

# Ortogonalidad y aplicaciones

Renato Álvarez Nodarse

Dpto. Análisis Matemático, Universidad de Sevilla.

17 de noviembre de 2016

<http://euler.us.es/~renato/>

## Índice

<b>1. Procesos aleatorios y cadenas de Markov</b>	<b>3</b>
1.1. Introducción y primeras propiedades . . . . .	3
1.2. Cadenas de Markov discretas . . . . .	5
1.3. El caso de dos estados . . . . .	7
1.4. La cadena infinita: la caminata del borracho . . . . .	10
1.5. La ecuación de renuevo . . . . .	12
1.6. Polinomios ortogonales . . . . .	14
1.7. Caminatas aleatorias y polinomios ortogonales . . . . .	19
1.8. Ejemplos . . . . .	24
1.8.1. Ejemplo 1 . . . . .	24
1.8.2. Ejemplo 2 . . . . .	25
1.9. El modelo de urnas de Ehrenfest . . . . .	26
<b>2. Resolución de modelos cuánticos</b>	<b>31</b>
2.1. Breve introducción histórica . . . . .	31
2.2. Las matemáticas de la Mecánica Cuántica . . . . .	37
2.2.1. Los postulados de la Mecánica Cuántica en $L^2$ . . . . .	38
<b>3. Resolviendo la ecuación de Schrödinger: El método de Nikiforov-Uvarov</b>	<b>39</b>
3.1. La ecuación hipergeométrica generalizada . . . . .	39
3.2. La ecuación diferencial hipergeométrica . . . . .	41
3.2.1. La propiedad de hipergeometricidad y la fórmula de Rodrigues	41

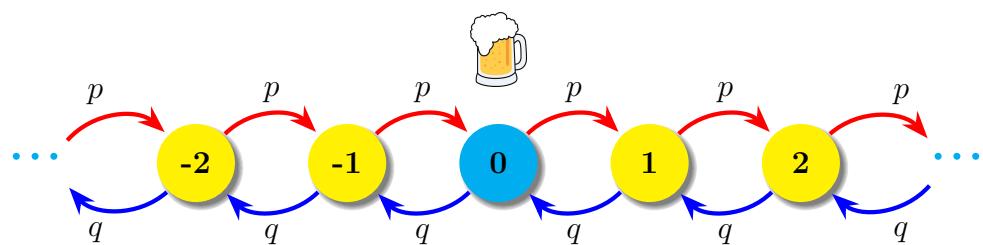
3.2.2.	Ortogonalidad y relación de recurrencia . . . . .	44
3.2.3.	Consecuencias de la fórmula de Rodrigues . . . . .	48
3.2.4.	Los polinomios de Hermite, Laguerre y Jacobi . . . . .	50
3.3.	Resolución de la ecuación de Schrödinger . . . . .	54
3.3.1.	El oscilador armónico cuántico unidimensional . . . . .	54
3.3.2.	La ecuación de Schrödinger en un potencial central . . . . .	55
3.3.3.	Resolviendo la parte radial de la ecuación de Schrödinger . . . . .	56
3.3.4.	El átomo de hidrógeno . . . . .	57

**Agradecimientos:** La primera parte de estas notas (sección 1) no hubiese sido posible sin la inestimable ayuda de mis colegas Alberto Grünbaum (U. California, Berkeley) y Luis Velázquez Campoy (U. Zaragoza).

# 1. Procesos aleatorios y cadenas de Markov

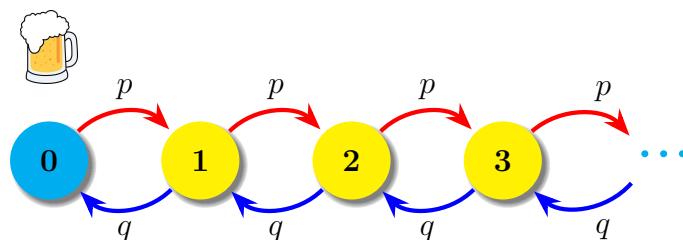
## 1.1. Introducción y primeras propiedades

¿Qué es un camino aleatorio (random walk)? Comenzaremos con ejemplo muy simple. Imaginemos una calle (finita o infinita) con bares colocados cada cierta distancia. Supongamos que una persona (que llamaremos cariñosamente “borracho”) está en su bar favorito que denominaremos por el número 0. Nuestro borracho se bebe una cerveza en el bar cero y luego con cierta probabilidad  $p$  se mueve al bar de la derecha y con probabilidad  $q$  se mueve al bar de la izquierda, de forma que  $p + q = 1$ . Así cada cierto tiempo  $\Delta t$  nuestro borracho estará en su siguiente bar. La siguiente figura ilustra el proceso en el caso de una calle doblemente infinita:

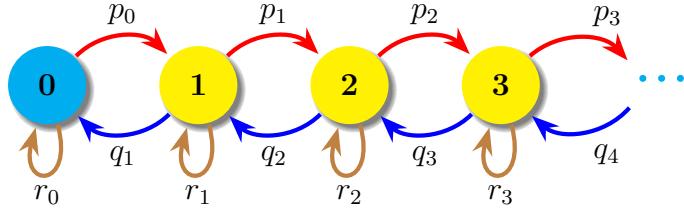


El proceso anterior es una *caminata aleatoria*. Para nuestro borracho las siguientes preguntas aparecen de forma natural: ¿qué probabilidad hay de que pueda regresar, digamos en  $k$  pasos, al bar inicial? ¿qué probabilidad hay de regresar en algún momento de tiempo? Si esta última probabilidad es 1, ¿cuál es el tiempo que debe esperar el borracho para regresar a su bar favorito?

¿Qué ocurrirá si la calle es semi infinita? Es decir, el siguiente caso:

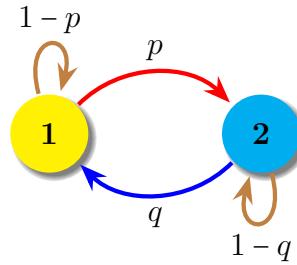


Otro proceso interesante que generaliza la caminata anterior es cuando asignamos cierta probabilidad de quedarse en cada estado (en el caso de nuestro borracho dicha probabilidad era siempre cero) cuya representación diagramática, en el caso semiinfinito, es la siguiente:



Vamos a intentar responder a las preguntas antes formuladas para este tipo de problemas de varias formas, en particular, usando la teoría de polinomios ortogonales.

Nuestro ejemplo de partida será uno mucho más simple. Asumamos que tenemos dos estados posibles 1 y 2 y que las transiciones entre ambos se representan por el diagrama



Vamos a hacerle corresponder a cada diagrama una matriz  $\mathbb{P}$  de la siguiente forma: la entrada  $p_{ij}$  de  $\mathbb{P}$  representa la probabilidad de pasar del estado  $i$  al  $j$ . Así para nuestro proceso con dos estados tenemos

$$\mathbb{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 1 - p & p \\ q & 1 - q \end{bmatrix}.$$

Dicha matriz  $\mathbb{P}$  la denominaremos *matriz de probabilidades de trascisión*, y en general toma las formas

$$\mathbb{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \dots & p_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \dots & p_{nn} \end{bmatrix}, \quad \sum_{j=1}^n p_{ij} = 1, \quad \forall i \geq 1,$$

para un proceso con un número finito de estados (cadena finita),

$$\mathbb{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} & \cdots \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \sum_{j \geq 1} p_{ij} = 1, \quad \forall i \geq 1,$$

para el caso de una cadena semiinfinita, o

$$\mathbb{P} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & p_{-10} & p_{-11} & p_{-12} & p_{-13} & \cdots & p_{-1n} & \cdots \\ \cdots & p_{00} & p_{01} & p_{02} & p_{03} & \cdots & p_{0n} & \cdots \\ \cdots & p_{10} & p_{11} & p_{12} & p_{13} & \cdots & p_{1n} & \cdots \\ \cdots & p_{20} & p_{21} & p_{22} & p_{23} & \cdots & p_{2n} & \cdots \\ \cdots & p_{30} & p_{31} & p_{32} & p_{33} & \cdots & p_{3n} & \cdots \\ \cdots & \vdots & \vdots & \vdots & \vdots & \ddots & \cdots & \cdots \\ \cdots & p_{n0} & p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} & \cdots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \sum_{j \in \mathbb{Z}} p_{ij} = 1, \quad \forall i \geq 1,$$

si la cadena es doblemente infinita. Dichas matrices se conocen por el nombre de matrices estocásticas.

**Ejercicio 1.1.** Prueba que si  $\mathbb{P}$  es una matriz estocástica, entonces  $\mathbb{P}^n$ ,  $\forall n \in \mathbb{N}$  también lo es.

## 1.2. Cadenas de Markov discretas

Vamos ahora a conectar el concepto de cadena de Markov discreta (tiempo discreto) con los esquemas del apartado anterior vía la matriz de probabilidades de trasición  $\mathbb{P}$ .

Sea un conjunto numerable  $\mathbb{I}$  y sea  $\lambda$  una medida sobre  $\mathbb{I}$ , i.e.,  $\lambda = \{\lambda_i, i \in \mathbb{I}\}$ ,  $0 \leq \lambda_i < +\infty$  para todo  $i \in \mathbb{I}\}$ . Asumiremos que  $\lambda$  es una medida de probabilidad, i.e.,  $\sum_{i \in \mathbb{I}} \lambda_i = 1$ . En adelante  $\mathbb{I}$  será un conjunto finito de valores enteros, el conjunto de los números naturales  $\mathbb{N}$  o el de los enteros  $\mathbb{Z}$ .

Sea  $X$  una variable aleatoria con valores sobre  $\mathbb{I}$ . Definiremos<sup>1</sup>  $\lambda_i = P(X = i)$  como la probabilidad de que la variable  $X$  tome el valor  $i \in \mathbb{I}$ . Como  $\sum_{i \in \mathbb{I}} P(X = i) = 1$ ,  $\lambda$  es una medida de probabilidad que denominaremos distribución de  $X$ .

---

<sup>1</sup>Para una definición más rigurosa ver [6, pág. 1].

Diremos que  $(X_n)_{n \geq 0}$  es una cadena de Markov discreta con una distribución inicial  $\lambda_0$  si

$$\begin{aligned} P(X_0) &= \lambda_0, \\ P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) &= p_{i_n i_{n+1}} = \mathbb{P}_{i_n i_{n+1}}, \end{aligned}$$

es decir,  $X_0$  tiene la distribución  $\lambda_0$  y la probabilidad de que  $X_{n+1} = i_{n+1}$  condicionada a que  $X_0 = i_0, \dots, X_n = i_n$  sólo depende del valor de  $X_n$  en el *paso* anterior (o sea, depende del último suceso y no de los anteriores). Además dicha probabilidad está determinada justamente por la matriz de probabilidades de transición  $\mathbb{P}$ , siendo su distribución la fila  $i$ -ésima, i.e.,  $(p_{ij})_{j \in \mathbb{I}}$ .

De lo anterior se deduce que

$$P(X_0 = i_0, X_1 = i_1) = \lambda_0 p_{i_0 i_1},$$

y, en general, que, para todos  $i_0, i_1, \dots, i_n \in \mathbb{I}$ ,

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \lambda_0 p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}. \quad (1.1)$$

La propiedad anterior caracteriza a los procesos de Markov [6, Teorema 1.1.1]. Más aún, a partir de la definición de proceso de Markov y la propiedad anterior se deduce el siguiente [6]

**Teorema 1.2.** *Sea  $(X_n)_{n \geq 0}$  una cadena de Markov con distribución inicial  $\lambda$ . Entonces la cadena  $(X_{m+n})_{n \geq 0}$  condicionada a que  $X_m = i$  es una cadena de Markov con distribución inicial  $\delta_i = \{\delta_{ij}, j \in \mathbb{I}\}$ , donde  $\delta_{kl} = 1$  si  $k = l$  y 0 en otro caso.*

Escribamos la distribución inicial  $\lambda$  mediante el vector  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n \dots)$  y sea  $\mathbb{P}$  la matriz de probabilidades de transición, que sabemos que es estocástica. Definiremos de la forma habitual para el caso de dimensión finita los productos  $\lambda\mathbb{P}$ ,  $\mathbb{P} \cdot \mathbb{P} = \mathbb{P}^2, \dots, \mathbb{P}^n$ , y lo extenderemos de manera *obvia* al caso de dimensión infinita. En adelante denotaremos por  $p_{ij}^{(n)}$  los elementos de la matriz  $\mathbb{P}^n$ .

El siguiente teorema se puede interpretar como la *falta de memoria* de un proceso de Markov.

**Teorema 1.3.** *Sea  $(X_n)_{n \geq 0}$  una cadena de Markov con distribución inicial  $\lambda$  y  $\mathbb{P}$  una matriz de probabilidades de transición. Entonces, para todos  $n, m \geq 0$*

- (i)  $P(X_n = j) = (\lambda \mathbb{P}^n)_j$
- (ii)  $P(X_{n+m} = j | X_m = i) = p_{ij}^{(n)}.$

*Demostración.* Para probar (i) usamos la propiedad de Markov (1.1)

$$\begin{aligned} P(X_n = j) &= \sum_{i_0, i_1, \dots, i_{n-1}} P(X_0 = i_0, X_1 = i_1, \dots, X_n = j) \\ &= \sum_{i_0, i_1, \dots, i_{n-1}} \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} = \sum_{i_0} \lambda_{i_0} \sum_{i_1, \dots, i_{n-1}} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \\ &= \sum_{i_0} \lambda_{i_0} p_{i_0 j}^{(n)} = (\lambda \mathbb{P}^n)_j. \end{aligned}$$

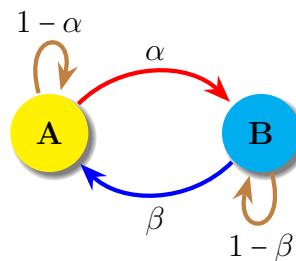
Para probar (ii) basta usar el teorema 1.2 de donde se sigue que  $\lambda = \delta_i$  que sustituyendo en (i) nos da el resultado.  $\square$

En efecto, el teorema asegura que si la variable tomó el valor  $i$  en cierto paso  $m$  la probabilidad de que dicha variable tome el valor  $j$  tras  $n$  pasos es independiente de lo que haya ocurrido antes del paso  $m$ .

Nótese además que del teorema anterior se deduce que la probabilidad de encontrar el sistema en el estado  $j$  habiendo partido del estado  $i$  en  $n$  pasos es justo  $p_{ij}^{(n)}$ .

