Probabilidades en cadena en los espacios de Hilbert. Aplicaciones físicas

Por DARÍO MARAVALL CASESNOVES

Recibido: 7 de marzo de 1984

Summary

The an author's theory about chain probabilities in the Euclidean and Hilbert spaces and about the vectors superposition in the spaces of Hilbert and Banach is developed. A distinction between first and second types of indeterminacy is made and it is shown the impossibility of passing from the quantum mechanics stochastic processes to those of the brownian movement.

A probabilistic interpretation of the polarization of the light, a theory of the Markov chains associated with the reduction of the wave packet in quantum mechanics and two ergodics theorems of the latter are presented.

FUENTES

Las investigaciones aquí expuestas son complemento y prolongación de otras anteriores, que ya había publicado en libros: *Estadística Teórica y Aplicada*, editada por la UNED en 1976; *Fundamentos de Mecánica Cuántica*, editado por la Universidad Politécnica de Madrid en 1979; la segunda edición del *Diccionario de Matemática Moderna*, editado en 1982 por la Editora Nacional, y en el discurso de apertura del curso 1975-1976, y en mis dos conferencias en 1975 del Cincuentenario de la Mecánica Cuántica, editados los tres por la Real Academia de Ciencias.

CONCLUSIONES

Se desarrolla la teoría de las probabilidades en cadena en C^m y en el espacio de Hilbert clásico (separable, de dimensión infinita). Se define y demuestran las propiedades de la superposición de vectores en espacios de Hilbert y de Banach, que llevan como consecuencia los fenómenos de interferencias de probabilidades. Estas estructuras matemáticas son sugeridas por la Mecánica Cuántica.

Se da una interpretación probabilística de la polarización de la luz y se construyen las cadenas de Markov propias de la Mecánica Cuántica, y se demuestran dos teoremas ergódicos de la misma.

Hay casos en que estos problemas se pueden resolver directamente en R^m o en el espacio de Hilbert clásico real, sin salirse de R a C.

1. PROBABILIDADES EN CADENA EN ESPACIOS EUCLIDEOS Y DE HILBERT. INDETERMINISMO DE PRIMERA Y DE SEGUNDA ESPECIE

Es sabido que en R^m a todo vector $\vec{p}$ normalizado y de componentes no negativas: $p_1, ..., p_m$, corresponde una distribución de probabilidad multinomial, por lo que el vector recibe el nombre de vector de probabilidad y es el elemento fundamental de las cadenas de Markov.

Pues bien, en C^m a todo vector unitario $\vec{x}$ (normalizado) de componentes $x_1, ..., x_m$ en una base ortonormal de C^m, corresponde un vector de probabilidad $\vec{p}$ de R^m definido por:

$$p_1 = |x_1|^2, ..., p_m = |x_m|^2 \tag{1}$$

Al vector $\vec{x}$ (véanse fuentes) le denominamos raíz cuadrada del vector $\vec{p}$, y a éste cuadrado del anterior. Es obvio que a cada vector de probabilidad $\vec{p}$ corresponden infinitas raíces cuadradas definidas por:

$$x_1 = |\sqrt{p_1}|e^{i\alpha_1}, ..., x_m = |\sqrt{p_m}|e^{i\alpha_m} \tag{2}$$

siendo las α constantes reales arbitrarias. Por el contrario, a todo vector unitario $\vec{x}$ de C^m corresponde un solo vector de R^m, que sea el cuadrado del anterior.

Si M es una matriz unitaria, entonces $M\vec{x}$ es también un vector unitario que define un nuevo vector de probabilidad. Se puede definir en C^m un proceso estocástico (de probabilidades en cadena) discreto en el tiempo, tal que si $\vec{x}_0$ es el vector unitario inicial, el vector unitario $\vec{x}_n$ en el instante enésimo es el:

$$\vec{x}_n = M\vec{x}_{n-1} \quad ; \quad \vec{x}_n = M^n\vec{x}_0 \tag{3}$$

Estas cadenas de probabilidad gozan de propiedades muy distintas de las de Markov, por de pronto, no se cumple la propiedad ergódica.

