

$$xP_n(x) = \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad (5.1.22)$$

donde

$$\alpha_n = \frac{a_n}{a_{n+1}}, \quad \beta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}, \quad \gamma_n = \frac{c_n - \alpha_n c_{n+1}}{a_{n-1}} - \frac{b_n}{a_{n-1}} \beta_n = \frac{a_{n-1}}{a_n} \frac{d_n^2}{d_{n-1}^2}, \quad (5.1.23)$$

donde  $a_n$ ,  $b_n$  y  $c_n$  son los coeficientes del desarrollo  $P_n(x) = a_n x^n + b_n x^{n-1} + c_n x^{n-2} + \dots$ , y  $d_n$  es la norma de los polinomios.

Generalmente se impone que  $P_{-1}(x) = 0$  y  $P_0(x) = 1$ , con lo que la sucesión de polinomios ortogonales queda determinada de forma única conocidas las sucesiones  $(\alpha_n)_n$ ,  $(\beta_n)_n$  y  $(\gamma_n)_n$ .

Demostración: Utilizando que la sucesión  $(P_n)_n$  es una base del espacio de los polinomios, tenemos que el polinomio  $xP_n(x)$  de grado  $n+1$  se puede desarrollar en la base  $(P_n)_n$

$$xP_n(x) = \sum_{k=0}^{n+1} c_{nk} P_k(x), \quad c_{nk} = \frac{\int_a^b P_n(x) [xP_k(x)] \rho(x) dx}{d_n^2}.$$

Pero como el grado de  $xP_k(x)$  es  $k+1$ , entonces  $c_{nk} = 0$  para todo  $0 \leq k < n-1$ , de donde se concluye que la SPO satisface una relación (5.1.22). Además, los coeficientes  $\alpha_n$ ,  $\beta_n$ , y  $\gamma_n$  se expresan mediante las fórmulas:

$$\begin{aligned} \alpha_n &= \frac{1}{d_{n+1}^2} \int_a^b xP_n(x)P_{n+1}(x)\rho(x)dx, & \beta_n &= \frac{1}{d_n^2} \int_a^b xP_n(x)P_n(x)\rho(x)dx, \\ \gamma_n &= \frac{1}{d_{n-1}^2} \int_a^b xP_n(x)P_{n-1}(x)\rho(x)dx. \end{aligned} \quad (5.1.24)$$

Para probar (5.1.23) basta sustituir la expresión  $P_n(x) = a_n x^n + b_n x^{n-1} + c_n x^{n-2} + \dots$  en (5.1.22) e igualar las potencias  $x^{n+1}$ ,  $x^n$  y  $x^{n-1}$ . Finalmente como  $xP_{n-1} = \alpha_{n-1}P_n + \beta_{n-1}P_{n-1} + \gamma_{n-1}P_{n-2}$ ,

$$\gamma_n = \frac{1}{d_{n-1}^2} \int_a^b P_n(x)[xP_{n-1}(x)]\rho(x)dx = \frac{1}{d_{n-1}^2} \alpha_{n-1} \int_a^b P_n^2(x)\rho(x)dx,$$

de donde se sigue el resultado. ■

Nótese que del resultado anterior se deduce que  $\alpha_n \neq 0$  para todo  $n \in \mathbb{N} \cup \{0\}$  así como que  $\gamma_n \neq 0$  para todo  $n \in \mathbb{N}$ . Además, si  $\rho(x) \geq 0$  en  $(a, b)$ , entonces  $\alpha_{n-1}\gamma_n > 0$  para todo  $n \in \mathbb{N}$ . Resulta que el recíproco también es cierto.

**Teorema 5.1.4** *Sea  $(\beta_n)_{n=0}^\infty$  y  $(\gamma_n)_{n=0}^\infty$  dos sucesiones cualesquiera de números reales con  $\gamma_n > 0$  para todo  $n \in \mathbb{N}$  y sea  $(P_n)_{n=0}^\infty$  una sucesión de polinomios mónicos definidos mediante la relación*

$$P_n(x) = (x - \beta_n)P_n(x) - \gamma_n P_{n-1}(x), \quad n \in \mathbb{N}, \quad (5.1.25)$$

donde  $P_{-1} = 0$  y  $P_0(x) = 1$ . Entonces, dichos polinomios  $P_n$  son ortonormales para cierta medida positiva sobre la recta real.

El teorema anterior se conoce como Teorema de Favard, aunque había sido demostrado antes por O. Perron (1929), A. Wintner (1929) y M. H. Stone (1932), J. Sherman (1935) y I. P. Natanson (1935) indistintamente.

Una consecuencia inmediata de la RRTT es el siguiente teorema cuya prueba dejamos como ejercicio al lector

**Teorema 5.1.5** *Si  $(P_n)_n$  es una sucesión de polinomios ortogonales que satisface la relación de recurrencia a tres términos (5.1.22). Entonces se cumple que*

$$\text{Ker}_n(x, y) \equiv \sum_{m=0}^n \frac{P_m(x)P_m(y)}{d_m^2} = \frac{\alpha_n}{d_n^2} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{x - y}, \quad n \geq 1. \quad (5.1.26)$$

Si hacemos tender  $y \rightarrow x$  en la fórmula anterior obtenemos la fórmula *confluyente* de Christoffel-Darboux:

$$\text{Ker}_n(x, x) \equiv \sum_{m=0}^n \frac{P_m^2(x)}{d_m^2} = \frac{\alpha_n}{d_n^2} [P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)] \quad n \geq 1. \quad (5.1.27)$$

Una propiedad muy importante de los polinomios ortogonales está relacionada con los ceros de los mismos. Así, se tiene el siguiente teorema fundamental cuya demostración omitiremos (ver e.g. [1]):

**Teorema 5.1.6** *Supongamos que  $\rho(x)$  es positiva en el interior del intervalo  $(a, b)$ . Entonces:*

1. *Todos los ceros de  $P_n$  son reales, simples y están localizados en  $(a, b)$ .*
2. *Dos polinomios consecutivos  $P_n$  y  $P_{n+1}$  no pueden tener ningún cero en común.*
3. *Denotemos por  $x_{n,j}$  a los ceros del polinomio  $P_n$ , (consideraremos en adelante que  $x_{n,1} < x_{n,2} < \dots < x_{n,n}$ ). Entonces:*

$$x_{n+1,j} < x_{n,j} < x_{n+1,j+1},$$

*es decir, los ceros de  $P_n$  y  $P_{n+1}$  entrelazan unos con otros.*

Calculemos ahora la relación de recurrencia (5.1.22) a tres que satisfacen los polinomios clásicos. Para calcular los coeficientes usando las expresiones de (5.1.23) tenemos antes que encontrar una expresión general para los coeficientes principales  $a_n$  y  $b_n$  del polinomio  $P_n$ .

Para calcular  $a_n$  usamos que, por un lado  $P_n^{(n)}(x) = n!a_n$  y por el otro, utilizando la fórmula de Rodrigues (5.1.12)  $P_n^{(n)}(x) = B_n A_{nn}$ , por tanto,

$$a_n = \frac{B_n A_{nn}}{n!} = B_n \prod_{k=0}^{n-1} [\tau' + \frac{1}{2}(n+k-1)\sigma'']. \quad (5.1.28)$$

Para calcular  $b_n$  utilizaremos la fórmula de Rodrigues para la  $n-1$ -ésima derivada de  $P_n$ :  $P_n^{(n-1)}(x) = A_{nn-1} B_n \tau_{n-1}(x)$ , de donde obtenemos la igualdad

$$P_n^{(n-1)}(x) = n!a_n x + (n-1)!b_n = A_{nn-1} B_n \tau_{n-1}(x).$$

Luego,

$$b_n = \frac{n\tau_{n-1}(0)}{\tau'_{n-1}} a_n. \quad (5.1.29)$$

Obsérvese que, al ser  $\tau'_n \neq 0$ ,  $b_n$  está definido para cualquier  $n$ .

Usando las expresiones (5.1.23) así como (5.1.29) deducimos

$$\alpha_n = \frac{a_n}{a_{n+1}} = \frac{B_n}{B_{n+1}} \frac{\tau' + (n-1)\frac{\sigma''}{2}}{(\tau' + (2n-1)\frac{\sigma''}{2})(\tau' + (2n)\frac{\sigma''}{2})} = \frac{B_n}{B_{n+1}} \frac{\lambda_n}{n} \frac{2n}{\lambda_{2n}} \frac{2n+1}{\lambda_{2n+1}}$$

$$\beta_n = \frac{n\tau_{n-1}(0)}{\tau'_{n-1}} - \frac{(n+1)\tau_n(0)}{\tau'_n}.$$

Vamos a dar una expresión alternativa para  $\gamma_n$  sin usar la norma de los polinomios. Para ello igualamos los coeficientes de  $x^{n-2}$  en la ecuación diferencial (5.1.7). Ello nos conduce a la expresión

$$c_n = -\frac{(n-1)[\tau(0) + (n-2)\sigma'(0)]b_n + n(n-1)\sigma(0)a_n}{(n-2)[\tau' + (n-3)\frac{\sigma''}{2}] + \lambda_n}$$

$$= -\frac{n(n-1)[\tau_{n-2}(0)\tau_{n-1}(0) + \sigma(0)\tau'_{n-1}]}{\tau'_{n-1}(\lambda_n - \lambda_{n-2})} a_n. \quad (5.1.30)$$

Luego nos resta sustituir la expresión anterior en la fórmula<sup>2</sup> (5.1.24)

$$\gamma_n = \frac{c_n - \alpha_n c_{n+1}}{a_{n-1}} - \frac{b_n}{a_{n-1}} \beta_n.$$

### Consecuencias de la fórmula de Rodrigues

La primera consecuencia inmediata de la fórmula de Rodrigues es que  $\tau(x)$  debe ser necesariamente un polinomio de grado exactamente uno. En efecto, si calculamos el polinomio de grado 1 utilizando la fórmula de Rodrigues (5.1.17) obtenemos

$$P_1(x) = \frac{B_1}{\rho(x)} [\sigma(x)\rho(x)]' = B_1\tau(x), \quad (5.1.31)$$

y por tanto  $\tau$  es un polinomio de grado exactamente uno. Nótese que esta fórmula es equivalente a la fórmula de Pearson (5.1.10).

Si escribimos la fórmula de Rodrigues (5.1.12) para las derivadas<sup>3</sup>  $P_{n+m}^{(m)}$  con  $n = 1$  tenemos

$$P_{1+m}^{(m)}(x) = \frac{A_{m+1m}}{\rho_m(x)} [\rho_{m+1}(x)]' = \frac{A_{m+1m}}{\rho_m(x)} [\sigma(x)\rho_m(x)]' = A_{m+1m}\tau_m(x),$$

es decir,  $\tau_m$  es de grado exactamente uno (pues los polinomios  $P_{n+m}^{(m)}$  son ortogonales). Por tanto  $\tau'_m \neq 0$  para todos  $m \in \mathbb{N}$  lo cual es la condición de regularidad (existencia de la SPO) que ya mencionamos.

<sup>2</sup>Recordar que  $P_n(x) = a_n x^n + b_n x^{n-1} + c_n x^{n-2} + \dots$ .

<sup>3</sup>Hemos usado  $P_{n+m}^{(m)}$  en vez de  $P_n^{(m)}$  pues estos son polinomios de grado exactamente  $n$  en  $x$  mientras que los últimos no. Obviamente ellos también son solución de la ecuación (5.1.8) y satisfacen la fórmula de Rodrigues (5.1.12) cambiando  $n$  por  $n+m$ .

Tomemos ahora  $m = 1$  en la fórmula (5.1.12). Realizando unos cálculos directos deducimos que

$$P'_n(x) = \frac{A_{n1}B_n}{\rho_1(x)} \frac{d^{n-1}}{dx^{n-1}}[\rho_n(x)] = \frac{-\lambda_n B_n}{\rho_1(x)} \frac{d^{n-1}}{dx^{n-1}}[\rho_{1_{n-1}}(x)].$$

Luego

$$P'_n(x) = \frac{-\lambda_n B_n}{\bar{B}_{n-1}} \bar{P}_{n-1}(x), \quad (5.1.32)$$

donde  $\bar{P}_{n-1}$  denota al polinomio ortogonal respecto a la función peso  $\rho_1(x) = \sigma(x)\rho(x)$ . O sea, si  $P_n$  es ortogonal  $P'_n(x)$  también lo será.

Si escribimos la fórmula de Rodrigues (5.1.12) para el polinomio de grado  $n + 1$ , utilizando la ecuación de Pearson  $[\sigma(x)\rho_n(x)]' = \tau_n(x)\rho_n(x)$  vemos que

$$\begin{aligned} P_{n+1}(x) &= \frac{B_{n+1}}{\rho(x)} \frac{d^{n+1}}{dx^{n+1}}[\sigma^{n+1}(x)\rho(x)] = \frac{B_{n+1}}{\rho(x)} \frac{d^n}{dx^n}[\tau_n(x)\rho_n(x)] \\ &= \frac{B_{n+1}}{\rho(x)} \left[ \tau_n(x) \frac{d^n \rho_n(x)}{dx^n} + n\tau'_n \frac{d^{n-1} \rho_n(x)}{dx^{n-1}} \right]. \end{aligned}$$

Utilizando ahora que  $P'_n(x) = \frac{-\lambda_n B_n}{\sigma(x)\rho(x)} \frac{d^{n-1} \rho_n(x)}{dx^{n-1}}$ , obtenemos la fórmula de diferenciación, comúnmente denominada caracterización de Al-Salam & Chihara,

$$\sigma(x)P'_n(x) = \frac{\lambda_n}{n\tau'_n} \left[ \tau_n(x)P_n(x) - \frac{B_n}{B_{n+1}}P_{n+1}(x) \right], \quad (5.1.33)$$

o, equivalentemente,

$$\left[ \frac{\lambda_n}{n\tau'_n} \mathbf{I} - \sigma(x)\mathcal{D} \right] P_n(x) = \frac{\lambda_{2n}}{2n} \alpha_n P_{n+1}(x). \quad (5.1.34)$$

Sustituyendo (5.1.22) en (5.1.34) obtenemos

$$\left[ \left( -\frac{\lambda_n}{n\tau'_n} + \frac{\lambda_{2n}}{2n}(x - \beta_n) \right) \mathbf{I} + \sigma(x)\mathcal{D} \right] P_n(x) = \frac{\lambda_{2n}}{2n} \gamma_n P_{n-1}(x). \quad (5.1.35)$$

Si ahora en la fórmula (5.1.33) desarrollamos  $\tau_n$  y utilizamos la relación de recurrencia (5.1.22) para descomponer los sumandos de la forma  $xP_n$  obtenemos el siguiente teorema

**Teorema 5.1.7** *Los polinomios ortogonales  $P_n(x)$ , soluciones de la ecuación (5.1.7), satisfacen la siguiente relación de estructura*

$$\sigma(x)P'_n(x) = \tilde{\alpha}_n P_{n+1}(x) + \tilde{\beta}_n P_n(x) + \tilde{\gamma}_n P_{n-1}(x), \quad n \geq 0, \quad (5.1.36)$$

donde

$$\tilde{\alpha}_n = \frac{\lambda_n}{n\tau'_n} \left[ \alpha_n \tau'_n - \frac{B_n}{B_{n+1}} \right], \quad \tilde{\beta}_n = \frac{\lambda_n}{n\tau'_n} [\beta_n \tau'_n + \tau_n(0)], \quad \tilde{\gamma}_n = \frac{\lambda_n \gamma_n}{n} \neq 0. \quad (5.1.37)$$

Las expresiones (5.1.37) anteriores para los coeficientes de la relación de estructura pueden reescribirse usando las fórmulas explícitas para los coeficientes de la relación de recurrencia de la siguiente forma

$$\tilde{\alpha}_n = n \frac{\sigma''}{2} \alpha_n, \quad \tilde{\beta}_n = \frac{\lambda_n}{\tau'_n \tau'_{n-1}} [\tau(0)\sigma'' - \tau' \sigma'(0)] = \frac{\lambda_n}{n \tau'_n} \tau_n(\beta_n), \quad \tilde{\gamma}_n = \frac{\lambda_n \gamma_n}{n}. \quad (5.1.38)$$

Finalmente enunciaremos un teorema de gran importancia en lo que sigue.

**Teorema 5.1.8** *Los polinomios de tipo hipergeométricos (clásicos)  $p_n(x)$  son las únicas soluciones de la ecuación hipergeométrica  $\sigma(z)y'' + \tau(z)y' + \lambda y = 0$ , tales que las funciones  $\psi_n(x) = \sqrt{\rho(x)}p_n(x)$ , donde  $\rho(x)$  es la función peso con respecto a la cual los  $p_n$  son ortogonales, son acotadas y de cuadrado integrable en  $(a, b)$ , siendo  $(a, b)$  el soporte de la función peso.*

**Nota 5.1.9** *Si queremos que la función peso sea positiva e integrable en el interior del intervalo de ortogonalidad y suponemos que  $\sigma(x) > 0$  en dicho intervalo se puede comprobar que el polinomio  $\tau$  ha de cumplir dos propiedades importantes:*

1. *En primer lugar la derivada de  $\tau$  ha de ser negativa. Esto es particularmente importante en los casos cuando  $\sigma = 1$  y  $\sigma = x$ , que corresponden a intervalos de ortogonalidad no acotados —véase el próximo apartado—,*
2.  *$\tau$  ha de anularse en el interior del intervalo de ortogonalidad. Ello es consecuencia de (5.1.31) y del teorema 5.1.6 que asegura que  $P_1$  ha de anularse en el interior del intervalo de ortogonalidad.*

### 5.1.3. Los polinomios de Hermite, Laguerre y Jacobi

#### Parámetros principales

Comenzaremos escribiendo los principales parámetros de las sucesiones de polinomios ortogonales mónicos clásicos (SPOMC). Como ya hemos visto los polinomios ortogonales en la recta real, solución de una ecuación del tipo (5.1.7), se pueden clasificar en tres grandes familias en función del grado del polinomio  $\sigma$  ( $\tau$  siempre es un polinomio de grado 1). Cuando  $\sigma$  es un polinomio de grado cero los polinomios correspondientes se denominan *polinomios de Hermite*  $H_n(x)$ , cuando  $\sigma$  es de grado 1, *polinomios de Laguerre*  $L_n^\alpha(x)$  y cuando  $\sigma$  es de grado 2 con dos raíces simples, *polinomios de Jacobi*  $P_n^{\alpha, \beta}(x)$ , respectivamente. En las tablas 5.1 y 5.2 están representados los principales parámetros de dichas familias, en las cuales  $(a)_n$  denota al símbolo de Pochhammer

$$(a)_0 = 1, \quad (a)_k = a(a+1) \cdots (a+k-1), \quad k = 1, 2, 3, \dots \quad (5.1.39)$$

Para los polinomios  $\sigma$  se han escogido las llamadas *formas canónicas*.

Cuadro 5.1: Clasificación de las SPO Clásicas.

$P_n(x)$	$H_n(x)$	$L_n^\alpha(x)$	$P_n^{\alpha,\beta}(x)$
$\sigma(x)$	1	$x$	$1 - x^2$
$\tau(x)$	$-2x$	$-x + \alpha + 1$	$-(\alpha + \beta + 2)x + \beta - \alpha$
$\lambda_n$	$2n$	$n$	$n(n + \alpha + \beta + 1)$
$\rho(x)$	$e^{-x^2}$	$x^\alpha e^{-x}$ $\alpha > -1$	$(1 - x)^\alpha (1 + x)^\beta$ $\alpha, \beta > -1$
$\rho_n(x)$	$e^{-x^2}$	$x^{n+\alpha} e^{-x}$	$(1 - x)^{n+\alpha} (1 + x)^{n+\beta}$

### Representación hipergeométrica

De la fórmula de Rodrigues<sup>4</sup> (5.1.12) se puede obtener la representación de los polinomios de Hermite, Laguerre y Jacobi en términos de la función hipergeométrica de Gauss  ${}_2F_1$  definida en el caso más general de forma

$${}_pF_q \left( \begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix} \middle| x \right) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{x^k}{k!}. \quad (5.1.40)$$

De esta manera encontramos que

$$H_n(x) = \begin{cases} (-1)^m \left(\frac{1}{2}\right)_m {}_1F_1 \left( \begin{matrix} -m \\ \frac{1}{2} \end{matrix} \middle| x^2 \right), & n = 2m \\ (-1)^m \left(\frac{3}{2}\right)_m x {}_1F_1 \left( \begin{matrix} -m \\ \frac{3}{2} \end{matrix} \middle| x^2 \right), & n = 2m + 1 \end{cases} \quad (5.1.41)$$

$$L_n^\alpha(x) = \frac{(-1)^n \Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1)} {}_1F_1 \left( \begin{matrix} -n \\ \alpha + 1 \end{matrix} \middle| x \right), \quad (5.1.42)$$

$$P_n^{\alpha,\beta}(x) = \frac{2^n (\alpha + 1)_n}{(n + \alpha + \beta + 1)_n} {}_2F_1 \left( \begin{matrix} -n, n + \alpha + \beta + 1 \\ \alpha + 1 \end{matrix} \middle| \frac{1 - x}{2} \right). \quad (5.1.43)$$

Como consecuencia de las fórmulas anteriores podemos obtener los valores de los polinomios en los extremos del intervalo de ortogonalidad. Estos valores pueden ser obtenidos también a partir de la fórmula de Rodrigues (5.1.12) aplicando la regla de Leibniz para calcular la  $n$ -ésima

<sup>4</sup>Para ello basta usar la regla de Leibniz para calcular la  $n$ -ésima derivada de un producto  $(f \cdot g)^{(n)} = \sum_{k=0}^n p_{\bar{i}} n k f^{(k)} \cdot g^{(n-k)}$ . Otra posibilidad es usar series de potencias y el método de coeficientes indeterminados de Euler (ver e.g. [6, 21]).

Cuadro 5.2: Parámetros de las SPO Mónicas ( $a_n = 1$ ).

$P_n(x)$	$H_n(x)$	$L_n^\alpha(x)$	$P_n^{\alpha,\beta}(x)$
$B_n$	$\frac{(-1)^n}{2^n}$	$(-1)^n$	$\frac{(-1)^n}{(n + \alpha + \beta + 1)_n}$
$b_n$	0	$-n(n + \alpha)$	$\frac{n(\alpha - \beta)}{2n + \alpha + \beta}$
$d_n^2$	$\frac{n!\sqrt{\pi}}{2^n}$	$\Gamma(n + \alpha + 1)n!$	$\frac{2^{\alpha+\beta+2n+1}n!\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+\alpha+\beta+1)(2n+\alpha+\beta+1)(n+\alpha+\beta+1)_n^2}$
$\alpha_n$	1	1	1
$\beta_n$	0	$2n + \alpha + 1$	$\frac{\beta^2 - \alpha^2}{(2n+\alpha+\beta)(2n+2+\alpha+\beta)}$
$\gamma_n$	$\frac{n}{2}$	$n(n + \alpha)$	$\frac{4n(n+\alpha)(n+\beta)(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)}$
$\tilde{\alpha}_n$	0	0	$-n$
$\tilde{\beta}_n$	0	$n$	$\frac{2(\alpha - \beta)n(n + \alpha + \beta + 1)}{(2n + \alpha + \beta)(2n + 2 + \alpha + \beta)}$
$\tilde{\gamma}_n$	$n$	$n(n + \alpha)$	$\frac{4n(n + \alpha)(n + \beta)(n + \alpha + \beta)(n + \alpha + \beta + 1)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)^2(2n + \alpha + \beta + 1)}$

derivada de un producto de funciones.

$$H_n(0) = \begin{cases} \frac{(-1)^m(2m)!}{2^{2m}m!}, & n = 2m \\ 0, & n = 2m + 1 \end{cases}, \quad L_n^\alpha(0) = \frac{(-1)^n\Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1)}, \quad (5.1.44)$$

$$P_n^{\alpha,\beta}(1) = \frac{2^n(\alpha + 1)_n}{(n + \alpha + \beta + 1)_n}, \quad P_n^{\alpha,\beta}(-1) = \frac{(-1)^n 2^n(\beta + 1)_n}{(n + \alpha + \beta + 1)_n}.$$

### Casos particulares

1. Los polinomios de Legendre  $P_n(x) = P_n^{0,0}(x)$ .

2. Los polinomios de Chebyshev de primera especie  $T_n(x)$

$$T_n(x) = P_n^{-\frac{1}{2},-\frac{1}{2}}(x) = \frac{1}{2^{n-1}} \cos[n \arccos(x)].$$

3. Los polinomios de Chebyshev de segunda especie  $U_n(x)$

$$U_n(x) = P_n^{\frac{1}{2},\frac{1}{2}}(x) = \frac{1}{2^n} \frac{\operatorname{sen}[(n+1) \arccos(x)]}{\operatorname{sen}[\arccos(x)]}.$$

4. Los polinomios de Gegenbauer  $G_n^\lambda(x) = P_n^{\lambda-\frac{1}{2}, \lambda-\frac{1}{2}}(x)$ ,  $\lambda > -\frac{1}{2}$ .

Utilizando la fórmula (5.1.32) se obtienen las ecuaciones ( $\nu = 1, 2, 3, \dots$ ,  $n = 0, 1, 2, \dots$ ):

$$(H_n(x))^{(\nu)} = \frac{n!}{(n-\nu)!} H_{n-\nu}(x), \quad (5.1.45)$$

$$(L_n^\alpha(x))^{(\nu)} = \frac{n!}{(n-\nu)!} L_{n-\nu}^{\alpha+\nu}(x), \quad (5.1.46)$$

$$(P_n^{\alpha,\beta}(x))^{(\nu)} = \frac{n!}{(n-\nu)!} P_{n-\nu}^{\alpha+\nu, \beta+\nu}(x), \quad (5.1.47)$$

donde  $(P_n(x))^{(\nu)}$  denota la  $\nu$ -ésima derivada de  $P_n(x)$ .

## 5.2. Resolución de la ecuación de Schrödinger

Veamos ahora algunos ejemplos sencillos de cómo se usa el método de Nikiforov y Uvarov para resolver algunas ecuaciones de la Mecánica cuántica.

### 5.2.1. El oscilador armónico cuántico unidimensional

Como ejemplo apliquemos la técnica anterior al caso del oscilador armónico cuántico.