### 1.3. El caso de dos estados

Vamos a ver un ejemplo muy simple que se puede resolver analíticamente. Sea un sistema constituido con dos estados  $A$  y  $B$  cuyo diagrama es el siguiente:



y cuya correspondiente matriz de probabilidades de transición tiene la forma

$$\mathbb{P} = \begin{bmatrix} p_{AA} & p_{AB} \\ p_{BA} & p_{BB} \end{bmatrix} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}, \quad \alpha, \beta \in [0, 1]. \quad (1.2)$$

Vamos a calcular  $\mathbb{P}^n$ , la matriz de probabilidades de transición en  $n$ -pasos. Para ello notamos que

$$\mathbb{P}^{n+1} = \mathbb{P}^n \cdot \mathbb{P} = \begin{bmatrix} p_{AA}^{(n)} & p_{AB}^{(n)} \\ p_{BA}^{(n)} & p_{BB}^{(n)} \end{bmatrix} \begin{bmatrix} p_{AA} & p_{AB} \\ p_{BA} & p_{BB} \end{bmatrix}.$$

Calculemos, por ejemplo,  $p_{AA}^{(n)}$ . De la ecuación matricial se sigue que

$$p_{AA}^{(n+1)} = (1 - \alpha)p_{AA}^{(n)} + \beta p_{AB}^{(n)}.$$

Usando que  $p_{AB}^{(n)} = 1 - p_{AA}^{(n)}$ , pues  $\mathbb{P}^n$  es una matriz estocástica, tenemos

$$p_{AA}^{(n+1)} = (1 - \alpha - \beta)p_{AA}^{(n)} + \beta, \quad p_{AA}^{(0)} = p_{AA} = 1 - \alpha.$$

La ecuación anterior es una ecuación en diferencias lineal con respecto a  $n$ . Su solución general es

$$p_{AA}^{(n)} = \frac{\beta}{\alpha + \beta} + C(1 - \alpha - \beta)^n, \quad n \geq 0,$$

que, usando la condición inicial  $p_{AA}^{(0)} = p_{AA} = 1 - \alpha$  nos da

$$p_{AA}^{(n)} = \frac{\beta + \alpha(1 - \alpha - \beta)^n}{\alpha + \beta}, \quad n \geq 0. \quad (1.3)$$

**Ejercicio 1.4.** *Calcula los restantes elementos de  $\mathbb{P}^n$ .*

El valor que hemos encontrado  $p_{AA}^{(n)}$  nos da la probabilidad de regresar al estado  $A$  tras  $n$  pasos. No obstante está claro que esta probabilidad incluye la opción de haber regresado muchas veces antes, es por ello que nos interesa calcular la probabilidad de regresar *por primera vez* al estado  $A$  tras  $n$  pasos. Este cálculo no es inmediato como en anterior, pero en nuestro ejemplo se puede realizar sin muchos problemas. Vamos a denotar por  $q_{AA}^{(n)}$  la probabilidad de regresar al estado  $A$  por primera vez en  $n$  pasos. Es fácil comprobar que

$$q_{AA}^{(1)} = 1 - \alpha, \quad q_{AA}^{(2)} = \alpha\beta, \quad q_{AA}^{(3)} = \alpha(1 - \beta)\beta, \quad q_{AA}^{(4)} = \alpha(1 - \beta)^2\beta, \dots$$

La probabilidad de regresar por primera vez a  $A$  será entonces

$$q_{AA} = \sum_{n=1}^{\infty} q_{AA}^{(n)} = 1 - \alpha + \alpha\beta + \alpha(1 - \beta)\beta + \alpha(1 - \beta)^2\beta + \dots = 1.$$

**Definición 1.5.** *Un estado  $i$  de un proceso de Markov se llama recurrente si la probabilidad de regresar a dicho estado (en algún momento) es igual a 1. En caso contrario se llama transitorio.*

Dado un estado recurrente, una pregunta natural es saber el tiempo de espera para regresar al mismo. Como nuestro sistema es en *tiempo discreto* (vamos paso a paso) entonces el tiempo medio de retorno a cierto estado  $i$  es, por definición

$$\tau_{ii} = \sum_{n=1}^{\infty} nq_{ii}^{(n)} + \infty(1 - q_{ii}). \quad (1.4)$$

En esta definición está implícito que el tiempo  $\tau_{AA}$  siempre es infinito en el caso cuando la probabilidad de primeros retornos no es 1 (¿por qué?).

Así, en nuestro ejemplo

$$\tau_{AA} = \sum_{n=1}^{\infty} nq_{AA}^{(n)} + \overbrace{\infty(1-q_{AA})}^{=0},$$

luego

$$\tau_{AA} = \sum_{n=1}^{\infty} nq_{AA}^{(n)} = (1-\alpha) + 2\alpha\beta + 3\alpha(1-\beta)\beta + 4\alpha(1-\beta)^2\beta + \dots.$$

Para calcular la suma anterior usamos que

$$1 + x + x^2 + x^3 + \dots = \frac{1}{1-x} \implies 1 + 2x + 3x^2 + 4x^3 + \dots = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2},$$

de donde se sigue que  $\tau_{AA} = 1 + \frac{\alpha}{\beta}$ . Nótese que en principio  $\tau_{AA} \in [1, \infty)$ .

**Definición 1.6.** *Un estado recurrente con tiempo medio de retorno finito se denomina positivamente recurrente.*

Vamos a intentar resolver el problema del retorno a  $A$  usando únicamente de las matrices  $\mathbb{P}$ . Denotemos por  $p_{11}^{(n)} = p_{AA}^{(n)}$  (a  $A$  le corresponderá en índice 1 y a  $B$  el 2). Entonces

$$p_{11}^{(n)} = \sum_{i_1, i_2, \dots, i_{n-1}=1,2} p_{1i_1} p_{i_1 i_2} \cdots p_{i_{n-1} 1}$$

mientras que la probabilidad de los primeros retornos a  $A$  será

$$q_{11}^{(n)} = \sum_{i_1, i_2, \dots, i_{n-1} \neq 1} p_{1i_1} p_{i_1 i_2} \cdots p_{i_{n-1} 1}.$$

¿Cómo podemos realizar la última operación mediante la multiplicación de matrices?

Está claro que el elemento  $p_{11}^{(2)}$  de  $\mathbb{P}^2$  nos da la probabilidad de retornar a 1 en dos pasos, cuyo cálculo directo nos da  $p_{11}^2 + p_{12}p_{21}$ . Ahora bien, si queremos tener la probabilidad del primer retorno hay que excluir el sumando  $p_{11}^2$ . Como

$$q_{11}^{(2)} = \sum_{i_1 \neq 1} p_{1i_1} p_{i_1 1} = p_{12}p_{21} = \alpha\beta,$$

podemos usar la matriz  $Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ . En efecto si multiplicamos

$$\mathbb{P} \cdot Q = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & p_{12} \\ 0 & p_{22} \end{bmatrix}$$

nos queda una matriz donde toda la primera columna es cero, luego si hacemos el producto  $\mathbb{P} \cdot Q \cdot \mathbb{P}$  habremos excluido la posibilidad de retornar a 1 en el primer paso (es decir, de habernos quedarnos en el estado 1)

$$\mathbb{P} \cdot Q \cdot \mathbb{P} = \begin{pmatrix} p_{12}p_{21} & p_{12}p_{22} \\ p_{21}p_{22} & p_{22}^2 \end{pmatrix}$$

pues el término  $p_{11}$  ha sido excluido. Continuando este razonamiento es fácil comprobar que para calcular  $q_{11}^{(n)}$  tenemos que calcular el primer elemento de la matriz

$$\underbrace{\mathbb{P} \cdot Q \cdot \mathbb{P} \cdot Q \cdots \mathbb{P} \cdot Q \cdot \mathbb{P}}_{n-1 \text{ veces}} = \widetilde{\mathbb{P}}^{n-1} \cdot \mathbb{P}, \quad \widetilde{\mathbb{P}} = \mathbb{P} \cdot Q.$$

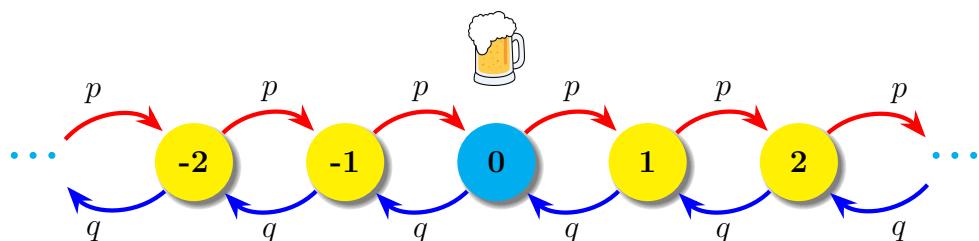
En notación matricial,  $q_{11}^{(n)} = (\widetilde{\mathbb{P}}^{n-1} \cdot \mathbb{P})_{11}$ . Desde el punto de vista del álgebra lineal al multiplicar  $Q$  por un vector lo que estamos encontrando es la proyección de cualquier vector sobre el estado 2.

**Ejercicio 1.7.** Realiza el mismo estudio para el caso de la transición  $A \rightarrow B$ , es decir con que probabilidad llegamos a  $B$  partiendo de  $A$ . En el caso de que dicha probabilidad sea 1, estudia y si lo en un tiempo finito. ¿Se puede deducir que  $q_{AB} = 0$  conocido que  $q_{AA} = 1$ ? Reproduce los resultados usando la matriz  $\mathbb{P}$  y una matriz  $Q$  adecuada.

**Comentario 1.8.** Como hemos visto  $\mathbb{P}$  es la matriz de probabilidades de trasición en un paso, mientras que  $\mathbb{P}^n$  es la matriz de probabilidades de trasición en  $n$  pasos. Por tanto el problema de estudio de muchas propiedades de una cadena de Markov, y en particular las propiedades de retorno y el tiempo de retorno, se reducen al cálculo de las potencias  $n$ -ésimas de una matriz (que puede ser infinita), problema, en general nada trivial.

## 1.4. La cadena infinita: la caminata del borracho

Recordemos el diagrama de la cadena doblemente infinita:





Definimos  $Q_j = \mathbb{I} - \Pi_j$ , entonces

$$p_{ij}^{(n)} = \mathbb{P}_{ij}^n, \quad q_{ij}^{(n)} = [(\mathbb{P} \cdot Q_j)^{n-1} \cdot \mathbb{P}]_{ij},$$

donde  $p_{ij}^{(n)}$  es la probabilidad de pasar del estado  $i$  al  $j$  en  $n$  pasos y  $q_{ij}^{(n)}$  es la probabilidad de pasar del estado  $i$  al  $j$  por primera vez en  $n$  pasos.

Es obvio que calcular los valores de  $q_{ij}^{(n)}$  mediante la multiplicación es altamente no trivial. Un método alternativo pasa por el uso de la técnica de las funciones generatrices.

## 1.5. La ecuación de renuevo

Supongamos que tenemos una cadena de Markov (finita o infinita) y comenzamos nuestro *paseo* en cierto estado inicial  $i$ . Denotemos por  $p_n$  la probabilidad de retornar a dicho estado en  $n$  pasos y por  $q_n$  la propiedad de retornar por primera vez a  $i$  en  $n$  pasos. Es obvio que  $p_n \geq q_n$  para todo  $n$ . Vamos a asumir que el proceso es de Markov, lo que implica que dado cierto instante (paso)  $n$  el futuro (lo que ocurrirá después del paso  $n$ ) es independiente del pasado (lo que ha ocurrido antes del paso  $n$ ). En estas condiciones existe una relación muy sencilla entre las cantidades  $(p_n)_n$  y  $(q_n)_n$  definidas anteriormente.

Por comodidad asumiremos que  $p_0 = 1$  (la probabilidad de regresar al estado de partida en ningún paso es 1, pues ya estamos allí) y que  $q_0 = 0$ . Entonces

$$\begin{aligned} p_1 &= q_1 \underbrace{p_0}_{=1} + \underbrace{q_0}_{=0} p_1 = q_1, & p_2 &= q_2 \underbrace{p_0}_{=1} + q_1 p_1 + \underbrace{q_0}_{=0} p_2 = q_2 + q_1 p_1, & \dots \\ p_n &= q_n p_0 + q_{n-1} p_1 + q_{n-2} p_2 + \dots + q_1 p_{n-1} + q_0 p_n, \end{aligned}$$

que podemos interpretar probabilísticamente de forma muy sencilla: la probabilidad de regresar el  $n$  pasos es la suma de la probabilidad de regresar por primera vez en  $n$  pasos  $q_n$  mas la probabilidad de regresar por primera vez en  $n-1$  pasos  $q_{n-1}$  multiplicada por la probabilidad de regresar en un paso (no necesariamente por primera vez)  $p_1$  mas a probabilidad de regresar por primera vez en  $n-2$  pasos  $q_{n-2}$  multiplicada por la probabilidad de regresar en dos pasos (no necesariamente por primera vez)  $p_2$ , etc.

Definamos las funciones (recordemos que  $p_0 = 1$ )

$$p(z) = 1 + p_1 z + p_2 z^2 + \dots = \sum_{n \geq 0} p_n z^n, \quad q(z) = q_1 z + q_2 z^2 + \dots = \sum_{n \geq 1} q_n z^n,$$

donde  $p_n \in (0, 1)$ ,  $q_n \in (0, 1)$ . En esas condiciones tanto  $p$  como  $q$  son funciones analíticas en  $|z| \leq 1$  y además  $q$  converge en  $z = 1$ .

Si multiplicamos  $p(z)q(z)$  es fácil comprobar que

$$\begin{aligned} p(z)q(z) &= q_1z + (q_2 + q_1p_1)z^2 + (q_3 + q_2p_1 + q_1p_2)z^3 + \dots + \\ &\quad + (q_n + q_{n-1}p_1 + q_{n-2}p_2 + \dots + q_1p_{n-1})z^n + \dots \\ &= p_1z + p_2z^2 + p_3z^3 + \dots + p_nz^n + \dots = p(z) - 1. \end{aligned}$$

Luego, para las funciones  $p(z)$  y  $q(z)$  se tiene que

$$p(z)q(z) = p(z) - 1 \iff q(z) = 1 - \frac{1}{p(z)}. \quad (1.7)$$

**Definición 1.9.** Las funciones  $p(z)$  y  $q(z)$  anteriores se denominan función generatriz de los retornos y función generatriz de los primeros retornos, respectivamente. La ecuación (1.7) se denomina ecuación de renuevo (renewal equation).