Si en cada instante n consideramos el vector unitario $\vec{y}_n$ transformado del $\vec{x}_n$ por la matriz unitaria T:

$$\vec{y}_n = T\vec{x}_n \tag{4}$$

en el tiempo, varía $\vec{y}_n$ de acuerdo con la ecuación:

$$\vec{y}_n = TM^nT^{-1}\vec{y}_0 = (TMT^{-1})^n\vec{y}_0 \tag{5}$$

que es un proceso estocástico paralelo al anterior. También son dos procesos estocásticos paralelos los definidos por la variación en el tiempo dada por [3], de las componentes de un mismo vector $\vec{x}$ de C^m en dos bases ortonormales distintas.

También se puede definir (véanse fuentes) un proceso de probabilidad en cadena en C^m continuo en el tiempo, de acuerdo con el sistema lineal de ecuaciones diferenciales de coeficientes constantes:

$$i \frac{d\vec{x}}{dt} = H\vec{x} \qquad [6]$$

donde H es una matriz hermítica con sus m valores propios distintos, porque:

$$||\vec{x}(0)||^2 = 1 \Rightarrow ||\vec{x}(t)||^2 = 1 \qquad [7]$$

lo que se demuestra multiplicando escalarmente a la derecha y a la izquierda ambos miembros de [6] por $\vec{x}$:

$$\langle i \frac{d\vec{x}}{dt}, \vec{x}\rangle = \langle H\vec{x}, \vec{x}\rangle \quad ; \quad \langle \vec{x}, i \frac{d\vec{x}}{dt} \rangle = \langle \vec{x}, H\vec{x}\rangle \qquad [8]$$

y restando queda:

$$i\left(\langle \frac{d\vec{x}}{dt}, x\rangle + \langle \vec{x}, \frac{d\vec{x}}{dt}\rangle\right) = i \frac{d}{dt} ||\vec{x}||^2 =$$
$$= \langle H\vec{x}, \vec{x}\rangle - \langle \vec{x}, H\vec{x}\rangle = \langle \vec{x}, (H^+ - H)\vec{x}\rangle = 0 \qquad [9]$$

donde por H^+ representamos la matriz adjunta de H, que por ser hermítica es igual a H.

Se puede pasar del proceso discreto al continuo en el tiempo, efectuando en [3] los siguientes cambios:

$$n \longrightarrow t \quad ; \quad n + 1 \longrightarrow t + dt \quad ; \quad M \longrightarrow I - iH\,dt \qquad [10]$$

en la que M es unitaria en primera aproximación, porque al ser I la matriz unidad y H hermítica, es:

$$M^{-1} = I + iH\,dt = M^+ \qquad [11]$$

la [3] se transforma en la:

$$\vec{x}(t + dt) = (I - iH\,dt)\vec{x}(t) \Rightarrow [6] \qquad [12]$$

En [6] podemos suponer la matriz H variable con el tiempo de modo continuo, sin que cambie nada de lo anterior y en [3[podemos suponer que la matriz M varía de instante en instante y la [3] se transforma en la:

$$\vec{x}_n = M_n\vec{x}_{n-1} \Rightarrow \vec{x}_n = M_n M_{n-1} \cdots M_1\vec{x}_0 \qquad [13]$$

Si las matrices M son reales, entonces son ortogonales, y se puede sustituir C^m por R^m. En el caso continuo de matrices reales, se puede sustituir la [6] por la:

$$\frac{d\vec{x}}{dt} = Q\vec{x} \qquad [14]$$

siendo Q una matriz antisimétrica, y se puede sustituir también C^m por R^m. En este caso se puede demostrar la normalización en el tiempo de $\vec{x}$ a partir de [14] multiplicando escalarmente (internamente) por $\vec{x}$ ambos miembros de [14], con lo que se obtiene:

$$\langle \vec{x}, \frac{d\vec{x}}{dt} \rangle = \frac{d}{dt} \frac{\|\vec{x}\|^2}{2} = \langle \vec{x}, Q\vec{x} \rangle = 0 \qquad [15]$$

por ser Q antisimétrica.