Partimos de la ecuación de Schrödinger para el oscilador armónico

$$-\frac{\hbar^2}{2m}\Psi''(x) + \frac{1}{2}m\omega^2 x^2 \Psi(x) = E\Psi(x), \quad x \in \mathbb{R}.$$

Haciendo el cambio  $x = x_0\xi$ ,  $x_0 = \sqrt{\hbar/(m\omega)}$ ,  $E = \hbar\omega\varepsilon/2$ , se transforma en la ecuación

$$\Psi''(\xi) + (\varepsilon - \xi^2)\Psi(\xi) = 0,$$

que obviamente es del tipo (5.1.1) con  $\tilde{\tau}(\xi) = 0$ ,  $\sigma(\xi) = 1$  y  $\tilde{\sigma}(\xi) = \varepsilon - \xi^2$ . Para  $\pi(\xi)$ , (5.1.5) nos da

$$\pi(\xi) = \pm\sqrt{\xi^2 + (k - \varepsilon)}.$$

Como el polinomio  $\xi^2 + (k - \varepsilon)$  ha de ser un cuadrado perfecto, entonces  $k = \varepsilon$  y, por tanto,  $\pi(\xi) = \pm\xi$ , luego

$$\begin{aligned} \pi(\xi) = \xi, & \quad \pi'(\xi) = 1, & \quad \lambda = \varepsilon + 1, & \quad \tau(\xi) = 2\xi, \\ \pi(\xi) = -\xi, & \quad \pi'(\xi) = -1, & \quad \lambda = \varepsilon - 1, & \quad \tau(\xi) = -2\xi, \end{aligned}$$

que nos conducen a las ecuaciones

$$y''(\xi) + 2\xi y'(\xi) + (\varepsilon + 1)y(\xi) = 0, \quad y''(\xi) - 2\xi y'(\xi) + (\varepsilon - 1)y(\xi) = 0,$$

respectivamente. En cada caso la función  $\phi(\xi)$  es la solución de las ecuaciones  $\phi'/\phi = \xi$  y  $\phi'/\phi = -\xi$ , que conducen a las funciones

$$\phi(\xi) = e^{\xi^2/2}, \quad \text{y} \quad \phi(\xi) = e^{-\xi^2/2},$$

respectivamente. Finalmente, la ecuación  $y''(\xi) - 2\xi y'(\xi) + (\varepsilon - 1)y(\xi) = 0$  corresponde a la ecuación hipergeométrica de los polinomios de Hermite, por tanto tenemos  $\varepsilon - 1 = 2n$ ,  $n = 0, 1, 2, \dots$  y las soluciones normalizadas de nuestra ecuación original serán

$$\Psi(\xi) = N_n e^{-\xi^2/2} H_n(\xi), \quad \varepsilon = 2n + 1, \quad n = 0, 1, 2, \dots \quad (5.2.1)$$

Para calcular  $N_n$  notamos que

$$\int_{-\infty}^{\infty} \Psi(\xi) \overline{\Psi(\xi)} d\xi = N_n^2 \int_{-\infty}^{\infty} H_n^2(\xi) e^{-\xi^2} d\xi = N_n^2 d_n^2, \quad d_n^2 = \frac{\sqrt{\pi n!}}{2^n}.$$

Entonces de la condición de normalización

$$1 = \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = x_0 \int_{-\infty}^{\infty} \Psi(\xi) \overline{\Psi(\xi)} d\xi,$$

$$\text{luego } N_n = \sqrt{\frac{2^n}{x_0 \sqrt{\pi n!}}}, \quad x_0 = \sqrt{\hbar/(m\omega)}.$$

Es fácil ver que la otra ecuación tiene como soluciones los polinomios  $H_n(-\xi)$ , por lo que sus soluciones  $\Psi(\xi) = e^{\xi^2/2} H_n(-\xi)$  no son de cuadrado integrable en  $\mathbb{R}$ , y por tanto no tienen sentido físico (esto se podía predecir si tenemos en cuenta la nota 5.1.9 del apartado anterior). De esta forma las únicas soluciones estacionarias del oscilador armónico son las funciones (5.2.1) anteriores.

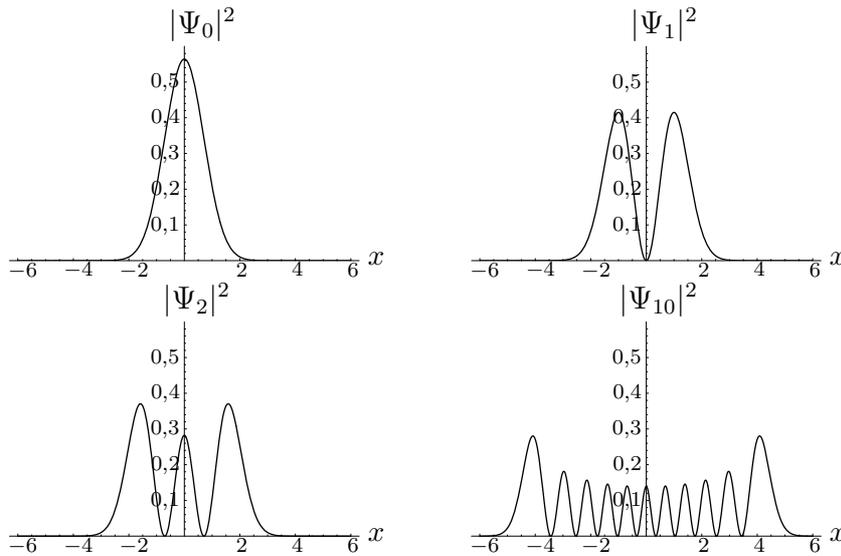


Figura 5.1: Estado fundamental  $\Psi_0$  y excitados  $\Psi_n$ ,  $n = 1, 2$  y  $10$  del oscilador armónico.

Así pues, a diferencia del oscilador clásico, el oscilador cuántico tiene una energía discreta definida por la expresión  $E_n = \hbar\omega(n + 1/2)$ ,  $n = 0, 1, 2, \dots$  correspondiente al estado

$$\Psi(x) = \sqrt{\frac{2^n}{x_0 \sqrt{\pi n!}}} e^{-\frac{1}{2}\left(\frac{x}{x_0}\right)^2} H_n\left(\frac{x}{x_0}\right), \quad n = 0, 1, 2, \dots, \quad x_0 = \sqrt{\hbar/(m\omega)}.$$

Finalmente, mencionemos que las autofunciones del oscilador definen una base ortogonal completa en el espacio de las funciones de cuadrado integrable y por tanto las podemos usar para desarrollar en la misma cualquier función de este espacio.

### 5.2.2. La ecuación de Schrödinger en un potencial central

Partiremos de la ecuación de Schrödinger estacionaria para el átomo de hidrógeno en coordenadas esféricas<sup>5</sup>

$$-\frac{\hbar^2}{2m}\Delta\Psi(r, \theta, \phi) + V(r)\Psi(r, \theta, \phi) = E\Psi(r, \theta, \phi), \quad (5.2.2)$$

donde  $m$  es la masa del electrón (que se supone despreciable respecto a la masa del núcleo),  $\phi \in [0, 2\pi)$ ,  $\theta \in [0, \pi]$ , y el laplaciano en coordenadas esféricas tiene la forma

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right), \quad (5.2.3)$$

o bien

$$\Delta = \Delta_r + \frac{1}{r^2} \Delta_{\sphericalangle},$$

donde

$$\Delta_r = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right), \quad \Delta_{\sphericalangle} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

denotan a los laplacianos radial y angular respectivamente.

Por simplicidad vamos a reescribir la ecuación anterior en la forma

$$\left[ \Delta_r + \frac{1}{r^2} \Delta_{\sphericalangle} \right] \Psi(r, \theta, \phi) + [\varepsilon - v(r)] \Psi(r, \theta, \phi) = 0, \quad (5.2.4)$$

donde  $v(r) = 2m/\hbar^2 V(r)$  y  $\varepsilon = -2m/\hbar^2 E$ .

Separando las variables  $\Psi(r, \theta, \phi) = F(r)Y(\theta, \phi)$  obtenemos las ecuaciones

$$\begin{aligned} \Delta_{\sphericalangle} Y(\theta, \phi) + \mu Y(\theta, \phi) &= 0, \\ \Delta_r F(r) + \left[ \varepsilon - v(r) - \frac{\mu}{r^2} \right] F(r) &= 0, \end{aligned} \quad (5.2.5)$$

donde  $\mu$  es cierta constante a determinar.

Nótese además que la condición de normalización

$$\int_0^\infty \int_0^{2\pi} \int_0^\pi |\Psi(r, \theta, \phi)|^2 r^2 \sin \theta d\theta d\phi dr = 1,$$

se transforma en

$$\int_0^{2\pi} \int_0^\pi |Y(\theta, \phi)|^2 \sin \theta d\theta d\phi = 1 \quad \text{y} \quad \int_0^\infty |F(r)|^2 r^2 dr = 1. \quad (5.2.6)$$

---

<sup>5</sup>Al ser el potencial de interacción  $V(r)$  un potencial central, i.e., sólo depende del radio, es más sencillo resolver la ecuación en coordenadas esféricas  $r, \theta, \phi$ .

### 5.2.3. Los armónicos esféricos

Comencemos por la primera de las dos ecuaciones anteriores. Separando variables  $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$  obtenemos las siguientes dos ecuaciones

$$\begin{aligned} \Phi''(\phi) + \nu\Phi(\phi) &= 0, \\ \text{sen } \theta \frac{\partial}{\partial \theta} \left( \text{sen } \theta \frac{\partial \Theta(\theta)}{\partial \theta} \right) + [\mu \text{sen}^2 \theta - \nu]\Theta(\theta) &= 0, \end{aligned} \quad (5.2.7)$$

donde  $\nu$  es cierta constante.

Asumamos que la función  $\Phi$  es univalente, entonces, de la condición de periodicidad  $\Phi(\phi + 2\pi) = \Phi(\phi)$  se sigue que  $\nu = m^2$ ,  $m \in \mathbb{Z}$ . Luego

$$\Phi_m(\phi) = C_m e^{im\phi}, \quad m \in \mathbb{Z}.$$

Nótese que las funciones  $\Phi_m(\phi)$  son ortogonales

$$\int_0^{2\pi} \overline{\Phi_m(\phi)} \Phi_{m'}(\phi) d\phi = \Xi_n \delta_{mm'},$$

con  $\Xi_n = 2\pi C_m^2$ . Si queremos que sean normalizadas a la unidad entonces  $C_m = 1/\sqrt{2\pi}$ . Así tenemos las soluciones

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}, \quad m \in \mathbb{Z}. \quad (5.2.8)$$

La segunda ecuación en (5.2.7) se transforma entonces en

$$\text{sen } \theta \frac{\partial}{\partial \theta} \left( \text{sen } \theta \frac{\partial \Theta(\theta)}{\partial \theta} \right) + [\mu \text{sen}^2 \theta - m^2]\Theta(\theta) = 0,$$

Haciendo el cambio  $x = \cos \theta$  obtenemos<sup>6</sup>

$$(1-x^2) \frac{d}{dx} \left( (1-x^2) \frac{\partial \Theta(x)}{\partial x} \right) + [\mu(1-x^2) - m^2]\Theta(x) = 0, \quad (5.2.9)$$

o, equivalentemente,

$$\begin{aligned} \frac{d}{dx} \left( (1-x^2) \frac{\partial \Theta(x)}{\partial x} \right) + \left[ \mu - \frac{m^2}{1-x^2} \right] \Theta(x) &= 0 \implies \\ \Theta''(x) - \frac{2x}{1-x^2} \Theta'(x) + \frac{\mu(1-x^2) - m^2}{(1-x^2)^2} \Theta(x) &= 0. \end{aligned}$$

Esta ecuación del tipo (5.1.1) con

$$\tilde{\tau}(x) = -2x, \quad \tilde{\sigma}(x) = \mu(1-x^2) - m^2, \quad \sigma = 1-x^2,$$

luego el polinomio (5.1.6)

$$\Upsilon(x) = (\mu - k)x^2 + (k - \mu) + m^2$$

es un cuadrado perfecto si:

---

<sup>6</sup>Nótese que si  $x = \cos \theta$ , entonces  $d/d\theta = dx/d\theta \cdot d/dx = -\text{sen } \theta d/dx$ , y  $\text{sen } \theta = \sqrt{1-x^2}$ .

1.  $k = \mu \implies \pi(x) = \pm m \implies \tau(x) = -2(x \mp m), m = 0, 1, 2, \dots,$
2.  $k = \mu - m^2 \implies \pi(x) = \pm mx \implies \tau(x) = -2(\mp m + 1)x, m = 0, 1, 2, \dots$

Si ahora tenemos en cuenta la nota 5.1.9 tenemos que de las cuatro opciones debemos escoger  $\tau(x) = -2(m+1)x$  correspondiente a  $\pi(x) = -mx, k = \mu - m^2, \lambda = k + \pi' = \mu - m(m+1).$

Para calcular<sup>7</sup>  $\phi$  usamos (5.1.2) que nos da

$$\frac{\phi'}{\phi} = -\frac{mx}{1-x^2} \implies \phi(x) = (1-x^2)^{m/2}.$$

La ecuación de tipo hipergeométrico que obtenemos es, por tanto,

$$(1-x^2)y'' - 2(m+1)xy' + \lambda y = 0,$$

que tiene soluciones polinómicas según (5.1.15) cuando  $\lambda = n(2m+n+1)$ , luego

$$\mu = (m+n)(n+m+1), \quad n, m = 0, 1, 2, \dots$$

En adelante definiremos  $l = m+n, l = 0, 1, 2, \dots$ . Entonces,  $n = l - m \geq 0$ , y por tanto<sup>8</sup>

$$y(x) = P_{l-m}^{m,m}(x),$$

donde  $P_{l-m}^{m,m}$  son los correspondientes polinomios de Jacobi. Así, para  $m \geq 0$  la solución de (5.2.9) tiene la forma

$$\Theta_{lm}(x) = C_{lm}(1-x^2)^{m/2}P_{l-m}^{m,m}(x).$$

Nótese que

$$\int_{-1}^1 \Theta_{lm}(x)\Theta_{l'm}(x)dx = C_{lm}C_{l'm} \int_{-1}^1 P_{l-m}^{m,m}(x)P_{l'-m}^{m,m}(x)(1-x^2)^m dx = 0, \quad l \neq l',$$

además, si  $l = l'$ , y haciendo  $n = l - m$

$$\int_{-1}^1 [\Theta_{lm}(x)]^2 dx = \int_{-1}^1 [P_n^{m,m}(x)]^2 (1-x^2)^m dx = \frac{2^{2l+1}(l-m)!(l+m)!(l!)^2}{(2l+1)!(2l)!},$$

luego si queremos que  $\Theta_{lm}(x)$  sean ortonormales, es decir que se cumpla la segunda condición de (5.2.6) basta definir

$$C_{lm} = \frac{(2l)!}{l!} \sqrt{\frac{2l+1}{2^{2l+1}(l-m)!(l+m)!}}.$$

Así

$$\Theta_{lm}(x) = \frac{(2l)!}{l!} \sqrt{\frac{2l+1}{2^{2l+1}(l-m)!(l+m)!}} (1-x^2)^{m/2} P_{l-m}^{m,m}(x), \quad l \geq 0, \quad m = 0, 1, \dots, l. \quad (5.2.10)$$

<sup>7</sup>No confundir la función  $\phi(x)$  con el ángulo  $\phi$  en coordenadas esféricas. Abusando de la paciencia del lector y para no introducir otra notación diferente hemos optado por mantener la notación original de [16].

<sup>8</sup>Nótese que  $l \geq m \geq 0$ . Es decir, si fijamos  $l$ , entonces  $m = 0, 1, 2, \dots, l$ .

Sólo nos queda pendiente analizar qué ocurre si  $m < 0$  (ver (5.2.8)). Para ello escribiremos (5.2.10) usando la expresión analítica de  $P_{l-m}^{(m,m)}(x)$  mediante la fórmula de Rodrigues (5.1.17)

$$\Theta_{lm}(x) = \frac{(-1)^{l-m}}{l!} \sqrt{\frac{(2l+1)(l+m)!}{2^{2l+1}(l-m)!}} (1-x^2)^{-m/2} \frac{d^{l-m}}{d^{l-m}x} [(1-x^2)^l] \quad (5.2.11)$$

Si ahora usamos la ecuación (5.1.47) tenemos

$$P_{l-m}^{m,m}(x) = \frac{(l-m)!}{l!} \frac{d^m}{d^m x} P_l(x),$$

donde  $P_l(x)$  son los polinomios de Legendre de grado  $l$ , expresados analíticamente mediante la fórmula de Rodrigues

$$P_l(x) = \frac{(-1)^l l!}{(2l)!} \frac{d^l}{d^l x} [(1-x^2)^l]$$

Combinado las dos expresiones anteriores con (5.2.10) tenemos otra representación de las funciones  $\Theta_{lm}(x)$

$$\Theta_{lm}(x) = \frac{(-1)^l}{l!} \sqrt{\frac{(2l+1)(l-m)!}{2^{2l+1}(l+m)!}} (1-x^2)^{m/2} \frac{d^{l+m}}{d^{l+m}x} [(1-x^2)^l]. \quad (5.2.12)$$

Comparando (5.2.11) y (5.2.12) se deduce que

$$\Theta_{l-m}(x) = (-1)^m \Theta_{lm}(x), \quad m = 0, 1, 2, \dots, l.$$

La expresión anterior permite definir la función  $\Theta_{lm}(x)$  para  $m < 0$ . Así, para  $\mu = l(l+1)$ ,  $l = 0, 1, 2, \dots$  y  $m = -l, -l+1, \dots, l$ , tenemos que las funciones  $Y(\theta, \phi) := Y_{lm}(\theta, \phi)$  se define por la expresión

$$Y_{lm}(\theta, \phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \Theta_{lm}(\cos \theta), \quad l = 0, 1, 2, \dots, \quad m = -l, -l+1, \dots, l. \quad (5.2.13)$$

Estas funciones se denominan *armónicos esféricos*.

#### 5.2.4. Resolviendo la parte radial de la ecuación de Schrödinger

De los resultados del apartado anterior se sigue que (5.2.5) se transforma en

$$\Delta_r F(r) + \left[ \varepsilon - v(r) - \frac{l(l+1)}{r^2} \right] F(r) = 0.$$

Para resolver esta ecuación hacemos el cambio<sup>9</sup>  $F(r) = R(r)/r$  que nos conduce a la ecuación

$$R''(r) + \left[ \varepsilon - v(r) - \frac{l(l+1)}{r^2} \right] R(r) = 0, \quad (5.2.14)$$

donde ahora la primera de las condiciones de contorno (5.2.6) se transforma en

$$\int_0^\infty |R(r)|^2 dr = 1.$$

La ecuación (5.2.14) será nuestro punto de partida para resolver dos casos de extrema importancia en las aplicaciones: el átomo de hidrógeno y el oscilador armónico tridimensional.

<sup>9</sup>La razón fundamental es que  $\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial F(r)}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} [rF(r)]$ .

### El átomo de hidrógeno

Puesto que para el átomo de hidrógeno  $V(r) = -\alpha/r$ , (5.2.14) tiene la forma

$$R''(r) + \left[ \frac{2m}{\hbar^2} \left( E + \frac{\alpha}{r} \right) - \frac{l(l+1)}{r^2} \right] R(r) = 0.$$

Haciendo el cambio  $\zeta = r/a_0$ , donde  $a_0 = \hbar^2/(m\alpha)$ , y  $\varepsilon = a_0 E/\alpha$ , la ecuación anterior se transforma en la ecuación *adimensional*

$$R''(\zeta) + \left[ 2 \left( \varepsilon + \frac{1}{\zeta} \right) - \frac{l(l+1)}{\zeta^2} \right] R(\zeta) = 0.$$

Esta ecuación es del tipo (5.1.1) con

$$\sigma(\zeta) = \zeta, \quad \tilde{\sigma}(\zeta) = 2\varepsilon\zeta^2 + 2\zeta - l(l+1), \quad \tilde{\tau}(\zeta) = 0.$$

Por tanto, tenemos

$$\pi(\zeta) = \frac{1}{2} \pm \sqrt{\frac{1}{4} - 2\varepsilon\zeta^2 - 2\zeta + l(l+1) + k\zeta}.$$

Como  $\Upsilon(\zeta) = 1/4 - 2\varepsilon\zeta^2 - 2\zeta + l(l+1) + k\zeta$  a de ser un cuadrado perfecto (en la variable  $\zeta$ ), tenemos que  $k = 2 \pm \sqrt{-2\varepsilon(2l+1)}$ , luego para  $\pi(\zeta)$  tenemos las siguientes cuatro opciones

$$\pi(\zeta) = \frac{1}{2} + \sqrt{-2\varepsilon}\zeta \pm \left( l + \frac{1}{2} \right), \quad \pi(\zeta) = \frac{1}{2} - \sqrt{-2\varepsilon}\zeta \pm \left( l + \frac{1}{2} \right).$$

Como  $\tau(\zeta) = 2\pi(\zeta)$ , tenemos, usando la nota 5.1.9 que  $\tau' < 0$ , luego  $\pi' < 0$  de donde eliminamos las dos primeras. De las dos posibilidades restantes para el polinomio  $\pi(\zeta)$  seleccionamos, siguiendo nuevamente la nota 5.1.9, la que conduce a una función  $\tau$  que se anule para algún  $\zeta > 0$ , es decir

$$\pi(\zeta) = -\sqrt{-2\varepsilon}\zeta + (l+1).$$

Dicha solución corresponde a  $k = 2 - \sqrt{-2\varepsilon}(2l+1)$ , luego —la otra posibilidad, como se puede comprobar, conduce a una función no integrable en  $(0, +\infty)$ —

$$\tau(\zeta) = 2(l+1 - \sqrt{-2\varepsilon}\zeta) \quad \Rightarrow \quad \lambda = k + \pi'(\zeta) = 2(1 - (l+1)\sqrt{-2\varepsilon}).$$

Usando (5.1.2) tenemos

$$\frac{\phi'(\zeta)}{\phi(\zeta)} = \frac{l+1 - \sqrt{-2\varepsilon}\zeta}{\zeta} \quad \Rightarrow \quad \phi(\zeta) = \zeta^{l+1} e^{-\sqrt{-2\varepsilon}\zeta}.$$

Entonces la solución de nuestra ecuación es del tipo  $R(\zeta) = \zeta^{l+1} e^{-\sqrt{-2\varepsilon}\zeta} y(\zeta)$ , siendo  $y$  la solución de la ecuación

$$\zeta y''(\zeta) + [2(l+1) - \sqrt{-2\varepsilon}\zeta] y'(\zeta) + \lambda y(\zeta) = 0.$$

El cambio lineal  $x = 2\sqrt{-2\varepsilon}\zeta$  nos transforma la ecuación anterior en la ecuación

$$xy''(x) + [(2l+1) + 1 - x]y'(x) + \tilde{\lambda}y(x) = 0, \quad \tilde{\lambda} = \frac{\lambda}{2\sqrt{-2\varepsilon}},$$

que corresponde a los polinomios de Laguerre  $L_n^{2l+1}(x)$ . Además, como  $\tilde{\lambda} = n$ , entonces  $\lambda = 2n\sqrt{-2\varepsilon}$ . Por otro lado,  $\lambda = k + \pi'(\zeta)$ , luego

$$\varepsilon := \varepsilon_{n,l} = -\frac{1}{2(n+l+1)^2}, \quad n, l = 0, 1, 2, \dots$$

Así, fijados  $n$  y  $l$ ,  $n, l = 0, 1, 2, \dots$  la solución es

$$R_{nl}(\zeta) = N_{n,l} x^{l+1} e^{-x/2} L_n^{2l+1}(x), \quad x = 2\sqrt{-2\varepsilon_{n,l}}\zeta,$$

con  $N_{n,l}$  tal que

$$\int_0^\infty R_{nl}^2(r) dr = 1 \implies a_0 \int_0^\infty R_{nl}^2(\zeta) d\zeta = 1 \implies \frac{a_0(n+l+1)}{2} \int_0^\infty R_{nl}^2(x) dx = 1.$$

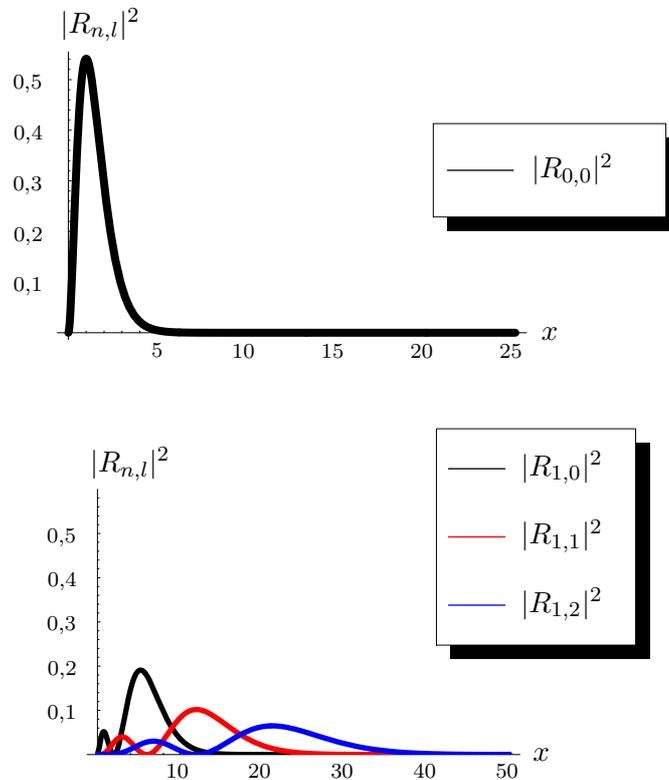


Figura 5.2: Estado fundamental (arriba) y excitados  $n = 1$ ,  $l = 0, 1, 2$  (abajo) del átomo de Hidrógeno.

Ahora bien,

$$\begin{aligned} \int_0^\infty R_{nl}^2(x) dx &= N_{n,l}^2 \int_0^\infty x^{2l+2} e^{-x} (L_n^{2l+1}(x))^2 dx = N_{n,l}^2 \int_0^\infty \rho(x) x (L_n^{2l+1}(x))^2 dx \\ &= N_{n,l}^2 \beta_n d_n^2 = N_{n,l}^2 2(n+l+1)n! \Gamma(n+2l+2), \end{aligned}$$

donde hemos usado la relación de recurrencia (5.1.22) para el producto  $x L_n^{2l+1}(x)$  y luego la ortogonalidad. Por tanto

$$N_{n,l} = \sqrt{\frac{1}{a_0(n+l+1)^2 n! (n+2l+1)!}}.$$

Entonces teniendo en cuenta que la parte radial es  $F(r) = R(r)/r$ , las autofunciones de la ecuación de Schrödinger para el átomo de hidrógeno (5.2.2) tienen la forma

$$\Psi_{n,l,m}(r, \theta, \phi) = \frac{2x^l e^{-x/2}}{a_0(n+l+1)^2 \sqrt{a_0 n! (n+2l+1)!}} L_n^{2l+1}(x) Y_{lm}(\theta, \phi), \quad n, l = 0, 1, 2, \dots$$

donde  $x = \frac{2r}{a_0(n+l+1)}$ ,  $a_0 = \hbar^2/(m\alpha)$  y sus correspondientes autoenergías  $E$  son por tanto

$$E_{n,l} = -\frac{m\alpha^2}{2\hbar^2} \frac{1}{(n+l+1)^2} = -\frac{m\alpha^2}{2\hbar^2} \frac{1}{n'^2}, \quad n' = 1, 2, 3, \dots$$

que, teniendo en cuenta que  $\alpha = e^2$ , nos conduce a la misma expresión<sup>10</sup> de Bohr (2.3.1).

### 5.2.5. El oscilador armónico tridimensional

Para el oscilador armónico tridimensional  $V(r) = \frac{1}{2}m\omega^2 r^2$ , así que (5.2.14) se convierte en

$$R''(r) + \left[ \frac{2m}{\hbar^2} \left( E - \frac{1}{2}m\omega^2 r^2 \right) - \frac{l(l+1)}{r^2} \right] R(r) = 0.$$

Haciendo el cambio  $\zeta = r/r_0$ , donde  $r_0 = \sqrt{\hbar/(m\omega)}$ , y  $E = (\hbar\omega/2)\varepsilon$  la ecuación anterior se transforma en la ecuación *adimensional*

$$R''(\zeta) + \left[ \varepsilon - \zeta^2 - \frac{l(l+1)}{\zeta^2} \right] R(\zeta) = 0.$$

Para convertir esta ecuación en una del tipo (5.1.1) hacemos el cambio  $z = \zeta^2$ , así tendremos

$$R''(z) + \frac{1}{2z} R'(z) + \frac{\varepsilon z - z^2 - l(l+1)}{4z^2} R(z) = 0.$$

donde tendremos

$$\sigma(z) = 2z, \quad \tilde{\sigma}(z) = \varepsilon z - z^2 - l(l+1), \quad \tilde{\tau}(\zeta) = 1.$$

Por tanto, tenemos

$$\pi(z) = \frac{1}{2} \pm \sqrt{\Upsilon(z)}, \quad \Upsilon(z) = z^2 + (2k - \varepsilon)z + \left( l + \frac{1}{2} \right)^2,$$

y para que este sea un cuadrado perfecto en  $z$ ,  $k$  debe tomar los valores  $k = \varepsilon/2 \pm (l + \frac{1}{2})$ , luego para  $\pi(z)$  tenemos las siguientes dos opciones

$$\pi(z) = \pm z + l + 1.$$

<sup>10</sup>En nuestra fórmula  $n$  no representa al número cuántico principal. Este correspondería al valor  $n' = n + l + 1$ .