Está claro que la probabilidad  $q_{ii}$  de retorno al estado inicial será

$$q_{ii} = q_1 + q_2 + \dots + q_n + \dots = q(1). \quad (1.8)$$

De lo anterior se sigue que nuestro *estado* inicial es *recurrente* si  $q(1) = 1$ , o, equivalentemente, usando (1.7),  $p(1) = \infty$ .

Además, si  $q(1) = 1$  (estado recurrente) entonces podemos preguntarnos si es positivamente recurrente, para ello calculamos el tiempo medio de retorno

$$\tau_{ii} = \sum_{n \geq 1} nq_n = q_1 + 2q_2 + \dots + nq_n + \dots = \lim_{z \rightarrow 1^-} \frac{q(z) - q(1)}{z - 1} = \lim_{z \rightarrow 1^-} \frac{q(z) - 1}{z - 1} = q'(1-). \quad (1.9)$$

**Ejemplo 1.10.** Sea nuestro ejemplo con dos estados definido por la matriz (1.2). Usando (1.3) tenemos

$$p_{AA}(z) = \sum_{n=0}^{\infty} \frac{\beta + \alpha(1 - \alpha - \beta)^n}{\alpha + \beta} z^n,$$

de donde comprobamos que  $p_{AA}(1) = \infty$  (la serie es divergente para todos  $\alpha, \beta \in [0, 1]$ ) luego la ecuación de renuevo nos da  $q_{AA}(1) = 1$ . Obviamente la serie anterior se puede sumar

$$p_{AA}(z) = \frac{\alpha}{(\alpha + \beta)(\alpha z + \beta z - z + 1)} + \frac{\beta}{(\alpha + \beta)(1 - z)},$$

luego

$$q_{AA}(z) = 1 - \frac{1}{p_{AA}(z)} = \frac{(\alpha + \beta - 1)z^2 + (1 - \alpha)z}{(\beta - 1)z + 1},$$

de donde se sigue, nuevamente, que  $q_{AA}(1) = 1$  y además, para el tiempo esperado de retorno tenemos

$$\tau_{AA} = \lim_{z \rightarrow 1^-} q'_{AA}(z) = \lim_{z \rightarrow 1^-} \left[ 1 - \alpha + \frac{2\alpha\beta z}{1 - (1 - \beta)z} - \frac{\alpha(\beta - 1)\beta z^2}{(1 - (1 - \beta)z)^2} \right] = 1 + \frac{\alpha}{\beta}.$$

**Ejemplo 1.11.** Veamos el caso de la caminata aleatoria infinita definida por la matriz (1.5). En este caso sabemos la expresión de  $p_{00}^{(n)}$  (1.6), luego

$$p_{00}(z) = \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n z^{2n}.$$

La serie anterior es divergente en  $z = 1$ , para ello basta usar la fórmula de Stirling  $n! \sim \sqrt{2\pi n} e^{-n} n^n$ . No obstante vamos a sumarla para dar una solución completa al problema del retorno al estado inicial. Para ello usaremos la identidad

$$(1 - z^2)^{-1/2} = \sum_{n=0}^{\infty} \binom{2n}{n} \frac{z^{2n}}{2^{2n}},$$

luego

$$p_{00}(z) = \frac{1}{\sqrt{1 - \sqrt{4pq}z^2}} \Rightarrow q_{00}(z) = 1 - \sqrt{1 - 2\sqrt{pq}z^2}, \quad (1.10)$$

de donde obtenemos

$$p_{00}(1) = \frac{1}{\sqrt{1 - \sqrt{4pq}}} \Rightarrow q_{00}(1) = 1 - \sqrt{1 - 2\sqrt{pq}}.$$

Por tanto  $q(1) = 1$  si y sólo si  $p = q = 1/2$ . Además, en ese caso

$$\tau_{00} = \lim_{z \rightarrow 1^-} q'_{00}(z) = \lim_{z \rightarrow 1^-} \frac{z}{\sqrt{1 - z^2}} = \infty,$$

es decir, el proceso es recurrente pero no positivamente recurrente.

**Ejercicio 1.12.** Desarrollando en series la función generatriz de primeros retornos  $q_{00}(z)$  obtenida en (1.10) obtener el número de caminos de longitud  $2n$  de retornar por primera vez al estado inicial y compararlo con el valor de  $p_{00}^{(2n)}$  (1.6).

## 1.6. Polinomios ortogonales

Antes de continuar conviene hacer una breve revisión de algunos conceptos de la teoría de polinomios ortogonales en la recta real.

Supongamos que  $\alpha$  es cierta una función no decreciente en  $(a, b)$  ( $\alpha(x) \neq \text{const}$ ) tal que si el intervalo  $(a, b)$  es no acotado, o sea si  $a = -\infty$ , entonces  $\lim_{x \rightarrow -\infty} \alpha(x) > -\infty$  y si  $b = \infty$ , entonces  $\lim_{x \rightarrow \infty} \alpha(x) < \infty$ . Diremos que una función  $f$  pertenece al espacio  $L_\alpha^p[a, b]$  si  $\int_a^b |f(x)|^p d\alpha(x) < \infty$ . Nos interesará el caso  $p = 2$ , i.e.,  $f \in L_\alpha^2[a, b]$ .

Definiremos el producto escalar de dos funciones  $f$  y  $g$  pertenecientes a  $L_\alpha^2[a, b]$  como la integral de Stieltjes-Lebesgue

$$\langle f, g \rangle = \int_a^b f(x)g(x)d\alpha(x). \quad (1.11)$$

Para una función  $\alpha$  prefijada de antemano, la ortogonalidad respecto a la distribución  $d\alpha$  vendrá definida por la relación  $\langle f, g \rangle = 0$ , y diremos que  $f$  y  $g$  son ortogonales, o que,  $f$  es ortogonal a  $g$  respecto a la distribución  $d\alpha$ . Si  $\alpha$  es absolutamente continua en el intervalo  $(a, b)$ , el producto escalar (1.11) se puede reescribir como la integral de Lebesgue:

$$\langle f, g \rangle = \int_a^b f(x)g(x)\rho(x)dx, \quad (1.12)$$

donde  $\rho$  es una función medible no negativa tal que  $0 < \int_a^b \rho(x)dx < \infty$ . A la función  $\rho$  se le denomina función peso.

Consideremos ahora el *espacio vectorial*  $L_\alpha^2(a, b)$ . Definamos en este espacio el producto escalar (1.11) y la norma de un vector mediante la expresión  $\|f\| = \sqrt{\langle f, f \rangle}$ . Si  $\|f\| = 0$  diremos que  $f$  es el vector nulo. Si  $\|f\| = 1$  diremos que  $f$  es un vector normalizado. Si  $f$  no es nula entonces para cierto valor  $\lambda \neq 0$  el vector  $\lambda f$  es un vector normalizado.

**Comentario 1.13.** En general se puede considerar el producto escalar (1.11) con respecto a una medida normalizada a la unidad cuya forma general es

$$d\mu(x) = \rho(x)dx + \sum_k m_k \delta(x - x_k)dx + d\mu_{sc}(x),$$

siendo el primer sumando la parte absolutamente continua y los dos restantes la parte singular de la medida. Para ello la medida  $d\mu(x)$  ha de estar soportada en un número infinito de puntos y deben existir todos los momentos  $\mu_n = \int_{\text{supp } d\mu} x^n d\mu(x) < +\infty$ .

**Definición 1.14.** Dada una sucesión de polinomios  $(P_n)_n$ , diremos que  $(P_n)_n$  es una sucesión de polinomios ortogonales con respecto a la medida  $d\mu(x)$  si se cumple que:

1.  $P_n$  es un polinomio de grado  $n$ ,

2.  $\langle P_n, P_m \rangle = 0, \quad m \neq n,$  para todos  $n, m = 0, 1, 2, \dots,$
3.  $d_n^2 := \langle P_n, P_n \rangle > 0,$  para todo  $n = 0, 1, 2, \dots .$

Nótese que en el caso  $d\mu(x) = \rho(x)dx,$  siendo  $\rho$  una función continua y no negativa en todo  $\mathbb{R},$  la tercera condición es inmediata.

**Comentario 1.15.** La tercera condición de la definición (1.14) corresponde al caso definido positivo (los valores  $d_n^2$  son realmente normas), no obstante se puede trabajar en el caso quasi-definido donde simplemente se necesita que  $d_n^2 \neq 0.$  Nosotros nos restringiremos al caso definido positivo.

**Definición 1.16.** Si una sucesión de polinomios cumple las condiciones de la definición (1.14) la denotaremos mediante SPO. Diremos que la sucesión es ortonormal si, para todo  $n,$  la norma  $d_n^2 = 1,$  i.e.,  $\int_{\text{supp } d\mu} p_i(x)p_j(x)d\mu(x) = \delta_{ij}.$  Diremos que la sucesión es una sucesión de polinomios ortogonales mónicos (SPOM) si el coeficiente principal  $a_n$  de  $P_n(x) = a_nx^n + \dots,$  es igual a uno, o sea, para todo  $n \geq 0,$   $P_n(x) = x^n + b_nx^{n-1} + \dots.$

Una de las principales características de las SPO es que satisfacen una relación de recurrencia a tres términos (RRTT). Ello es una consecuencia inmediata de la ortogonalidad.

**Teorema 1.17.** Sea  $(P_n)_n$  una SPO con respecto a cierta medida  $d\mu(x).$  Entonces la SPO  $(P_n)_n$  satisface una relación de recurrencia a tres términos de la forma:

$$xP_n(x) = \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad n = 0, 1, 2, \dots. \quad (1.13)$$

Generalmente se suele imponer que  $P_{-1}(x) = 0$  y  $P_0(x) = 1,$  con lo que una SPO queda determinada de forma única conocidas las sucesiones  $(\alpha_n)_n,$   $(\beta_n)_n$  y  $(\gamma_n)_n.$

En el caso particular cuando los polinomios son ortonormales la RRTT (1.13) se transforma en

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1} p_{n-1}(x), \quad a_n > 0, \quad n \geq 0, \quad (1.14)$$

donde  $b_n = \beta_n$  y  $a_n = \sqrt{\alpha_{n+1}\gamma_n}.$

El recíproco también es cierto y se conoce como Teorema de Favard.

**Teorema 1.18.** Supongamos que una sucesión de polinomios mónicos  $(P_n)_n$  satisface una relación de recurrencia a tres términos de la forma:

$$xP_n(x) = P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad n \geq 0 ,$$

con  $\gamma_n > 0$  y  $\beta_k \in \mathbb{R}$  ( $k = 0, 1, 2, \dots$ ) y las condiciones iniciales  $P_{-1}(x) = 0$  y  $P_0(x) = 1.$  Entonces, dichos polinomios  $P_n$  son ortonormales en  $L^2(\mu),$  siendo  $d\mu(x)$  cierta medida positiva sobre la recta real.

Dado que los polinomios ortogonales quedan biunivocamente determinados salvo un factor multiplicativo, el teorema de Favard es directamente extendible al caso de RRTT del tipo (1.14).

Nótese que la RRTT (1.13) se puede escribir en forma matricial

$$x \mathbf{P}_n(x) = J_{n+1} \mathbf{P}_n(x) + \alpha_n P_{n+1}(x) \mathbf{e}_n, \quad (1.15)$$

donde  $\mathbf{P}_n(x) = [P_0(x), \dots, P_n(x)]^T$ ,  $\mathbf{e}_n = [0, 0, \dots, 0, 1]^T \in \mathbb{R}^{n+1}$  y  $J_{n+1}$  es la matriz tridiagonal de orden  $n+1$

$$J_{n+1} = \begin{bmatrix} \beta_0 & \alpha_0 & 0 & 0 & \dots & 0 & 0 \\ \gamma_1 & \beta_1 & \alpha_1 & 0 & \dots & 0 & 0 \\ 0 & \gamma_2 & \beta_2 & \alpha_2 & \dots & 0 & 0 \\ 0 & 0 & \gamma_3 & \beta_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \beta_{n-1} & \alpha_{n-1} \\ 0 & 0 & 0 & 0 & \dots & \gamma_n & \beta_n \end{bmatrix}.$$

Una de las consecuencias de lo anterior es que si denotamos por  $\{x_{nj}\}_{1 \leq j \leq n}$  a los ceros del polinomio  $P_n$ , entonces de (1.15) se sigue que cada  $x_{nj}$  es un autovalor de la matriz tridiagonal  $J_n$  y que  $\mathbf{P}_{n-1}(x_{nj}) := [P_0(x_{nj}), \dots, P_{n-1}(x_{nj})]^T$  es el correspondiente autovector.

Si tomamos  $\alpha_n = 1$  y  $\gamma_n > 0$  para todo  $n = 1, 2, \dots$  la sucesión de polinomios monómicos  $(P_n)_{n \geq 0}$  definida por la RRTT (1.13) aparece como los denominadores de los aproximantes de Padé de la fracción continua

$$\cfrac{1}{x - \beta_0 - \cfrac{\gamma_1}{x - \beta_1 - \cfrac{\gamma_2}{x - \beta_2 - \dots - \cfrac{\gamma_{n-1}}{x - \beta_{n-1} - \cfrac{\gamma_n}{x - \beta_n - \ddots}}}}}.$$

Por otro lado, los numeradores de los aproximantes de Padé, que denotaremos por  $(P_n^{(1)})_{n \geq 0}$ , satisfacen una RRTT similar (desplazada en el índice  $n$ )

$$x P_n^{(1)}(x) = P_{n+1}^{(1)}(x) + \beta_{n+1} P_n^{(1)}(x) + \gamma_{n+1} P_{n-1}^{(1)}(x), \quad n \geq 0,$$

con las condiciones iniciales  $P_{-1}^{(1)} = 0$  y  $P_0^{(1)} = 1$ . La fracción continua anterior converge a cierta función  $F(z; \sigma)$  que es analítica en  $\mathbb{C} \setminus \text{supp}(d\mu(x))$ , i.e., se tiene el

Teorema de Markov-Stieltjes

$$-\lim_{n \rightarrow +\infty} \frac{\mu_0 P_{n-1}^{(1)}(z)}{P_n(z)} = F(z; \mu) := \int_{-\infty}^{+\infty} \frac{d\mu(x)}{x - z}, \quad z \in \mathbb{C} \setminus \text{supp}(d\mu(x)), \quad (1.16)$$

donde  $\mu_0 = \int_{-\infty}^{+\infty} d\mu(x)$  es el primer momento de la medida  $d\mu(x)$  y  $F$  es la función de Stieltjes asociada a la medida  $d\mu(x)$ . De hecho, la medida  $d\mu(x)$  se puede recuperar a partir de la función de Stieltjes por la fórmula de inversión de Stieltjes:

$$\mu(x) - \mu(y) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \int_y^x \Im F(t + i\epsilon, \mu) dt, \quad (1.17)$$

que en caso de que la medida sea absolutamente continua  $d\mu(x) = \omega(x)dx$  se transforma en

$$\omega(x) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \Im F(x + i\epsilon, \mu),$$

y si hay un punto de masa entonces

$$\mu(\{x\}) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \epsilon \Im F(x + i\epsilon, \mu) > 0.$$

Aquí se asume que  $d\mu(x)$  está normalizada de forma que

$$\mu(-\infty) = 0, \quad \mu(x) = \frac{\mu(x+0) + \mu(x-0)}{2}$$

e  $\Im z$  denota la parte imaginaria de  $z$ .