Obsérvese que las [1] son los cuadrados de los módulos de las coordenadas de un punto situado sobre una esfera de radio la unidad y centro en el origen de coordenadas, del espacio euclídeo complejo de m dimensiones. Por tanto, si un punto se mueve sobre dicha esfera con un movimiento cierto o aleatorio, hay asociado a dicho movimiento un proceso estocástico que denominamos indeterminado de primera y de segunda especie, respectivamente (véanse fuentes), por las [1]. Se puede sustituir el espacio euclídeo complejo de m dimensiones, por un espacio real euclídeo de m dimensiones, y en ese caso las [1] son los cuadrados de las coordenadas. En el caso del plano y del espacio ordinario, se pueden visualizar y materializar los anteriores procesos estocásticos con facilidad sobre la circunferencia y la esfera ordinaria (véanse fuentes).

Todos los resultados anteriores se generalizan sin más de C^m al espacio de Hilbert, y se definen así probabilidades en cadena en los espacios de Hilbert (véanse fuentes)*.

Tanto en el caso en que el tiempo varía de manera discreta, como cuando varía de modo continuo, en todo instante al vector $\vec{x}_n$ y a toda matriz hermítica A (en el espacio de Hilbert operador hermítico completo) de C^m con sus m valores propios $\lambda_1, \ldots, \lambda_m$ distintos, se le puede asociar una variable aleatoria (v.a.) ξ_A definida por:

$$\text{Prob}\,(\xi_A = \lambda_1) = |\langle \vec{u}_1, \vec{x}_n \rangle|^2, \ldots, \text{Prob}\,(\xi_A = \lambda_m) = |\langle \vec{u}_m, \vec{x}_n \rangle|^2 \qquad [16]$$

siendo $\vec{u}_1, \ldots, \vec{u}_m$ los m vectores propios distintos de A. Denotamos esta correspondencia por:

$$\xi_A = A * \vec{x}_n \qquad [17]$$

que leemos producto de composición de la matriz hermítica A por el vector unitario $\vec{x}_n$ que también denominaremos vector de onda. Cuando $\vec{x}$ varía en el tiempo, ξ_A varía también en el tiempo, lo que define un proceso estocástico en ambos casos.

Se puede definir un fenómeno similar al de la reducción o contracción del paquete de ondas de la Mecánica Cuántica, de la siguiente manera: si B es otra matriz hermítica de valores propios distintos $\mu_1, \ldots, \mu_m$ y vectores propios

* Al espacio de Hilbert clásico, es decir, al espacio de Hilbert separable y de dimensión infinita.

$\vec{v}_1, ..., \vec{v}_m$, llamamos producto de composición de las matrices A y B, que denotamos por $B * A$ (que no coincide con el producto ordinario de matrices) a la v.a. bivariante:

$$\chi_{BA} = \eta_B \xi_A = (B * A) * \vec{x} \qquad [18]$$

definida por:

$$\text{Prob}\,(\eta_B = \mu_j, \xi_A = \lambda_i) = |\langle \vec{v}_j, \vec{u}_i \rangle|^2 |\langle \vec{u}_i, \vec{x} \rangle|^2 \qquad [19]$$

Se tiene para las distribuciones marginales de probabilidad:

$$\text{Prob}\,(\xi_A = \lambda_i) = |\langle \vec{u}_i, \vec{x} \rangle|^2 \sum_{j=1}^{m} |\langle \vec{v}_j, \vec{u}_i \rangle|^2 = |\langle \vec{u}_i, \vec{x} \rangle|^2 \qquad [20]$$

que coincide con ξ_A y:

$$\text{Prob}\,(\eta_B = \mu_j) = \sum_{i=1}^{m} |\langle \vec{v}_j, \vec{u}_i \rangle|^2 |\langle \vec{u}_i, \vec{x} \rangle|^2 \neq \text{Prob}\,(\xi_B = \mu_j) \qquad [21]$$

que es distinta de ξ_B.

Obsérvese que la v.a. bivariante:

$$\chi_{AB} = \eta_A \xi_B = (A * B) * \vec{x} \qquad [22]$$

viene definida por:

$$\text{Prob}\,(\eta_A = \lambda_i, \xi_B = \mu_j) = |\langle \vec{u}_i, \vec{v}_j \rangle|^2 |\langle \vec{v}_j, \vec{x} \rangle|^2 \qquad [23]$$

que muestra que en general son distintas χ_{BA} y χ_{AB}.