Como  $\tau(z) = 1 + 2\pi(z)$ , tenemos, usando la nota 5.1.9 que  $\tau' < 0$ , luego  $\pi' < 0$  de donde deducimos que

$$\pi(z) = -z + l + 1,$$

que corresponde a  $k = \varepsilon/2 - (l + \frac{1}{2})$ , luego

$$\tau(z) = -2z + 2l + 3 \quad \Rightarrow \quad \lambda = k + \pi'(z) = \frac{\varepsilon}{2} - l - \frac{3}{2}.$$

Usando (5.1.2) tenemos

$$\frac{\phi'(z)}{\phi(z)} = \frac{-z + l + 1}{2z} \quad \Rightarrow \quad \phi(z) = z^{(l+1)/2} e^{-z/2}.$$

Entonces la solución de nuestra ecuación es del tipo  $R(z) = z^{(l+1)/2} e^{-z/2} y(z)$ , siendo  $y$  la solución de la ecuación

$$zy''(z) + [-z + l + 3/2]y'(z) + \frac{\lambda}{2}y(z) = 0.$$

que corresponde a los polinomios de Laguerre  $L_n^{l+1/2}(z)$ . Además, como  $\lambda/2 = n$ , y  $\lambda = k + \pi'(z) = \frac{\varepsilon}{2} - l - \frac{3}{2}$ , luego

$$\varepsilon := \varepsilon_{n,l} = 2 \left( 2n + l + \frac{3}{2} \right), \quad n, l = 0, 1, 2, \dots$$

Así, fijados  $n$  y  $l$ , la parte radial es  $F_{nl}(r) = R_{nl}(r)/r$  donde

$$R_{nl}(z) = N_{n,l} z^{(l+1)/2} e^{-z/2} L_n^{l+1/2}(z), \quad z = \zeta^2 = \left( \frac{r}{r_0} \right)^2,$$

con  $N_{n,l}$  tal que

$$\int_0^\infty R_{nl}^2(r) dr = 1 \quad \Rightarrow \quad r_0 \int_0^\infty R_{nl}^2(\zeta) d\zeta = 1,$$

pero

$$\int_0^\infty R_{nl}^2(\zeta) d\zeta = N_{n,l}^2 \int_0^\infty \frac{R(z)^2}{2z} dz = \frac{N_{n,l}^2}{2} \int_0^\infty z^{l+1/2} e^{-z} (L_n^{l+1/2}(z))^2 dz = \frac{N_{n,l}^2}{2},$$

donde  $d_n^2$  es el cuadrado de la norma de los polinomios de Laguerre  $L_n^{l+1/2}$ , i.e.,  $d_n^2 = n! \Gamma(n + l + 3/2)$ . Luego

$$N_{n,l} = \sqrt{\frac{2}{r_0 n! \Gamma(n + l + 3/2)}}.$$

## 5.3. El método de factorización de Schrödinger

### 5.3.1. Introducción

El objetivo de este apartado es estudiar un método sencillo que permite resolver ecuaciones diferenciales de Sturm-Liouville. El método debe su “popularidad” fundamentalmente a Schrödinger que lo usó para resolver la ecuación de *Schrödinger* para muchos sistemas físicos reales.

Actualmente el método se le conoce como método de factorización de Infeld y Hull debido al estudio detallado que estos autores hicieron en [13].

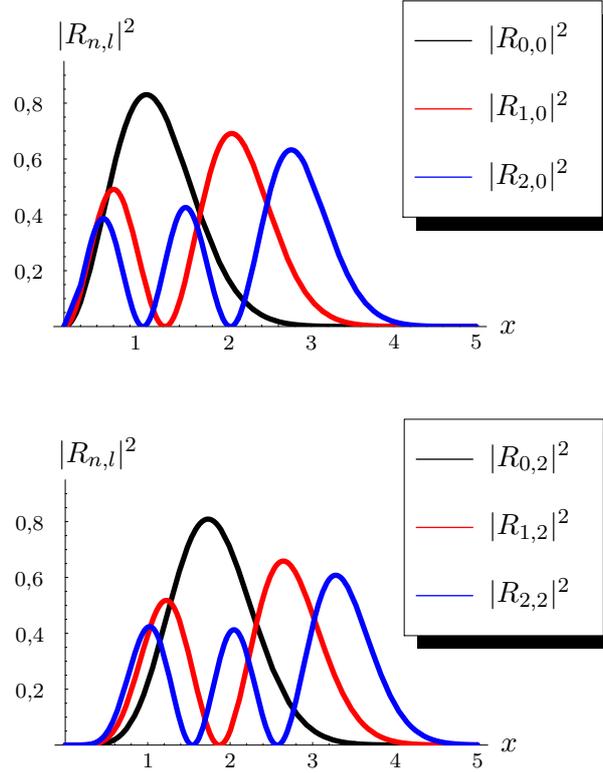


Figura 5.3: Estado con  $l = 0$ ,  $n = 0, 1, 2$  (arriba) y  $l = 2$ ,  $n = 0, 1, 2$  (abajo) del oscilador cuántico tridimensional.

### 5.3.2. El oscilador armónico

Sea la ecuación de Schrödinger para el oscilador armónico

$$\mathcal{H}\Psi_\lambda(x) := -\Psi_\lambda''(x) + x^2\Psi_\lambda(x) = \lambda\Psi_\lambda(x), \quad (5.3.1)$$

que escribiremos convenientemente en la forma

$$-(\mathcal{D}^2 - x^2\mathbf{I})\Psi_\lambda(x) = \lambda\Psi_\lambda(x), \quad \text{donde } \mathcal{D} := \frac{d}{dx}, \quad (5.3.2)$$

e  $\mathbf{I}$  es el operador identidad.

Si operamos “formalmente” podemos sustituir  $x^2 - \mathcal{D}^2$  por la correspondiente diferencia de cuadrados  $(x - \mathcal{D})(x + \mathcal{D})$  o  $(x + \mathcal{D})(x - \mathcal{D})$ . En la práctica esto no es del todo cierto pues  $x^2 - \mathcal{D}^2$  es un operador y obviamente no tiene por que tener lugar las expresiones anteriores como de hecho ocurre. No obstante

$$\begin{aligned} (x - \mathcal{D})(x + \mathcal{D})\Psi_\lambda(x) &= (x\mathbf{I} - \mathcal{D})(x\Psi_\lambda(x) + \Psi_\lambda'(x)) = -\Psi_\lambda''(x) + x^2\Psi_\lambda(x) - \Psi_\lambda(x) \\ &= \mathcal{H}\Psi_\lambda(x) - \Psi_\lambda(x) = (\lambda - 1)\Psi_\lambda(x), \end{aligned}$$

y

$$\begin{aligned} (x + \mathcal{D})(x - \mathcal{D})\Psi_\lambda(x) &= (x\mathbf{I} + \mathcal{D})(x\Psi_\lambda(x) - \Psi_\lambda'(x)) = -\Psi_\lambda''(x) + x^2\Psi_\lambda(x) + \Psi_\lambda(x) \\ &= \mathcal{H}\Psi_\lambda(x) + \Psi_\lambda(x) = (\lambda + 1)\Psi_\lambda(x). \end{aligned}$$

Lo anterior nos conduce a definir dos operadores

$$H_+ := x + \mathcal{D}, \quad H_- := x - \mathcal{D}$$

para los que se cumple

$$H_- H_+ \Psi_\lambda(x) = (\lambda - 1) \Psi_\lambda(x), \quad (5.3.3)$$

$$H_+ H_- \Psi_\lambda(x) = (\lambda + 1) \Psi_\lambda(x), \quad (5.3.4)$$

y

$$\mathcal{H} \Psi_\lambda(x) = (H_- H_+ + \mathbf{I}) \Psi_\lambda(x) = \lambda \Psi_\lambda(x). \quad (5.3.5)$$

Aplicando  $H_+$  a (5.3.3) tenemos

$$(H_+ H_-) H_+ \Psi_\lambda(x) = (\lambda - 1) H_+ \Psi_\lambda(x),$$

que al compararla con (5.3.4) nos conduce a<sup>11</sup>

$$H_+ \Psi_\lambda(x) = \alpha(\lambda) \Psi_{\lambda-2}(x). \quad (5.3.6)$$

Análogamente, aplicando  $H_-$  a (5.3.4) tenemos

$$(H_- H_+) H_- \Psi_\lambda(x) = (\lambda + 1) H_- \Psi_\lambda(x),$$

que al compararla con (5.3.3) nos da

$$H_- \Psi_\lambda(x) = \beta(\lambda) \Psi_{\lambda+2}(x). \quad (5.3.7)$$

De lo anterior se deduce que  $H_-$  es un operador “ascenso” (raising) y  $H_+$  es un operador “descenso” (lowering).<sup>12</sup>

Sea ahora el conjunto de las funciones  $\Psi(x)$  tales que  $\Psi(a) = \Psi(b) = 0$  ( $a, b \in \mathbb{R} \cup \pm\infty$ ) y definamos en dicho espacio el producto escalar<sup>13</sup>

$$\langle \Psi, \Phi \rangle = \int_a^b \Psi(x) \overline{\Phi(x)} dx.$$

Nótese que  $\overline{H_\pm} = H_\pm$ .

Los operadores  $H_+$  y  $H_-$  son uno el adjunto del otro. En efecto,

$$\begin{aligned} \int_a^b \Psi(x) \overline{H_+ \Phi(x)} dx &= \int_a^b \Psi(x) x \overline{\Phi(x)} dx + \int_a^b \Psi(x) \overline{\Phi'(x)} dx \\ &= \int_a^b \Psi(x) x \overline{\Phi(x)} dx + \Psi(x) \overline{\Phi(x)} \Big|_a^b - \int_a^b \Psi'(x) \overline{\Phi(x)} dx \\ &= \int_a^b (x - \mathcal{D}) \Psi(x) \overline{\Phi(x)} dx = \int_a^b H_- \Psi(x) \overline{\Phi(x)} dx \end{aligned}$$

<sup>11</sup>Nótese que de (5.3.4) se deduce que el operador  $H_+ H_-$  tiene como autovectores  $\Psi_\lambda$  y sus correspondientes autovalores son  $\lambda + 1$ , por tanto la expresión anterior indica que  $H_+ \Psi_\lambda$  son los autovectores correspondientes a  $\lambda - 2$  de donde se sigue que han de ser proporcionales a  $\Psi_{\lambda-2}$ .

<sup>12</sup>Pues aumentan o disminuyen el valor del autovalor, i.e., suben o bajan por el espectro de  $\mathcal{H}$ .

<sup>13</sup>Aquí la operación  $\bar{a}$  denota al complejo conjugado de  $a$ . En general trabajaremos en el espacio de funciones reales.

es decir  $\langle H_- \Psi, \Phi \rangle = \langle \Psi, H_+ \Phi \rangle$ .

Como consecuencia de lo anterior se sigue que  $\mathcal{H} := -D^2 + x^2 = H_- H_+ + I$  es autoadjunto y por tanto (ver Proposición 4.2.12) sus autovalores son reales. Además, si  $\lambda \neq \gamma$  entonces  $\Psi_\lambda$  es ortogonal a  $\Psi_\gamma$  (ver Proposición 4.2.13).

Supongamos que  $|\Psi_\lambda|^2$  es integrable. Calculemos

$$\begin{aligned} \|\Psi_{\lambda-2}\|^2 &:= \langle \Psi_{\lambda-2}, \Psi_{\lambda-2} \rangle = \alpha^{-2}(\lambda) \langle H_+ \Psi_\lambda, H_+ \Psi_\lambda \rangle = \alpha^{-2}(\lambda) \langle \Psi_\lambda, H_- H_+ \Psi_\lambda \rangle \\ &= \alpha^{-2}(\lambda)(\lambda-1) \langle \Psi_\lambda, \Psi_\lambda \rangle = \alpha^{-2}(\lambda)(\lambda-1) \|\Psi_\lambda\|^2, \end{aligned} \quad (5.3.8)$$

luego

$$\|\Psi_{\lambda-2n}\|^2 = \frac{(\lambda-1)}{\alpha^2(\lambda)} \frac{(\lambda-2)}{\alpha^2(\lambda-1)} \cdots \frac{(\lambda-n)}{\alpha^2(\lambda-n+1)} \|\Psi_\lambda\|^2.$$

Obviamente para cada  $\lambda$  fijo, existirá un  $n$  tal que  $\lambda - n + 1 > 0$ ,  $\lambda - n \leq 0$ , por tanto el proceso  $H_+ \cdots H_+ \Psi_\lambda$  debe de culminar en algún momento, i.e., debe existir una función que denotaremos por  $\Psi_0(x)$  tal que

$$H_+ \Psi_0(x) = 0,$$

o, equivalentemente, para cierto  $\lambda_0$ ,  $\alpha(\lambda_0) = 0$ .

En este caso la ecuación (5.3.5) nos da  $\mathcal{H}\Psi_0 = \lambda_0\Psi_0$ ,

$$(H_- H_+ + I)\Psi_0(x) = \lambda_0\Psi_0(x) \quad \Rightarrow \quad \lambda_0 = 1.$$

Así pues aplicando  $H_-$  consecutivamente a  $\Psi_0$  obtendremos los valores  $\Psi_\lambda$  para  $\lambda = 1, 3, 5, \dots, 2N+1, \dots$ . En general tenemos  $\lambda_n = 2n + 1$ ,  $n = 0, 1, 2, \dots$ .

Análogamente

$$\|\Psi_{\lambda+2}\|^2 = \beta^{-2}(\lambda)(\lambda+1) \|\Psi_\lambda\|^2. \quad (5.3.9)$$

Lo anterior nos indica que conviene escribir la función  $\Psi_{\lambda_n}$  como  $\Psi_n$  ( $\Psi_{\lambda_{n-2}}$  se convierte en  $\Psi_{n-1}$  y  $\Psi_{\lambda_{n+2}}$  se convierte en  $\Psi_{n+1}$ ). Luego tenemos las ecuaciones

$$H_+ \Psi_0(x) = 0, \quad H_+ \Psi_n = \alpha_n \Psi_{n-1}, \quad H_- \Psi_n = \beta_n \Psi_{n+1}, \quad n = 0, 1, 2, \dots,$$

donde hemos denotado por  $\alpha_n := \alpha(\lambda_n)$  y  $\beta_n := \beta(\lambda_n)$ .

Además (5.3.8) implica  $\|\Psi_{n-1}\|^2 = 2n\alpha_n^{-2} \|\Psi_n\|^2$  y (5.3.9) implica  $\|\Psi_{n+1}\|^2 = (2n+2)\beta_n^{-2} \|\Psi_n\|^2$ . Si queremos que los operadores  $H_+$  y  $H_-$  mantengan invariante la norma de las funciones  $\Psi_n$  entonces de las fórmulas anteriores se deduce que  $\alpha_n = \sqrt{2n}$  y  $\beta_n = \sqrt{2n+2}$ . Así tenemos

$$H_+ \Psi_n = \sqrt{2n} \Psi_{n-1}, \quad H_- \Psi_n = \sqrt{2n+2} \Psi_{n+1}, \quad n = 0, 1, 2, \dots$$

Finalmente, usando que  $H_+ \Psi_0(x) = 0$ ,  $\|\Psi_0\| = 1$ , obtenemos las soluciones normalizadas

$$x\Psi_0(x) + \Psi_0'(x) = 0 \quad \Rightarrow \quad \Psi_0(x) = \frac{1}{\sqrt[4]{\pi}} e^{-x^2/2},$$

y  $[H_-]^n \Psi_0(x) = \sqrt{(2n)!!} \Psi_n(x)$ , luego

$$\Psi_n(x) = \frac{1}{\pi^{1/4} \sqrt{(2n)!!}} [H_-]^n e^{-x^2/2} = \frac{1}{\pi^{1/4} \sqrt{(2n)!!}} [xI - \mathcal{D}]^n e^{-x^2/2}.$$

Esta forma de resolver el problema del oscilador armónico cuántico es totalmente análoga a la del apartado 5.2.1.

### 5.3.3. El método de factorización

En este apartado vamos a generalizar el ejemplo anterior a ecuaciones algo más generales. Partiremos de la ecuación diferencial lineal de segundo orden (de Sturm-Liouville)

$$y''(x) + r(x, m)y(x) + \lambda y(x) = 0, \quad m = 0, 1, 2, \dots \quad (5.3.10)$$

o, equivalentemente<sup>14</sup>,

$$\mathcal{H}^m y(x, m, \lambda) = \lambda y(x, m, \lambda), \quad \mathcal{H}^m := -\mathcal{D}^2 - r(x, m)\mathbf{I}. \quad (5.3.11)$$

La función  $r(x, m)$  se suele denominar función potencial y a  $\mathcal{H}^m$ , hamiltoniano. Aquí  $m$  es un parámetro y  $\lambda$  es el correspondiente autovalor. En adelante asumiremos que  $\lambda$  y  $m$  son independientes.

**Definición 5.3.1** Diremos que la ecuación diferencial ordinaria (EDO) (5.3.10) es factorizable si la podemos cambiar por cada una de las siguientes ecuaciones

$$H_+^{m+1} H_-^{m+1} y(x, \lambda, m) = [\lambda - L(m+1)]y(x, \lambda, m), \quad (5.3.12)$$

$$H_-^m H_+^m y(x, \lambda, m) = [\lambda - L(m)]y(x, \lambda, m), \quad (5.3.13)$$

donde  $L(m)$  son ciertas constantes y

$$H_-^m = k(x, m)\mathbf{I} - \frac{d}{dx}, \quad H_+^m = k(x, m)\mathbf{I} + \frac{d}{dx}. \quad (5.3.14)$$

Los operadores  $H_\pm^m$  se conocen como operadores “escalera”. Nótese además que ambos operadores  $H_\pm^m$  son de orden 1.

**Teorema 5.3.2** Si  $y(x, \lambda, m)$  es solución de  $y''(x) + r(x, m)y(x) + \lambda y(x) = 0$ , entonces

$$H_-^{m+1} y(x, \lambda, m) = \alpha(\lambda, m)y(x, \lambda, m+1), \quad (5.3.15)$$

$$H_+^m y(x, \lambda, m) = \beta(\lambda, m)y(x, \lambda, m-1). \quad (5.3.16)$$

Demostración: Aplicando  $H_-^{m+1}$  a (5.3.12), obtenemos

$$H_-^{m+1} H_+^{m+1} [H_-^{m+1} y(\lambda, m)] = [\lambda - L(m+1)][H_-^{m+1} y(\lambda, m)],$$

y comparando el resultado con (5.3.13) deducimos (5.3.15). Análogamente, si aplicamos  $H_+^{m+1}$  a (5.3.13), obtenemos

$$H_+^m H_-^m [H_+^m y(\lambda, m)] = [\lambda - L(m)][H_+^m y(\lambda, m)],$$

que al compararlo con (5.3.12), nos conduce a (5.3.16).  $\square$

En adelante, denotaremos  $\alpha_m := \alpha(\lambda, m)$ ,  $\beta_m := \beta(\lambda, m)$ .

Como en el caso del oscilador armónico vamos a considerar el conjunto de las funciones  $\Psi(x)$  tales que  $\Psi(a) = \Psi(b) = 0$  ( $a, b \in \mathbb{R} \cup \pm\infty$ ) y en dicho espacio definiremos el producto escalar

$$\langle \Psi, \Phi \rangle = \int_a^b \Psi(x) \overline{\Phi(x)} dx.$$

<sup>14</sup>Hemos decidido mantener en la medida de lo posible la notación original de [13], por lo que aquí  $\mathcal{H}^m$  denota al operador  $\frac{d^2}{dx^2} + r(x, m)\mathbf{I}$ , y no a la potencia  $m$ -ésima de  $\mathcal{H}$ .

**Teorema 5.3.3** *Los operadores  $H_+^m$  y  $H_-^m$  son adjuntos entre sí, i.e.,*

$$\langle H_-^m \phi, \psi \rangle = \int_a^b H_-^m \phi \cdot \psi dx = \int_a^b \phi \cdot H_+^m \psi dx = \langle \phi, H_+^m \psi \rangle \quad (5.3.17)$$

Demostración: Ante todo notemos que si  $k(x, m)$  es una función real entonces  $\overline{H_\pm^m} = H_\pm^m$ . En adelante asumiremos que  $k(x, m)$  es real. Entonces

$$\begin{aligned} \int_a^b \Psi(x) \overline{H_-^m \Phi(x)} dx &= \int_a^b \Psi(x) k(x, m) \overline{\Phi(x)} dx - \int_a^b \Psi(x) \overline{\Phi'(x)} dx \\ &= \int_a^b \Psi(x) k(x, m) \overline{\Phi(x)} dx - \Psi(x) \overline{\Phi(x)} \Big|_a^b + \int_a^b \Psi'(x) \overline{\Phi(x)} dx \\ &= \int_a^b (k(x, m) + \mathcal{D}) \Psi(x) \overline{\Phi(x)} dx = \int_a^b H_+^m \Psi(x) \overline{\Phi(x)} dx, \end{aligned}$$

es decir  $\langle H_-^m \Psi, \Phi \rangle = \langle \Psi, H_+^m \Phi \rangle$ . □

Nótese también que el hamiltoniano  $\mathcal{H}^m$  es un operador autoadjunto. Ello se deduce de las expresiones

$$\mathcal{H}^m = H_+^{m+1} H_-^{m+1} + L(m+1)I = H_-^m H_+^m + L(m)I$$

o directamente de (5.3.11). Además, como es autoadjunto, entonces sus autovalores son reales y las autofunciones correspondientes a diferentes autovalores son ortogonales.

El objetivo fundamental del método es poder resolver EDOs de una forma sencilla y “razonable” obteniendo soluciones “buenas”. Así, sería deseable que si partiésemos de una función  $y(x, \lambda, m)$  de cuadrado integrable las operaciones  $H_-^{m+1} y(x, \lambda, m)$  y  $H_+^m y(x, \lambda, m)$  dieran como resultado funciones de cuadrado integrable.

Esta condición la asumiremos (su demostración es bastante engorrosa), no obstante en los ejemplos veremos fácilmente que se cumple, es decir que partiendo de una  $y(x, \lambda, m)$  de cuadrado integrable  $y(x, \lambda, m \pm 1)$  también lo será. Además, nos restringiremos a EDOs cuyos coeficientes tienen singularidades a lo sumo en el extremo del intervalo  $(a, b)$  lo cual reduce el análisis de la integrabilidad cuadrática de las soluciones a su comportamiento en dichos extremos.

**Teorema 5.3.4 (Clase I)** *Sea  $L(m)$  una función creciente en  $0 < m \leq M$  ( $M$  puede ser  $+\infty$ ),  $m \in \mathbb{Z}$  y supongamos que*

$$\lambda \leq \max[L(M), L(M+1)].$$

*Entonces para que la EDO  $y''(x) + r(x, m)y(x) + \lambda y(x) = 0$  tenga una sucesión de soluciones integrables es necesario que exista un  $l \in \mathbb{N}$  tal que*

$$\lambda = \lambda_l = L(l+1).$$

*Además para ese  $l$ , y denotando  $y_m^l(x) := y(x, L(l+1), m)$  tenemos*

$$H_-^{l+1} y_m^l(x) = 0 \quad \Rightarrow \quad y_m^l(x) = e^{\int k(x, l+1) dx},$$

$y$ , entonces para  $m = 0, 1, \dots, l-1, l$ ,

$$H_+^m y_m^l(x) = \sqrt{L(l+1) - L(m)} y_{m-1}^l(x), \quad (5.3.18)$$

y, por tanto,

$$y_m^l(x) = \left[ \prod_{k=m+1}^l \frac{1}{\sqrt{L(l+1) - L(k)}} \right] H_+^{m+1} H_+^{m+2} \dots H_+^l y_l^l(x), \quad m < l. \quad (5.3.19)$$

Nótese que, si  $\|y_l^l\| = 1$ , entonces las funciones  $y_m^l$  son ortonormales.

Demostración: Asumamos que  $y(x, \lambda, m)$  es de cuadrado integrable en  $(a, b)$ . Entonces,  $y(x, \lambda, m+1)$  también lo será. En efecto como

$$H_-^{m+1} y(x, \lambda, m) = \alpha_m y(x, \lambda, m+1),$$

tenemos

$$\begin{aligned} \int_a^b [y(x, \lambda, m+1)]^2 dx &= \alpha_m^{-2} \int_a^b [H_-^{m+1} y(x, \lambda, m)]^2 dx \\ &= \alpha_m^{-2} \int_a^b y(x, \lambda, m) H_+^{m+1} H_-^{m+1} y(x, \lambda, m) dx \\ &= \frac{\lambda - L(m+1)}{\alpha_m^2} \int_a^b [y(x, \lambda, m)]^2 dx. \end{aligned}$$

Luego para  $l > m$ ,

$$\int_a^b [y(x, \lambda, l+1)]^2 dx = \frac{\lambda - L(m+1)}{\alpha_m^2} \dots \frac{\lambda - L(l+1)}{\alpha_m^2} \int_a^b [y(x, \lambda, m)]^2 dx.$$

Ahora bien, como  $L(m)$  es creciente entonces existirá un  $l$  tal que  $L(l+1) > \lambda$ , lo cual implica que  $\lambda = L(l+1)$  o  $y(x, \lambda, l+1) \equiv 0$ . En caso que  $y(x, \lambda, l+1) \equiv 0$ , usando (5.3.15) tenemos

$$H_-^{l+1} y(x, \lambda, l) = 0,$$

y por tanto, asumiendo que  $y(x, \lambda, l) \neq 0$ , (5.3.12) nos da

$$H_+^{l+1} \underbrace{H_-^{l+1} y(x, \lambda, l)}_{=0} = [\lambda - L(l+1)] \underbrace{y(x, \lambda, l)}_{\neq 0} \Rightarrow \lambda - L(l+1) = 0.$$

Además, si queremos que el operador  $H_-^{m+1}$  mantenga la normalización entonces  $\alpha_m = \sqrt{L(l+1) - L(m+1)}$ . De forma análoga pero usando (5.3.16) y (5.3.13) tenemos que

$$\|y(x, \lambda, m-1)\|^2 = \frac{\lambda - L(m)}{\beta_m^2} \|y(x, \lambda, m)\|^2 = \frac{L(l+1) - L(m)}{\beta_m^2} \|y(x, \lambda, m)\|^2,$$

de donde se sigue que  $\beta_m = \sqrt{L(l+1) - L(m)}$ .

Así pues, denotando por  $y_m^l(x) := y(x, L(l+1), m)$  tenemos que (5.3.15) se transforma en

$$H_-^{m+1} y_m^l(x) = \sqrt{L(l+1) - L(m+1)} y_{m+1}^l(x), \quad m < l, \quad (5.3.20)$$

y, en particular,

$$H_-^{l+1}y_l^l(x) = 0,$$

así que usando (5.3.16)

$$H_+^m y_m^l(x) = \sqrt{L(l+1) - L(m)} y_{m-1}^l(x), \quad (5.3.21)$$

de donde, por inducción se deduce (5.3.19).  $\square$

De forma totalmente análoga se tiene el siguiente

**Teorema 5.3.5 (Clase II)** *Sea  $L(m)$  una función decreciente en  $0 \leq m \leq M$  ( $M$  puede ser  $+\infty$ ,  $m \in \mathbb{Z}$  y supongamos que  $\lambda \leq L(0)$ ). Entonces para que la EDO  $y''(x) + r(x, m)y(x) + \lambda y(x) = 0$  tenga una sucesión de soluciones integrables es necesario que exista un  $l \in \mathbb{N}$  tal que*

$$\lambda := \lambda_l = L(l),$$

Además, denotando por  $y_m^l(x) := y(x, L(l), m)$ , tenemos

$$H_+^m y_m^l(x) = \sqrt{L(l) - L(m)} y_{m-1}^l(x), \quad (5.3.22)$$

y, en particular,

$$H_+^l y_l^l(x) = 0 \quad \Rightarrow \quad y_l^l(x) = e^{-\int k(x, l) dx}$$

Además, para  $m = l, l+1, \dots$  tenemos

$$H_-^{m+1} y_m^l(x) = \sqrt{L(l) - L(m+1)} y_{m+1}^l(x), \quad (5.3.23)$$

y por tanto,

$$y_m^l(x) = \left[ \prod_{k=l}^{m-1} \frac{1}{\sqrt{L(l) - L(k+1)}} \right] H_-^m H_-^{m-1} \dots H_-^{l+1} y_l^l(x), \quad m > l. \quad (5.3.24)$$

La demostración es análoga y la omitiremos.