Nótese que tomando  $n \rightarrow \infty$  (con la métrica adecuada) en (1.15) tenemos, para la SPO la ecuación matricial

$$x\mathbf{P}(x) = J\mathbf{P}, \quad J = \begin{bmatrix} \beta_0 & \alpha_0 & 0 & 0 & \dots & 0 & 0 & \dots \\ \gamma_1 & \beta_1 & \alpha_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \gamma_2 & \beta_2 & \alpha_2 & \dots & 0 & 0 & \dots \\ 0 & 0 & \gamma_3 & \beta_3 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \dots \\ 0 & 0 & 0 & 0 & \dots & \beta_{n-1} & \alpha_{n-1} & \dots \\ 0 & 0 & 0 & 0 & \dots & \gamma_n & \beta_n & \dots \\ \vdots & \ddots \end{bmatrix}, \quad (1.18)$$

donde  $\mathbf{P}(x) = [P_0(x), \dots, P_n(x), \dots]^T$  y  $J$  es cierta matriz asociada con el operador de Jacobi. De hecho, si trabajamos con sucesiones de polinomios ortonormales lo

anterior se convierte en

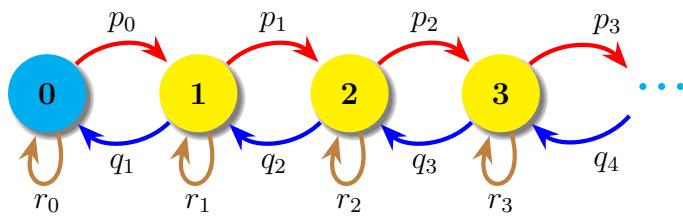
$$x \begin{bmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ p_3(x) \\ \vdots \end{bmatrix} = \begin{bmatrix} b_0 & a_0 & 0 & 0 & \dots \\ a_0 & b_1 & a_1 & 0 & \dots \\ 0 & a_1 & b_2 & a_2 & \dots \\ 0 & 0 & a_2 & b_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} p_0(x) \\ p_1(x) \\ p_2(x) \\ p_3(x) \\ \vdots \end{bmatrix} \iff x \mathbf{p}(x) = \mathbb{J} \mathbf{p}(x) \quad (1.19)$$

donde  $\mathbf{p}(x) = [p_0(x), \dots, p_n(x), \dots]^T$ , y  $\mathbb{J}$  es la matriz tridiagonal, real y simétrica (comúnmente denominada matriz de Jacobi)

$$\mathbb{J} = \begin{bmatrix} b_0 & a_0 & 0 & 0 & \dots \\ a_0 & b_1 & a_1 & 0 & \dots \\ 0 & a_1 & b_2 & a_2 & \dots \\ 0 & 0 & a_2 & b_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

## 1.7. Caminatas aleatorias y polinomios ortogonales

Vamos ahora a considerar la siguiente caminata aleatoria



cuya matriz  $\mathbb{P}$  de probabilidades de transición es

$$\mathbb{P} = \begin{bmatrix} r_0 & p_0 & 0 & 0 & 0 & 0 & \dots \\ q_1 & r_1 & p_1 & p & 0 & 0 & \dots \\ 0 & q_2 & r_2 & p_2 & 0 & 0 & \dots \\ 0 & 0 & q_3 & r_3 & p_3 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad r_0 + p_0 = 1, \quad p_i + q_i + r_i = 1, \quad \forall i \geq 1, \quad (1.20)$$

y donde se asume que los  $p_i$  y  $q_i$  son todos distintos de cero.

Está claro que la matriz anterior es similar a la matriz (1.18) anterior, por lo tanto la teoría de polinomios ortogonales puede ser útil para estudiar la caminata

definida por (1.20). Definiendo la matriz

$$\Pi = \begin{bmatrix} \pi_0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & \pi_1 & 0 & p & 0 & 0 & \dots \\ 0 & 0 & \pi_2 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \pi_3 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \frac{\pi_i}{\pi_{i-1}} = \sqrt{\frac{p_{i-1}}{q_i}}, \quad i \geq 1, \quad (1.21)$$

tenemos

$$\Pi \cdot \mathbb{P} \cdot \Pi^{-1} = \begin{bmatrix} r_0 & a_0 & 0 & 0 & 0 & 0 & \dots \\ a_0 & r_1 & a_1 & p & 0 & 0 & \dots \\ 0 & a_1 & r_2 & a_2 & 0 & 0 & \dots \\ 0 & 0 & a_2 & r_3 & a_3 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \mathbb{J}, \quad a_i = \sqrt{p_i q_{i+1}} > 0, \quad i \geq 0,$$

donde  $\mathbb{J}$  es una matriz de Jacobi (tridiagonal, real y simétrica) asociada a la correspondiente sucesión de polinomios ortonormales  $(p_n)_n$ .

A partir de (1.19) podemos probar, usando inducción, que para los polinomios ortonormales

$$x^n \mathbf{p}(x) = \mathbb{J}^n \mathbf{p}(x), \quad (1.22)$$

donde los elementos de la matriz  $\mathbb{J}^n$  se calculan por la fórmula

$$\mathbb{J}_{ij}^n = \int_{\text{supp } d\mu} x^n p_i(x) p_j(x) d\mu(x), \quad i, j \geq 0, \quad (1.23)$$

La fórmula anterior se denomina fórmula de Karling y McGregor.

Para probar formalmente la fórmula de Karling y McGregor escribimos (1.22) término a término:

$$x^n p_i(x) = \sum_{j=i-n}^{i+n} \mathbb{J}_{ij}^n p_j(x) = \sum_{j=0}^{i+n} \mathbb{J}_{ij}^n p_j(x)$$

donde, en la primera suma, se asume que  $p_j(x) \equiv 0$  si  $j < 0$ . Multiplicando ambas partes de la ecuación anterior por  $p_i(x) d\mu(x)$  e integrando sobre el soporte de la medida de ortogonalidad  $d\mu(x)$  se obtiene el resultado. Conviene mencionar que existe una prueba más estricta usando análisis funcional (ver e.g. [4]).

Dado que todos los elementos de la matriz  $\mathbb{P}$  son positivos y menores que uno, podemos definir un operador  $\hat{\mathbb{P}} : \ell^1 \mapsto \ell^1$ ,  $v \mapsto \hat{\mathbb{P}}v = v \cdot \mathbb{P}$ , asociado a  $\mathbb{P}$  que actúa sobre el espacio de Banach  $\ell^1$  de las sucesiones  $(x_n)_n$  tales que  $\sum_{n \geq 0} |x_n| < +\infty$ . Dicho operador es acotado en norma  $\ell^1$  pues

$$\|v \cdot \mathbb{P}\| = \sum_{j \geq 0} \left| \sum_{i \geq 0} v_i p_{ij} \right| \leq \sum_{j \geq 0} \sum_{i \geq 0} |v_i| p_{ij} = \sum_{i \geq 0} |v_i| \sum_{j \geq 0} p_{ij} = \sum_{i \geq 0} |v_i| \cdot 1 = \|v\|,$$

donde las sumas se intercambian al ser ambas series absolutamente convergentes. Luego  $\|\hat{\mathbb{P}}\| \leq 1$ . Ahora bien, dado un operador acotado  $\hat{\mathbb{P}}$ , es conocido que su espectro  $\sigma(\hat{\mathbb{P}}) \in \{z \in \mathbb{C}; |z| \leq \|\hat{\mathbb{P}}\|\}$  entonces el espectro de nuestro operador está contenido en el círculo unidad. Por otro lado, el Teorema de Favard nos asegura que la familia de PO asociada a  $\mathbb{J}$  (y por tanto a  $\mathbb{P}$ ) es ortogonal respecto a una medida  $d\mu(x)$  soportada en el recta real, luego para nuestra caminata aleatoria  $\text{supp}(d\mu(x)) \in [-1, 1]$ .

Como  $\mathbb{J}$  está muy relacionada con  $\mathbb{P}$  ( $\mathbb{J} = \Pi \cdot \mathbb{P} \cdot \Pi^{-1}$ ) podemos usarla para el estudio de las propiedades de retorno de nuestra caminata aleatoria.

Sea  $f(x)$  una función analítica en el soporte  $\text{supp } d\mu(x)$  de la medida  $d\mu(x)$

$$f(x) = \sum_{k \geq 0} c_n x^n.$$

Entonces

$$\begin{aligned} \int_{\text{supp } d\mu} f(x) p_i(x) p_j(x) d\mu(x) &= \int_{\text{supp } d\mu} \sum_{k \geq 0} c_n x^n p_i(x) p_j(x) d\mu(x) \\ &= \sum_{k \geq 0} c_n \int_{\text{supp } d\mu} x^n p_i(x) p_j(x) d\mu(x) \\ &= \sum_{k \geq 0} c_n \mathbb{J}_{ij}^n = (f(\mathbb{J}))_{ij}. \end{aligned}$$

Vamos a suponer que el soporte de  $d\mu(x)$  es acotado (recordemos que en nuestro caso  $\text{supp}(d\mu(x)) \in [-1, 1]$ ) y que  $z$  es lo suficientemente pequeño. Entonces

$$f(x) = \frac{1}{1 - zx}, \quad z \in \mathbb{C} \setminus \text{supp}(d\mu(x)),$$

es analítica en  $\text{supp}(d\mu(x))$  por lo que aplicando la fórmula anterior obtenemos

$$\int_{\text{supp } d\mu} \frac{p_i(x) p_j(x)}{1 - xz} d\mu(x) = (\mathbb{I} - z\mathbb{J})_{ij}^{-1}.$$

Consideremos el caso especial  $i = j$ . En este caso

$$(\mathbb{I} - z\mathbb{J})_{ii}^{-1} = \int_{\text{supp } d\mu} \frac{p_i^2(x) d\mu(x)}{1 - xz} = \int_{\text{supp } d\mu} \frac{d\mu_i(x)}{1 - xz}, \quad d\mu_i(x) = p_i^2(x) d\mu(x),$$

donde  $d\mu_i(x)$  se denomina medida espectral de la sucesión  $(p_n)_n$ . Desarrollando en serie la función  $(1 - zx)^{-1}$  obtenemos

$$\int_{\text{supp } d\mu} \frac{d\mu_i(x)}{1 - xz} = \int_{\text{supp } d\mu} \sum_{n \geq 0} (zx)^n d\mu_i(x) = \sum_{n \geq 0} \mu_n^{(i)} z^n = \widehat{\mu}_i(z),$$

donde  $\mu_n^{(i)}$  son los momentos de orden  $n$  de la medida espectral  $d\mu_i(x)$  y  $\widehat{\mu}_i(z)$  es la función generatriz de los correspondientes momentos. Luego la función  $g(z) = (\mathbb{I} - z\mathbb{J})_{ii}^{-1}$  es la función generatriz de los momentos  $\widehat{\mu}_i(z)$  de la medida espectral de  $(p_n)_n$  y en el caso especial  $i = 0$ , el de la medida original  $d\mu(x)$ .

Veamos ahora las propiedades de retorno. Comenzamos con la función generatriz de los retornos

$$p_i(z) = \sum_{n \geq 0} p_{ii}^{(n)} z^n = \sum_{n \geq 0} \mathbb{P}_{ii}^n z^n = \sum_{n \geq 0} (\Pi \cdot \mathbb{J}^n \cdot \Pi^{-1})_{ii} z^n$$

donde  $\Pi$  es la matriz diagonal (1.21), luego

$$p_i(z) = \sum_{n \geq 0} (\Pi \cdot \mathbb{J}^n \cdot \Pi^{-1})_{ii} z^n = \sum_{n \geq 0} \pi_i \mathbb{J}_{ii}^n \pi_i^{-1} z^n = \sum_{n \geq 0} \mathbb{J}_{ii}^n z^n = (I - z\mathbb{J})_{ii}^{-1}. \quad (1.24)$$

Pero usando la fórmula de Karlin-Mcgregor (1.23)  $\mathbb{J}_{ii}^n = \mu_n^{(i)}$ , donde  $\mu_n^{(i)}$  son los momentos de orden  $n$  de la medida espectral de los polinomios ortonormales asociados a  $\mathbb{J}$ . Luego, teniendo en cuenta que el soporte de nuestra medida es  $[-1, 1]$ , tenemos

$$p_i(z) = \sum_{n \geq 0} \mu_n^{(i)} z^n = \int_{-1}^1 \frac{d\mu_i(x)}{1 - xz} = s_i(z), \quad (1.25)$$

Recordemos que la probabilidad de retorno al estado  $i$  es  $p_i(1) = s_i(1)$  y por tanto, la probabilidad de primeros retornos, según la ecuación de renuevo (1.7), es  $q(1) = 1 - 1/p(1)$ . Así el proceso será recurrente si la función  $s_i(1) = \infty$ .

Para decidir si el proceso es positivamente recurrente tenemos que calcular el tiempo esperado de retorno (1.9)

$$\tau_{ii} = \lim_{z \rightarrow 1^-} \frac{q(z) - 1}{z - 1} = \lim_{z \rightarrow 1^-} \frac{1}{(1 - z)s_i(z)} = \lim_{z \rightarrow 1^-} \frac{1}{(1 - z)p_i(z)}. \quad (1.26)$$

Si  $\tau_{ii}$  es finito, entonces el proceso será positivamente recurrente o *ergódico*.

Este último límite está estrechamente relacionado con que la medida espectral asociada a  $\mathbb{P}$  tenga un punto de masa en  $x = 1$ .