En el caso en que las dos matrices A y B son conmutativas, entonces sus vectores propios son iguales:

$$\vec{v}_i = \vec{u}_i, \quad i = 1, 2, ..., m \qquad [24]$$

y las [19] y [23] se transforman en las:

$$\text{Prob}\,(\eta_B = \mu_j, \xi_A = \lambda_i) = 0 \quad ; \quad \text{Prob}\,(\eta_A = \lambda_i, \xi_B = \mu_j) = 0, \quad i \neq j$$
$$\text{Prob}\,(\eta_B = \mu_i, \xi_A = \lambda_i) = \text{Prob}\,(\eta_A = \lambda_i, \xi_B = \mu_i) = |\langle \vec{u}_i, \vec{x} \rangle|^2 \qquad [25]$$

y de aquí:

$$AB = BA \Rightarrow (A * B)\vec{x} = (B * A)\vec{x} \Leftrightarrow \chi_{AB} = \chi_{BA} \qquad [26]$$

donde el producto de matrices en el primer miembro de la implicación es el ordinario.

Se puede demostrar que la implicación [26] es una equivalencia:

$$A * B = B * A \Leftrightarrow AB = BA \qquad [27]$$

porque la igualdad de las [19-23], consecuencia de la implicación [26] en sentido inverso, arrastran las [24] que implican la conmutatividad de A y B. Se tiene, por tanto, que:

$$AB = BA \Leftrightarrow \text{Prob}\,(\xi_A = \lambda_i) = \text{Prob}\,(\xi_B = \mu_i) \qquad [28]$$

además de [27].

Los resultados anteriores se extienden sin más al espacio de Hilbert, sustituyendo las matrices hermíticas por operadores hermíticos y las distribuciones bivariantes de probabilidad finitas por otras infinitas*.

En el caso de la circunferencia y de la esfera ordinaria, a las v.a. asociadas a matrices hermíticas (simétricas por ser reales) corresponden las distribuciones de probabilidad binomial o trinomial definidas por los cuadrados de las coordenadas de un punto móvil sobre una circunferencia o una esfera, respecto a los infinitos pares o ternas de ejes ortogonales cartesianos. También es fácil visualizar el fenómeno de la contracción del paquete de ondas, antes definido de forma general para espacios euclídeos complejos de cualquier número de dimensiones, o espacios de Hilbert.

Vamos a demostrar que el valor medio de $\bar{\xi}_A$, que también designamos por $\bar{A}$, viene dado por:

$$\bar{A} = \bar{\xi}_A = \langle x, Ax \rangle \qquad [29]$$

En efecto, se tiene que:

$$\vec{x} = \sum_{i=1}^{m} \langle \vec{x}, \vec{u}_i \rangle \vec{u}_i \quad ; \quad A\vec{x} = \sum_{i=1}^{m} \lambda_i \langle \vec{x}, \vec{u}_i \rangle u_i \qquad [30]$$

y de aquí:

$$\langle \vec{x}, A\vec{x} \rangle = \langle \Sigma \langle \vec{x}, \vec{u}_j \rangle \vec{u}_j, \Sigma \lambda_i \langle \vec{x}, \vec{u}_i \rangle \vec{u}_i \rangle = \Sigma \lambda_i |\langle \vec{x}, \vec{u}_i \rangle|^2 \qquad [31]$$

y análogamente para cualquier potencia de orden r de A:

$$\overline{\xi^r_A} = \overline{A^r} = \langle \vec{x}, A^r \vec{x} \rangle \qquad [32]$$

La suma de los productos de composición de dos matrices hermíticas por un vector unitario no tiene sentido, pero por [31] se cumple que:

$$\overline{(A + B) * \vec{x}} = \overline{A * \vec{x}} + \overline{B * \vec{x}} \qquad [33]$$

Para una matriz diagonal:

$$\Delta = ||\lambda_i \delta_{ij}|| \qquad [34]$$

* Operadores de espectro discreto, con todos los valores propios distintos y cuyos vectores propios forman una base del espacio de Hilbert.