Veamos ahora bajo que condiciones podemos factorizar la ecuación (5.3.10).

Partimos de la expresión (5.3.12)

$$\begin{aligned} (k(x, m+1) + \mathcal{D})(k(x, m+1) - \mathcal{D})y(x) &= [\lambda - L(m+1)]y(x) \quad \Rightarrow \\ (k(x, m+1) + \mathcal{D})(k(x, m+1)y(x) - y'(x)) &= [\lambda - L(m+1)]y(x) \quad \Rightarrow \\ k^2(x, m+1)y(x) + k'(x, m+1)y(x) - y''(x) &= [\lambda - L(m+1)]y(x). \end{aligned}$$

Si eliminamos  $y''$  usando la ecuación diferencial (5.3.10) y cambiamos  $m$  por  $m-1$  obtenemos

$$k^2(x, m) + k'(x, m) + L(m) = -r(x, m-1). \quad (5.3.25)$$

Si ahora usamos (5.3.13)

$$\begin{aligned} (k(x, m) - \mathcal{D})(k(x, m) + \mathcal{D})y(x) &= [\lambda - L(m)]y(x) \quad \Rightarrow \\ (k(x, m) - \mathcal{D})(k(x, m)y(x) + y'(x)) &= [\lambda - L(m)]y(x) \quad \Rightarrow \\ k^2(x, m)y(x) - k'(x, m)y(x) - y''(x) &= [\lambda - L(m)]y(x), \end{aligned}$$

y nuevamente eliminamos  $y''$  usando la ecuación diferencial (5.3.10) obtenemos

$$k^2(x, m) - k'(x, m) + L(m) = -r(x, m). \quad (5.3.26)$$

Restando y sumando (5.3.25) y (5.3.26) obtenemos, respectivamente

$$L(m) = -\frac{r(x, m) + r(x, m-1)}{2} - k^2(x, m),$$

$$k'(x, m) = \frac{r(x, m) - r(x, m-1)}{2}.$$

Derivando la primera de las expresiones respecto a  $x$  y sustituyendo el resultado en la segunda deducimos que<sup>15</sup>

$$k(x, m) = -\frac{r'(x, m) + r'(x, m-1)}{2(r(x, m) - r(x, m-1))}.$$

Ahora bien, para que la EDO (5.3.10) sea factorizable  $L(m)$  ha de ser independiente de  $x$ . Así pues hemos probado el siguiente

**Teorema 5.3.6** *La EDO  $y''(x) + r(x, m)y(x) + \lambda y(x) = 0$  es factorizable en el sentido de la definición 5.3.1 si y sólo si*

$$k(x, m) = -\frac{r'(x, m) + r'(x, m-1)}{2(r(x, m) - r(x, m-1))}, \quad (5.3.27)$$

y

$$L(m) = -\frac{r(x, m) + r(x, m-1)}{2} - k^2(x, m), \quad (5.3.28)$$

es independiente de  $x$ .

### 5.3.4. Ejemplos

Antes de pasar a ver algunos ejemplos debemos destacar que el método descrito es válido para una gran cantidad de ecuaciones diferenciales. Es fácil comprobar que si tenemos la EDO

$$\frac{d}{dt} \left( p(t) \frac{dz(t)}{dt} \right) + q(t)z(t) + \lambda \rho(t)z(t) = 0,$$

entonces el cambio

$$y(x) = \sqrt[4]{p(t)\rho(t)}z(t), \quad dx = \sqrt{\frac{\rho(t)}{p(t)}} dt$$

la transforma en una ecuación del tipo (5.3.10)

$$y''(x) + r(x, m)y(x) + \lambda y(x) = 0.$$

En el razonamiento anterior se supone que  $p(x)\rho(x) \geq 0$  en el intervalo a considerar.

<sup>15</sup>Nótese que si  $r(x, m)$  es una función real, entonces  $k(x, m)$  también lo será.

**Los armónicos esféricos**

Comenzaremos por los armónicos esféricos  $Y_m^l(\theta)$ . Estos satisfacen la ecuación

$$\frac{1}{\operatorname{sen} \theta} \left( \frac{d}{d\theta} \left( \operatorname{sen} \theta \frac{dY}{d\theta} \right) \right) - \frac{m^2}{\operatorname{sen}^2 \theta} Y + \lambda Y = 0, \quad 0 \leq \theta \leq \pi.$$

Hacemos el cambio  $y(x) = \sqrt{\operatorname{sen} \theta} Y(\theta)$ ,  $x = \theta$  y obtenemos la forma canónica (5.3.10)

$$y'' - \frac{m^2 - \frac{1}{4}}{\operatorname{sen}^2 x} y + \mu y = 0, \quad \mu = \lambda + \frac{1}{4},$$

por tanto

$$k(x, m) = \left( m - \frac{1}{2} \right) \cot \theta, \quad L(m) = m(m-1) + \frac{1}{4} = \left( m - \frac{1}{2} \right)^2.$$

Como  $L(m)$  es creciente tenemos el caso I, así que el teorema 5.3.4 nos da

$$\mu = \lambda + \frac{1}{4} = L(l+1) \Rightarrow \lambda = l(l+1), \quad l = 0, 1, 2, \dots, \quad m = 0, 1, 2, \dots, l.$$

Usando (5.3.20) tenemos

$$H_-^{l+1} y_l^l(\theta) = 0 \Rightarrow \left[ \left( l + \frac{1}{2} \right) \cot \theta - \frac{d}{d\theta} \right] y_l^l(\theta) = 0 \Rightarrow$$

$$y_l^l(\theta) = C \operatorname{sen}^{l+1/2} \theta.$$

Calculamos  $C$  para que  $\|y_l^l\| = 1$ :

$$1 = C^2 \int_0^\pi \operatorname{sen}^{2l+1} \theta d\theta = C^2 \frac{2^{l+1} l!}{(2l+1)!!},$$

luego

$$y_l^l(\theta) = \sqrt{\frac{(2l+1)!!}{2^{l+1} l!}} \operatorname{sen}^{l+1/2} \theta.$$

Finalmente, como  $L(l+1) - L(m) = (l+m)(l-m+1)$  y  $L(l+1) - L(m+1) = (l-m)(l+m+1)$ , entonces (5.3.20) y (5.3.21) nos dan

$$\left[ \left( m - \frac{1}{2} \right) \cot \theta - \frac{d}{d\theta} \right] y_m^l(\theta) = \sqrt{(l-m)(l+m+1)} y_{m+1}^l(\theta),$$

$$\left[ \left( m - \frac{1}{2} \right) \cot \theta + \frac{d}{d\theta} \right] y_m^l(\theta) = \sqrt{(l+m)(l-m+1)} y_{m-1}^l(\theta),$$

respectivamente.

**Potencial de Poschl-Teller 1**

Sea

$$y'' + \left( -\frac{\alpha^2(m+g)(m+g+1)}{\sin^2 \alpha(x-x_0)} - \frac{\alpha^2(m-g)(m-g-1)}{\cos^2 \alpha(x-x_0)} \right) y + \lambda y = 0,$$

con  $x \in (x_0, x_0 + \pi/(2\alpha))$ . Entonces el teorema 5.3.6 nos da

$$\begin{aligned} k(x, m) &= 2\alpha(g + m \cos 2\alpha(x-x_0)) \csc 2\alpha(x-x_0) \\ &= (g+m)\alpha \cot \alpha(x-x_0) + (g-m)\alpha \tan \alpha(x-x_0), \end{aligned}$$

$$L(m) = 4m^2\alpha^2.$$

Al ser  $L(m)$  creciente usamos el teorema 5.3.4 y por tanto

$$\lambda = 4\alpha^2(l+1)^2, \quad l = 0, 1, 2, \dots, \quad m = 0, 1, 2, \dots, l.$$

En este caso

$$\begin{aligned} H_-^{l+1} y_l^l(x) = 0 &\Rightarrow [k(x, l+1) - \mathcal{D}] y_l^l(x) = 0 \Rightarrow \\ y_l^l(x) &= C \cos^{l-g+1} \alpha(x-x_0) \sin^{l+g+1} \alpha(x-x_0). \end{aligned}$$

Calculamos  $C$  para que  $\|y_l^l\| = 1^{16}$

$$\begin{aligned} 1 &= C^2 \int_{r_0}^{r_0 + \frac{\pi}{2\alpha}} \cos^{2(l-g+1)} \alpha(x-x_0) \sin^{2(l+g+1)} \alpha(x-x_0) dx \\ &= C^2 \frac{\Gamma(l-g+\frac{3}{2})\Gamma(l+g+\frac{3}{2})}{2\alpha\Gamma(2l+3)}. \end{aligned}$$

luego

$$y_l^l(x) = \sqrt{\frac{2\alpha\Gamma(2l+3)}{\Gamma(l-g+\frac{3}{2})\Gamma(l+g+\frac{3}{2})}} \cos^{l-g+1} \alpha(x-x_0) \sin^{l+g+1} \alpha(x-x_0).$$

Finalmente, como

$$\begin{aligned} L(l+1) - L(m+1) &= 4\alpha^2(l-m)(l+m+2), \\ L(l+1) - L(m) &= 4\alpha^2(l-m+1)(l+m+1), \end{aligned}$$

entonces (5.3.20) y (5.3.21) nos conducen a las expresiones

$$\begin{aligned} \left( k(x, m+1) - \frac{d}{dx} \right) y_m^l(x) &= 2\alpha\sqrt{(l-m)(l+m+2)} y_{m+1}^l(x), \quad 0 \leq m \leq l, \\ \left( k(x, m) + \frac{d}{dx} \right) y_m^l(x) &= 2\alpha\sqrt{(l-m+1)(l+m+1)} y_{m-1}^l(x), \quad 0 \leq m \leq l, \end{aligned}$$

respectivamente.

---

<sup>16</sup>Hemos usado que

$$\int_0^{\frac{\pi}{2}} \cos^a x \sin^b x dx = \frac{\Gamma(\frac{a+1}{2})\Gamma(\frac{b+1}{2})}{2\Gamma(\frac{a+b+2}{2})}, \quad \Re a > -1, \Re b > -1.$$

**Potencial de Poschl-Teller 2**

Sea

$$y'' + \left( -\frac{\alpha^2(m+g)(m+g+1)}{\sinh^2 \alpha(x-x_0)} + \frac{\alpha^2(m-g)(m-g+1)}{\cosh^2 \alpha(x-x_0)} \right) y + \lambda y = 0,$$

con  $x \in (r_0, \infty)$ . Entonces el teorema 5.3.6 nos da

$$\begin{aligned} k(x, m) &= 2\alpha(g+m \cosh 2\alpha(x-x_0)) \operatorname{csch} 2\alpha(x-x_0) \\ &= (g+m)\alpha \operatorname{coth} \alpha(x-x_0) + (m-g)\alpha \tanh \alpha(x-x_0) \end{aligned}$$

$$L(m) = -4m^2\alpha^2.$$

Al ser  $L(m)$  decreciente usamos el teorema (5.3.5) y por tanto

$$\lambda = -4\alpha^2 l^2, \quad l = 0, 1, 2, \dots, \quad m = l, l+1, l+2, \dots$$

En este caso

$$H_+^l y_l^l(x) = 0 \quad \Rightarrow \quad [k(x, l) + \mathcal{D}] y_l^l(x) = 0 \quad \Rightarrow$$

y por tanto<sup>17</sup> para todo  $g < \frac{1}{2} - l$

$$y_l^l(x) = \sqrt{\frac{2\alpha\Gamma(l-g+\frac{1}{2})}{\Gamma(-l-g+\frac{1}{2})\Gamma(2l)}} \cosh^{g-l} \alpha(x-x_0) \sinh^{-l-g} \alpha(x-x_0).$$

Finalmente, como

$$L(l) - L(m) = 4\alpha^2(m-l)(m+l),$$

$$L(l) - L(m+1) = 4\alpha^2(m+1-l)(l+m+1),$$

entonces (5.3.20) y (5.3.21) nos dan

$$\left( k(x, m+1) - \frac{d}{dx} \right) y_m^l(x) = 2\alpha\sqrt{(m+1-l)(l+m+1)} y_{m+1}^l(x), \quad m \geq l,$$

$$\left( k(x, m) + \frac{d}{dx} \right) y_m^l(x) = 2\alpha\sqrt{(m-l)(m+l)} y_{m-1}^l(x),$$

respectivamente.

**Potencial de Morse**

Sea ahora la ecuación

$$y'' + (ae^{-2u} + be^{-u})y + \lambda y = 0, \quad a, b \in \mathbb{C}, \quad 0 \leq u < +\infty.$$

Si hacemos el cambio  $u = x + \alpha$  de forma que  $e^{2\alpha}a = -1/4$ ,  $m = be^\alpha - \frac{1}{2}$  obtenemos

---

<sup>17</sup>Hemos usado que

$$\int_0^\infty \cosh^a x \sinh^b x dx = \frac{\Gamma(-\frac{a+b}{2})\Gamma(\frac{b+1}{2})}{2\Gamma(\frac{1-a}{2})}, \quad \Re b > -1, \quad \Re a + \Re b < 0.$$

$$y'' + \left(-\frac{1}{4}e^{-2x} + \left(m + \frac{1}{2}\right)e^{-x}\right)y + \lambda y = 0,$$

con  $x \in \mathbb{R}$ . Entonces el teorema 5.3.6 nos da

$$k(x, m) = -\frac{e^{-x}}{2} + m, \quad L(m) = -m^2.$$

Al ser  $L(m)$  decreciente usamos el teorema 5.3.5 y por tanto

$$\lambda = -l^2, \quad l = 1, 2, \dots, \quad m = l, l + 1, l + 2, \dots$$

En este caso

$$H_+^l y_l^l(x) = 0 \quad \Rightarrow \quad [k(x, l) + \mathcal{D}] y_l^l(x) = 0 \quad \Rightarrow$$

y por tanto<sup>18</sup>

$$y_l^l(x) = \sqrt{\frac{1}{\Gamma(2l)}} e^{-\exp(-x)/2 - lx}.$$

Finalmente, como

$$L(l) - L(m) = (m - l)(m + l),$$

$$L(l) - L(m + 1) = (m + 1 - l)(l + m + 1),$$

entonces usando (5.3.20) y (5.3.21) tenemos

$$\left(-\frac{e^{-x}}{2} + m + 1 - \frac{d}{dx}\right) y_m^l(x) = \sqrt{(m + 1 - l)(l + m + 1)} y_{m+1}^l(x), \quad m \geq l,$$

$$\left(-\frac{e^{-x}}{2} + m + \frac{d}{dx}\right) y_m^l(x) = \sqrt{(m - l)(m + l)} y_{m-1}^l(x), \quad m \geq l,$$

respectivamente.

### La ecuación de Whittaker & Watson

Sea

$$W'' + \left(-\frac{1}{4} + \frac{m + \frac{1}{2}}{z} + \frac{\frac{1}{4} + \lambda}{z^2}\right) W = 0, \quad 0 \leq z < +\infty.$$

Haciendo el cambio  $z = e^x$ ,  $W(z) = e^{\frac{x}{2}} y$  obtenemos

$$y'' + \left(-\frac{1}{4}e^{2x} + \left(m + \frac{1}{2}\right)e^x\right)y + \lambda y = 0,$$

que esencialmente la misma ecuación diferencial del ejemplo anterior pero cambiando  $x$  por  $-x$ . También corresponde al caso II con

$$k(x, m) = \frac{e^x}{2} - m, \quad L(m) = -m^2,$$

<sup>18</sup>Hemos usado que

$$\int_{-\infty}^{\infty} e^{-\exp(-x) - 2lx} dx = \Gamma(2l), \quad l > 0.$$

así que

$$\lambda = -l^2, \quad l = 0, 1, 2, \dots, \quad m = l, l+1, l+2, \dots,$$

$$H_+^l y_l^l(x) = 0 \quad \Rightarrow [k(x, l) + \mathcal{D}] y_l^l(x) = 0 \quad \Rightarrow y_l^l(x) = \sqrt{\frac{1}{\Gamma(2l)}} e^{-\exp(x)/2+lx}.$$

Finalmente, (5.3.20) y (5.3.21) nos dan

$$\left(\frac{e^x}{2} - m - 1 - \frac{d}{dx}\right) y_m^l(x) = \sqrt{(m+1-l)(l+m+1)} y_{m+1}^l(x), \quad m \geq l,$$

$$\left(\frac{e^x}{2} - m + \frac{d}{dx}\right) y_m^l(x) = \sqrt{(m-l)(m+l)} y_{m-1}^l(x), \quad m \geq l,$$

respectivamente.

### La ecuación de Bessel

Sea

$$y'' - \frac{m^2 - \frac{1}{4}}{z^2} y + \lambda y = 0, \quad z \in \mathbb{R}.$$

Aplicando el teorema 5.3.6 tenemos

$$k(x, m) = \frac{m - \frac{1}{2}}{x}, \quad L(m) = 0.$$

Como  $L(m)$  es constante no podemos usar ninguno de nuestros resultados por lo que no tenemos ninguna expresión para  $\lambda$ . No obstante si podemos seguir usando las expresiones (5.3.20) y (5.3.21) que en este caso dan

$$\left(\frac{m + \frac{1}{2}}{x} - \frac{d}{dx}\right) y_m(x) = \sqrt{\lambda} y_{m+1}(x),$$

$$\left(\frac{m - \frac{1}{2}}{x} + \frac{d}{dx}\right) y_m(x) = \sqrt{\lambda} y_{m-1}(x).$$

Ambas expresiones dan sendas relaciones de recurrencia para las funciones de Bessel.

Para resolver la ecuación tenemos que usar el método de Frobenius (ver e.g [6, 21]). La solución general en este caso da

$$y(x) = \sqrt{x} \left( c_1 J_m(\sqrt{\lambda}x) + c_2 Y_m(\sqrt{\lambda}x) \right),$$

donde  $J_m(z)$  y  $Y_m(z)$  son las funciones de Bessel de primera y segunda especie, respectivamente.

### El átomo de hidrógeno

Sea la ecuación radial de Schrödinger para el átomo de hidrógeno

$$R'' + \frac{2}{r}R' + \frac{2}{r}R - \frac{m(m+1)}{r^2}R + \lambda R = 0, \quad r \geq 0.$$

Si hacemos el cambio  $y(r) = rR(r)$ , tenemos

$$y'' + \left( \frac{2}{r} - \frac{m(m+1)}{r^2} \right) y + \lambda y = 0, \quad r \geq 0.$$

Aplicando el teorema 5.3.6 obtenemos

$$k(r, m) = \frac{m}{r} - \frac{1}{m}, \quad L(m) = -\frac{1}{m^2}.$$

Como  $L(m)$  crece tenemos el caso I. Así,

$$\lambda = -\frac{1}{(l+1)^2}, \quad l = 0, 1, 2, \dots, \quad m = 1, 2, \dots, l.$$

En este caso

$$\begin{aligned} H_-^{l+1} y_l^l(r) = 0 &\Rightarrow \left[ \frac{l+1}{r} - \frac{1}{l+1} - \mathcal{D} \right] y_l^l(r) = 0 \Rightarrow \\ &y_l^l(x) = Cr^{l+1} e^{-\frac{r}{l+1}}. \end{aligned}$$

Calculamos  $C$  para que  $\|y_l^l\| = 1$ <sup>19</sup>

$$y_l^l(r) = \sqrt{\frac{1}{(l+1)^{l+3} \Gamma(l+1)}} r^{l+1} e^{-\frac{r}{l+1}}.$$

Finalmente, como

$$\begin{aligned} L(l+1) - L(m+1) &= \frac{1}{(m+1)^2} - \frac{1}{(l+1)^2}, \\ L(l+1) - L(m) &= \frac{1}{m^2} - \frac{1}{(l+1)^2}, \end{aligned}$$

entonces de (5.3.20) y (5.3.21) obtenemos

$$\left( \frac{m+1}{r} - \frac{1}{m+1} - \frac{d}{dr} \right) y_m^l(r) = \sqrt{\frac{1}{(m+1)^2} - \frac{1}{(l+1)^2}} y_{m+1}^l(r), \quad 1 \leq m \leq l,$$

$$\left( \frac{m}{r} - \frac{1}{m} + \frac{d}{dr} \right) y_m^l(r) = \sqrt{\frac{1}{m^2} - \frac{1}{(l+1)^2}} y_{m-1}^l(r), \quad 1 \leq m \leq l,$$

respectivamente.

---

<sup>19</sup>Hemos usado que

$$\int_0^\infty x^\beta e^{-\alpha x} dx = \Gamma(\beta+1) \alpha^{-1-\beta}, \quad \Re \alpha > 0, \quad \Re \beta > -1.$$

## 5.4. Factorización de la EDO hipergeométrica

En este apartado vamos a probar que la ecuación hipergeométrica (5.1.7) admite una factorización tipo Infeld-Hull. Vamos a seguir el trabajo de Lorente [14]

### 5.4.1. El *hamiltoniano* y los operadores *escalera*

Sean las funciones

$$\varphi_n(s) = \sqrt{\rho(s)/d_n^2} P_n(x), \quad (5.4.1)$$

Usando (5.1.22) y (5.1.7), obtenemos las expresiones

$$\alpha_n \frac{d_{n+1}}{d_n} \varphi_{n+1}(s) + \gamma_n \frac{d_{n-1}}{d_n} \varphi_{n-1}(s) + (\beta_n - x(s)) \varphi_n(s) = 0, \quad (5.4.2)$$

$$\sigma(x) \varphi_n''(x) + \sigma'(x) \varphi_n'(x) + \nu(x) \varphi_n(x) + \lambda_n \varphi_n(x) = 0, \quad (5.4.3)$$

donde

$$\nu(x) = -\frac{1}{4} \frac{(\tau(x) - \sigma'(x))^2}{\sigma(x)} - \frac{1}{2} (\tau' - \sigma'').$$

En particular la EDO anterior induce a definir el siguiente operador  $\mathcal{H}(x, n)$  orden 2

$$\mathcal{H}(x, n) := \sigma(x) \mathcal{D}^2 + \sigma'(x) \mathcal{D} + [\nu(x) + \lambda_n] \mathbf{I}, \quad \mathcal{H}(x, n) \varphi_n(x) = 0. \quad (5.4.4)$$

En adelante diremos que  $\mathcal{H}(x, n)$  es el hamiltoniano asociado a la EDO (5.1.7).

Si ahora usamos las expresiones (5.1.34) y (5.1.35) obtenemos

$$L^+(x, n) \varphi_n(s) := [f(x, n) - \sigma(x) \mathcal{D}] \varphi_n(x) = \alpha_n \frac{\lambda_{2n}}{2n} \frac{d_{n+1}}{d_n} \varphi_{n+1}(x), \quad (5.4.5)$$

donde

$$L^+(x, n) = \underbrace{\left( \frac{\lambda_n \tau_n(x)}{n \tau'_n} + \frac{1}{2} (\tau(x) - \sigma'(x)) \right)}_{f(x, n)} \mathbf{I} - \sigma(x) \mathcal{D}, \quad (5.4.6)$$

y

$$L^-(x, n) \varphi_n(x) := [g(x, n) + \sigma(x) \mathcal{D}] \varphi_n(x) = \gamma_n \frac{\lambda_{2n}}{2n} \frac{d_{n-1}}{d_n} \varphi_{n-1}(x), \quad (5.4.7)$$

donde

$$L^-(x, n) = \underbrace{\left( -\frac{\lambda_n \tau_n(x)}{n \tau'_n} + \frac{\lambda_{2n}}{2n} (x - \beta_n) - \frac{1}{2} (\tau(x) - \sigma'(x)) \right)}_{g(x, n)} \mathbf{I} + \sigma(x) \mathcal{D}. \quad (5.4.8)$$

Los operadores anteriores  $L^+(x, n)$  y  $L^-(x, n)$  constituyen los operadores “escalera” de ascenso (raising) y descenso (lowering), respectivamente.

Si usamos las expresiones explícitas de  $\tau_n$  y  $\lambda_n$  (ver apartado 5.1.2)

$$\lambda_n = -n \left( \tau' + (n-1) \frac{\sigma''}{2} \right), \quad \tau_n(x) = \tau(x) + n\sigma'(x),$$

y

$$\beta_n = \frac{n\tau_{n-1}(0)}{\tau'_{n-1}} - \frac{(n+1)\tau_n(0)}{\tau'_n},$$

se comprueba que

$$f(x, n-1) = g(x, n). \quad (5.4.9)$$

Además, para todos  $n, k \in \mathbb{N} \cup \{0\}$ ,

$$\int_a^b \varphi_n(x)[L^+(x, n)\varphi_k(x)] dx = \int_a^b [L^-(x, n)\varphi_n(x)]\varphi_k(x) dx,$$

en particular

$$\begin{aligned} \int_a^b \varphi_{n+1}(x)[L^+(x, n)\varphi_n(x)] dx &= \alpha_n \frac{\lambda_{2n}}{2n} \frac{d_{n-1}}{d_n}, \\ \int_a^b [L^-(x, n+1)\varphi_{n+1}(x)]\varphi_n(x) dx &= \gamma_{n+1} \frac{\lambda_{2n+2}}{2n+2} \frac{d_n}{d_{n+1}}. \end{aligned}$$

De lo anterior se sigue que  $L^+$  y  $L^-$  son el adjunto el uno del otro respecto al producto escalar<sup>20</sup>

$$\langle \Psi, \Phi \rangle = \int_a^b \Psi(x)\overline{\Phi(x)} dx.$$

Nótese que el operador  $\mathcal{H}(x, n)$  es autoadjunto.