Vamos a estudiar dicha conexión. Para ello conviene hacer notar que la función  $s_i(z)$  obtenida en (1.25) es una función estrechamente relacionada con la función de Stieltjes  $F(z; \mu_i)$  asociada a la matriz  $\mathbb{J}$ . De hecho se tiene que para todo  $z \neq 0$

$$s_i(z) = -z^{-1}F(z^{-1}; \mu_i), \quad s_0(z) := -s(z) = z^{-1}F(z^{-1}; \mu).$$

Notemos que la medida espectral  $d\mu_i(x) = p_i^2(x)d\mu(x)$  tiene un punto de masa en  $x = 1$  si y sólo si  $d\mu(x)$  lo tiene. Dado que nos interesará el caso cuando  $z \rightarrow 1$ ,

vamos mostrar como se puede recuperar el punto de masa en  $z = 1$  a partir de la función de Stieltjes. Para ello descompongamos nuestra medida de la siguiente forma

$$d\mu(x) = \mu(\{1\})\delta(x - 1)dx + d\tilde{\mu}(x),$$

donde  $\tilde{\mu}(x)$  es la medida que se obtiene al eliminar el punto de masa (si lo hay) en  $x = 1$ . Entonces

$$F(\zeta; \mu) = \frac{\mu(\{1\})}{\zeta - 1} + \int_{-1}^1 \frac{d\tilde{\mu}(x)}{x - \zeta}, \quad \zeta \in \mathbb{C} \setminus [-1, 1].$$

Luego

$$(\zeta - 1)F(\zeta; \mu) = \mu(\{1\}) + \int_{-1}^1 \frac{\zeta - 1}{\zeta - x} d\tilde{\mu}(x).$$

Ahora bien, para todo  $\zeta \in (1, +\infty)$ ,  $|\frac{\zeta - 1}{\zeta - x}| < 1$ , por tanto usando el teorema de la convergencia dominada tenemos que (recordemos que  $x \in [-1, 1]$  y que  $\zeta \in \mathbb{C} \setminus [-1, 1]$ )

$$\lim_{\zeta \rightarrow 1^+} \int_{-1}^1 \frac{\zeta - 1}{\zeta - x} d\tilde{\mu}(x) = \int_{-1}^1 \lim_{\zeta \rightarrow 1^+} \frac{\zeta - 1}{\zeta - x} d\tilde{\mu}(x) = \int_{-1}^1 0 d\tilde{\mu}(x) = 0.$$

Por tanto,

$$\lim_{\zeta \rightarrow 1^+} (\zeta - 1)F(\zeta; \mu) = \mu(\{1\}). \quad (1.27)$$

Usando ahora la relación  $s(z) = -z^{-1}F(z^{-1}; \mu)$ , (1.27) y (1.26) obtenemos

$$\tau_{ii} = \lim_{z \rightarrow 1^-} \frac{1}{(1 - z)s_i(z)} = \frac{1}{\mu(\{1\})}.$$

Luego, si la medida de ortogonalidad de los correspondientes polinomios ortonormales tiene una masa distinta de cero en  $x = 1$ , el proceso será positivamente recurrente.

Lo anterior se puede resumir en el siguiente teorema:

**Teorema 1.19.** *Sea una caminata aleatoria con matriz de transición de la forma (1.20) y sea  $F(z; \mu_i)$  la función de Stieltjes asociada a la correspondiente familia de polinomios ortonormales  $(p_n)_n$ . Entonces el estado  $i$  de dicha caminata es recurrente si  $F(1; \mu_1) = -\infty$  y positivamente recurrente si  $\mu_i(\{1\}) > 0$ , siendo la medida  $d\mu_i(x) = p_i(x)^2 d\mu(x)$  la medida espectral asociada a la familia  $(p_n)_n$  y  $d\mu(x)$  la medida de ortogonalidad de  $(p_n)_n$ .*

## 1.8. Ejemplos

Comenzaremos con un simple ejercicio de álgebra lineal elemental. Sea la matriz por bloques

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad \text{entonces } \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & * \\ * & * \end{bmatrix} \quad (1.28)$$

Usando (1.24) tenemos  $p(z) := p_{00}(z) = (I - z\mathbb{J})_{00}^{-1}$ . Sea

$$\mathbb{J} = \left[ \begin{array}{c|ccccccc} r_0 & a_0 & 0 & 0 & 0 & 0 & \dots \\ \hline a_0 & r_1 & a_1 & 0 & 0 & 0 & \dots \\ 0 & a_1 & r_2 & a_2 & 0 & 0 & \dots \\ 0 & 0 & a_2 & r_3 & a_3 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \end{array} \right] = \left[ \begin{array}{c|ccccc} r_0 & a_0 & 0 & 0 & 0 & \dots \\ \hline a_0 & 0 & & & & \\ 0 & & & & & \mathbb{J}^{(1)} \\ 0 & & & & & \\ \vdots & & & & & \end{array} \right]$$

Sea  $p^{(1)}(z) = (I - z\mathbb{J}^{(1)})_{00}^{-1}$ , entonces usando (1.28) tenemos

$$p(z) = (1 - r_0 z - a_0^2 z^2 p^{(1)}(z))^{-1}. \quad (1.29)$$

Veamos dos ejemplos interesantes.

### 1.8.1. Ejemplo 1

El primero corresponde a la matriz

$$\mathbb{J} = \left[ \begin{array}{c|ccccccc} 0 & a & 0 & 0 & 0 & 0 & \dots \\ \hline a & 0 & a & 0 & 0 & 0 & \dots \\ 0 & a & 0 & a & 0 & 0 & \dots \\ 0 & 0 & a & 0 & a & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \right].$$

En este caso  $p(z) = p^{(1)}(z)$  luego (1.29) nos conduce a

$$p(z) = \frac{1}{1 - a^2 z^2 p^{(1)}(z)} \Rightarrow p(z) = \frac{1 - \sqrt{1 - 4a^2 z^2}}{2a^2 z^2}, \quad (1.30)$$

donde hemos escogido el signo menos para que  $p(z)$  sea analítica en  $z = 0$ .

Recuperemos la medida a partir de la función de Stieltjes  $F(z; \mu) = -z^{-1}p(z^{-1})$  mediante la fórmula (1.17). Está claro que al no tener polos la función de Stieltjes no habrá puntos de masa. Un cálculo directo, para la parte de la medida absolutamente continua nos da

$$d\mu(x) = \omega(x)dx = \frac{1}{\pi} \frac{\sqrt{4a^2 - x^2}}{2a^2},$$

y no hay parte singular.

### 1.8.2. Ejemplo 2

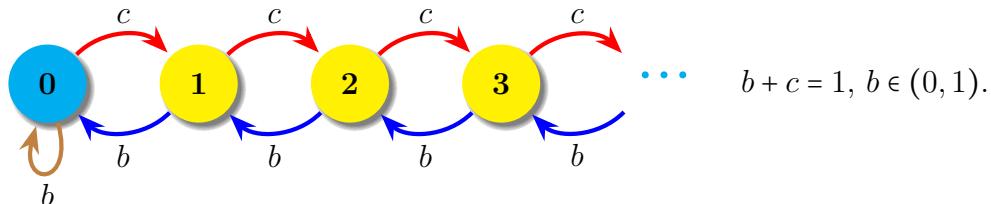
El segundo ejemplo corresponde a la matriz

$$\mathbb{J} = \left[ \begin{array}{c|cccccc} b & a & 0 & 0 & 0 & 0 & \dots \\ \hline a & 0 & a & 0 & 0 & 0 & \dots \\ 0 & a & 0 & a & 0 & 0 & \dots \\ 0 & 0 & a & 0 & a & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \right].$$

Usando (1.29) y teniendo en cuenta que en este caso  $p^{(1)}(z)$  es la función  $p(z)$  del ejemplo anterior se obtiene

$$p(z) = \frac{2}{1 - 2bz + \sqrt{1 - 4a^2z^2}}. \quad (1.31)$$

Consideremos el caso especial cuando  $a = \sqrt{b(1-b)}$ ,  $b \in (0, 1)$ . Entonces se obtiene una matriz de Jacobi que corresponde a la caminata aleatoria



En este caso (1.31) se convierte en

$$p(z) = \frac{2}{1 - 2bz + \sqrt{1 - 4a^2z^2}}, \quad (1.32)$$

luego, la ecuación de renuevo nos da

$$q(z) = 1 - \frac{1}{p(z)} = bz + \frac{1 - \sqrt{1 - 4a^2z^2}}{2}.$$

La función de Stieltjes es

$$F(z; \mu) = -z^{-1}p(z^{-1}) = \frac{1}{2b} \frac{z - 2b - \sqrt{z^2 + 4b^2 - 4b}}{z - 1}. \quad (1.33)$$

Calculamos

$$q(1) = bz + \frac{1 - \sqrt{(1 - 2b)^2}}{2} = \begin{cases} 1, & b \geq 1/2, \\ 2b, & b < 1/2. \end{cases}$$

Para el tiempo medio, si  $b \geq 1/2$ ,

$$\tau = \frac{dq(z)}{dz} \Big|_{z=1} = b + \frac{2(1-b)bz}{\sqrt{1-4(1-b)bz}} \Big|_{z=1} = \frac{b}{2b-1},$$

luego es positivamente recurrente si  $b > 1/2$ .

**Ejercicio 1.20.** Calcula, usando la función de Stieltjes (1.33), la medida de ortogonalidad y comprueba que tiene una masa en el cero si  $b > 1/2$ .

## 1.9. El modelo de urnas de Ehrenfest

La mecánica estadística estudia el comportamiento de sistemas con muchas partículas mediante el uso, entre otras muchas otras herramientas, de la teoría de las probabilidades.

Veamos un ejemplo. Imaginemos que tenemos una habitación con un tabique y todas las moléculas están en una esquina. ¿Qué ocurre si quitamos el tabique? La figura 1 nos muestra el resultado.

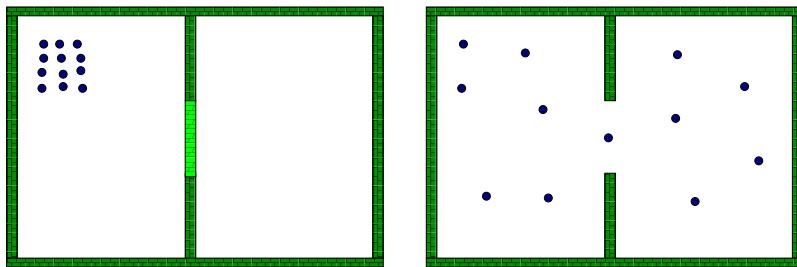


Figura 1: Izquierda: Gas concentrado en una esquina de la habitación. Derecha: El gas después de un tiempo se dispersa por toda la habitación.

La mecánica estadística tendría que dar como resultado que el proceso es reversible, pues esta se basa en las leyes de Newton que son reversibles (podemos ir hacia delante o hacia atrás en el tiempo) pero la segunda ley de la termodinámica nos dice que es irreversible. ¿Cómo resolver esta aparente contradicción?

Paul y Tatiana Ehrenfest propusieron en 1907 el siguiente modelo para explicar el dicha contradicción. Imaginemos que nos dan  $2N$  bolas numeradas colocadas en dos urnas,  $i$  en la urna de la izquierda y  $2N - i$  en la de la derecha (ver figura 2 izquierda). Se elige al azar un número del 1 al  $2N$  y la bola a la que le corresponde ese número se pasa a la otra urna. El estado del sistema queda fijado por el número de bolas  $i$  en la urna de la izquierda.

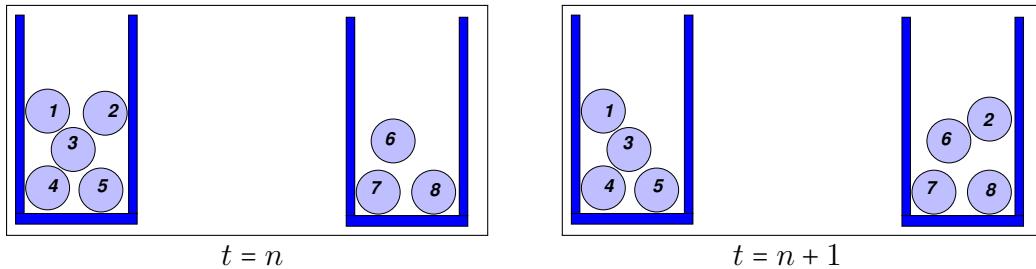
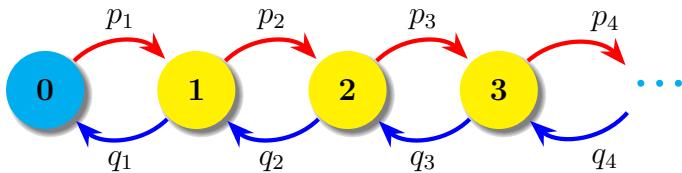


Figura 2: El modelo de urnas de los Ehrenfest.

Si sale el número 2, la bola 2 pasa de la urna izquierda a la urna derecha (ver figura 2 derecha).

Es obvio que es más probable que al elegir un número al azar, este corresponda a una bola de la urna que tenga más bolas, por lo que el estado de equilibrio sería que en cada urna haya  $\approx N$  bolas. Pero, ¿y si partimos del estado inicial  $2N$  bolas en la urna de la izquierda, qué probabilidad hay de que regresemos a dicho estado después de  $n$  pasos?

El proceso anterior se puede modelar mediante una *cadena de Markov* del tipo



Supongamos que en cierto momento estamos en el estado  $E_i$  ( $i$  bolas en la urna de la izquierda). Entonces el próximo estado en el que podemos estar es o bien el  $E_{i-1}$  o  $E_{i+1}$ . Además la probabilidad de estar en dichos estados es  $k/2N$  y  $1 - k/2N$ , respectivamente. Así, la matriz de probabilidades de transición será

$$\mathbb{P} = \begin{pmatrix} 0 & 1 & & & \\ \frac{1}{2N} & 0 & \frac{2N-1}{2N} & & \\ \frac{2}{2N} & \frac{2}{2N} & 0 & \frac{2N-2}{2N} & \\ & \ddots & 0 & \ddots & \ddots \\ & & \ddots & \ddots & \ddots \\ & & & 0 & \frac{1}{2N} \\ & & & \frac{2N}{2N} & 0 \end{pmatrix}$$

Para este caso resulta que existe una familia muy conocida de polinomios ortogonales (en este caso finita) que nos permite resolver el problema de forma analítica.