como sus vectores propios son los $\vec{e}_i$:

$$\vec{e}_1 = ||1, 0, ..., 0||; \vec{e}_2 = ||0, 1, 0, .., 0||; ...; \vec{e}_m = ||0, 0, ..., 0, 1|| \quad [35]$$

se tiene que ξ_Δ viene definida por:

$$\text{Prob}\ (\xi_\Delta = \lambda_i) = |\langle\vec{x}, \vec{e}_i\rangle|^2 = |\vec{x}_i|^2 \quad [36]$$

La teoría se puede modificar y adaptar a matrices hermíticas convalores propios iguales, e incluso para aquellas que no son completas, es decir, que con sus vectores propios no se puede formar una base ortonormal de C^m o del espacio de Hilbert. Las matrices múltiñlos de la unidad λI (I es la matriz unidad) están asociadas a constantes, porque con dicha modificación es:

$$\text{Prob}\ (\xi \neq \lambda) = 0 \quad [37]$$

Si C es otra matriz hermítica de m valores propios distintos $v_1, ..., v_m$ y vectores propios $w_1, ..., w_m$, al producto de composición $C * B * A$ y al vector unitario $\vec{x}$ se le asocia una v.a. trivariante:

$$\chi_{CBA} = \tau_C \eta_B \xi_A \quad [38]$$

definida por:

$$\text{Prob}\ (\tau_C = v_k, \eta_B = \mu_j, \xi_A = \lambda_i) = |\langle\vec{w}_k, \vec{v}_j\rangle|^2 |\langle\vec{v}_j, \vec{u}_i\rangle|^2 |\langle\vec{u}_i, \vec{x}\rangle|^2 \quad [39]$$

Obsérvese que las distribuciones de probabilidad marginales de la v.a. bivariante B, A y univariante A, son las correspondientes a χ_{BA} y ξ_A, respectivamente.

Vamos a calcular la derivada respecto al tiempo del valor medio ξ_A. Se tiene que:

$$\frac{d\bar{A}}{dt} = \frac{d\xi_A}{dt} = \langle \frac{d\vec{x}}{dt}, A\vec{x}\rangle + \langle\vec{x}, \frac{dA}{dt}\vec{x}\rangle + \langle\vec{x}, A\frac{d\vec{x}}{dt}\rangle \quad [40]$$

y teniendo en cuenta la [6] se sigue de la anterior que:

$$\frac{d\bar{A}}{dt} = \langle -iH\vec{x}, A\vec{x}\rangle + \langle\vec{x}, \frac{dA}{dt}\vec{x}\rangle + \langle\vec{x}, -iAH\vec{x}\rangle \quad [41]$$

luego:

$$\frac{d\bar{A}}{dt} = \langle\vec{x}, \left(\frac{dA}{dt} + i(HA - AH)\right)\vec{x}\rangle = \overline{\frac{dA}{dt}} + \overline{i(HA - AH)} \quad [42]$$

Por el mismo cálculo se demuestra que cualquiera que sea la potencia errésima de A es:

$$\frac{d\overline{A^r}}{dt} = \overline{\frac{dA^r}{dt}} + \overline{i(HA^r - A^rH)} \quad [43]$$

Si A es independiente del tiempo, y conmuta con H, como:

$$AH = HA \Rightarrow A^r H = HA^r \qquad [44]$$

de [43] se sigue que:

$$\frac{d\overline{A^r}}{dt} = \frac{\overline{dA^r}}{dt} = 0 \qquad [45]$$

lo que implica que ξ_A es independiente del tiempo. Como:

$$-\frac{dA}{dt} = i(HA - AH) \Rightarrow \frac{dA^r}{dt} = i(A^r H - HA^r) \qquad [46]$$

si A depende del tiempo, pero se cumple la primera [46], entonces de [42] y [43] se sigue que ξ_A es independiente del tiempo. Cuando ξ_A es independiente del tiempo se dice que es una integral primera de este proceso estocástico.