### 5.4.2. Factorización de $\mathcal{H}(x, n)$

Si ahora calculamos

$$\begin{aligned} L^-(x, n+1)L^+(x, n) &= g(x, n+1)f(x, n)\mathbf{I} + \sigma(x) \underbrace{[f(x, n) - g(x, n+1)]}_{=0} \mathcal{D} \\ &\quad + \sigma(x)[f'(x, n) - \sigma'(x)\mathcal{D} - \sigma(x)\mathcal{D}^2] \end{aligned}$$

o, equivalentemente,

$$\begin{aligned} L^-(x, n+1)L^+(x, n) &= [g(x, n+1)f(x, n) + \sigma(x)(f'(x, n) + \nu(x) + \lambda_n)]\mathbf{I} \\ &\quad - \sigma(x)\mathcal{H}(x, n). \end{aligned}$$

Ahora bien,

$$\begin{aligned} &g(x, n+1)f(x, n) + \sigma(x)(f'(x, n) + \nu(x) + \lambda_n) \\ &= - \frac{(n+1)(\tau' + \tau'_{n-1})(\tau(0)^2\sigma'' - 2\tau(0)\sigma'(0)\tau' + 2\sigma(0)\tau_n'^2 - n\sigma'^2(0)(\tau' + \tau'_n))}{4\tau_n'^2} \end{aligned}$$

no depende de  $x$ . Dicha cantidad la denotaremos por  $\mu_n$ . Además, si aplicamos el operador

$$L^-(x, n+1)L^+(x, n) = \mu_n\mathbf{I} - \sigma(x)\mathcal{H}(x, n)$$

<sup>20</sup>Estamos usando que las funciones  $(\varphi_n)_n$  son una base ortonormal completa del espacio de las funciones de cuadrado integrable y que  $\varphi_n(\pm a) = 0$ .

a  $\varphi_n(x)$  y usamos (5.4.5), (5.4.7) y (5.4.4) obtenemos la expresión

$$\mu_n = \frac{\lambda_{2n}}{2n} \frac{\lambda_{2n+2}}{2n+2} \alpha_n \gamma_{n+1}.$$

Análogamente se tiene

$$\begin{aligned} L^+(x, n-1)L^-(x, n) = & g(x, n)f(x, n-1)\mathbf{I} + \sigma(x) \underbrace{[f(x, n-1) - g(x, n)]}_{=0} \mathcal{D} \\ & - \sigma(x)[g'(x, n) + \sigma'(x)\mathcal{D} + \sigma(x)\mathcal{D}^2] \end{aligned}$$

o, equivalentemente,

$$\begin{aligned} L^+(x, n-1)L^-(x, n) = & [g(x, n)f(x, n-1) - \sigma(x)(g'(x, n) - \nu(x) - \lambda_n)]\mathbf{I} \\ & - \sigma(x)\mathcal{H}(x, n). \end{aligned}$$

Ahora bien, como antes,

$$\begin{aligned} & g(x, n)f(x, n-1) - \sigma(x)(g'(x, n) - \nu(x) - \lambda_n) \\ = & - \frac{n(\tau' + \tau'_{n-2})(\tau(0)^2\sigma'' - 2\tau(0)\sigma'(0)\tau' + 2\sigma(0)\tau'_{n-1}{}^2 - (n-1)\sigma'^2(0)(\tau' + \tau'_{n-1}))}{4\tau'_{n-1}{}^2} \end{aligned}$$

no depende de  $x$ . Nótese que la expresión anterior coincide con  $\mu_{n-1}$ . Otra forma de comprobarlo es aplicar el operador

$$L^+(x, n-1)L^-(x, n) = M(n)\mathbf{I} - \sigma(x)\mathcal{H}(x, n)$$

a  $\varphi_n(x)$  y usar, como antes (5.4.5), (5.4.7) y (5.4.4). Lo anterior nos conduce a la expresión

$$M(n) = \frac{\lambda_{2n}}{2n} \frac{\lambda_{2n-2}}{2n-2} \alpha_{n-1} \gamma_n = \mu_{n-1}.$$

Así, hemos probado el siguiente

**Teorema 5.4.1** *Para la ecuación hipergeométrica (5.1.7), haciendo el cambio (5.4.1), se obtiene la ecuación (5.4.4), para la cual se cumple la siguiente factorización del tipo Infeld y Hull*

$$\begin{aligned} L^-(x, n+1)L^+(x, n) &= \mu_n \mathbf{I} - \sigma(x)\mathcal{H}(x, n), \\ L^+(x, n-1)L^-(x, n) &= \mu_{n-1} \mathbf{I} - \sigma(x)\mathcal{H}(x, n), \end{aligned} \tag{5.4.10}$$

donde los operadores  $L^+(x, n)$  y  $L^-(x, n)$  son los operadores de ascenso (raising) y descenso (lowering), respectivamente definidos por

$$\begin{aligned} L^+(x, n) &= f(x, n)\mathbf{I} - \sigma(x)\mathcal{D}, \quad L^-(x, n) = f(x, n-1)\mathbf{I} + \sigma(x)\mathcal{D}, \\ f(x, n) &= \frac{\lambda_n}{n} \frac{\tau_n(x)}{\tau'_n} + \frac{1}{2}(\tau(x) - \sigma'(x)) \end{aligned}$$

Además,

$$L_+\varphi_n(x) = -\tau'_n \sqrt{\alpha_n \gamma_{n+1}} \varphi_{n+1}(x), \quad L_-\varphi_n(x) = -\tau'_n \sqrt{\alpha_{n-1} \gamma_n} \varphi_{n-1}(x).$$

y, por tanto,

$$L^-(x, 0)\varphi_0(x) = 0 \quad \Rightarrow \quad \varphi_0(x) = N_0 \exp\left(-\int f(x, n-1)dx\right), \quad \|\varphi_0\|^2 = 1,$$

$$\varphi_n(x) = \frac{1}{\prod_{k=0}^{n-1} (-\tau'_k \alpha_k \gamma_{k+1})} L^+(x, n-1) \cdots L^+(x, 1) L^+(x, 0) \varphi_0(x), \quad n \geq 1.$$

### 5.4.3. Ejemplos

Como ejemplo consideraremos el caso correspondiente a los polinomios de Hermite y Laguerre.

#### Caso Hermite

En el primer caso como  $\sigma(x) = 1$  y  $\tau(x) = -2x$  tenemos<sup>21</sup>

$$\varphi_n(x) = \frac{2^{n/2}}{\sqrt{n!}\sqrt{\pi}} e^{-x^2/2} H_n(x),$$

$$\mathcal{H}(x, n) = \mathcal{D}^2 + (1 - x^2 + \lambda_n)\mathbf{I}, \quad \lambda_n = 2n,$$

entonces  $f(x, n) = g(x, n) = x$ , luego

$$L^+(x, n)\varphi_n(x) = (x\mathbf{I} - \mathcal{D})\varphi_n(x) = \sqrt{2(n+1)}\varphi_{n+1}(x),$$

$$L^-(x, n)\varphi_n(x) = (x\mathbf{I} + \mathcal{D})\varphi_n(x) = \sqrt{2n}\varphi_{n-1}(x).$$

De lo anterior se sigue

$$\varphi_0(x) = \frac{1}{\sqrt[4]{\pi}} e^{-x^2/2}, \quad \varphi_n(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^n n!}} (x\mathbf{I} - \mathcal{D})^n e^{-x^2/2}.$$

Finalmente

$$L^+(x, n-1)L^-(x, n) = 2n\mathbf{I} - \mathcal{H}(x, n), \quad L^-(x, n+1)L^+(x, n) = (2n+2)\mathbf{I} - \mathcal{H}(x, n).$$

#### Caso Laguerre

En este caso  $\sigma(x) = x$  y  $\tau(x) = -x + \alpha + 1$  tenemos

$$\mathcal{H}(x, n) = x\mathcal{D}^2 + \mathcal{D} + \left( \frac{1}{2} - \frac{(x-\alpha)^2}{4x} \lambda_n \right) \mathbf{I}, \quad \lambda_n = n,$$

y<sup>22</sup>

$$\varphi_n(x) = \frac{1}{\sqrt{n!\Gamma(\alpha+n+1)}} e^{-x/2} x^{\alpha/2} L_n^\alpha(x)$$

Entonces

$$f(x, n) = \frac{+x - \alpha - 2 - 2n}{2}, \quad g(x, n) = \frac{x - \alpha - 2n}{2} \Rightarrow$$

$$L^+(x, n)\varphi_n(x) = \left( \frac{x - \alpha - 2 - 2n}{2} \mathbf{I} - x\mathcal{D} \right) \varphi_n(x) = \sqrt{(n+1)(n+\alpha+1)} \varphi_{n+1}(x),$$

$$L^-(x, n)\varphi_n(x) = \left( \frac{x - \alpha - 2n}{2} \mathbf{I} + x\mathcal{D} \right) \varphi_n(x) = \sqrt{n(n+\alpha)} \varphi_{n-1}(x).$$

<sup>21</sup>Estamos usando polinomios mónicos de Hermite.

<sup>22</sup>Estamos usando polinomios mónicos de Laguerre.

De lo anterior se sigue

$$\varphi_0(x) = \frac{1}{\sqrt{\Gamma(\alpha + 1)}} e^{-x/2} x^{\alpha/2},$$

$$\varphi_n(x) = \frac{1}{\sqrt{n! \Gamma(\alpha + n + 1)}} L^+(x, n-1) \cdots L^+(x, 1) L^+(x, 0) e^{-x/2} x^{\alpha/2}.$$

Finalmente

$$L^+(x, n-1) L^-(x, n) = (n)(n + \alpha) I - \mathcal{H}(x, n),$$

$$L^-(x, n+1) L^+(x, n) = (n+1)(n+1 + \alpha) I - \mathcal{H}(x, n).$$

## 5.5. Problemas

**Problema 5.5.1** Prueba la relación de ortogonalidad (5.1.20) y calcula una expresión para la norma de las derivadas  $P_n^{(k)}$ .

**Problema 5.5.2** Calcula todas las características de los polinomios clásicos que aparecen en la tabla 5.2.

**Problema 5.5.3** Prueba que los polinomios ortogonales mónicos  $P_n$ , soluciones de la ecuación (5.1.7), satisfacen la siguiente relación de estructura

$$P_n(x) = Q_n + \delta_n Q_{n-1} + \epsilon_n Q_{n-2}, \quad (5.5.1)$$

donde  $Q_n(x) \equiv P'_{n+1}(x)/(n+1)$ . Encuentra una expresión de los coeficientes en función de los coeficientes calculados en el apartado 5.1.2.

**Problema 5.5.4** Probar que para los polinomios clásicos se tienen las siguientes identidades.

$$\sum_{n=0}^{\infty} \frac{2^n}{n!} H_n(x) t^n = e^{2xt-t^2}, \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} L_n^\alpha(x) t^n = \frac{e^{-\frac{tx}{1-t}}}{(1-t)^{\alpha+1}}, \quad (5.5.2)$$

$$\sum_{n=0}^{\infty} \frac{(-1)^n (\alpha + \beta + 1)_n}{n!} P_n^{\alpha, \beta}(x) t^n = \frac{2^{\alpha+\beta}}{R(1-2t+R)^\alpha (1+2t+R)^\beta}, \quad (5.5.3)$$

con  $R = \sqrt{1 + 4t(t+x)}$ .

**Problema 5.5.5** Sea la ecuación de Schrödinger estacionaria unidimensional

$$-\frac{\hbar^2}{2m} \Psi'' - \frac{U_0}{\cosh^2(\alpha x)} \Psi = E \Psi, \quad x \in \mathbb{R},$$

con  $E < 0$ . Hacer el cambio  $s = \tanh x$  y obtener la EDO

$$\Psi'' - \frac{2s}{1-s^2} \Psi' + \frac{\gamma^2(1-s^2) - \beta^2}{(1-s^2)^2}, \quad \beta^2 = -\frac{2mE}{\hbar^2 \alpha^2}, \quad \gamma^2 = \frac{2mU_0}{\hbar^2 \alpha^2}, \quad \gamma, \beta > 0.$$

Resuelve la EDO anterior y calcula las soluciones estacionarias de la ecuación. El potencial anterior se conoce como potencial de Posch-Teller y modeliza la interacción molecular.

**Problema 5.5.6** Resolver la ecuación de Schrödinger tridimensional

$$-\frac{\hbar^2}{2m}\Delta\Psi + D(e^{-2\alpha x} - 2e^{-\alpha x})\Psi = E\Psi, \quad x = \frac{r - r_0}{r_0}.$$

Pasa a unidades adimensionales y haz el cambio  $y = \frac{2\gamma}{\alpha}e^{-\alpha x}$  donde  $\gamma^2 = \frac{2mDr_0^2}{\hbar^2}$ . Resuelve primero el caso  $l = 0$ . El potencial anterior se conoce como potencial de Morse y también modeliza la interacción molecular.

**Problema 5.5.7** Para el átomo de hidrógeno calcula el valor medio de la posición del electrón en el estado definido por los valores  $n$ ,  $l$  y  $m$   $\langle r \rangle_{n,l,m} = \langle R_{nl} | r | R_{nl} \rangle$ . Prueba que si  $n = l = 0$ , entonces  $\langle r \rangle_{0,0,0} = a_0 = \hbar^2 / (2\alpha)$ . Prueba además que en el estado definido por los valores  $n$ ,  $l$  y  $m$  la media de la energía potencial

$$\langle V(r) \rangle_{n,l,m} = \langle R_{nl} | V(r) | R_{nl} \rangle$$

es la mitad de la energía total  $E_{n,l}$ .

Calcula  $\langle r \rangle_{n,l,m}$  y  $\langle V(r) \rangle_{n,l,m}$  para el oscilador armónico tridimensional.

**Problema 5.5.8** La ecuación de Klein-Gordon

$$\Delta\psi_{KG} + \left[ \left( E_{KG} + \frac{\mu}{r} \right)^2 - 1 \right] \psi_{KG} = 0,$$

se puede resolver usando el método de separación de variables obteniéndose para sus soluciones la expresión  $\psi_{KG}(r, \theta, \phi) = R(r)r^{-1}Y_{l,m}(\theta, \phi)$ , siendo  $Y_{l,m}$  los armónicos esféricos y  $R(r)$  las soluciones de

$$R'' + \left[ \left( E + \frac{\mu}{r} \right)^2 - 1 - \frac{l(l+1)}{r^2} \right] R = 0.$$

Demuestra usando el método descrito al final de capítulo que las soluciones se expresan mediante la fórmula

$$R(x) = N_{l,n} e^{-\frac{x}{2}} r^{\nu+1} L_n^{2\nu+1}(2ar), \quad \nu = -\frac{1}{2} + \sqrt{\left( l + \frac{1}{2} \right)^2 - \mu^2},$$

donde  $L_n^\alpha$  son los polinomios de Laguerre y determina el factor de normalización  $N_{l,n}$ .

**Problema 5.5.9** Aplica el método de factorización de Schrödinger para resolver las siguientes EDOs

1. Los esféricos armónicos generalizados  $\Phi$ , definidos por la expresión  $y(\theta) = \sin^\gamma \theta \Phi(\theta)$  donde  $y$  satisface la EDO

$$y'' - \frac{(m+\gamma)(m+\gamma-1)}{\sin^2 \theta} y + (\lambda + \gamma^2) y = 0.$$

2. La ecuación de onda para una "peonza" simétrica. La función de onda  $\Psi(\theta, \phi, \psi) = \Theta(\theta)e^{iK\phi}e^{iM\psi}$ , donde la función  $\Theta$  es tal que  $y(\theta) = \sin^{\frac{1}{2}} \theta \Theta(\theta)$ , e  $y$  satisface la EDO

$$y'' - \frac{(M - \frac{1}{2})(M + \frac{1}{2}) + K^2 - 2MK \cos \theta}{\sin^2 \theta} y + \left( \lambda + K^2 + \frac{1}{4} \right) y = 0.$$

3. La parte radial de la ecuación de Schrödinger para  $s$  osciladores de Plank unidimensionales

$$\Psi'' + \frac{s-1}{r}\Psi' - \left[ \frac{n(n+s-2)}{r^2} + r^2 \right] \Psi + \lambda\Psi = 0, \quad n = 0, 1, 2, \dots,$$

donde  $r^2 = \sum_{k=1}^s x_k^2$ . Si hacemos el cambio  $\Psi(r) = r^{(1-s)/2}y(r)$ ,  $y$  satisface la EDO

$$y'' - \left[ \frac{(n + \frac{s}{2} - \frac{1}{2})(n + \frac{s}{2} - \frac{3}{2})}{r^2} + r^2 \right] y + \lambda y = 0.$$

4. El problema de Kepler generalizado. En este caso  $R$  satisface la EDO

$$R'' + \frac{2}{r}R' + \frac{2}{r}R - \frac{(l+\gamma)(l+\gamma+1)}{r^2}y + \lambda y = 0, \quad r \geq 0, \quad \lambda = -\frac{1}{(n+\gamma)^2}.$$

**Problema 5.5.10** Usando el método de factorización del apartado 5.4 obtén las soluciones para el caso de los polinomios de Legendre, Gegenbauer y Jacobi.



# Bibliografía

- [1] R. Álvarez-Nodarse, *Polinomios hipergométricos y  $q$ -polinomios*. Monografías del Seminario Matemático “García de Galdeano” Vol. **26**. Prensas Universitarias de Zaragoza, Zaragoza, 2003.
- [2] S. Albeverio (with Appendix written by P. Exner) *Solvable Models In Quantum Mechanics*. AMS Chelsea Publishing, 2004.
- [3] G. Auletta, *Foundations and interpretation of quantum mechanics. In the light of a critical-historical analysis of the problems and of a synthesis of the results*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.
- [4] J-L. Basdevant, J. Dalibard, *Quantum mechanics*. Springer Verlag, 2002.
- [5] J.S. Bell, *Lo decible y lo indecible en mecánica cuántica*. Alianza Universidad, 1991.
- [6] M. Braun, *Differential Ecuations and their Applications*. Springer Verlag, 1993.
- [7] L. Debnath y P. Mikusinsk. Introduction to Hilbert spaces with applications, Eselvier Academic Press, 2005.
- [8] S. Deligeorges (Ed.), *El mundo cuántico*, Alianza Universidad, 1985.
- [9] П. В. Елютин, В. Д. Кривченко *Квантовая Механика*, Наука 1975.
- [10] R. Feynman, R. B. Leighton y M. Sands, *The Feynman Lectures on Physics*. Vol. III. *Mecánica Cuántica*. Fondo Educativo Interamericano 1971.
- [11] R. Feynman, *El carácter de la Ley física*. Metatemáticas 65, 2000.
- [12] J. Gribbin, *En busca del gato de Schrödinger. La fascinante historia de la Mecánica Cuántica*. Biblioteca Científica Salvat, 1994.
- [13] L. Infeld and T. E. Hull, The factorization method. *Rev. Modern Physics* **23** (1951), 21–68.
- [14] M. Lorente, Raising and lowering operators, factorization method and differential/difference operators of hypergeometric type. *J. Phys. A: Math. Gen.* **34** (2001), 569–588.
- [15] A. Messiah, *Mecánica Cuántica*. Vol. I y II. Ed. Tecnos. (Hay una edición inglesa reciente *Quantum Mechanics*, Dover, New York, 1999)

- [16] A. F. Nikiforov y V. B. Uvarov, *Special Functions of Mathematical Physics*. Birkhäuser Verlag, Basilea, 1988.
- [17] E. Prugovecki, *Quantum Mechanics in Hilbert space*. Academic Press, New York, 1971.
- [18] H. Rechenberg, Quanta and Quantum Mechanics, *Twentieth century physics*. Vol. I, 143–248, Inst. Phys., Bristol, 1995.
- [19] I. V. Savéliev, *Curso de física general*. Mir, Moscú, 1989.
- [20] J. Schwinger, *Quantum Kinematic and Dynamics*, Frontiers in Physics. W.A. Benjamin Inc. Publishers, 1970.
- [21] G. F. Simmons, *Ecuaciones diferenciales: con aplicaciones y notas históricas*. McGraw-Hill, 1993.
- [22] A. A. Sokolov, I. M. Ternov, V. Ch. Zhukovskii, *Quantum Mechanics*. Mir, Moscú, 1994.
- [23] Steven Weinberg, *Lectures on Quantum Mechanics*. Cambridge University Press, 2012.
- [24] Una historia interactiva de la Mecánica Cuántica en la WWW:  
[http://wayback.archive-it.org/3635/20130720124405/  
http://library.thinkquest.org/C005775/frameset.html](http://wayback.archive-it.org/3635/20130720124405/http://library.thinkquest.org/C005775/frameset.html)
- [25] Cuántica sin fórmulas en la WWW: <http://eltamiz.com/cuantica-sin-formulas/>

# Anexo A: Breve introducción al análisis funcional

## A.1. Introducción: Espacios métricos y espacios normados

**Definición A.1.1** Un espacio métrico es un par  $(\mathbb{X}, \vec{r})$  donde  $\mathbb{X}$  es un conjunto y  $\vec{r} := \vec{r}(x, y)$  es una función real (univaluada) no negativa definida para todos  $x, y, z \in \mathbb{X}$  tal que

1.  $\vec{r}(x, y) = 0 \iff x = y$ ,
2.  $\vec{r}(x, y) = \vec{r}(y, x)$ ,
3.  $\vec{r}(x, z) \leq \vec{r}(x, y) + \vec{r}(y, z)$ .

Normalmente se dice que  $\mathbb{X}$  es el espacio y  $\vec{r}$  su métrica.

Por ejemplo, el espacio  $\mathbb{X} = \mathbb{R}^n$  de las  $n$ -tuplas  $x = (x_1, x_2, \dots, x_n)$  con la métrica

$$\vec{r}(x, y) = \left( \sum_{k=1}^n |x_k - y_k|^p \right)^{1/p}, \quad p \geq 1,$$

es un espacio métrico. Otro ejemplo de especial importancia es el espacio de todas las sucesiones reales (o complejas)  $x = (x_1, x_2, \dots, x_n, \dots)$  tales que  $\sum_{k=1}^{\infty} |x_k|^p < +\infty$  con la métrica

$$\vec{r}(x, y) = \left( \sum_{k=1}^{\infty} |x_k - y_k|^p \right)^{1/p}, \quad p \geq 1.$$

Dicho espacio lo denotaremos por  $l^p$ . Un caso particular de estos espacios corresponde al caso  $p = 2$  que conduce al espacio métrico  $l^2$  de Hilbert.

Como último ejemplo consideraremos el espacio  $\mathbb{X} = C_{[a,b]}$ , es decir, el espacio de las funciones continuas definidas sobre el segmento  $[a, b]$ . Definamos en  $\mathbb{X}$  la métrica

$$\rho(f, g) = \left( \int_a^b |f(x) - g(x)|^p \right)^{1/p}, \quad p \geq 1.$$

El par obtenido  $C_p([a, b])$  es un espacio métrico. Como caso particular (de especial relevancia) tenemos el caso  $p = 2$ , i.e.,  $\mathbb{X} = C_{[a,b]}$  y  $\vec{r}$  es la función

$$\rho(f, g) = \sqrt{\int_a^b |f(x) - g(x)|^2}.$$

El espacio  $L^2(a, b)$  de las (clase de equivalencia de las) funciones de cuadrado integrable en  $[a, b]$  que vimos en el capítulo 3 es una generalización de este espacio.

**Definición A.1.2** Sea  $\mathbb{X}$  un espacio métrico,  $x_0 \in \mathbb{X}$  y  $r > 0$ . Definiremos la bola abierta  $B(x_0, r)$  al conjunto

$$B(x_0, r) = \{x \in \mathbb{X}; \vec{r}(x_0, x) < r\},$$

bola o esfera cerrada  $S(x_0, r)$  al conjunto

$$S(x_0, r) = \{x \in \mathbb{X}; \vec{r}(x_0, x) \leq r\}.$$

**Definición A.1.3** Se dice que el conjunto  $M \subset \mathbb{X}$  es abierto en  $\mathbb{X}$  si todos sus puntos (elementos) se pueden encerrar en una bola abierta contenida completamente en  $M$ . Un conjunto  $M \subset \mathbb{X}$  es cerrado en  $\mathbb{X}$  si es su complementario en  $\mathbb{X}$ ,  $\mathbb{X} \setminus M$  es abierto.

Las bolas abiertas  $B(x_0, \varepsilon)$  se suelen denominar  $\varepsilon$ -vecindades (o entornos) de  $x_0$ . Es evidente que toda  $\varepsilon$ -vecindad de  $x_0$  contiene al propio  $x_0$ .

**Definición A.1.4** Un punto  $x_0$  se denomina punto interior del conjunto  $M \subset \mathbb{X}$  si existe un  $\varepsilon > 0$  tal que  $B(x_0, \varepsilon) \subset M$ .

De lo anterior se deduce que el conjunto  $M \subset \mathbb{X}$  es abierto si y sólo si todos sus puntos son interiores.

**Definición A.1.5** Por aplicación (operador) o función entenderemos una regla  $T$  que le hace corresponder a cada elemento del subconjunto  $\mathcal{D}(T) \subset \mathbb{X}$  un único elemento del espacio métrico  $\mathbb{Y}$ . Así,  $T : \mathbb{X} \mapsto \mathbb{Y}$ ,  $y = Tx$  o  $y = T(x)$ , donde  $x \in \mathcal{D}(T) \subset \mathbb{X}$  e  $y \in \mathbb{Y}$ . Al conjunto  $\mathcal{D}(T) \subset \mathbb{X}$  se le denomina dominio de la aplicación. Si a cada  $x \in \mathcal{D}(T)$  le corresponde un valor  $y = Tx \in \mathbb{Y}$  diremos que  $Tx$  es la imagen de  $x$  según  $T$ . Al conjunto de todas las imágenes  $Tx$  le denominaremos imagen de  $T$  y le denotaremos por  $I(T)$ .

**Definición A.1.6** Una aplicación  $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$  se llama sobreyectiva si todo elemento  $y$  de  $\mathbb{Y}$  es imagen de algún elemento  $x$  del dominio. Una función se llama inyectiva si todo elemento  $y$  de la imagen de  $T$  es imagen a lo sumo de uno y sólo un elemento  $x$  del dominio. Una aplicación inyectiva y sobreyectiva se denomina biyectiva.

Es decir una aplicación es sobreyectiva si, para todo  $y \in \mathbb{Y}$ , la ecuación  $Tx = y$  tiene al menos una solución, e inyectiva si la ecuación anterior tiene o bien una única solución, o bien no tiene solución. Así mismo,  $T$  es biyectiva si para todo  $y \in \mathbb{Y}$ , la ecuación  $Tx = y$  tiene una y sólo una solución.

Para las funciones inyectivas se puede definir la aplicación inversa.

**Definición A.1.7** Sea  $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$  una aplicación inyectiva. Definiremos su aplicación inversa  $T^{-1}$  a la aplicación  $T^{-1} : I(T) \subset \mathbb{Y} \mapsto \mathcal{D}(T) \subset \mathbb{X}$  tal que a cada elemento  $y \in I(T)$  le hace corresponder un único  $x \in \mathcal{D}(T)$  tal que  $Tx = y$ .

**Definición A.1.8** La restricción de una aplicación  $T : \mathcal{D}(T) \subset \mathbb{X} \longrightarrow \mathbb{Y}$  a un subconjunto  $B \subset \mathcal{D}(T)$  es la aplicación  $T|_B$  que se obtiene de  $T$  cuando  $x$  se restringe al conjunto  $B \subset \mathcal{D}(T)$ .

**Definición A.1.9** *La extensión de una aplicación  $T : \mathcal{D}(T) \subset \mathbb{X} \rightarrow \mathbb{Y}$  a un subconjunto  $C \supset \mathcal{D}(T)$  es la aplicación  $\tilde{T}$  tal que  $\tilde{T}|_{\mathcal{D}(T)} = T$ , i.e.,  $\tilde{T}x = Tx$  para todo  $x \in \mathcal{D}(T)$ .*

**Definición A.1.10** *Una aplicación  $T : \mathcal{D}(T) \subset \mathbb{X} \rightarrow \mathbb{Y}$  es continua en  $x_0 \in \mathcal{D}(T)$  si para todo  $\varepsilon > 0$ , existe un  $\delta > 0$  tal que  $\forall x \in \mathcal{D}(T)$  con  $\tilde{r}(x, x_0) < \delta$  es tal que  $\sigma(Tx, Tx_0) < \varepsilon$ . Se dice que  $T$  es continua en todo  $M \subset \mathcal{D}(T)$  si  $T$  es continua en todo  $x \in M$ .*

**Definición A.1.11** *La imagen inversa de  $y \in \mathbb{Y}$  es el conjunto de todas las  $x \in \mathcal{D}(T)$  tales que  $Tx = y$ . La imagen inversa de un subconjunto  $N \subset \mathbb{Y}$  es el conjunto de todas las  $x \in \mathcal{D}(T)$  tales que  $Tx = y$  para todos  $y \in N$ .*

La imagen inversa de un elemento  $y \in \mathbb{Y}$  puede ser el conjunto vacío, un único punto (elemento) de  $\mathcal{D}(T)$  o un subconjunto  $M \subset \mathcal{D}(T)$ .