Vamos a considerar un caso particular de los polinomios de Kravchuk definidos por (ver, por ejemplo, [1])

$$K_i(x) = {}_2\tilde{F}_1\left(\begin{matrix} -i, -x \\ -2N \end{matrix}; 2\right), \quad x = 0, 1, \dots, 2N; \quad i = 0, 1, \dots, 2N,$$

donde  ${}_2\tilde{F}_1$  es la función hipergeométrica de Gauss

$${}_2\tilde{F}_1\left(\begin{matrix} a, b \\ c \end{matrix}; z\right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n! (c)_n} z^n, \quad (a)_n \equiv a(a+1)\dots(a+n-1), \quad (a)_0 = 1.$$

Estos satisfacen la propiedad de ortogonalidad

$$\sum_{x=0}^{2N} K_i(x) K_j(x) \frac{x}{2^{2N}} = \frac{(-1)^i i!}{(-2N)_i} \delta_{ij} \equiv d_i^2 \delta_{ij} \quad 0 \leq i, j \leq 2N,$$

de donde se sigue la siguiente RRTT

$$\frac{1}{2}(2N-i)K_{i+1}(x) - \frac{1}{2}2N K_i(x) + \frac{i}{2}K_{i-1}(x) = -xK_i(x)$$

que, como ya hemos visto, podemos escribir en forma matricial

$$\begin{pmatrix} 0 & 1 & & & & \\ \frac{1}{2N} & 0 & \frac{2N-1}{2N} & & & \\ & \frac{2}{2N} & 0 & \frac{2N-2}{2N} & & \\ & & \ddots & 0 & \ddots & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & 0 & \frac{1}{2N} \\ & & & & & \frac{2N}{2N} & 0 \end{pmatrix} \begin{pmatrix} K_0(x) \\ K_1(x) \\ K_2(x) \\ \vdots \\ K_{2N}(x) \end{pmatrix} = \left(1 - \frac{x}{N}\right) \begin{pmatrix} K_0(x) \\ K_1(x) \\ K_2(x) \\ \vdots \\ K_{2N}(x) \end{pmatrix},$$

siendo la matriz la misma que la de nuestra cadena de Markov. La fórmula de Karlin-McGregor nos conduce a (el coeficiente  $d_i$  corresponde a la norma, necesario para *ortonormalizar* a la familia  $K_n$ )

$$(\mathbb{P}^n)_{ij} = d_j^{-2} \sum_{x=0}^{2N} \left(1 - \frac{x}{N}\right)^n K_i(x) K_j(x) \frac{x}{2^{2N}}.$$

A partir de lo anterior tenemos

$$(\mathbb{P}^n)_{00} = \sum_{x=0}^{2N} \left(1 - \frac{x}{N}\right)^n \frac{x}{2^{2N}} \binom{2N}{x}$$

luego la función generatriz  $p_{00}(z) = p(z)$  se convierte en

$$p(z) \equiv \sum_{n=0}^{\infty} \mathbb{P}_{00}^n z^n = \sum_{x=0}^{2N} \frac{N}{N(1-z) + xz} \frac{\binom{2N}{x}}{2^{2N}}.$$

Es fácil comprobar que  $p(1) = \infty$  y por tanto la ecuación de renuevo nos dice que  $q(1) = 1 - 1/p(1) = 1$ , es decir el estado  $E_0$  (ninguna bola en la urna de la izquierda) es recurrente.

Vamos a calcular ahora el tiempo medio de espera. Para ello usamos la fórmula

$$\tau = q'(z) = \frac{p'(z)}{p^2(z)}.$$

Un cálculo directo nos da

$$q'(z) = \sum_{x=0}^{2N} \frac{N(N-x)}{(N(1-z) + xz)^2} \frac{\binom{2N}{x}}{2^{2N}}$$

luego  $q'(1) = 2^{2N}$ . Es decir, el estado  $E_0$  es positivamente recurrente.

**Ejercicio 1.21.** Comprueba que cualquier estado  $E_i$  ( $i$  bolas en la urna de la izquierda) también es positivamente recurrente y el tiempo medio de espera es

$$\tau_i = \frac{2^{2N}}{\binom{2N}{i}}.$$

Resumiendo:

1. Con probabilidad 1 se volvemos a cualquier estado inicial  $i$  del que hayamos partido, luego cualquier estado es recurrente
2. El tiempo esperado (o medio) de retorno al estado  $i$  dado es (en unidades de tiempo)

$$T_{ii} = \frac{2^{2N}}{\binom{2N}{i}} \quad \binom{2N}{i} = \frac{(2N)!}{(2N-i)!i!}$$

O sea, cualquier estado es positivamente recurrente.

Escojamos por ejemplo  $N = 1000$  (dos mil bolas) y repitamos nuestro experimento cada  $\tau = 10^{-9}$  segundos (una simulación con un ordenador cuyo procesador tenga una velocidad de 1GHz)

Si partimos del *equilibrio*  $i = N$  (mitad y mitad), el tiempo  $T_{NN}$  esperado para que el sistema regrese a él es de  $5,6 \times 10^{-8}$  segundos. Pero, si partimos de  $i = 0$  o  $i = 2N$  (todas en una urna),  $T_{00} \approx 3,6 \times 10^{585}$  años. Así pues, aunque podamos regresar al estado inicial el tiempo medio es desorbitado. Como dato interesante debemos decir que si tenemos en cuenta que la edad del Universo se estima en  $1,4 \times 10^{10}$  años, y que la vida en la Tierra desaparecerá en unos  $5 \times 10^9$  años podemos asegurar que no ha ninguna posibilidad de ver regresar las 2000 moléculas a la misma esquina.

## 2. Resolución de modelos cuánticos

### 2.1. Breve introducción histórica

Uno de los problemas que los físicos de finales del siglo XIX se planteaban era ¿cómo emiten la luz los cuerpos al calentarse? Para ello construyeron un modelo que llamaron *cuerpo negro*. 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.

El primero en establecer una ley empírica para la radiación del cuerpo negro fue Wien quien obtuvo la fórmula  $\alpha\omega^3e^{-\beta\omega/T}$ ,  $\alpha$  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 que establecía que la densidad por unidad de frecuencia de la radiación era

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

donde  $c$  es la velocidad de la luz y  $k$  es la constante de Boltzmann. Nótese que para frecuencias muy altas (zona de frecuencias ultravioletas) esta densidad es muy grande lo cual está en contradicción con las mediciones experimentales de Wein. Además, si calculamos la cantidad total de energía

$$U(T) = \int_0^\infty u(\omega, T) d\omega \tag{2.1}$$

obtenemos  $U(T) = \infty$ , lo cual no tiene sentido. Esta fue la llamada catástrofe ultravioleta.

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 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.

Para explicar su fórmula Plank, 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  $E = \hbar\omega$ .

Debemos destacar que la fórmula de Plank encajaba perfectamente en el engranaje de la física clásica de principios del siglo XX. Incluso daba una explicación a la fórmula para la densidad o luminosidad de la energía de radiación  $U(T) = \sigma T^4$ , donde  $\sigma$  es una constante conocida también como constante de Boltzmann para la densidad o luminosidad de la energía de radiación.

Años después (en el otoño de 1911), Henri Poincaré prueba que que la distribución

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

sólo se podía obtener bajo la suposición de que la energía está cuantizada, 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. Había nacido la física cuántica.

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. Sus resultados los publicó en uno de sus tres famosos artículos de 1905.

El próximo paso en la historia lo dio el danés Niels Bohr. Para intentar explicar las líneas espectrales del átomo de hidrógeno obligó 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.

O sea, una mezcla de las ideas de Plank y Einstein, que curiosamente coincidían espectacularmente con los datos experimentales. La conclusión era desconcertante: la energía de un átomo de hidrógeno dejaba de ser una cantidad continua y se convertía en un conjunto discreto de valores.

Antes de poder describir los fundamentos de la Mecánica Cuántica tenemos que detenernos en un personaje singular: el francés Louis de Broglie. Si Einstein había recuperado la propiedad corpuscular para la luz, De Broglie postuló la propiedad ondulatoria para la materia.

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).

Estos, y otros muchos trabajos generaron un gran *caos* en la física clásica pues eran incomprensibles para la mayoría de los científicos.

Los primeros pasos para encontrar una explicación satisfactoria los dieron, independientemente los físicos Werner Heisenberg (mecánica matricial) por un lado y Erwin Schrödinger (mecánica ondulatoria) por el otro. Aunque ambas formulaciones resultaron equivalentes, fue la de Schrödinger la que se impuso por su sencillez.

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, 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,

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

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

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi + V(x) \Psi(x, t) = E \Psi(x, t). \quad (2.2)$$

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

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi(x, y, z) - \frac{e^2}{r} \Psi(x, y, z) = E \Psi(x, y, z). \quad (2.3)$$

Schrödinger sabía como tratar este tipo de ecuaciones y la resolvió. En particular Schrödinger obtuvo para los valores de la energía en el estado estacionario del átomo de hidrógeno la fórmula

$$E_n = -\frac{me^4}{2\hbar^2 n^2}, \quad (2.4)$$

que era la misma que había obtenido Bohr con su modelo atómico.

La serie de trabajos que publicó Schrödinger en 1926, incluido la descripción del espectro del átomo de hidrógeno, calaron muy hondo en los físicos de la época. En apariencia se perdía la cualidad discreta del modelo de Bohr al aparecer nuevamente las “ondas” continuas, 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 pero no tuvo éxito en su intento.

Fue Max Born, colega y amigo de Heisenberg, quién rápidamente intuyó una interpretación plausible. 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 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.*

Ese mismo año, Heisenberg descubre una de las consecuencias más importantes y controvertidas de la Mecánica Cuántica: el principio de incertidumbre.

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

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

es decir<sup>2</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

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

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

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

---

<sup>2</sup>La operación  $\bar{a}$  denota el complejo conjugado de  $a$ .

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

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

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.

El principio de incertidumbre 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. Pero ¿de dónde sale esta incertidumbre que no existe en la mecánica clásica? Desde el punto de vista físico, este principio es consecuencia de la dualidad onda-partícula. De hecho, el 1929 el físico G.P. Thomson por un lado y los físicos C.J. Davisson y L.H. Germer por el otro, observaron de la *difracción* de electrones en dos experimentos independientes, lo que confirmó experimentalmente la hipótesis de De Broglie.

Para terminar con esta breve introducción histórica conviene comentar que 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.*

## 2.2. 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. 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.
- IV. 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.

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 respecto a la norma  $\|x\|$  inducida por el producto escalar<sup>3</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  $\hat{A}$  se dice hermítico si  $\langle a | \hat{A}b \rangle = \langle \hat{A}a | b \rangle$ . 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.

---

<sup>3</sup>Para el que usaremos la notación de Dirac: el producto escalar de  $x$  e  $y$  se denominará por  $\langle x | y \rangle$ .

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.

En el caso que nos ocupa nos interesará estudiar la ecuación de Shrödinger de una partícula material, por lo que resumiremos los postulados más importantes en el contexto de las funciones de cuadrado integrable.

Para más detalles se puede consultar [2, 8].

### 2.2.1. Los postulados de la Mecánica Cuántica en $L^2$

**Postulado 2.1.** *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  $d^3P(\vec{r}) = |\Psi(\vec{r}, t)|^2 d^3\vec{r}$ .*

Para que el postulado 2.1 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 está normalizada a la unidad. Al espacio vectorial de las funciones de cuadrado integrable lo denotaremos por  $L^2(\Omega)$  o simplemente  $L^2$ .

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 2.2.** *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>4</sup>*

---

<sup>4</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 está 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 2.3** (Ecuación dinámica). *La función de onda viene determinada por la ecuación de Schrödinger*

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(\vec{r}, t) + V(\vec{r}, t) \Psi(\vec{r}, t), \quad (2.5)$$

donde  $\Delta$  es el laplaciano<sup>5</sup> en  $\mathbb{R}^3$ , y  $V$  es el potencial al que está sometida la partícula.

La ecuación anterior se suele escribir de la forma

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = \hat{\mathcal{H}} \Psi(\vec{r}, t), \quad (2.6)$$

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.

Veamos ahora como podemos resolver la ecuación de Schrödinger (2.5) de forma sencilla.

### 3. Resolviendo la ecuación de Schrödinger: El método de Nikiforov-Uvarov

#### 3.1. La ecuación hipergeométrica generalizada

La ecuación hipergeométrica generalizada es una ecuación lineal de segundo orden de la forma

$$u''(z) + \frac{\tilde{\tau}(z)}{\sigma(z)} u'(z) + \frac{\tilde{\sigma}(z)}{\sigma^2(z)} u(z) = 0, \quad (3.1)$$

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.

---

<sup>5</sup>En coordenadas cartesianas es  $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ .

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 (3.1), así que al menos debemos tener

$$2\frac{\phi'(z)}{\phi(z)} + \frac{\tilde{\tau}(z)}{\sigma(z)} = \frac{\tau(z)}{\sigma(z)}, \quad \text{o} \quad \frac{\phi'(z)}{\phi(z)} = \frac{\tau(z) - \tilde{\tau}(z)}{2\sigma(z)} = \frac{\pi(z)}{\sigma(z)}, \quad (3.2)$$

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 (3.1) en la siguiente

$$\begin{aligned} y''(z) + \frac{\tau(z)}{\sigma(z)}y'(z) + \frac{\bar{\sigma}(z)}{\sigma^2(z)}y(z) &= 0, \\ \tau(z) &= \tilde{\tau}(z) + 2\pi(z), \end{aligned} \quad (3.3)$$

$$\bar{\sigma}(z) = \tilde{\sigma}(z) + \pi^2(z) + \pi(z)[\tilde{\tau}(z) - \sigma'(z)] + \pi'(z)\sigma(z).$$

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

$$\sigma(z)y'' + \tau(z)y' + \lambda y = 0. \quad (3.4)$$

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

$$\pi(z) = \frac{\sigma'(z) - \tilde{\tau}(z)}{2} \pm \sqrt{\left(\frac{\sigma'(z) - \tilde{\tau}(z)}{2}\right)^2 - \tilde{\sigma}(z) + k\sigma(z)}, \quad (3.5)$$

pero  $\pi(z)$  ha de ser un polinomio de grado a lo sumo uno, por tanto el polinomio

$$\Upsilon(z) = \left( \frac{\sigma'(z) - \tilde{\tau}(z)}{2} \right)^2 - \tilde{\sigma}(z) + k\sigma(z) \quad (3.6)$$

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 (3.5) 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 (3.5).