Si A es independiente del tiempo, de [43] se sigue que:

$$\frac{d\overline{A^r}}{dt} = \overline{i(HA^r - A^r H)} \qquad [47]$$

Si A conmuta con H se tiene que:

$$\frac{d\overline{A^r}}{dt} = \frac{\overline{dA^r}}{dt} \qquad [48]$$

En el caso en que el tiempo varía de modo discreto es:

$$\overline{\xi_n^r} = \langle \vec{x}_n, A_n^r \vec{x}_n \rangle = \langle M_n \vec{x}_{n-1} A_n^r M_n \vec{x}_{n-1} \rangle = \langle \vec{x}_{n-1}, M_n^{-1} A_n^r M_n \vec{x}_{n-1} \rangle$$

y:

$$\overline{\xi_n^r} = \overline{\xi_{n-1}^r} \Leftrightarrow M_n^{-1} A_n^r M_n = A_{n-1}^r \qquad [49]$$

pero como:

$$M_n^{-1} A_n M_n = A_{n-1} \Leftrightarrow M_n^{-1} A_n^r M_n = A_{n-1}^r \qquad [50]$$

se sigue que la condición necesaria y suficiente para que ξ_n sea independiente del tiempo es:

$$M_n^{-1} A_n M_n = A_{n-1} \qquad [51]$$

que si M y A son independientes del tiempo se simplifica en la conmutatividad de M y A.

Si en [51] pasamos de lo discreto a lo continuo, haciendo los cambios [10] se obtiene:

$$(I + iH\,dt)\left(A + \frac{dA}{dt}dt\right)(I - iH\,dt) = A \Rightarrow$$

$$\Rightarrow \frac{dA}{dt} + i(HA - AH) = 0 = [46] \qquad [52]$$

Como H es hermítica con sus m valores propios distintos, tiene una doble función: es la matriz de la evolución del vector de onda [6] y a ella está también asociada una v.a. ξ_H.

Como H conmuta consigo misma, por [43] es:

$$\frac{d\overline{H^r}}{dt} = \overline{\frac{dH^r}{dt}}, \quad \forall\, r \in N \qquad [53]$$

que muestra que las derivadas respecto al tiempo de los valores medios de las potencias de H son iguales a los valores medios de las potencias de la derivada respecto al tiempo de H. Las operaciones de derivación respecto al tiempo y de cálculo del valor medio son conmutativas. Si H es independiente del tiempo, por lo visto anteriormente, es una integral primera.

Si se conoce a priori la distribución de probabilidad de ξ_A se conocen todas las:

$$|\langle \vec{u}_i, \vec{x} \rangle|^2 = c_i^2 \qquad [54]$$

en virtud de [16] y, por tanto, se conocen las distribuciones de probabilidad de χ_{BA} [19] y de χ_{CBA} [39], pero no se conoce exactamente el vector $\vec{x}$, porque éste, para que se cumpla [54], basta conque sea de la forma:

$$\vec{x} = \sum_{j=1}^{m} |c_j| e^{\alpha_j i} \vec{u}_j \qquad [55]$$

donde las α son números reales cualesquiera.

Análogamente, si se conoce la distribución marginal de η_B [20] se pueden calcular las [54]. Y lo mismo puede decirse si se conoce la distribución marginal de τ_C en [39].

Las distribuciones de probabilidad [19] y [39] de χ_{BA} y de χ_{CBA} resultan cuando se aplican los operadores B, o el B y el C inmediatamente después del A, sin haber dejado evolucionar en el tiempo al vector $\vec{x}$ entre los instantes de aplicación de dichos operadores. Si se deja pasar un intervalo de tiempo entre la aplicación de A y de B, la reducción del paquete de ondas actúa de la siguiente manera: $\vec{x}(t)$, evoluciona en el tiempo de acuerdo con [6] a partir de un valor inicial $\vec{x}(0)$; si en el instante t_1 se aplica el operador A, $\vec{x}(t_1)$ pasa bruscamente a valer uno cualquiera de sus vectores propios $\vec{u}_i(t_1)$ con la probabilidad $p_i(*)$. Si pasa a valer $\vec{u}_i(t_1)$, por ejemplo, $\vec{x}(t)$ evoluciona de acuerdo con [6] a partir del tiempo t_1 con el valor inicial para t_1 igual a $\vec{u}_i(t_1)$, y si en el instante t_2 se aplica el operador B, $\vec{x}(t_2)$ pasa bruscamente a valer uno cualquiera de los vectores propios de B: $\vec{v}_j(t_2)$, con las probabilidades:

$$|\langle \vec{v}_j(t_2), \vec{x}(t_2)\rangle|^2 \quad ; \quad \vec{x}(t_1) = \vec{u}_i(t_1) \qquad [56]$$

si es el $\vec{v}_j(t_2)$ en concreto, entonces $\vec{x}(t)$ evoluciona a partir de t_2 de acuerdo con [6] y el valor inicial $\vec{x}(t_2) = \vec{v}_j(t_2)$. Las mismas consideraciones son válidas, cuando el tiempo varía de manera discreta.