**Proposición A.1.12** *Una aplicación  $T : \mathcal{D}(T) \subset \mathbb{X} \rightarrow \mathbb{Y}$  es continua si y sólo si la imagen inversa de cualquier subconjunto abierto (cerrado) de  $\mathbb{Y}$  es un subconjunto abierto (cerrado) de  $\mathbb{X}$ .*

Demostración:  $\Rightarrow$  Sea  $T$  continua y sea  $S \subset \mathbb{Y}$  un abierto. Sea  $S_0$  la imagen inversa de  $S$ . Si  $S_0 = \emptyset$  la proposición es trivial así que asumiremos que  $S_0 \neq \emptyset$ . Sea  $x_0 \in S_0$  cualquiera y sea  $y_0 = Tx_0 \in S$  su imagen. Como  $S$  es abierto existe una bola  $B(y_0, \varepsilon) \subset S$ . Pero  $T$  es continua así que existe una bola  $B(x_0, \delta)$  tal que  $T(B(x_0, \delta)) \subset B(y_0, \varepsilon)$ . Pero como  $B(y_0, \varepsilon) \subset S$ , entonces necesariamente  $B(x_0, \delta) \subset S_0$ , luego  $x_0 \in S_0$  es interior y en virtud de que  $x_0 \in S_0$  es arbitrario,  $S_0$  es abierto.

$\Leftarrow$  Sea  $T$  una aplicación tal que imagen inversa de cualquier subconjunto abierto de  $\mathbb{Y}$  es un subconjunto abierto de  $\mathbb{X}$ . Sea  $x_0 \in S_0$  cualquiera e  $y_0 = Tx_0 \in S$ . Sea  $B(y_0, \varepsilon) \subset S$  una bola de radio  $\varepsilon$  arbitrario. Sea  $S_0$  la imagen inversa de dicha bola  $B(y_0, \varepsilon)$  que es un abierto. Entonces existe una bola  $B(x_0, \delta) \subset S_0$  tal que  $T(B(x_0, \delta)) \subset B(y_0, \varepsilon) \subset S$ , i.e.,  $T$  es continua en  $x_0$ . Como  $x_0$  era arbitrario, entonces  $T$  es continua en  $\mathcal{D}(T)$ .  $\square$

**Definición A.1.13** *Sea  $M \subset \mathbb{X}$ . Diremos que  $x \in \mathbb{X}$  es un punto de contacto (o adherente) de  $M$  si en cualquier bola  $B(x, \varepsilon)$ ,  $\varepsilon > 0$  hay al menos un elemento de  $M$ . Así mismo, diremos que  $x$  es un punto de acumulación (o punto límite) de  $M$  si en cualquier bola  $B(x, \varepsilon)$ ,  $\varepsilon > 0$  hay al menos un elemento de  $M$  distinto de  $x$ , o equivalentemente, en cada bola  $B(x, \varepsilon)$ ,  $\varepsilon > 0$  hay infinitos elementos de  $M$ . Un punto  $x$  se denomina aislado de  $M$  si existe una bola  $B(x, \varepsilon)$ ,  $\varepsilon > 0$  que no contiene ningún elemento  $M$  excepto el propio  $x$ .*

Es fácil ver que si  $M$  solo contiene puntos aislados entonces  $M$  es cerrado (pues  $\mathbb{X} \setminus M$  es abierto). De lo anterior se deduce además que los puntos de contacto de  $M$  o bien son puntos límites, o bien son aislados.

**Definición A.1.14** *Dado un subconjunto  $M \in \mathbb{X}$ , se denomina clausura de  $M$  al conjunto  $\overline{M}$  de los elementos de  $M$  y sus puntos de contacto.*

De lo anterior se sigue que  $\overline{M} = M \cup \{\text{conjunto de sus puntos límites}\}$ .

Por ejemplo, si  $\mathbb{X} = \mathbb{Q}$ , entonces  $\overline{\mathbb{Q}} = \mathbb{R}$  pues todo  $x \in \mathbb{R}$  es un punto límite de  $\mathbb{Q}$  (¿por qué?).

---

<sup>23</sup>Aquí  $\tilde{r}$  denota la métrica de  $\mathbb{X}$  y  $\sigma$  la de  $\mathbb{Y}$ .

**Proposición A.1.15** *Un subconjunto  $M \in \mathbb{X}$  es cerrado si y sólo si  $M = \overline{M}$ . Así, como  $M \subset \overline{M}$ , entonces  $\overline{M}$  es el menor conjunto cerrado que contiene a  $M$ .*

Demostración:  $\Rightarrow$  Sea  $M$  cerrado. Como  $M \subset \overline{M}$  basta probar que  $M \supset \overline{M}$ . Como  $M$  es cerrado, entonces  $\mathbb{X} \setminus M$  es abierto así que para todo  $x \in \mathbb{X} \setminus M$  existe una bola  $B(x, r)$  completamente contenida en  $\mathbb{X} \setminus M$ , i.e.,  $B(x, r)$  no contiene puntos de  $M$ , luego  $x \notin \overline{M}$ , es decir,  $x \in \mathbb{X} \setminus \overline{M}$ , así que  $\mathbb{X} \setminus M \subset \mathbb{X} \setminus \overline{M}$ , i.e.  $\overline{M} \subset M$ .

$\Leftarrow$  Asumamos que  $M = \overline{M}$ . Probaremos que  $\mathbb{X} \setminus M$  es abierto. Sea  $x \notin M$ , entonces  $x \notin \overline{M}$ . Entonces  $x \in \mathbb{X} \setminus M$ . Como  $x \notin \overline{M}$  entonces  $x$  no es un punto adherente de  $M$  así que debe existir al menos una bola  $B(x, r)$  que no contiene a ningún elemento de  $M$ , es decir  $x \in \mathbb{X} \setminus M$  es un punto interior del conjunto  $\in \mathbb{X} \setminus M$ . Como  $x$  es arbitrario,  $\mathbb{X} \setminus M$  es abierto luego  $M$  es cerrado.  $\square$

**Definición A.1.16** *Un subconjunto  $M \subset \mathbb{X}$  es acotado si su diámetro  $d(M) = \sup_{x, y \in M} \vec{r}(x, y)$  es finito.*

**Definición A.1.17** *Dada una sucesión  $(x_n)_n$  de elementos de  $\mathbb{X}$ , diremos que  $(x_n)_n$  es acotada si existe un subconjunto  $M \subset \mathbb{X}$  acotado tal que  $x_n \in M$  para todo  $n \in \mathbb{N}$ .*

Lo anterior es equivalente a que exista un  $x \in \mathbb{X}$  y un número  $K > 0$  tal que  $\vec{r}(x, x_n) < K$  para todo  $n \in \mathbb{N}$ .

**Definición A.1.18** *Una sucesión  $(x_n)_n$  de elementos de  $\mathbb{X}$  es convergente, y lo denotaremos por  $\lim_{n \rightarrow \infty} x_n = x$ , si existe un  $x \in \mathbb{X}$  tal que para todo  $\varepsilon > 0$  existe un  $N \in \mathbb{N}$  tal que para todo  $n > N$ ,  $\vec{r}(x, x_n) < \varepsilon$ . En caso contrario diremos que  $(x_n)_n$  es divergente.*

Nótese que en la propia definición de límite está explícito que el límite ha de ser un elemento de  $\mathbb{X}$ . Por ejemplo, sea  $\mathbb{X}$  el intervalo abierto  $(0, 1)$  con la métrica habitual de  $\mathbb{R}$ . La sucesión  $x_n = 1/(n+1)$  no tiene límite en  $\mathbb{X}$  ya que claramente  $1/(n+1) \xrightarrow{n \rightarrow \infty} 0$  pero  $0 \notin (0, 1)$ .

**Definición A.1.19** *Una sucesión  $(x_n)_n$  de elementos de  $\mathbb{X}$  se denomina de Cauchy o fundamental si existe para todo  $\varepsilon > 0$  existe un  $N \in \mathbb{N}$  tal que para todo  $n > N$  y todo  $p \in \mathbb{N}$ ,  $\vec{r}(x_n, x_{n+p}) < \varepsilon$ .*

En  $\mathbb{R}$  toda sucesión es convergente si y sólo si es de Cauchy. Esta propiedad fundamental de  $\mathbb{R}$  no es cierta para cualquier espacio métrico  $\mathbb{X}$ . Por ejemplo, si escogemos nuevamente  $\mathbb{X}$  como el intervalo abierto  $(0, 1)$  con la métrica habitual de  $\mathbb{R}$ , la sucesión  $x_n = 1/(n+1)$ , que es de Cauchy (¿por qué?) no tiene límite en  $\mathbb{X}$ .

**Definición A.1.20** *Un espacio métrico  $\mathbb{X}$  se denomina completo si y sólo si toda sucesión de Cauchy de elementos de  $\mathbb{X}$  converge (a un elemento de  $\mathbb{X}$ ).*

Por ejemplo, el espacio  $\mathbb{X} = \mathbb{R}$  con la métrica usual de  $\mathbb{R}$ , es completo. También lo es  $\mathbb{X} = \mathbb{C}$  con la métrica usual de  $\mathbb{C}$ . Sin embargo  $\mathbb{Q}$ , el conjunto de los números racionales, es incompleto (¿por qué?), y el conjunto  $\mathbb{X} = (0, 1)$  de antes también lo es. El espacio métrico (de Hilbert)  $l^2$  es completo no así el espacio  $C_2([a, b])$ .

**Teorema A.1.21** *Un subespacio  $M$  de un espacio métrico completo  $\mathbb{X}$  es completo si y sólo si es cerrado en  $\mathbb{X}$ .*

**Demostración:**  $\Rightarrow$  Sea  $M$  completo. Probaremos que  $M = \overline{M}$ . Sea  $x \in \overline{M}$  cualquiera, entonces existe (¿por qué?) una sucesión de elementos de  $M$  que converge a  $x$ . Entonces  $(x_n)_n$  es de Cauchy (pues es convergente) pero  $M$  es completo, luego  $x \in M$ , i.e.,  $M = \overline{M}$ .

$\Leftarrow$  Sea  $M$  cerrado (i.e.,  $M = \overline{M}$ ) y  $(x_n)_n$  una sucesión de Cauchy en  $M$ . Entonces  $\lim_{n \rightarrow \infty} x_n = x$  con  $x \in \mathbb{X}$  ( $\mathbb{X}$  es completo). Pero entonces  $x \in \mathbb{X}$  es un punto de acumulación de  $M$ , i.e.,  $x \in \overline{M}$  y como  $M$  es cerrado  $x \in M$ , i.e., toda sucesión de Cauchy en  $M$  tiene límite en  $M$ , i.e.,  $M$  es completo.  $\square$

**Definición A.1.22** *Un subconjunto  $M \subset \mathbb{X}$  es denso en  $\mathbb{X}$  si su clausura  $\overline{M} = \mathbb{X}$ .*

De la definición anterior se infiere que si  $M$  es denso  $\mathbb{X}$  entonces cualquiera sea la bola  $B(x, \varepsilon)$  (por pequeño que sea  $\varepsilon > 0$ ) siempre contiene puntos de  $M$ . En otras palabras, cualquiera sea  $x \in \mathbb{X}$ , siempre tiene elementos de  $M$  tan cerca como se quiera.

Por ejemplo  $\mathbb{Q}$  es denso en  $\mathbb{R}$  pues como ya hemos visto  $\overline{\mathbb{Q}} = \mathbb{R}$ .

**Definición A.1.23** *Un espacio métrico  $\mathbb{X}$  es separable si contiene un subespacio numerable<sup>24</sup>  $M \subset \mathbb{X}$  denso en  $\mathbb{X}$ .*

Así pues,  $\mathbb{R}$  es separable pues  $\mathbb{Q}$  es numerable y denso en  $\mathbb{R}$ . Usando la separabilidad de  $\mathbb{R}$  se puede probar que  $l^2$  también es separable.

**Definición A.1.24** *Un espacio métrico  $\mathbb{X}$  se denomina compacto si cualquier sucesión  $(x_n)_n$  de elementos de  $\mathbb{X}$  tiene una subsucesión convergente.*

Entenderemos que  $M \subset \mathbb{X}$  es compacto si  $M$  es compacto como subconjunto de  $\mathbb{X}$ , i.e., cualquier  $(x_n)_n$  de elementos de  $M$  tiene una subsucesión convergente en  $M$ .

**Lema A.1.25** *Si  $M \subset \mathbb{X}$  es compacto, entonces  $M$  es cerrado y acotado.*

**Demostración:** Sea  $M$  compacto y sea  $x \in \overline{M}$  cualquiera. Como  $x \in \overline{M}$  entonces existe una sucesión  $(x_n)_n$  en  $M$  tal que  $x_n \xrightarrow{n \rightarrow \infty} x \in \mathbb{X}$ . Como  $M$  es compacto, entonces  $x \in M$ , luego  $M = \overline{M}$  por lo que  $M$  es cerrado.

Supongamos que  $M$  es no acotado. Entonces existe al menos una sucesión  $(x_n)_n$  de elementos de  $M$  tal que, fijado un  $b \in M$  arbitrario, se tiene que  $\vec{r}(x_n, b) > n$  (¿por qué?). Dicha sucesión obviamente no puede tener ninguna subsucesión convergente (pues caso que la tuviera ésta fuese acotada) y por tanto  $M$  no puede ser compacto.  $\square$

El recíproco es falso. Por ejemplo escojamos  $\mathbb{X} = l^2$  y sea el conjunto  $M$  de los vectores  $e_k = \delta_{k,i}$ , i.e., vectores que todas las coordenadas son cero excepto la  $k$ -ésima que es 1. Obviamente  $\|e_k\| = 1$ . Además todos los puntos de  $M$  son aislados (¿por qué?), por tanto  $M$  es cerrado. Ahora bien, como  $M$  solo tiene puntos aislados,  $M$  no tiene ningún punto de acumulación por lo tanto ninguna sucesión que escojamos de elementos *distintos* de  $M$  contiene una subsucesión convergente.

<sup>24</sup>Un conjunto  $M$  cualquiera se denomina *numerable* si se puede poner en correspondencia biunívoca con  $\mathbb{N} = \{1, 2, 3, \dots\}$ . Es decir, existe una correspondencia biunívoca entre los elementos de  $M$  y los números naturales. Por ejemplo,  $\mathbb{Q}$  es numerable, pero  $\mathbb{R}$  no lo es.

**Definición A.1.26** *Un espacio vectorial  $\mathbb{X}$  se denomina espacio normado si para todo  $x \in \mathbb{X}$  existe un número real denominado norma, y que denotaremos por  $\|x\|$ , que cumple con las condiciones*

1. Para todo  $x \in \mathbb{X}$ ,  $\|x\| \geq 0$  y si  $\|x\| = 0$  entonces  $x = 0$ .
2. Para todo  $x \in \mathbb{X}$  y  $\lambda \in \mathbb{R}$ ,  $\|\lambda x\| = |\lambda|\|x\|$ ,
3. Para todos  $x, y \in \mathbb{X}$  se tiene la desigualdad triangular

$$\|x + y\| \leq \|x\| + \|y\|. \quad (\text{A.1.1})$$

Es evidente que si en un espacio normado  $\mathbb{X}$  definimos la función  $\vec{r}(x, y) = \|x - y\|$ , esta satisface los axiomas de la definición A.1.1, i.e., todo espacio normado es un espacio métrico. La función  $\vec{r}$  anterior se denomina métrica inducida por la norma.

**Definición A.1.27** *Un espacio normado completo (en la métrica inducida por la norma) se denomina espacio de Banach.*

Así, el espacio  $l^2$ , de todas las sucesiones  $x = (x_1, x_2, \dots, x_n, \dots)$  reales (o complejas) tales que  $\sum_{k=1}^{\infty} |x_k|^2 < +\infty$  con la norma  $\|x\| = (\sum_{k=1}^{\infty} |x_k|^2)^{1/2}$ , es un espacio de Banach, pero el espacio de las funciones continuas en  $[a, b]$  con la norma  $\|f\| = \left(\int_a^b |f(x)|^2\right)^{1/2}$  es un espacio normado pero no de Banach (¿por qué?).

Está claro que todo espacio normado es un espacio métrico con la métrica inducida por la norma, pero no a la inversa (construir un contra ejemplo como ejercicio).

Obviamente en los espacios normados podemos definir la convergencia de sucesiones, sucesiones de Cauchy, etc.. Basta considerarlos como espacios métricos con la métrica  $\vec{r}$  inducida por la norma:  $\vec{r}(x, y) = \|x - y\|$ .

Consideremos ahora un caso particular de las aplicaciones introducidas en la definición A.1.5.

**Definición A.1.28** *Una aplicación (operador) es lineal si*

1. El dominio de  $T$ ,  $\mathcal{D}(T)$ , y la imagen de  $T$ ,  $I(T)$ , son ambos espacios vectoriales sobre el mismo cuerpo  $\mathbb{K}$  ( $\mathbb{R}$  o  $\mathbb{Z}$ ).
2.  $\forall \alpha, \beta \in \mathbb{K}, \quad \forall x, y \in \mathcal{D}(T), \quad T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ .

**Definición A.1.29** *Sean  $\mathbb{X}$  e  $\mathbb{Y}$  dos espacios normados y sea el operador  $T : \mathcal{D}(T) \mapsto \mathbb{Y}$  lineal.  $T$  es acotado si existe  $c \geq 0$  tal que<sup>25</sup>*

$$\|Tx\| \leq c\|x\|, \quad \forall x \in \mathcal{D}(T). \quad (\text{A.1.2})$$

<sup>25</sup>Se sobrentiende que  $\|x\|$  es la norma en  $\mathbb{X}$  y  $\|Tx\|$  es en  $\mathbb{Y}$ .

De lo anterior se sigue que si  $T$  es acotado, entonces para todo  $x \neq 0$ ,

$$\frac{\|Tx\|}{\|x\|} \leq c, \quad \forall x \in \mathcal{D}(T), x \neq 0. \quad (\text{A.1.3})$$

El menor valor de  $c$  para el cual (A.1.2) se cumple lo denotaremos por  $\|T\|$  y se denomina *norma del operador* lineal  $T$ . De hecho se tiene que

$$\|T\| = \sup_{x \in \mathbb{X} \setminus \{0\}} \frac{\|Tx\|}{\|x\|}. \quad (\text{A.1.4})$$

Si  $T = 0$  obviamente  $\|T\| = 0$ . Además de (A.1.2), tomando ínfimos en  $c$  obtenemos

$$\forall y \in \mathbb{X}, \quad \frac{\|Ty\|}{\|y\|} \leq \|T\| \iff \|Ty\| \leq \|T\|\|y\|.$$

Es fácil probar que  $\|T\|$  es una norma, es decir se cumplen los axiomas de la definición.

**Teorema A.1.30** *Sea  $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$  una aplicación lineal de un espacio normado  $\mathbb{X}$  a otro espacio normado  $\mathbb{Y}$ . Entonces*

1.  $T$  es continuo si y sólo si  $T$  es acotado.
2. Si  $T$  es continuo en algún  $x_0 \in \mathcal{D}(T)$ ,  $T$  es continuo en  $\mathcal{D}(T)$ .

Demostración: Asumiremos que  $T$  no es el operador nulo.

1.  $\Rightarrow$  Sea  $T$  acotado y sea  $x_0 \in \mathcal{D}(T)$  cualquiera. Como  $T$  es lineal y acotado, entonces

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\|.$$

Entonces, para todo  $\varepsilon > 0$ , existe un  $\delta = \varepsilon/\|T\| > 0$  tal que, para todo  $x$  con  $\|x - x_0\| < \delta$ ,  $\|Tx - Tx_0\| < \varepsilon$ , i.e.,  $T$  es continuo en  $\mathcal{D}(T)$ .

$\Leftarrow$  Sea  $T$  lineal y continuo en  $x_0 \in \mathcal{D}(T)$  cualquiera. Entonces para todo  $\varepsilon > 0$ , existe un  $\delta > 0$  tal que, para todo  $x$  con  $\|x - x_0\| < \delta$ ,  $\|Tx - Tx_0\| < \varepsilon$ . Sea  $y \neq 0$  en  $\mathcal{D}(T)$  cualquiera. Escojamos  $x$  tal que

$$x = x_0 + \frac{\delta}{2\|y\|}y \implies x - x_0 = \frac{\delta}{2\|y\|}y \implies \|x - x_0\| < \delta \implies \|Tx - Tx_0\| < \varepsilon.$$

Además, para dichos  $x$  tenemos, usando la linealidad de  $T$ , que

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\| T \frac{\delta}{2\|y\|}y \right\| = \frac{\delta}{2\|y\|} \|Ty\| \leq \varepsilon \implies \|Ty\| \leq \frac{2\varepsilon}{\delta} \|y\|,$$

luego  $T$  es acotado.

2. Nótese que en la prueba anterior se probó que si  $T$  era continuo en un punto  $x_0 \in \mathcal{D}(T)$ , entonces era acotado en  $\mathcal{D}(T)$ . Pero entonces por 1, al ser acotado en  $\mathcal{D}(T)$ , es continuo en  $\mathcal{D}(T)$ .  $\square$

Finalmente concluiremos con una sección sobre los espacios de Hilbert separables que complementa los apartados 4.1 y 4.2.

## A.2. Espacios de Hilbert separables

**Definición A.2.1** Dado un vector  $x \in \mathbb{H}$  definiremos la serie de Fourier respecto al sistema ortonormal  $(\phi_n)_{n=1}^{\infty}$  a la serie

$$s := \sum_{n=1}^{\infty} c_n \phi_n, \quad (\text{A.2.1})$$

donde los coeficientes vienen dados por las expresiones

$$c_n = \langle x, \phi_n \rangle, \quad \forall n \geq 1. \quad (\text{A.2.2})$$

**Teorema A.2.2** Sea  $H$  el subespacio lineal de  $\mathbb{H}$  generado por los vectores  $\phi_1, \phi_2, \dots, \phi_n$ ,  $n \in \mathbb{N}$ , i.e.,  $H = \text{span}(\phi_1, \phi_2, \dots, \phi_n)$ . Entonces

$$\min_{q \in H} \|x - q\|^2 = \|x\|^2 - \sum_{k=1}^n |c_k|^2$$

donde  $c_k$  son los coeficientes definidos en (A.2.2) y se alcanza cuando  $q$  es la suma parcial de la serie de Fourier (A.2.1)

$$q = s_n := \sum_{k=1}^n c_k \phi_k.$$

Demostración: Sea  $g_n = \sum_{k=1}^n a_k \phi_k$ . Calculamos

$$\langle x - g_n, x - g_n \rangle = \|x\|^2 - \sum_{k=1}^n c_k^2 \|\phi_k\|^2 + \sum_{k=1}^n (a_k - c_k)^2 \|\phi_k\|^2.$$

Obviamente la expresión anterior es mínima si y sólo si  $a_k = c_k$  para todo  $k \in \mathbb{N}$ , i.e.,  $g_n = s_n$ .  $\square$

Nótese que

$$\begin{aligned} \|x - s_n\|^2 &= \left\langle x - \sum_{k=1}^n \frac{\langle x, \phi_k \rangle}{\|\phi_k\|^2} \phi_k, x - \sum_{k=1}^n \frac{\langle x, \phi_k \rangle}{\|\phi_k\|^2} \phi_k \right\rangle = \|x\|^2 - 2 \sum_{k=1}^n \frac{|\langle x, \phi_k \rangle|^2}{\|\phi_k\|^2} \\ &+ \sum_{k=1}^n \sum_{m=1}^n \frac{|\langle x, \phi_k \rangle|^2}{\|\phi_k\|^2} \frac{\langle \phi_m, \phi_k \rangle}{\|\phi_m\|^2} = \|x\|^2 - \sum_{k=1}^n \frac{|\langle x, \phi_k \rangle|^2}{\|\phi_k\|^2} = \|x\|^2 - \sum_{k=1}^n |c_k|^2. \end{aligned} \quad (\text{A.2.3})$$

Como corolario de lo anterior tenemos que

$$I_n = \|x - s_n\|^2 = \|x\|^2 - \sum_{k=1}^n |c_k|^2 \geq 0, \quad \forall n \in \mathbb{N},$$

luego

$$\sum_{k=1}^{\infty} |c_k|^2 \leq \|x\|^2, \quad (\text{A.2.4})$$

por lo que la serie  $\sum_{k=1}^{\infty} |c_k|^2$  converge (¿por qué?) y por tanto

$$\lim_{n \rightarrow \infty} |c_n| = 0 \implies \lim_{n \rightarrow \infty} \langle x, \phi_n \rangle = \lim_{n \rightarrow \infty} \langle \phi_n, x \rangle = 0. \quad (\text{A.2.5})$$

La desigualdad (A.2.4) se conoce como **desigualdad de Bessel**. Nótese que una condición necesaria y suficiente para que la serie de Fourier (A.2.1) converja a  $x$  (en norma) es que

$$\|x\|^2 = \sum_{k=1}^{\infty} |c_k|^2 = \sum_{k=1}^{\infty} |\langle x, \phi_k \rangle|^2.$$

Esta igualdad se denomina comúnmente **igualdad de Parseval** y es, en general, muy complicada de comprobar.

**Definición A.2.3** Se dice que un sistema de vectores linealmente independientes  $(\phi_n)_n$  es completo en  $\mathbb{X} \subset \mathbb{H}$  si para todo vector  $x \in \mathbb{X} \subset \mathbb{H}$  y cualquiera sea  $\epsilon > 0$  existe una combinación lineal

$$l_n = \sum_{k=1}^n \alpha_k \phi_k \quad \text{tal que} \quad \|x - l_n\| < \epsilon.$$

En otras palabras cualquier vector  $x \in \mathbb{X} \subset \mathbb{H}$  se puede aproximar en norma tanto como se quiera mediante alguna combinación finita de vectores del sistema  $(\phi_n)_n$ . Esta definición es equivalente a decir que  $\mathbb{H}$  es el menor subespacio vectorial cerrado que contiene al conjunto  $\phi_1, \phi_2, \dots$  ( $(\phi_n)_n$  genera a todo  $\mathbb{H}$ ).

**Definición A.2.4** Un sistema ortogonal (ortonormal) completo de  $\mathbb{X} \subset \mathbb{H}$  se denomina base ortogonal (ortonormal) de  $\mathbb{X} \subset \mathbb{H}$ .

Por ejemplo, los sistemas  $(e_k)_k$  definidos por (A.2.6) y (A.2.7) son bases ortogonales completas de  $\mathbb{C}^n$  y  $l^2$ , respectivamente

$$e_1 = (1, 0, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad e_n = (0, 0, 0, \dots, 1), \quad (\text{A.2.6})$$

$$e_1 = (1, 0, 0, 0, \dots), \quad e_2 = (0, 1, 0, 0, \dots), \quad e_n = (0, 0, 1, 0, \dots), \quad \dots \quad (\text{A.2.7})$$

**Teorema A.2.5** Sea  $\mathbb{H}$  un espacio de Hilbert y sea el sistema ortonormal de vectores  $(\phi_n)_{n=1}^{\infty}$  de  $\mathbb{H}$ . Entonces las siguientes condiciones son equivalentes:

1.  $(\phi_n)_n$  es completo en  $\mathbb{X} \subset \mathbb{H}$ .

2. Para todo  $x \in \mathbb{X} \subset \mathbb{H}$ ,  $x = \sum_{k=1}^{\infty} \langle x, \phi_k \rangle \phi_k$ .

3. Para todo  $x \in \mathbb{X} \subset \mathbb{H}$ , se cumple la igualdad de Parseval

$$\|x\|^2 = \sum_{k=1}^{\infty} |\langle x, \phi_k \rangle|^2.$$

4. Si  $\langle x, \phi_k \rangle = 0$  para todo  $k \in \mathbb{N}$  entonces  $x = 0$ .

Demostración: 1)  $\leftrightarrow$  2) Cualquiera sea  $x \in \mathbb{H}$  construimos la serie de Fourier (A.2.1) y sean  $s_n = \sum_{k=1}^n c_k \phi_k$  sus sumas parciales. Usando 1) y el teorema A.2.2 tenemos

$$\forall \epsilon > 0, \quad \exists N \in \mathbb{N} \text{ tal que } \|x - s_N\| < \epsilon.$$

Pero como para todo  $n > N$ ,  $\|x - s_n\| \leq \|x - s_N\|$  (basta usar la identidad (A.2.3)), entonces para todo  $n > N$ ,  $\|x - s_n\| \leq \epsilon$ , es decir,  $\lim_{n \rightarrow \infty} s_n = x$  de donde se sigue 2). Obviamente 2) implica 1) (¿por qué?).