### 3.2. La ecuación diferencial hipergeométrica

Como hemos visto en el apartado anterior el estudio de la ecuación (3.1) se puede reducir al de la ecuación hipergeométrica (3.4) por lo que nos centraremos en el estudio de ésta última. Aquí nos restringiremos a estudiar las soluciones polinómicas de (3.4). Para el caso general remitimos al lector a [5].

#### 3.2.1. La propiedad de hipergeometricidad y la fórmula de Rodrigues

Pasemos a continuación a estudiar la ecuación diferencial

$$\sigma(x)y'' + \tau(x)y' + \lambda y = 0, \quad (3.7)$$

donde  $\sigma$  y  $\tau$  son polinomios de grados a lo sumo 2 y 1, respectivamente.

La ecuación (3.7) 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 (3.7) son tales que sus  $m$ -ésimas derivadas  $y^{(m)} := y_m$  satisfacen una ecuación del mismo tipo. En efecto, si derivamos (3.7)  $m$  veces obtenemos que  $y_m$  satisface una ecuación de la forma

$$\sigma(x)y_m'' + \tau_m(x)y_m' + \mu_m y_m = 0,$$

$$\tau_m(x) = \tau(x) + m\sigma'(x), \quad (3.8)$$

$$\mu_m = \lambda + \sum_{i=0}^{m-1} \tau_i'(x) = \lambda + m\tau'(x) + m(m-1)\frac{\sigma''(x)}{2}.$$

Es evidente que grado  $\tau_m \leq 1$  y que  $\mu_m$  es una constante. Además, toda solución de (3.8) es necesariamente de la forma  $y_m = y^{(m)}$  siendo  $y$  solución de (3.7). La demostración es por inducción y la omitiremos.

Vamos a intentar encontrar las soluciones polinómicas de (3.7). Para encontrarlas comenzaremos escribiendo (3.7) y (3.8) en su forma *simétrica o autoconjuguada*

$$\begin{aligned} [\sigma(x)\rho(x)y']' + \lambda\rho(x)y &= 0, \\ [\sigma(x)\rho_m(x)y'_m]' + \mu_m\rho_m(x)y_m &= 0, \end{aligned} \quad (3.9)$$

donde  $\rho$  y  $\rho_m$  son funciones de simetrización que satisfacen las ecuaciones diferenciales de primer orden (conocidas como ecuaciones de Pearson)

$$\begin{aligned} [\sigma(x)\rho(x)]' &= \tau(x)\rho(x), \\ [\sigma(x)\rho_m(x)]' &= \tau_m(x)\rho_m(x). \end{aligned} \quad (3.10)$$

Si  $\rho$  es conocida entonces, utilizando las ecuaciones anteriores, obtenemos para  $\rho_m$  la expresión

$$\rho_m(x) = \sigma^m(x)\rho(x). \quad (3.11)$$

**Teorema 3.1.** *Las soluciones polinómicas de la ecuación (3.8) se expresan mediante la fórmula de Rodrigues*

$$P_n^{(m)}(x) = \frac{A_{nm}B_n}{\rho_m(x)} \frac{d^{n-m}}{dx^{n-m}}[\rho_n(x)], \quad (3.12)$$

donde  $B_n = P_n^{(n)}/A_{nn}$  y

$$A_{nm} = A_m(\lambda) \Big|_{\lambda=\lambda_n} = \frac{n!}{(n-m)!} \prod_{k=0}^{m-1} [\tau' + \frac{1}{2}(n+k-1)\sigma'']. \quad (3.13)$$

Además, el autovalor  $\mu_m$  de (3.8) es

$$\mu_m = \mu_m(\lambda_n) = -(n-m)[\tau' + \frac{1}{2}(n+m-1)\sigma'']. \quad (3.14)$$

Demostración: Para demostrar el teorema vamos a escribir la ecuación autoconjuguada 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  $A_{nm} = A_m(\lambda)|_{\lambda=\lambda_n}$  y  $B_n = P_n^{(n)}/A_{nn}$ . Como  $P_n^{(n)}$  es una constante, de (3.8) obtenemos que  $\mu_n = 0$ , luego, usando la expresión (3.8)

$$\mu_n = \lambda_n + n\tau'(x) + n(n-1)\sigma''(x)/2 = 0,$$

deducimos que el valor de  $\lambda_n$  en (3.7) se expresa mediante la fórmula<sup>6</sup>

$$\lambda := \lambda_n = -n\tau' - \frac{n(n-1)}{2}\sigma''. \quad (3.15)$$

Sustituyendo (3.15) en (3.8) obtenemos el valor de  $\mu_{nm} = \mu_m(\lambda_n)$

$$\mu_{nm} = \mu_m(\lambda_n) = -(n-m)[\tau' + \frac{1}{2}(n+m-1)\sigma''], \quad (3.16)$$

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 (3.16) 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 (3.12) se convierte en la conocida fórmula de Rodrigues para los polinomios clásicos

$$P_n(x) = \frac{B_n}{\rho(x)} \frac{d^n}{dx^n} [\sigma^n(x)\rho(x)], \quad n = 0, 1, 2, \dots \quad (3.17)$$

La fórmula (3.15) determina los autovalores  $\lambda_n$  de (3.7) y es conocida como condición de *hipergeometricidad*.

---

<sup>6</sup>Usando la expresión (3.15) podemos obtener una expresión alternativa  $A_{nm} = (-n)_m \prod_{k=0}^{m-1} \frac{\lambda_{n+k}}{(n+k)}$ .

### 3.2.2. Ortogonalidad y relación de recurrencia

Veamos ahora cómo a partir de las ecuaciones diferenciales simetrizadas (3.9) podemos demostrar la ortogonalidad de las soluciones polinómicas respecto a la función peso  $\rho$ .

**Teorema 3.2.** *Supongamos que*

$$x^k \sigma(x) \rho(x) \Big|_a^b = 0, \quad \text{para todo } k \geq 0. \quad (3.18)$$

*Entonces las soluciones polinómicas  $P_n$  de la ecuación (3.7) 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*

$$\int_a^b P_n(x) P_m(x) \rho(x) dx = \delta_{n,m} d_n^2, \quad (3.19)$$

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 (3.7). 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 (3.18) 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 (3.9) 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

$$\int_a^b P_n^{(k)}(x) P_m^{(k)}(x) \rho(x) dx = \delta_{n,m} d_{kn}^2. \quad (3.20)$$

Finalmente, para calcular la norma  $d_n$  de los polinomios podemos utilizar la fórmula de Rodrigues. En efecto, sustituyendo (3.17) en (3.19) 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

$$d_n^2 = B_n (-1)^n n! a_n \int_a^b \sigma^n(x) \rho(x) dx. \quad (3.21)$$

**Teorema 3.3.** *Los polinomios ortogonales satisfacen una relación de recurrencia a tres términos de la forma*

$$xP_n(x) = \alpha_n P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad (3.22)$$

donde

$$\alpha_n = \frac{a_n}{a_{n+1}}, \quad \beta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}, \quad \gamma_n = \frac{c_n - \alpha_n c_{n+1}}{a_{n-1}} - \frac{b_n}{a_{n-1}} \beta_n = \frac{a_{n-1}}{a_n} \frac{d_n^2}{d_{n-1}^2}, \quad (3.23)$$

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 (3.22). Además, los coeficientes

$\alpha_n$ ,  $\beta_n$ , y  $\gamma_n$  se expresan mediante las fórmulas:

$$\begin{aligned}\alpha_n &= \frac{1}{d_{n+1}^2} \int_a^b x P_n(x) P_{n+1}(x) \rho(x) dx, & \beta_n &= \frac{1}{d_n^2} \int_a^b x P_n(x) P_n(x) \rho(x) dx, \\ \gamma_n &= \frac{1}{d_{n-1}^2} \int_a^b x P_n(x) P_{n-1}(x) \rho(x) dx.\end{aligned}\tag{3.24}$$

Para probar (3.23) basta sustituir la expresión  $P_n(x) = a_n x^n + b_n x^{n-1} + c_n x^{n-2} + \dots$  en (3.22) e igualar las potencias  $x^{n+1}$ ,  $x^n$  y  $x^{n-1}$ . Finalmente como  $x P_{n-1} = \alpha_{n-1} P_n + \beta_{n-1} P_{n-1} + \gamma_{n-1} P_{n-2}$ ,

$$\gamma_n = \frac{1}{d_{n-1}^2} \int_a^b P_n(x) [x P_{n-1}(x)] \rho(x) dx = \frac{1}{d_{n-1}^2} \alpha_{n-1} \int_a^b P_n^2(x) \rho(x) dx,$$

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 3.4.** *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*

$$P_n(x) = (x - \beta_n) P_{n-1}(x) - \gamma_n P_{n-2}(x), \quad n = \infty \in \mathbb{N}, \tag{3.25}$$

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 3.5.** *Si  $(P_n)_n$  es una sucesión de polinomios ortogonales que satisface la relación de recurrencia a tres términos (3.22). Entonces se cumple que*

$$\text{Ker}_n(x, y) \equiv \sum_{m=0}^n \frac{P_m(x) P_m(y)}{d_m^2} = \frac{\alpha_n}{d_n^2} \frac{P_{n+1}(x) P_n(y) - P_{n+1}(y) P_n(x)}{x - y}, \quad n \geq 1. \tag{3.26}$$

Si hacemos tender  $y \rightarrow x$  en la fórmula anterior obtenemos la fórmula *confluente* de Christoffel-Darboux:

$$\text{Ker}_n(x, x) \equiv \sum_{m=0}^n \frac{P_m^2(x)}{d_m^2} = \frac{\alpha_n}{d_n^2} [P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)] \quad n \geq 1. \quad (3.27)$$

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 3.6.** *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 (3.22) a tres que satisfacen los polinomios clásicos. Para calcular los coeficientes usando las expresiones de (3.23) 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 (3.12)  $P_n^{(n)}(x) = B_n A_{nn}$ , por tanto,

$$a_n = \frac{B_n A_{nn}}{n!} = B_n \prod_{k=0}^{n-1} [\tau' + \frac{1}{2}(n+k-1)\sigma'']. \quad (3.28)$$

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,

$$b_n = \frac{n\tau_{n-1}(0)}{\tau'_{n-1}} a_n. \quad (3.29)$$

Obsérvese que, al ser  $\tau'_n \neq 0$ ,  $b_n$  está definido para cualquier  $n$ .

Usando las expresiones (3.23) así como (3.29) 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 (3.7). Ello nos conduce a la expresión

$$c_n = -\frac{(n-1)[\tau(0) + (n-2)\sigma'(0)]b_n + n(n-1)\sigma(0)a_n}{(n-2)[\tau' + (n-3)\frac{\sigma''}{2}] + \lambda_n} \quad (3.30)$$

$$= -\frac{n(n-1)[\tau_{n-2}(0)\tau_{n-1}(0) + \sigma(0)\tau'_{n-1}]}{\tau'_{n-1}(\lambda_n - \lambda_{n-2})} a_n.$$

Luego nos resta sustituir la expresión anterior en la fórmula<sup>7</sup> (3.24)

$$\gamma_n = \frac{c_n - \alpha_n c_{n+1}}{a_{n-1}} - \frac{b_n}{a_{n-1}} \beta_n.$$

### 3.2.3. 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 (3.17) obtenemos

$$P_1(x) = \frac{B_1}{\rho(x)} [\sigma(x)\rho(x)]' = B_1 \tau(x), \quad (3.31)$$

y por tanto  $\tau$  es un polinomio de grado exactamente uno. Nótese que esta fórmula es equivalente a la fórmula de Pearson (3.10).

Si escribimos la fórmula de Rodrigues (3.12) para las derivadas<sup>8</sup>  $P_{n+m}^{(m)}$  con  $n = 1$  tenemos

$$P_{1+m}^{(m)}(x) = \frac{A_{m+1m}}{\rho_m(x)} [\rho_{m+1}(x)]' = \frac{A_{m+1m}}{\rho_m(x)} [\sigma(x)\rho_m(x)]' = A_{m+1m} \tau_m(x),$$

---

<sup>7</sup>Recordar que  $P_n(x) = a_n x^n + b_n x^{n-1} + c_n x^{n-2} + \dots$ .

<sup>8</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 (3.8) y satisfacen la fórmula de Rodrigues (3.12) cambiando  $n$  por  $n+m$ .

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.

Tomemos ahora  $m = 1$  en la fórmula (3.12). 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

$$P'_n(x) = \frac{-\lambda_n B_n}{\bar{B}_{n-1}} \bar{P}_{n-1}(x), \quad (3.32)$$

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 (3.12) 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,

$$\sigma(x)P'_n(x) = \frac{\lambda_n}{n\tau'_n} \left[ \tau_n(x)P_n(x) - \frac{B_n}{B_{n+1}} P_{n+1}(x) \right], \quad (3.33)$$

o, equivalentemente,

$$\left[ \frac{\lambda_n}{n\tau'_n} \mathbb{I} - \sigma(x)\mathcal{D} \right] P_n(x) = \frac{\lambda_{2n}}{2n} \alpha_n P_{n+1}(x). \quad (3.34)$$

Sustituyendo (3.22) en (3.34) obtenemos

$$\left[ \left( -\frac{\lambda_n}{n\tau'_n} + \frac{\lambda_{2n}}{2n} (x - \beta_n) \right) \mathbb{I} + \sigma(x)\mathcal{D} \right] P_n(x) = \frac{\lambda_{2n}}{2n} \gamma_n P_{n-1}(x). \quad (3.35)$$

Si ahora en la fórmula (3.33) desarrollamos  $\tau_n$  y utilizamos la relación de recurrencia (3.22) para descomponer los sumandos de la forma  $xP_n$  obtenemos el siguiente teorema

**Teorema 3.7.** Los polinomios ortogonales  $P_n(x)$ , soluciones de la ecuación (3.7), satisfacen la siguiente relación de estructura

$$\sigma(x)P'_n(x) = \tilde{\alpha}_n P_{n+1}(x) + \tilde{\beta}_n P_n(x) + \tilde{\gamma}_n P_{n-1}(x), \quad n \geq 0, \quad (3.36)$$

donde

$$\tilde{\alpha}_n = \frac{\lambda_n}{n\tau'_n} \left[ \alpha_n \tau'_n - \frac{B_n}{B_{n+1}} \right], \quad \tilde{\beta}_n = \frac{\lambda_n}{n\tau'_n} [\beta_n \tau'_n + \tau_n(0)], \quad \tilde{\gamma}_n = \frac{\lambda_n \gamma_n}{n} \neq 0. \quad (3.37)$$

Las expresiones (3.37) 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

$$\tilde{\alpha}_n = n \frac{\sigma''}{2} \alpha_n, \quad \tilde{\beta}_n = \frac{\lambda_n}{\tau'_n \tau'_{n-1}} [\tau(0)\sigma'' - \tau'\sigma'(0)] = \frac{\lambda_n}{n\tau'_n} \tau_n(\beta_n), \quad \tilde{\gamma}_n = \frac{\lambda_n \gamma_n}{n}. \quad (3.38)$$

Finalmente enunciaremos un teorema de gran importancia en lo que sigue.