Si T es una matriz unitaria, el vector $\vec{y}$:

$$\vec{y} = T\vec{x} \qquad [57]$$

evoluciona en el tiempo de acuerdo con la ecuación:

$$i\frac{d}{dt}(T^{-1}\vec{y}) = HT^{-1}\vec{y} \qquad [58]$$

que si T es independiente del tiempo se simplifica en la:

$$i\frac{d\vec{y}}{dt} = THT^{-1}\vec{y} \qquad [59]$$

Si H y T conmutan, entonces $\vec{x}$ e $\vec{y}$ evolucionan de acuerdo con la misma [6].

Si T depende del tiempo, la [58] se escribe:

$$iT^{-1}\frac{d\vec{y}}{dt} + i\frac{dT^{-1}}{dt}\vec{y} = HT^{-1}\vec{y} \Rightarrow i\frac{d\vec{y}}{dt} = T\left(HT^{-1} - i\frac{dT^{-1}}{dt}\right)\vec{y} \qquad [60]$$

y si:

$$i\frac{dT^{-1}}{dt} = HT^{-1} - T^{-1}H \qquad [61]$$

$\vec{x}$ e $\vec{y}$ evolucionan de acuerdo con la misma ecuación [6]. Aun cuando [46] y [61] son de la misma forma, se distinguen en que A de [46] es hermítica y T de [61] es unitaria, por tanto esta última no es integral primera, a menos que sea hermítica. También se cumple que:

$$\frac{dT}{dt} + i(HT - TH) = 0 \qquad [62]$$

porque [62] es la condición necesaria y suficiente tanto como [61] para que $\vec{x}$ e $\vec{y}$ evolucionen de acuerdo con la misma [6]. También se puede demostrar directamente, porque:

$$0 = \frac{d}{dt}(TT^{-1}) = \frac{dT}{dt}T^{-1} + T\frac{dT^{-1}}{dt} \qquad [63]$$

* $p_i = |\langle \vec{u}(t_1), \vec{x}(t_1)\rangle|^2$ [55 bis].

y multiplicando a la derecha por T se obtiene:

$$\frac{dT}{dt} + T\frac{dT^{-1}}{dt}T = \frac{dT}{dt} + Ti(T^{-1}H - HT^{-1})T = 0 \Rightarrow [62] \qquad [64]$$

En el caso en que el tiempo varía de modo discreto, es:

$$\vec{x}_n = M_n\vec{x}_{n-1} \quad ; \quad \vec{y}_n = T_n\vec{x}_n \qquad [65]$$

y la condición para que $\vec{x}$ e $\vec{y}$ evolucionen de acuerdo con la misma ecuación es:

$$\vec{y}_n = T_nM_nT_{n-1}^{-1}\vec{y}_{n-1} \Rightarrow T_nM_nT_{n-1}^{-1} = M_n \qquad [66]$$

que si M y T son independientes en el tiempo, se simplifica en que M y T conmutan. Se ontiene la [62] a partir de [66] haciendo los cambios:

$$M_n = I - iH\,dt \quad ; \quad T_n = T \quad ; \quad T_{n-1} = T - \frac{dT}{dt}dt \qquad [67]$$

en [66] que la transforman en la [62] despreciando lostérminos en dt^2.

Existe una diferencia muy importante entre el caso discreto [3] y el continuo [6] y es que mientras en este último H es hermítica y, por tanto, hay asociada a la misma una v.a. ξ_H, por el contrario las M son unitarias en vez de hermíticas, y, por tanto, no hay asociadas a las mismas v.a.