2)→3) Tomando el límite  $n \rightarrow \infty$  en (A.2.3) y usando 2) ( $\lim_{n \rightarrow \infty} \|x - s_n\| = 0$ ), se sigue 3).

3)→4) Si para todo  $k \in \mathbb{N}$   $\langle x, \phi_k \rangle = 0$ , entonces de 3) se deduce que  $\|x\| = 0$ , luego  $x = 0$ .

4)→2) Por sencillez consideremos que el sistema  $(\phi_n)_n$  es ortonormal. Sea  $y_n = \sum_{k=1}^n c_k \phi_k$ ,  $c_k = \langle x, \phi_k \rangle$ . Usando la desigualdad de Bessel (A.2.4) se sigue que la serie  $\sum_{m=1}^{\infty} |c_m|^2$  es una serie convergente y por tanto la cantidad

$$\|y_n - y_{n+p}\|^2 = \sum_{k=n+1}^{n+p} |c_k|^2,$$

se puede hacer tan pequeña como se quiera, i.e., la sucesión  $y_n$  es de Cauchy, luego es convergente. Sea  $y$  su límite. Probemos que  $y = x$ .

Usando la desigualdad de Cauchy-Schwarz

$$|\langle y, \phi_k \rangle - \langle y_n, \phi_k \rangle| = |\langle y - y_n, \phi_k \rangle| \leq \|y_n - y\| \|\phi_k\|,$$

deducimos que  $\lim_{n \rightarrow \infty} \langle y_n, \phi_k \rangle = \langle y, \phi_k \rangle$  para todo  $k \in \mathbb{N}$ . Pero

$$\langle y_n, \phi_k \rangle = \sum_{j=1}^n \langle x, \phi_j \rangle \langle \phi_k, \phi_j \rangle = \langle x, \phi_k \rangle, \quad \forall k \leq n,$$

luego  $\langle x - y_n, \phi_k \rangle = 0$  para todo  $k \leq n$  de donde, tomando  $n \rightarrow \infty$  deducimos que  $\langle x - y, \phi_k \rangle = 0$  para todo  $k \in \mathbb{N}$ , luego por 4)  $x - y = 0$ .  $\square$

**Nota A.2.6** La equivalencia entre 1 y 2, así como las implicaciones  $2 \rightarrow 3 \rightarrow 4$ , son también ciertas para espacios euclídeos cualesquiera (no necesariamente completos).

A partir del apartado 4 del Teorema A.2.5 se sigue el siguiente corolario:

**Corolario A.2.7** Sea el sistema ortonormal completo  $(\phi_n)_n$  y sean  $x, y \in \mathbb{X} \subset \mathbb{H}$  tales que  $\langle x, \phi_k \rangle = \langle y, \phi_k \rangle$  para todo  $k \in \mathbb{N}$ , entonces  $x = y$ .

En otras palabras, dos elementos de  $\mathbb{H}$  con iguales coeficientes de Fourier son iguales, por tanto cualquier vector de  $\mathbb{H}$  queda biunívocamente determinado por sus coeficientes de Fourier.

**Definición A.2.8** Se dice que un sistema ortonormal  $(\phi_n)_n$  es cerrado en un espacio euclídeo  $\mathbb{E}$  si para todo vector  $x \in \mathbb{E}$  se cumple la igualdad de Parseval

$$\sum_{k=1}^{\infty} |c_k|^2 = \sum_{k=1}^{\infty} |\langle x, \phi_k \rangle|^2 = \|x\|^2.$$

De la definición anterior y el Teorema A.2.5 se sigue que un sistema ortonormal  $(\phi_n)_n$  es completo en un espacio de Hilbert  $\mathbb{H}$  si y sólo si  $(\phi_n)_n$  es cerrado en  $\mathbb{H}$ .

**Teorema A.2.9** *Todo espacio de Hilbert  $\mathbb{H}$  separable tiene una base ortonormal.*

Demostración: Como  $\mathbb{H}$  es separable, existe un conjunto numerable de vectores  $(\phi_n)_n$  denso en  $\mathbb{H}$ . Si de dicho conjunto eliminamos aquellos vectores  $\phi_k$  que se pueden obtener como combinación lineal de los anteriores  $\phi_j, j < k$  obtenemos un sistema completo de vectores linealmente independientes de  $\mathbb{H}$ . La base ortonormal se obtiene al aplicar a dicho sistema el proceso de ortogonalización de Gram-Schmidt.  $\square$

El teorema anterior se puede generalizar a cualquier espacio euclídeo separable.

**Teorema A.2.10 (Riesz-Fischer)** *Sea  $(\phi_n)_n$  un sistema ortonormal en un espacio de Hilbert  $\mathbb{H}$  y sean los números  $c_1, c_2, \dots, c_n, \dots$  tales que*

$$\sum_{n=1}^{\infty} |c_n|^2 < +\infty.$$

*Entonces, existe un elemento  $x \in \mathbb{H}$  cuyos coeficientes de Fourier son precisamente los números  $c_1, c_2, \dots, c_n, \dots$ , i.e.,*

$$\sum_{n=1}^{\infty} |c_n|^2 = \|x\|^2, \quad c_n = \langle x, \phi_n \rangle.$$

Demostración: Sea  $x_n = \sum_{k=1}^n c_k \phi_k$ . Como vimos en la prueba del Teorema A.2.5 la sucesión anterior es de Cauchy y como  $\mathbb{H}$  es completo entonces  $x_n \xrightarrow{n \rightarrow \infty} x \in \mathbb{H}$ . Probemos que entonces  $c_k = \langle x, \phi_k \rangle$ ,  $k = 1, 2, \dots$ . Para ello notemos que

$$\langle x, \phi_k \rangle = \langle x_n, \phi_k \rangle + \langle x - x_n, \phi_k \rangle.$$

Pero entonces si  $n \geq k$ ,  $\langle x - x_n, \phi_k \rangle = 0$  y, por tanto,  $c_k = \langle x, \phi_k \rangle = \langle x_n, \phi_k \rangle$ , de donde se sigue, tomando el límite  $n \rightarrow \infty$  y usando la continuidad del producto escalar que para todo  $k = 1, 2, 3, \dots$ ,  $c_k = \langle x, \phi_k \rangle$ . Además, como  $x_n \xrightarrow{n \rightarrow \infty} x$ , usando (A.2.3), se tiene que

$$\|x\|^2 - \sum_{k=1}^n |c_k|^2 = \|x - x_n\|^2 \xrightarrow{n \rightarrow \infty} 0,$$

de donde se sigue el teorema.  $\square$

**Definición A.2.11** *Una aplicación  $U$  entre dos espacios de Hilbert  $\mathbb{H}$  y  $\mathbb{H}^*$  se denomina unitaria si  $U$  es lineal, biyectiva y preserva el producto escalar, i.e.<sup>26</sup>,*

$$\langle x, y \rangle = \langle Ux, Uy \rangle_* = \langle x^*, y^* \rangle_*.$$

*Los espacios  $\mathbb{H}$  y  $\mathbb{H}^*$  son isomorfos si existe una aplicación unitaria  $U : \mathbb{H} \mapsto \mathbb{H}^*$  tal que  $x^* = Ux$ , donde  $x \in \mathbb{H}$  y  $x^* \in \mathbb{H}^*$ .*

Como consecuencia de los Teoremas A.2.5, A.2.9 y A.2.10 se tiene el siguiente resultado:

**Teorema A.2.12 (del isomorfismo)** *Cualquier espacio de Hilbert separable  $\mathbb{H}$  es isomorfo a  $\mathbb{C}^n$  o a  $l^2$ .*

---

<sup>26</sup>Se entiende que  $\langle \cdot, \cdot \rangle_*$  denota el producto escalar en  $\mathbb{H}^*$  que no tiene por que ser el mismo que en  $\mathbb{H}$ .

**Demostración:** Como  $\mathbb{H}$  es separable, en  $\mathbb{H}$  existe una base ortonormal numerable (ver Teorema A.2.9) que denotaremos por  $(\phi_n)_{n \in I}$ , donde  $I$  es un conjunto numerable (finito o infinito). Asumiremos que  $I$  es infinito, ( $I = \mathbb{N}$ ) i.e., probaremos el caso cuando  $\mathbb{H}$  es isomorfo a  $l^2$  (el caso finito es totalmente análogo). Sea  $x \in \mathbb{H}$  y sea la aplicación  $U : \mathbb{H} \mapsto l^2$  definida por  $x^* = Ux = \sum_{k \in I} \langle x, \phi_k \rangle e_k = \sum_{k \in I} x_k e_k \in l^2$ , donde  $(e_n)_n$  denota la base ortonormal canónica de  $l^2$  que ya hemos visto antes. Está claro que  $U$  es biunívoco, pues dada cualquier sucesión  $(x_n)_n$  de  $l^2$  por el Teorema de Riesz-Fischer existe un  $x \in \mathbb{H}$  cuyos coeficientes de Fourier coinciden con dichos valores  $x_n$  y por el Corolario A.2.7 dicho elemento es único. Además, como el producto escalar  $\langle x, y \rangle$  es lineal respecto al elemento de la izquierda (i.e.,  $x$ ), está claro que  $U$  es lineal (probarlo como ejercicio). Finalmente, para probar que  $U$  es unitario usamos, por un lado, que

$$\langle x, y \rangle = \left\langle \sum_{k \in I} x_k \phi_k, \sum_{m \in I} y_m \phi_m \right\rangle = \sum_{k \in I} x_k \overline{y_k},$$

y, por el otro, que el producto escalar en  $l^2$  viene dado por  $\langle x^*, y^* \rangle_{l^2} = \sum_{k \in I} x_k \overline{y_k}$ , de donde se sigue que  $\langle x, y \rangle = \langle x^*, y^* \rangle_{l^2}$ .  $\square$

**Definición A.2.13** Sea  $M \subset \mathbb{H}$  un subespacio<sup>27</sup> cerrado del espacio de Hilbert  $\mathbb{H}$ . Denominaremos complemento ortogonal de  $M$ , y lo denotaremos por  $M^\perp$ , al conjunto

$$M^\perp = \{x \in \mathbb{H}; \langle x, y \rangle = 0, \forall y \in M\}.$$

Nótese que al ser  $M^\perp$  cerrado, es completo (pues  $\mathbb{H}$  es completo, y todo subespacio  $M$  de un espacio métrico completo  $\mathbb{H}$  es completo si y sólo si es cerrado en  $\mathbb{H}$ ).

**Teorema A.2.14** Sea  $M \subset \mathbb{H}$  un subespacio cerrado del espacio de Hilbert  $\mathbb{H}$  y  $M^\perp$  su complemento ortogonal. Entonces, todo vector  $x \in \mathbb{H}$  admite una única representación de la forma  $x = y + y^\perp$  donde  $y \in M$  e  $y^\perp \in M^\perp$ .

**Demostración:** Como  $\mathbb{H}$  es separable y  $M$  es cerrado en  $\mathbb{H}$ , entonces  $M$  es completo (recordemos nuevamente que todo subespacio  $M$  de un espacio métrico completo  $\mathbb{H}$  es completo si y sólo si es cerrado en  $\mathbb{H}$ ) y separable, i.e.,  $M$  es a su vez un espacio de Hilbert separable por lo que existe una base ortonormal completa  $(\phi_n)_n$  de  $M$ . Definamos  $y = \sum_n \langle x, \phi_n \rangle \phi_n$ . Obviamente  $y \in M$ . Definamos  $y^\perp = x - y$ . Entonces, para todo  $n$   $\langle y^\perp, \phi_n \rangle = 0$ . Ahora bien, cualquiera sea  $\tilde{y} \in M$ ,  $\tilde{y} = \sum_n a_n \phi_n$ , luego  $\langle y^\perp, \tilde{y} \rangle = 0$ , es decir  $y^\perp \in M^\perp$ . Luego cualquiera sea  $x \in \mathbb{H}$ , existen  $y \in M$  y  $y^\perp \in M^\perp$  tales que  $x = y + y^\perp$ . Probemos que esta descomposición es única. Para ello supongamos que no, i.e., supongamos que existen un  $y' \in M$ ,  $y' \neq y$  tal que  $x = y' + y'^\perp$ ,  $y'^\perp \in M^\perp$  (y obviamente distinta de  $y^\perp$ ). Entonces

$$\langle y', \phi_n \rangle = \langle x - y'^\perp, \phi_n \rangle = \langle x, \phi_n \rangle = \langle y, \phi_n \rangle \implies y' = y,$$

lo que es una contradicción.  $\square$

Si todo elemento de  $\mathbb{H}$  se puede escribir en la forma  $x = y + y^\perp$  donde  $y \in M$  e  $y^\perp \in M^\perp$ , entonces se dice que  $M$  es suma directa de  $M$  y  $M^\perp$  y se escribe como  $\mathbb{H} = M \oplus M^\perp$ . Es fácil ver que la noción de suma directa se puede extender al caso de un número finito o contable de subespacios  $M_1, M_2$ , etc. Nótese además que del teorema anterior se sigue que  $(M^\perp)^\perp = M$ .

Para terminar mencionaremos un teorema sobre funcionales lineales acotados.

<sup>27</sup>Se entenderá, como el caso de los espacios normados que  $M$  es un subespacio *lineal* de  $\mathbb{H}$ .

**Teorema A.2.15 (Riesz)** *Cualquier funcional lineal acotado  $T : \mathbb{H} \mapsto \mathbb{K}$  ( $\mathbb{K}$  es  $\mathbb{C}$  o  $\mathbb{R}$ ) se puede representar en términos de un producto escalar, i.e.,*

$$Tx = \langle x, z \rangle,$$

donde  $z$  depende de  $T$  y esta unívocamente determinado por  $T$  y su norma satisface la ecuación

$$\|z\| = \|T\|.$$

La prueba de este importante teorema se puede encontrar en [E. Kreyszig. Introductory Functional Analysis with Applications, pág. 188].

### A.2.1. Operadores en espacios de Hilbert

**Definición A.2.16** *Sea la aplicación (operador) lineal  $A : \mathbb{E} \mapsto \mathbb{E}'$ ,  $\mathbb{E}, \mathbb{E}'$  espacios euclídeos. Si existe el operador lineal  $A^* : \mathbb{E}' \mapsto \mathbb{E}$  tal que para todo  $x \in \mathbb{E}$  e  $y \in \mathbb{E}'$*

$$\langle Ax, y \rangle' = \langle x, A^*y \rangle,$$

lo denominaremos adjunto de  $A$ .

Por sencillez asumiremos  $\mathbb{E}' = \mathbb{E}$ .

Por ejemplo sea el operador  $S : l^2 \mapsto l^2$

$$S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots),$$

comúnmente denominado *operador desplazamiento* (shift). Entonces su adjunto  $S^*$  es el operador  $S : l^2 \mapsto l^2$

$$S^*(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots),$$

Para los espacios euclídeos la existencia del operador adjunto no está garantizada en general, no obstante si que lo está en el caso de los espacios de Hilbert. De hecho, como consecuencia del Teorema de representación de Riesz A.2.15 se tiene el siguiente resultado:

**Teorema A.2.17** *Sea la aplicación (operador) lineal  $A : \mathbb{H} \mapsto \mathbb{H}'$ ,  $\mathbb{H}, \mathbb{H}'$  espacios de Hilbert. Entonces existe un único operador  $A^* : \mathbb{H}' \mapsto \mathbb{H}$  adjunto a  $A$ . Además,  $A^*$  es lineal y  $\|A^*\| = \|A\|$ .*

Demostración: Definimos el funcional  $T_y x : \mathbb{H} \mapsto \mathbb{K}$ ,  $T_y x := \langle Ax, y \rangle'$ . Obviamente, para cada  $y$  fijo

$$|T_y x| \leq \|Ax\| \|y\| \leq \|A\| \|x\| \|y\| \leq K \|x\|, \quad \forall y \in \mathbb{H}',$$

i.e.,  $T_y$  es un funcional acotado, así que el teorema de Riesz A.2.15 nos asegura que existe un único vector  $y^*$  tal que  $T_y x = \langle x, y^* \rangle$ , para todo  $x \in \mathbb{H}$ . Así, el operador  $A : \mathbb{H} \mapsto \mathbb{H}'$  induce un operador  $A^* : \mathbb{H}' \mapsto \mathbb{H}$  con  $A^*y = y^*$ , pues  $\langle Ax, y \rangle' = \langle x, y^* \rangle = \langle x, A^*y \rangle$ . La linealidad de  $A^*$  se sigue de la linealidad de  $A$ . Probemos ahora que  $\|A^*\| = \|A\|$ . De la igualdad  $\langle Ax, y \rangle' = \langle x, A^*y \rangle$  y usando que  $A$  es acotado tenemos

$$|\langle x, A^*y \rangle| = |\langle Ax, y \rangle'| \leq \|A\| \|x\| \|y\|.$$

Escogiendo  $x = A^*y$  obtenemos

$$\|A^*y\|^2 \leq \|A\| \|A^*y\| \|y\| \implies \frac{\|A^*y\|}{\|y\|} \leq \|A\| \implies \|A^*\| \leq \|A\|.$$

De lo anterior se sigue que  $A^*$  es acotado. Pero entonces podemos aplicar el mismo razonamiento intercambiando  $A$  y  $A^*$  lo que nos conduce a la desigualdad contraria  $\|A\| \leq \|A^*\|$  por tanto  $\|A\| = \|A^*\|$ .  $\square$

En adelante asumiremos que  $\mathbb{H} = \mathbb{H}'$ . Supongamos que  $\mathbb{H}$  es un espacio de Hilbert de dimensión finita. Entonces como ya hemos visto,  $\mathbb{H}$  es isomorfo a  $\mathbb{C}^n$ . Sea  $(e_k)_k$  una base de  $\mathbb{H}$ . Entonces para todo  $x \in \mathbb{H}$

$$x = \sum_{k=1}^n x_k e_k \implies y = Ax = \sum_{k=1}^n x_k A e_k.$$

Si

$$A e_k = \sum_{i=1}^n a_{i,k} e_i \implies y = \sum_{i=1}^n \left( \sum_{k=1}^n a_{i,k} x_k \right) e_i = \sum_{i=1}^n y_i e_i \implies$$

$$y_i = \sum_{k=1}^n a_{i,k} x_k, \quad i = 1, 2, \dots, n.$$

Es decir, si consideramos los vectores  $x, y \in \mathbb{C}^n$  con coordenadas  $x_i, y_i, i = 1, \dots, n$ , respectivamente, entonces el operador  $A$  se puede representar como una matriz  $(a_{i,j})_{i,j=1}^n$ , i.e., tenemos la aplicación  $A : \mathbb{C}^n \mapsto \mathbb{C}^n, y = Ax$ , donde  $A$  es una matriz  $n \times n$

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{pmatrix}.$$

Obviamente la matriz de la aplicación identidad es la matriz identidad.

Nótese que lo anterior se puede generalizar al caso de dimensión infinita, sólo que en este caso la matriz sería una matriz infinita

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} & \cdots \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} & \cdots \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Definición A.2.18** Sea el operador lineal  $A : \mathbb{X} \mapsto \mathbb{Y}$ ,  $\mathbb{X}$  e  $\mathbb{Y}$  espacios de Banach. Diremos que  $A$  es invertible si existe un operador  $B, A : \mathbb{Y} \mapsto \mathbb{X}$  tal que  $AB = I_{\mathbb{Y}}, BA = I_{\mathbb{X}}$ .

Está claro que para el caso de dimensión finita, existirá el inverso de  $A$  si  $\dim \mathbb{X} = \dim \mathbb{Y}$  y la matriz correspondiente será la matriz inversa de  $A$ . En dimensión infinita la situación es algo más complicada. Por ejemplo, el operador desplazamiento  $S$  cumple con  $S^*S = I$  pero  $SS^* \neq I$ , luego  $S$  no tiene inverso.

**Teorema A.2.19** Sea el operador lineal  $A : \mathbb{X} \mapsto \mathbb{X}$ ,  $\mathbb{X}$  espacio de Banach. Si  $\|A\| < 1$ , entonces  $I - A$  es invertible y (en norma)

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k, \quad \text{donde } A^0 := I.$$

**Demostración:** Sea la sucesión de operadores  $(A_n)_n$ , definida por  $A_n x := X_n = (I + A + A^2 + \dots + A^n)x$ ,  $x \in \mathbb{X}$  (cualquiera). Obviamente  $A_n$  es un operador acotado (probarlo como ejercicio). Probemos que  $(X_n)_n$  es de Cauchy (en la norma de  $\mathbb{X}$ ):

$$\begin{aligned} \|X_{n+p} - X_n\| &= \|X_{n+1} + \dots + X_{n+p}\| = \|A^{n+1}x + \dots + A^{n+p}x\| \leq \|A^{n+1}x\| + \dots + \|A^{n+p}x\| \\ &\leq \|x\|(\|A^{n+1}\| + \dots + \|A^{n+p}\|) \leq \|x\|(\|A\|^{n+1} + \dots + \|A\|^{n+p}) \\ &\leq \|x\|(\|A\|^{n+1} + \dots + \|A\|^{n+p} + \dots) \leq \|x\| \frac{\|A\|^{n+1}}{1 - \|A\|} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Como  $\mathbb{X}$  es completo (es de Banach) entonces, para cada  $x \in \mathbb{X}$  la sucesión  $A_n x$  converge a un  $y \in \mathbb{X}$  que denotaremos por  $y = Tx$ . Así, tendremos el operador  $T : \mathbb{X} \mapsto \mathbb{X}$  bien definido. Como  $A$  es lineal,  $A_n$  lo será y por tanto  $T$  también. Para ello basta tomar el límite cuando  $n \rightarrow \infty$  en la igualdad  $A_n(\alpha x + \beta y) = \alpha A_n x + \beta A_n y$ ,  $\alpha, \beta \in \mathbb{K}$  y  $x, y \in \mathbb{X}$ . Si tomamos ahora el límite  $p \rightarrow \infty$  en la desigualdad de antes y usamos que  $X_{n+p} = A_{n+p}x \rightarrow Tx$ , obtenemos

$$\|Tx - A_n x\| \leq \|x\| \frac{\|A\|^{n+1}}{1 - \|A\|},$$

de donde se sigue que  $T - A_n$  es un operador lineal acotado, luego  $T$  lo será (¿por qué?). De la desigualdad anterior se sigue además que

$$\frac{\|Tx - A_n x\|}{\|x\|} \leq \frac{\|A\|^{n+1}}{1 - \|A\|} \implies \|T - A_n\| \leq \frac{\|A\|^{n+1}}{1 - \|A\|} \xrightarrow{n \rightarrow \infty} 0 \implies \lim_{n \rightarrow \infty} A_n = T.$$

Probemos ahora que  $T = (I - A)^{-1}$ . Ante todo notemos que  $I - A$  es un operador lineal y acotado y por tanto continuo<sup>28</sup>. Calculemos el producto

$$(I - A)Tx = (I - A) \lim_{n \rightarrow \infty} A_n x = \lim_{n \rightarrow \infty} (I - A)A_n x = \lim_{n \rightarrow \infty} (A_n x - A \cdot A_n x) = \lim_{n \rightarrow \infty} (x - A^{n+1}x).$$

Pero  $\|A^{n+1}x\| \leq \|A\|^{n+1}\|x\| \xrightarrow{n \rightarrow \infty} 0$ , luego  $A^{n+1}x \xrightarrow{n \rightarrow \infty} 0$ , y por tanto  $(I - A)Tx = x$  para todo  $x$ . La prueba de que  $T(I - A)x = x$  para todo  $x$  es análoga y la omitiremos.  $\square$

**Definición A.2.20** Denotaremos por  $\mathcal{L}(\mathbb{X})$  el conjunto de todos los operadores lineales  $A$  en  $\mathbb{X}$ , i.e.,  $A : \mathbb{X} \mapsto \mathbb{X}$ .

Nótese que la prueba del Teorema A.2.19 se puede adaptar para probar que toda sucesión de Cauchy de operadores  $T_n \in \mathcal{L}(\mathbb{X})$  es convergente. Si  $T_n$  es de Cauchy entonces  $\forall \varepsilon > 0$ , existe un  $N \in \mathbb{N}$  tal que  $\forall n \in \mathbb{N}$ ,  $n > N$  y  $\forall p \in \mathbb{N}$ ,  $\|T_{n+p} - T_n\| < \varepsilon$ . Luego la sucesión  $X_n := T_n x$  es de Cauchy

$$\|T_{n+p}x - T_n x\| \leq \|T_{n+p} - T_n\| \|x\| \leq \varepsilon \|x\| \implies \|T_{n+p}x - T_n x\| \xrightarrow{n \rightarrow \infty} 0.$$

Como  $\mathbb{X}$  es completo  $T_n x$  tendrá un límite  $y$  para cada  $x$ , i.e., podemos definir el operador  $T : \mathbb{X} \mapsto \mathbb{X}$  por  $Tx = y$ . De la misma forma que en la prueba del Teorema A.2.19 se sigue que  $T$  es lineal. Para probar que es acotado usamos la desigualdad anterior:

$$\|T_{n+p}x - T_n x\| \leq \|T_{n+p} - T_n\| \|x\| \implies \frac{\|T_{n+p}x - T_n x\|}{\|x\|} \leq \varepsilon \implies \frac{\|Tx - T_n x\|}{\|x\|} \leq \varepsilon$$

donde hemos tomando el límite  $p \rightarrow \infty$  y usado la continuidad de la norma. I.e.,  $T - T_n$  es acotado y por tanto  $T$  lo es ( $T_n$  es acotado). Además, de lo anterior se sigue también, tomando el supremo en  $x$ , que  $\|T - T_n\| \leq \varepsilon$  para todo  $n > N$ , i.e.,  $T_n \xrightarrow{n \rightarrow \infty} T$ . Lo anterior nos conduce al siguiente teorema:

<sup>28</sup>Aquí usamos el hecho conocido de que toda aplicación lineal  $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$  de un espacio normado (Hilbert)  $\mathbb{X}$  a otro espacio normado (Hilbert)  $\mathbb{Y}$  tiene las propiedades: 1.  $T$  es continua si y sólo si  $T$  es acotada y 2. Si  $T$  es continua en algún  $x_0 \in \mathcal{D}(T)$ ,  $T$  es continua en todo su dominio  $\mathcal{D}(T)$ .

**Teorema A.2.21** *El espacio  $\mathcal{L}(\mathbb{X})$  es un espacio de Banach respecto a la norma de los operadores.*

Una aplicación directa del Teorema A.2.19 es el siguiente resultado:

**Teorema A.2.22** *Sea  $\mathbb{X}$  un espacio de Banach y sea  $\mathcal{L}(\mathbb{X})$  el conjunto de todos los operadores lineales  $A : \mathbb{X} \mapsto \mathbb{X}$ . El conjunto  $E \subset \mathcal{L}(\mathbb{X})$  de los operadores invertibles en  $\mathbb{X}$  es abierto en  $\mathcal{L}(\mathbb{X})$ .*

Demostración: Sea  $A \in \mathcal{L}(\mathbb{X})$  un operador invertible. Definamos la bola  $B(A, 1/\|A^{-1}\|)$ . Cualquiera sea  $B \in B(A, 1/\|A^{-1}\|)$  tendremos que

$$\|B - A\| \leq \frac{1}{\|A^{-1}\|} \implies \|(B - A)A^{-1}\| \leq \|B - A\| \|A^{-1}\| < 1,$$

luego el operador  $I + (B - A)A^{-1} = BA^{-1}$  es invertible, por tanto el operador  $B = (BA^{-1})A$  lo será, i.e., todo operador  $B \in B(A, 1/\|A^{-1}\|)$  es invertible, por tanto para cualquiera sea  $A \in E$  existe una bola  $B(A, 1/\|A^{-1}\|) \subset E$  centrada en  $A$ , luego  $E$  es abierto.  $\square$

**Definición A.2.23** *Un operador lineal  $A : \mathbb{X} \mapsto \mathbb{Y}$ ,  $\mathbb{X}$  e  $\mathbb{Y}$  espacios de Banach es compacto si para toda sucesión acotada  $(x_n)_n$  de  $\mathbb{X}$ , la sucesión  $(Ax_n)_n$  de  $\mathbb{Y}$  tiene una subsucesión convergente.*

Nótese que si  $A$  es compacto,  $A$  es acotado pues en caso contrario existiría una sucesión acotada  $(x_n)_n$  tal que  $\|Ax_n\| \xrightarrow{n \rightarrow \infty} \infty$  y entonces la sucesión  $(Ax_n)_n$  no tendría una subsucesión convergente.