**Teorema 3.8.** 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.

**Comentario 3.9.** 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 (3.31) y del teorema 3.6 que asegura que  $P_1$  ha de anularse en el interior del intervalo de ortogonalidad.

### 3.2.4. 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 (3.7), 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 1 y 2 están representados los principales parámetros de dichas familias, en las cuales  $(a)_n$  denota al símbolo símbolo de Pochhammer

$$(a)_0 = 1, \quad (a)_k = a(a+1)\cdots(a+k-1), \quad k = 1, 2, 3, \dots . \quad (3.39)$$

Para los polinomios  $\sigma$  se han escogido las llamadas *formas canónicas*.

Cuadro 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>9</sup> (3.12) 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

$${}_pF_q\left(\begin{array}{l} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{array} \middle| x\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{x^k}{k!}. \quad (3.40)$$

---

<sup>9</sup>Para ello basta usar la regla de Leibnitz para calcular la  $n$ -ésima derivada de un producto  $(f \cdot g)^{(n)} = \sum_{k=0}^n p_k n k f^{(k)} \cdot g^{(n-k)}$ . Otra posibilidad es usar series de potencias y el método de coeficientes indeterminados de Euler (ver e.g. [?, ?]).

Cuadro 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	$n$	$\frac{2(\alpha - \beta)n(n + \alpha + \beta + 1)}{(2n + \alpha + \beta)(2n + 2 + \alpha + \beta)}$
$\tilde{\gamma}_n$	$n$	$n(n + \alpha)$	$\frac{4n(n + \alpha)(n + \beta)(n + \alpha + \beta)(n + \alpha + \beta + 1)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)^2(2n + \alpha + \beta + 1)}$

De esta manera encontramos que

$$H_n(x) = \begin{cases} (-1)^m \left(\frac{1}{2}\right)_m {}_1F_1\left(\begin{array}{c} -m \\ \frac{1}{2} \end{array} \middle| x^2\right), & n = 2m \\ (-1)^m \left(\frac{3}{2}\right)_m x {}_1F_1\left(\begin{array}{c} -m \\ \frac{3}{2} \end{array} \middle| x^2\right), & n = 2m + 1 \end{cases} \quad (3.41)$$

$$L_n^\alpha(x) = \frac{(-1)^n \Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1)} {}_1F_1\left(\begin{array}{c} -n \\ \alpha + 1 \end{array} \middle| x\right), \quad (3.42)$$

$$P_n^{\alpha,\beta}(x) = \frac{2^n (\alpha + 1)_n}{(n + \alpha + \beta + 1)_n} {}_2F_1\left(\begin{array}{c} -n, n + \alpha + \beta + 1 \\ \alpha + 1 \end{array} \middle| \frac{1-x}{2}\right). \quad (3.43)$$

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 (3.12) aplicando la regla de Leibniz para calcular la  $n$ -ésima derivada de un producto de funciones.

$$H_n(0) = \begin{cases} \frac{(-1)^m (2m)!}{2^{2m} m!}, & n = 2m \\ 0, & n = 2m + 1 \end{cases}, \quad L_n^\alpha(0) = \frac{(-1)^n \Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1)}, \quad (3.44)$$

$$P_n^{\alpha, \beta}(1) = \frac{2^n (\alpha + 1)_n}{(n + \alpha + \beta + 1)_n}, \quad P_n^{\alpha, \beta}(-1) = \frac{(-1)^n 2^n (\beta + 1)_n}{(n + \alpha + \beta + 1)_n}.$$

### 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  $U_n(x)$

$$U_n(x) = P_n^{\frac{1}{2}, \frac{1}{2}}(x) = \frac{1}{2^n} \frac{\sin[(n+1) \arccos(x)]}{\sin[\arccos(x)]}.$$

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 (3.32) se obtienen las ecuaciones ( $\nu = 1, 2, 3, \dots$ ,  $n = 0, 1, 2, \dots$ ):

$$(H_n(x))^{(\nu)} = \frac{n!}{(n-\nu)!} H_{n-\nu}(x), \quad (3.45)$$

$$(L_n^\alpha(x))^{(\nu)} = \frac{n!}{(n-\nu)!} L_{n-\nu}^{\alpha+\nu}(x), \quad (3.46)$$

$$(P_n^{\alpha, \beta}(x))^{(\nu)} = \frac{n!}{(n-\nu)!} P_{n-\nu}^{\alpha+\nu, \beta+\nu}(x), \quad (3.47)$$

donde  $(P_n(x))^{(\nu)}$  denota la  $\nu$ -ésima derivada de  $P_n(x)$ .

### 3.3. 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.

#### 3.3.1. 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

$$-\frac{\hbar^2}{2m}\Psi''(x) + \frac{1}{2}m\omega^2x^2\Psi(x) = E\Psi(x), \quad x \in \mathbb{R}.$$

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 (3.1) con  $\tilde{\tau}(\xi) = 0$ ,  $\sigma(\xi) = 1$  y  $\tilde{\sigma}(\xi) = \varepsilon - \xi^2$ . Para  $\pi(\xi)$ , (3.5) 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

$$\begin{aligned} \pi(\xi) &= \xi, & \pi'(\xi) &= 1, & \lambda &= \varepsilon + 1, & \tau(\xi) &= 2\xi, \\ \pi(\xi) &= -\xi, & \pi'(\xi) &= -1, & \lambda &= \varepsilon - 1, & \tau(\xi) &= -2\xi, \end{aligned}$$

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

$$\Psi(\xi) = N_n e^{-\xi^2/2} H_n(\xi), \quad \varepsilon = 2n + 1, \quad n = 0, 1, 2, \dots \quad (3.48)$$

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

$$1 = \int_{-\infty}^{\infty} |\Psi(x)|^2 dx = x_0 \int_{-\infty}^{\infty} \Psi(\xi) \overline{\Psi(\xi)} d\xi,$$

luego  $N_n = \sqrt{\frac{2^n}{x_0 \sqrt{\pi} n!}}$ ,  $x_0 = \sqrt{\hbar/(m\omega)}$ .

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 3.9 del apartado anterior). De esta forma las únicas soluciones estacionarias del oscilador armónico son las funciones (3.48) anteriores.

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.

### 3.3.2. La ecuación de Schrödinger en un potencial central

Partiremos de la ecuación de Schrödinger estacionaria para el átomo de hidrógeno en coordenadas esféricas<sup>10</sup>

$$-\frac{\hbar^2}{2m} \Delta \Psi(r, \theta, \phi) + V(r) \Psi(r, \theta, \phi) = E \Psi(r, \theta, \phi), \quad (3.49)$$

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

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right), \quad (3.50)$$

---

<sup>10</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  $r, \theta, \phi$ .

o bien

$$\Delta = \Delta_r + \frac{1}{r^2} \Delta_\infty,$$

donde

$$\Delta_r = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right), \quad \Delta_\infty = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

denotan a los laplacianos radial y angular respectivamente.

Por simplicidad vamos a reescribir la ecuación anterior en la forma

$$\left[ \Delta_r + \frac{1}{r^2} \Delta_\infty \right] \Psi(r, \theta, \phi) + [\varepsilon - v(r)] \Psi(r, \theta, \phi) = 0, \quad (3.51)$$

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

$$\begin{aligned} \Delta_\infty Y(\theta, \phi) + \mu Y(\theta, \phi) &= 0, \\ \Delta_r F(r) + \left[ \varepsilon - v(r) - \frac{\mu}{r^2} \right] F(r) &= 0, \end{aligned} \quad (3.52)$$

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

$$\int_0^{2\pi} \int_0^\pi |Y(\theta, \phi)|^2 \sin \theta d\theta d\phi = 1 \quad \text{y} \quad \int_0^\infty |F(r)|^2 r^2 dr = 1. \quad (3.53)$$

### 3.3.3. Resolviendo la parte radial de la ecuación de Schrödinger

De los resultados del apartado anterior se sigue que (3.52) 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>11</sup>  $F(r) = R(r)/r$  que nos conduce a la ecuación

$$R''(r) + \left[ \varepsilon - v(r) - \frac{l(l+1)}{r^2} \right] R(r) = 0, \quad (3.54)$$

donde ahora la primera de las condiciones de contorno (3.53) se transforma en

$$\int_0^\infty |R(r)|^2 dr = 1.$$

La ecuación (3.54) será nuestro punto de partida para resolver un caso de extrema importancia: el átomo de hidrógeno.

---

<sup>11</sup>La razón fundamental es que  $\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial F(r)}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} [rF(r)]$ .

### 3.3.4. El átomo de hidrógeno

Puesto que para el átomo de hidrógeno  $V(r) = -\alpha/r$ , (3.54) 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  $\varepsilon = a_0E/\alpha$ , 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 (3.1) 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

$$\pi(\zeta) = \frac{1}{2} \pm \sqrt{\frac{1}{4} - 2\varepsilon\zeta^2 - 2r + l(l+1) + k\zeta}.$$

Como  $\Upsilon(\zeta) = 1/4 - 2\varepsilon\zeta^2 - 2r + l(l+1) + k\zeta$  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 3.9 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 3.9, 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  $(0, +\infty)$ —

$$\tau(\zeta) = 2(l+1 - \sqrt{-2\varepsilon}\zeta) \Rightarrow \lambda = k + \pi'(\zeta) = 2(1 - (l+1)\sqrt{-2\varepsilon}).$$

Usando (3.2) tenemos

$$\frac{\phi'(\zeta)}{\phi(\zeta)} = \frac{l+1 - \sqrt{-2\varepsilon}\zeta}{\zeta} \implies \phi(\zeta) = \zeta^{l+1} e^{-\sqrt{-2\varepsilon}\zeta}.$$

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

$$R_{n,l}(\zeta) = N_{n,l} x^{l+1} e^{-x/2} L_n^{2l+1}(x), \quad x = 2\sqrt{-2\varepsilon_{n,l}}\zeta,$$

con  $N_{n,l}$  tal que

$$\int_0^\infty R_{n,l}^2(r) dr = 1 \implies a_0 \int_0^\infty R_{n,l}^2(\zeta) d\zeta = 1 \implies \frac{a_0(n+l+1)}{2} \int_0^\infty R_{n,l}^2(x) dx = 1.$$

Ahora bien,

$$\begin{aligned} \int_0^\infty R_{n,l}^2(x) dx &= N_{n,l}^2 \int_0^\infty x^{2l+2} e^{-x} (L_n^{2l+1}(x))^2 dx = N_{n,l}^2 \int_0^\infty \rho(x) x (L_n^{2l+1}(x))^2 dx \\ &= N_{n,l}^2 \beta_n d_n^2 = N_{n,l}^2 2(n+l+1)n! \Gamma(n+2l+2), \end{aligned}$$

donde hemos usado la relación de recurrencia (3.22) 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 (3.49) 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>12</sup> de Bohr (2.4).

**Ejercicio 3.10.** *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

$$\Psi'' - \frac{2s}{1-s^2}\Psi' + \frac{\gamma^2(1-s^2) - \beta^2}{(1-s^2)^2}\Psi = E\Psi, \quad \beta^2 = -\frac{2mE}{\hbar^2\alpha^2}, \quad \gamma^2 = \frac{2mU_0}{\hbar^2\alpha^2}, \quad \gamma, \beta > 0.$$

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.

**Ejercicio 3.11.** *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 el caso  $l = 0$ . El potencial anterior se conoce como potencial de Morse y también modeliza la interacción molecular.

**Ejercicio 3.12.**

$$\Delta\psi_{KG} + \left[ \left( E_{KG} + \frac{\mu}{r} \right)^2 - 1 \right] \psi_{KG} = 0,$$

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

$$R'' + \left[ \left( E + \frac{\mu}{r} \right)^2 - 1 - \frac{l(l+1)}{r^2} \right] R = 0.$$

---

<sup>12</sup>En nuestra fórmula  $n$  no representa al número cuántico principal. Este correspondería al valor  $n' = n + l + 1$ .

*Demuestra usando el método descrito al final de capítulo que las soluciones se expresan mediante la fórmula*

$$R(x) = N_{l,n} e^{-\frac{r}{2}} r^{\nu+1} L_n^{2\nu+1}(2ar), \quad \nu = -\frac{1}{2} + \sqrt{\left(l + \frac{1}{2}\right)^2 - \mu^2},$$

*donde  $L_n^\alpha$  son los polinomios de Laguerre y determina el factor de normalización  $N_{l,n}$ .*

## Referencias

- [1] R. Álvarez-Nodarse, *Polinomios hipergeométricos y  $q$ -polinomios*. Monografías del Seminario Matemático “García de Galdeano” Vol. **26**. Prensas Universitarias de Zaragoza, Zaragoza, 2003. (versión actualizada y corregida en la web <http://euler.us.es/~renato/q-libro/>)
- [2] R. Álvarez-Nodarse, *Una introducción a la Mecánica Cuántica para “no iniciados”*. Notas de un curso introductorio de Mecánica Cuántica, 2016. Disponible en <http://euler.us.es/~renato/papers/cuantica.pdf>
- [3] L. Debnath y P. Mikusinsk. *Introduction to Hilbert spaces with applications*, Elsevier Academic Press, 2005.
- [4] S. Karlin y J. McGregor, Random walks, *Illinois J. Math.*, **3** (1959), 66–81.
- [5] A. F. Nikiforov y V. B. Uvarov, *Special Functions of Mathematical Physics*. Birkhäuser Verlag, Basilea, 1988.
- [6] J. R. Norris, *Markov Chains*, Cambridge University Press, 1998.
- [7] D. W. Stroock, *An Introduction to Markov Processes*, Springer Berlin Heidelberg, 2005.
- [8] S. Weinberg, *Lectures on Quantum Mechanics*. Cambridge University Press, 2012.