Se puede probar que cualquier operador  $A : \mathbb{H} \mapsto \mathbb{H}$ , siendo  $\mathbb{H}$  un espacio de Banach de dimensión finita es compacto, no obstante la compacidad no es trivial en el caso infinito tal y como muestra el siguiente ejemplo: El operador identidad  $I : \mathbb{H} \mapsto \mathbb{H}$ ,  $\mathbb{H}$  espacio de Hilbert de dimensión infinita no es compacto. Para probarlo, escojamos una sucesión orthonormal  $(x_n)_n$  en  $\mathbb{H}$ . Como  $\|x_n - x_m\|^2 = 2$  para todos  $n, m \in \mathbb{N}$ , entonces la sucesión  $Ix_n = x_n$  no contiene subsucesiones de Cauchy y por tanto no tiene subsucesiones convergentes.

**Teorema A.2.24** *Sea  $A \in \mathcal{L}(\mathbb{H})$  un operador compacto y  $B \in \mathcal{L}(\mathbb{H})$  uno acotado,  $\mathbb{H}$  espacio de Hilbert. Entonces los operadores  $AB$  y  $BA$  son compactos.*

Demostración: Sea  $(x_n)_n$  una sucesión acotada de  $\mathbb{H}$ . Como  $B$  es acotado, entonces la sucesión  $(Bx_n)_n$  es acotada y como  $A$  es compacto, de la sucesión  $(ABx_n)_n$  se puede extraer una subsucesión convergente, luego  $AB$  es compacto.

Sea ahora  $(x_n)_n$  una sucesión acotada de  $\mathbb{H}$ . Como  $A$  es compacto, existe una subsucesión  $(Ax_{n_k})_k$  de  $(Ax_n)_n$  que converge. Ahora bien,  $B$  es acotado, luego es continuo (ver Teorema A.1.30), por lo que la subsucesión  $(BAx_{n_k})_k$  converge, luego  $BA$  es compacto.  $\square$

**Definición A.2.25** *Un operador  $A : \mathbb{H} \mapsto \mathbb{H}$ ,  $\mathbb{H}$  espacio de Hilbert, se llama hermítico o autoadjunto si  $A = A^*$ , i.e.,*

$$\langle Ax, y \rangle = \langle x, A^*y \rangle, \quad \forall x, y \in \mathbb{H}.$$

### A.2.2. Teoría Espectral de operadores compactos autoadjuntos

En un espacio de dimensión finita podemos definir el espectro de un operador como el conjunto de los autovalores de su correspondiente matriz (en alguna base<sup>29</sup>), i.e., es el conjunto de los números complejos  $\lambda$  tales que

$$Ax = \lambda x, \quad x \neq 0. \tag{A.2.8}$$

Puesto que para cualquier matriz  $n \times n$  existen  $n$  autovalores, en el caso finito es relativamente simple de estudiar. No así el caso infinito.

Por ejemplo el operador desplazamiento ya visto antes  $S : l^2 \mapsto l^2$ ,  $S(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$ , no tiene autovalores pues la igualdad  $Sx = \lambda x$  implica  $x = 0$ .

Así se precisa de una definición más general.

**Definición A.2.26** *Sea  $\mathbb{X}$  un espacio de Banach y  $A$  una aplicación lineal  $A : \mathbb{X} \mapsto \mathbb{X}$ . El espectro de  $A$ , que denotaremos por  $\sigma(A)$  es el conjunto de números complejos tales que el operador  $(\lambda I - A)$  es no invertible, i.e., no existe  $(\lambda I - A)^{-1}$ .*

De la definición anterior se sigue que en dimensión finita  $\sigma(A)$  es el conjunto de todos los autovalores de  $A$ .

**Teorema A.2.27**  *$\sigma(A)$  es un compacto de  $\mathbb{C}$  (conjunto cerrado y acotado de  $\mathbb{C}$ ) contenido en el interior del disco cerrado  $D = \{z; |z| \leq \|A\|\}$ .*

Demostración: Sea  $\mathcal{L}(\mathbb{X})$  el conjunto de todos los operadores lineales de  $\mathbb{X}$  en  $\mathbb{X}$ ,  $\mathbb{X}$  espacio de Banach. Definamos el operador  $F : \mathbb{C} \mapsto \mathcal{L}(\mathbb{X})$ ,  $F(\lambda) = \lambda I - A$ . Esta claro que  $F$  es lineal. Como para todo  $\mu, \lambda \in \mathbb{C}$ ,  $\|F(\mu) - F(\lambda)\| = |\mu - \lambda|$ , entonces  $F$  es acotado y por el Teorema A.1.30  $F$  es continuo (en  $\lambda$ ) en la norma de los operadores.

Sea  $E \subset \mathcal{L}(\mathbb{X})$  el espacio de las aplicaciones invertibles. Si  $\lambda \in \sigma(A)$ , entonces  $\lambda I - A$  no es invertible, luego para  $\lambda \in \sigma(A)$  no existe el inverso de  $F$ , i.e.,  $\sigma(A)$  es la imagen inversa  $F^{-1}$  de  $\mathcal{L}(\mathbb{X}) \setminus E$ ,  $\sigma(A) = F^{-1}(\mathcal{L}(\mathbb{X}) \setminus E)$ . Como  $E$  es abierto (Teorema A.2.22) entonces  $\mathcal{L}(\mathbb{X}) \setminus E$  es cerrado, y como  $F$  es continua entonces por la Proposición A.1.12,  $F^{-1}(\mathcal{L}(\mathbb{X}) \setminus E)$  es cerrado, luego  $\sigma(A)$  lo será. Escojamos ahora  $\lambda$  tal que  $|\lambda| > \|A\|$ , entonces  $\|\lambda^{-1}A\| < 1$  así que  $(I - \lambda^{-1}A)$  es invertible y por tanto  $\lambda I - A$  lo será. Entonces para dichos  $\lambda$  tendremos que  $\lambda \notin \sigma(A)$ , i.e.,  $\sigma(A) \subset B(0, \|A\|)$ . Es decir,  $\sigma(A)$  es cerrado y acotado, por lo tanto es compacto (en dimensión finita cerrado y acotado es equivalente a compacto) contenido en el interior de la bola cerrada  $B(0, \|A\|)$ .  $\square$

**Teorema A.2.28** *Sea  $\mathbb{H}$  un espacio de Hilbert y  $A$  una aplicación lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  hermítica (autoadjunta). Entonces todos los autovalores de  $A$  (si los tiene) son reales. Además los autovectores correspondientes a autovalores distintos son ortogonales.*

Demostración: Sea  $\lambda$  un autovalor de  $A$  y  $x$  su correspondiente autovector, que sin pérdida de generalidad asumiremos normalizado  $\|x\| = 1$ . Entonces, usando que  $Ax = \lambda x$  y que  $A$  es hermítico, tenemos

$$\langle Ax, x \rangle = \langle x, Ax \rangle \implies \lambda \|x\| = \bar{\lambda} \|x\| \implies \lambda = \bar{\lambda} \implies \lambda \in \mathbb{R}.$$

Sea  $Ax_1 = \lambda_1 x_1$  y  $Ax_2 = \lambda_2 x_2$ . Entonces como  $A$  es hermítico

$$\langle Ax_1, x_2 \rangle = \langle x_1, Ax_2 \rangle \implies \lambda_1 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle \implies (\lambda_1 - \lambda_2) \langle x_1, x_2 \rangle = 0,$$

de donde se sigue que si  $\lambda_1 \neq \lambda_2$  entonces  $x_1$  y  $x_2$  son ortogonales.  $\square$

<sup>29</sup>Es un hecho conocido del álgebra lineal que los autovalores de un operador no dependen de la base escogida.

**Teorema A.2.29** *Sea  $A$  un operador compacto en un espacio de Hilbert y  $(\phi_n)_n$  una sucesión ortonormal de  $\mathbb{H}$ . Entonces  $\lim_{n \rightarrow \infty} A\phi_n = 0$ .*

Demostración: Supongamos que el teorema es falso, entonces para algún  $\varepsilon > 0$  ha de existir una subsucesión  $(\phi_{k_n})_n$  tal que  $\|A\phi_{k_n}\| > \varepsilon$  para todo  $n \in \mathbb{N}$ . Como  $A$  es compacto y la sucesión  $(\phi_{k_n})_n$  es acotada, entonces hay al menos una subsucesión convergente  $(\phi_{m_n})_n$  tal que  $\lim_{n \rightarrow \infty} A\phi_{m_n} = \psi \neq 0$ . Entonces

$$0 \neq \|\psi\|^2 = \langle \psi, \psi \rangle = \lim_{n \rightarrow \infty} \langle A\phi_{m_n}, \psi \rangle = \lim_{n \rightarrow \infty} \langle \phi_{m_n}, A^*\psi \rangle = 0,$$

pues  $A^*\psi \in \mathbb{H}$  y por tanto  $\langle \phi_{m_n}, A^*\psi \rangle$  es, esencialmente, el  $m_n$  coeficiente de Fourier  $c_{m_n}$  de  $A^*\psi$  el cual sabemos que tiende a cero si  $n \rightarrow \infty$  (A.2.5).  $\square$

Como hemos visto en dimensión infinita un operador lineal  $A$  en general puede no tener autovalores. No ocurre así con los operadores compactos y autoadjuntos.

**Teorema A.2.30** *Sea  $\mathbb{H}$  un espacio de Hilbert y  $A$  un operador lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  autoadjunto (hermítico) y compacto. Entonces  $\lambda = \|A\|$  o  $\lambda = -\|A\|$  es un autovalor de  $A$ .*

Demostración: Ver, por ejemplo, L. Debnath & P. Mikusinski - Introduction to Hilbert spaces with applications, Academic Press, 1990. Teorema 4.9.8 pág. 182.  $\square$

Del teorema anterior se sigue que todo operador compacto y autoadjunto tiene siempre al menos un autovalor. De hecho se tiene el siguiente teorema:

**Teorema A.2.31** *Sea  $\mathbb{H}$  un espacio de Hilbert separable y  $A$  una aplicación lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  autoadjunta (hermítica) y compacta. Entonces  $A$  tiene un número finito de autovalores  $\lambda_n$  reales distintos o si es infinito, entonces, es numerable y si lo ordenamos de mayor a menor  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$ .*

Demostración: Del teorema A.2.30 se sigue que el conjunto de autovalores no es vacío, luego está compuesto por un número finito o infinito de elementos (que son números reales por el Teorema A.2.28). Sea  $\lambda_1$  tal que  $|\lambda_1| = \|A\|$  y sea  $x_1$  el correspondiente autovector (normalizado a la unidad, i.e.  $\|x_1\| = 1$ ). Sea  $\mathbb{H}_1 = \mathbb{H}$  y definamos  $\mathbb{H}_2$  el espacio de todos los vectores ortogonales a  $x_1$ , i.e.,

$$\mathbb{H}_2 = \{x \in \mathbb{H} \mid \langle x, x_1 \rangle = 0\},$$

es el complemento ortogonal de  $x_1$ . Pero, para todo  $x \in \mathbb{H}_2$ ,

$$\langle Ax, x_1 \rangle = \langle x, Ax_1 \rangle = \lambda_1 \langle x, x_1 \rangle = 0$$

es decir,  $\mathbb{H}_2$  es invariante respecto a la acción de  $A$ . Sea  $A|_{\mathbb{H}_2}$  la restricción de  $A$  al espacio  $\mathbb{H}_2$ . Si  $A|_{\mathbb{H}_2}$  no es el operador nulo, entonces podremos aplicar el mismo razonamiento de antes, por lo que existirá  $\lambda_2$  y  $x_2$  tales que  $|\lambda_2| = \|A|_{\mathbb{H}_2}\|$ , donde además es obvio que  $|\lambda_2| \leq |\lambda_1|$ . Así podemos seguir hasta que en cierto paso  $k$   $A|_{\mathbb{H}_k} = 0$  obteniendo la sucesión finita de autovalores  $(\lambda_k)_{k=1}^n$  con sus correspondientes autovectores normalizados<sup>30</sup>  $(x_k)_{k=1}^n$  tales que

$$|\lambda_1| = \|A|_{\mathbb{H}_1}\| \geq |\lambda_2| = \|A|_{\mathbb{H}_2}\| \geq \dots \geq |\lambda_n| = \|A|_{\mathbb{H}_n}\|.$$

<sup>30</sup>Es importante destacar que para cada  $\lambda_k$ ,  $k = 1, 2, 3, \dots$  puede haber más de un autovector.

Si  $A|_{\mathbb{H}_k} \neq 0$  para todo  $k \in \mathbb{N}$ , entonces existen infinitos autovalores distintos  $\lambda_n$  con sus correspondientes autovectores  $x_n$  normalizados (que pueden ser más de uno), que son numerables pues  $\mathbb{H}$  es separable (¿por qué?). Ahora bien, por el Teorema A.2.29 tenemos

$$0 = \lim_{n \rightarrow \infty} \|Ax_n\|^2 = \lim_{n \rightarrow \infty} \langle Ax_n, Ax_n \rangle = \lim_{n \rightarrow \infty} |\lambda_n|^2,$$

de donde se sigue que  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$ . □

Existe otra forma de probar que  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$ . En efecto, supongamos que hay infinitos  $\lambda_n$  distintos pero que  $\lambda_n \not\xrightarrow{n \rightarrow \infty} 0$ . Entonces existe un  $\varepsilon > 0$  tal que infinitos  $\lambda_{n_k}$  son tales que  $|\lambda_{n_k}| > \varepsilon$ . Construyamos con dichos elementos una sucesión que denotaremos por  $(\lambda_k)_k$ . Como todos los elementos de  $(\lambda_k)_k$  son distintos, el Teorema A.2.28 garantiza que sus correspondientes autovectores son ortogonales  $x_k$ , i.e.,  $\langle x_{k_1}, x_{k_2} \rangle = 0$ . Si calculamos ahora la norma

$$\|Ax_{k+p} - Ax_k\|^2 = \|\lambda_{k+p}x_{k+p} - \lambda_kx_k\|^2 = |\lambda_{k+p}|^2 + |\lambda_k|^2 > 2\varepsilon^2, \quad \forall k, p \in \mathbb{N}$$

es decir, la sucesión  $(Ax_k)_k$  no contiene ninguna subsucesión de Cauchy y por tanto no contiene ninguna sucesión convergente (¿por qué?) lo que contradice que  $A$  sea un operador compacto.

Del teorema anterior se sigue además que los espacios  $\ker(\lambda_k I - A)$ , para cada  $\lambda_k, k = 1, 2, 3, \dots$  son de dimensión finita, pues en caso contrario  $\lambda_n \not\xrightarrow{n \rightarrow \infty} 0$ .

**Teorema A.2.32 (Teorema espectral)** *Sea  $\mathbb{H}$  un espacio de Hilbert y  $A$  una aplicación lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  autoadjunta y compacta. Existe una sucesión numerable (finita o infinita) de autovectores ortonormales  $(x_n)_n$  de  $\mathbb{H}$  cuya correspondiente sucesión de autovalores denotaremos por  $(\lambda_n)_n$  tales que,*

$$Ax = \sum_n \lambda_n \langle x, x_n \rangle x_n, \quad \forall x \in \mathbb{H}, \tag{A.2.9}$$

donde se tiene que:

1. En (A.2.9) aparecen todos los autovalores de  $A$ .
2. Si la sucesión  $(\lambda_n)_n$  es infinita se puede reordenar de forma que  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$ .
3. Los correspondientes espacios  $\ker(\lambda_n I - A)$ , para todo  $n = 1, 2, 3, \dots$  son de dimensión finita, siendo la dimensión de estos el número de veces que aparece un mismo  $\lambda_k$  en la fórmula (A.2.9).

Demostración: Utilizando la misma construcción que usamos en la prueba del teorema A.2.31 obtenemos una sucesión de autovalores  $(\lambda_k)_{k=1}^n$  con sus correspondientes autovectores normalizados  $(x_k)_{k=1}^n$  (para cada  $\lambda_k$  puede haber más de un autovector, en función de la *multiplicidad* del mismo) tales que

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Durante el proceso hemos construido además la cadena de subespacios invariantes respecto a  $A$

$$\mathbb{H} = \mathbb{H}_1 \supseteq \mathbb{H}_2 \supseteq \dots \supseteq \mathbb{H}_n$$

donde  $\mathbb{H}_{k+1} = \{x \in \mathbb{H}_k \mid \langle x, x_j \rangle = 0, j = 1, 2, \dots, k\}$ .

Supongamos que  $A|_{\mathbb{H}_{n+1}} = 0$ , entonces el proceso acaba (caso finito). Sea, en este caso,  $y_n = x - \sum_{k=1}^n \langle x, x_k \rangle x_k$ . Entonces

$$\langle y_n, x_j \rangle = \langle x, x_j \rangle - \sum_{k=1}^n \underbrace{\langle x, x_k \rangle \langle x_k, x_j \rangle}_{\delta_{j,k}} = \langle x, x_j \rangle - \underbrace{\langle x, x_j \rangle}_{=1} \|x_j\| = 0.$$

Luego,  $y_n \in \mathbb{H}_{n+1}$  y por tanto  $Ay_n = 0$ , i.e.,

$$0 = Ay_n = Ax - \sum_{k=1}^n \langle x, x_k \rangle Ax_k \implies Ax = \sum_{k=1}^n \lambda_k \langle x, x_k \rangle x_k$$

como se quería probar.

El caso infinito es similar. Ante todo notemos que en este caso tenemos una sucesión  $\lambda_n$  tal que  $\lambda_n \xrightarrow{n \rightarrow \infty} 0$  (ver Teorema A.2.31). Definamos nuevamente el vector  $y_n = x - \sum_{k=1}^n \langle x, x_k \rangle x_k$ . Usando (A.2.3) se sigue que  $\|y_n\|^2 \leq \|x\|^2$ . Además, como ya vimos  $y_n \in \mathbb{H}_{n+1}$ , luego, usando que  $Ay_n = A|_{\mathbb{H}_{n+1}} y_n$ , y  $\|A|_{\mathbb{H}_{n+1}}\| = |\lambda_{n+1}|$  obtenemos

$$\|Ay_n\| \leq |\lambda_{n+1}| \|x\| \xrightarrow{n \rightarrow \infty} 0 \implies \lim_{n \rightarrow \infty} Ay_n = 0,$$

de donde se sigue la igualdad buscada

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, x_n \rangle x_n, \quad \forall x \in \mathbb{H}.$$

Comprobemos ahora que todos los autovalores de  $A$  están presentes en la fórmula (A.2.9). Esta cuestión es de suma importancia pues en la prueba hemos usado los autovalores obtenidos gracias al Teorema A.2.30 por lo que procede preguntarse si existen otros autovalores de  $A$  no nulos distintos a los anteriores. Comprobemos que eso es imposible. Supongamos que existe un autovalor  $\lambda \neq 0$  distinto de los autovalores  $\lambda_k$  que aparecen en la suma (A.2.9), i.e.,  $\lambda \neq \lambda_k$ , para todo  $k$ , y sea  $x$  su correspondiente autovector normalizado a la unidad. Por el Teorema A.2.28  $x$  es ortogonal a todos los  $x_k$ , i.e.,  $\langle x, x_k \rangle = 0$  para todo  $k$ . Pero entonces, aplicando la fórmula (A.2.9) a dicho autovector  $x$  tenemos

$$\lambda x = Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, x_n \rangle x_n = \sum_{n=1}^{\infty} \lambda_n 0 x_n = 0,$$

lo que es una contradicción. Lo anterior tiene una implicación importante: Formalmente el proceso de encontrar los autovalores de  $A$  basado en el Teorema A.2.30 permite encontrarlos **todos**.

Probemos finalmente que, en caso de que  $\lambda_k \neq 0$ ,  $k = 1, 2, 3, \dots$ , la dimensión de  $\ker(\lambda_k I - A) = p$ , donde  $p$  es el número de veces que aparece  $\lambda_k$  en la fórmula (A.2.9). Por simplicidad denotemos por  $x_j^{(1)}$ ,  $\dots$ ,  $x_j^{(p)}$   $p$  autovectores linealmente independientes asociados a  $\lambda_j$ . Supongamos que  $\dim \ker(\lambda_j I - A) > p$ , entonces existe al menos un  $x \in \ker(\lambda_j I - A)$  que es linealmente independiente de los vectores  $x_j^{(1)}$ ,  $\dots$ ,  $x_j^{(p)}$  y que asumiremos, sin pérdida de generalidad, que ortogonal a ellos (¿por qué?). Entonces está claro que dicho  $x$  es ortogonal a todos los vectores  $x_k$  que aparecen en la fórmula (A.2.9) (ya sea porque los  $x_k$  corresponden a autovalores distintos a  $\lambda_j$  o bien sean los  $x_j^{(1)}$ ,  $\dots$ ,  $x_j^{(p)}$  de antes), i.e.,  $\langle x, x_k \rangle = 0$  para todo  $k$ , luego usando (A.2.9) obtenemos

$$0 \neq \lambda_j x = Ax = \sum_k \lambda_k \langle x, x_k \rangle x_k = \sum_k \lambda_k 0 x_k = 0,$$

lo cual es una contradicción. □

**Corolario A.2.33** *Sea  $\mathbb{H}$  un espacio de Hilbert separable y  $A$  una aplicación lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  auto-adjunta y compacta. Entonces existe un sistema ortogonal completo (base ortonormal) de autovectores ortonormales  $(e_n)_n$  de  $\mathbb{H}$  consistente en los correspondientes autovectores de  $A$ . Además,*

$$Ax = \sum_n \lambda_n \langle x, e_n \rangle e_n, \quad \forall x \in \mathbb{H},$$

donde  $(\lambda_n)_n$  es la correspondiente sucesión de autovalores asociados a  $(e_n)_n$ .

Demostración: Del Teorema espectral A.2.32 se sigue que existe un conjunto de autovectores (finito o infinito)  $(\phi_n)_n$  tal que

$$Ax = \sum_n \lambda_n \langle x, \phi_n \rangle \phi_n, \quad \forall x \in \mathbb{H}. \quad (\text{A.2.10})$$

Asumamos que  $\lambda_n \neq 0$  (si hay algún  $\lambda_n = 0$  este se puede omitir de la suma). Como el núcleo de  $A$ ,  $\ker A \subset \mathbb{H}$  (que coincide con el subespacio vectorial generado por los autovectores correspondientes a  $\lambda = 0$ ) es a su vez un espacio de Hilbert separable (¿por qué?) existirá una sistema (numerable) ortogonal completo que denotaremos por  $(\psi_n)_n$ . Dicho sistema de vectores son autovectores correspondientes al autovalor 0. Sea ahora un autovector cualquiera  $\phi_m$  correspondiente al autovalor  $\lambda_m \neq 0$ . Entonces  $\phi_m$  será ortogonal a todos los  $\psi_n$  y el sistema  $(\phi_m)_m$  será ortogonal a  $(\psi_n)_n$ . Además de (A.2.10) se tiene que para todo  $x \in \mathbb{H}$

$$A \left( x - \sum_m \langle x, \phi_m \rangle \phi_m \right) = Ax - \sum_m \langle x, \phi_m \rangle A\phi_m = 0,$$

i.e.,  $x - \sum_m \langle x, \phi_m \rangle \phi_m \in \ker A$ , luego podemos desarrollarlo en serie de Fourier en la base de  $\ker A$   $(\psi_n)_n$  de forma que obtenemos

$$x - \sum_m \langle x, \phi_m \rangle \phi_m = \sum_n \langle x, \psi_n \rangle \psi_n \implies x = \sum_m \langle x, \phi_m \rangle \phi_m + \sum_n \langle x, \psi_n \rangle \psi_n \quad \forall x \in \mathbb{H},$$

i.e., el sistema  $(\phi_m)_m \cup (\psi_n)_n$  es completo (Teorema A.2.5) que podemos ortogonalizar utilizando el método de Gram-Schmidt obteniendo el sistema ortonormal (numerable)  $(e_n)_n$  (recordemos que para autovectores correspondientes a autovalores distintos ya tenemos la ortogonalidad, así que sólo es necesario aplicarlo a los autovectores correspondientes a un mismo autovalor), luego tendremos  $x = \sum_n \langle x, e_n \rangle e_n$  de donde se sigue el teorema.  $\square$

Como consecuencia del teorema anterior tenemos que todo operador lineal  $A : \mathbb{H} \mapsto \mathbb{H}$  autoadjunto y compacto en  $\mathbb{H}$ , espacio de Hilbert separable, se le puede hacer corresponder una matriz (finita o infinita) que además es diagonalizable y en cuya diagonal aparecen los correspondientes autovalores. Además se tiene el siguiente resultado:

**Teorema A.2.34** Sean  $A$  y  $B$  dos operadores autoadjuntos y compactos en un espacio de Hilbert separable. Si  $A$  y  $B$  conmutan, entonces tienen un sistema completo de autovectores común.

Demostración: Sea  $\lambda$  un autovalor de  $A$  y sea  $S$  el correspondiente subespacio lineal generado por los autovectores de  $A$ . Para todo  $x \in S$  tenemos  $ABx = BAx = \lambda Bx$ , i.e.,  $Bx$  es también un autovector de  $A$  correspondiente a  $\lambda$ ,  $Bx \in S$ , así que el espacio  $S$  es un subespacio vectorial de  $\mathbb{H}$  invariante respecto a  $B$  y es a su vez un espacio de Hilbert. Como  $B$  es autoadjunto y compacto, el Corolario anterior A.2.33 nos asegura que  $S$  tiene una base ortonormal de autovectores de  $B$ , que además son autovectores de  $A$  pues están en  $S$ . Repitiendo el proceso para cada uno de los subespacios  $S$  de  $A$  obtenemos para cada uno de dichos subespacios la correspondiente base de autovectores. La unión de todas ellas es la base común buscada.  $\square$



# Bibliografía

- [1] L. Debnath y P. Mikusinsk. Introduction to Hilbert spaces with applications, Academic Press, 1990.<sup>h31</sup>
- [2] Yuli Eidelman, Vitali D. Milman, Antonis Tzolomitis. Functional Analysis: An Introduction, Graduate Studies in Mathematics Vol. 66, AMS, 2004.<sup>m,h</sup>
- [3] A.N. Kolmogorov y A.V. Fomín. Elementos de la teoría de funciones y del análisis funcional. Editorial MIR, 1978. (Elements of the Theory of Functions and Functional Analysis. Dover, 1999.)<sup>m,h</sup>
- [4] E. Kreyszig. Introductory Functional Analysis with Applications. Wiley Classics Library Edition, 1989.<sup>m,n</sup>
- [5] K. Saxe. Beginning Functional Analysis. Springer, New York, 2002.<sup>h</sup>
- [6] J. Tinsley Oden y L.F. Demkowicz. Applied Functional Analysis. CRC Press, 1996.<sup>m,n</sup>
- [7] N. Young. An introduction to Hilbert Space. Cambridge University Press, 1988.<sup>n,h</sup>

---

<sup>31</sup>Se recomienda para el tema de espacios métricos, normados y de Hilbert los libros marcados con “m”, “n” y “h”, respectivamente.