Estas estructuras matemáticas y las que siguen me han sido sugeridas por la Mecánica Cuántica. Cuando a un fenómeno hay asociada una distribución de probabilidad que varía de manera determinística, le llamamos indeterminado de primera especie. Tal es el caso de la función de onda de la ecuación de Schródinger. Si la distribución de probabilidad asociada al fenómeno varía de manera probabilística, le llamamos indeterminado de segunda especie, tal es el caso de la distribución trinomial de probabilidad asociada a un punto de la esfera, tal como hemos descrito, cuando este punto tiene un movimiento browniano.

Aun cuando todos los desarrollos anteriores, los hemos hecho para C^m, como ya hemos dicho, son aplicables al espacio de Hilbert clásico, sustituyendo las matrices hermíticas y unitarias, por operadores hermíticos y unitarios de espectro discreto y cuyos vectores propios forman una base del espacio de Hilbert. Para los de espectro continuo hay que introducir modificaciones.

En el caso del espacio de Hilbert, la derivada respecto al tiempo del vector de onda en [6] puede ser parcial en vez de ordinaria.

2. SUPERPOSICION DE VECTORES EN ESPACIOS DE HILBERT Y DE BANACH. INTERFERENCIAS DE PROBABILIDADES

En la estructura matemática del parágrafo anterior (véanse fuentes) existen interferencias de probabilidades. Sean $\vec{x}$ e $\vec{y}$ dos vectores normalizados de un espacio de Hilbert E (o de C^m); llamamos superposición de $\vec{x}$ e $\vec{y}$ a su suma seguida de renormalización, es decir al vector:

$$\frac{\vec{x} + \vec{y}}{||\vec{x} + \vec{y}||} \tag{68}$$

La superposición se puede extender a cualquier número de vectores y además afectar a estos de pesos de superposición complejos, de modo que al conjunto de vectores $\vec{x}_1, ..., \vec{x}_n$ dotados de los pesos $c_1, ..., c_n$, les corresponde por superposición el vector normalizado:

$$\frac{c_1\vec{x}_1 + \cdots + c_n\vec{x}_n}{||c_1\vec{x}_1 + \cdots + c_n\vec{x}_n||} \tag{69}$$

Este concepto matemático de superposición es el que físicamente se traduce en interferencias de las probabilidades, porque, por ejemplo, el valor medio del operador A para el vector [68] da:

$$\frac{1}{||\vec{x} + \vec{y}||^2} \langle \vec{x} + \vec{y}, A(\vec{x} + \vec{y})\rangle \tag{70}$$

distinto del de la media aritmética:

$$\frac{1}{2}(\langle \vec{x}, A\vec{x}\rangle + \langle \vec{y}, A\vec{y}\rangle) \tag{71}$$

Vamos a demostrar que la superposición es la adición de un grupo conmutativo G, que vamos a construir ahora. Sea E' el conjunto de los vectores normalizados de E, $\vec{\theta}$ el vector nulo de E y hagamos el convenio:

$$\frac{\vec{\theta}}{0} = \vec{\theta} \tag{72}$$

sea R^+ el conjunto de los números reales positivos, entonces el conjunto G, que es el producto cartesiano directo $R^+ \cup 0 \times E' \cup \vec{\theta}$ con el convenio [72], tiene estructura de grupo abeliano para la adición definida por:

$$(a, \vec{x}) + (b, \vec{y}) = \left(||a\vec{x} + b\vec{y}||, \frac{a\vec{x} + b\vec{y}}{||a\vec{x} + b\vec{y}||}\right) \tag{73}$$

si consideramos equivalentes entre sí todos los elementos de la forma:

$$(a, \vec{\theta}) \quad ; \quad (0, \vec{x}), \quad \forall\, a \in R^+, \ \forall\, \vec{x} \in E' \tag{74}$$

que es el elemento neutro de G, porque:

$$(a, \vec{\theta}) + (b, \vec{y}) = (b, \vec{y}) \quad ; \quad (0, \vec{x}) + (b, \vec{y}) = (b, \vec{y}) \tag{75}$$

Se cumplen los cuatro axiomas del grupo: *a*) es cerrado ya que el segundo miembro de [73] pertenece a *G*; *b*) la ley de composición es asociativa, porque:

$$(a, \vec{x}) + [(b, \vec{y}) + (c, \vec{z})] = (a, \vec{x}) + \left(||b\vec{y} + c\vec{z}||, \frac{b\vec{y} + c\vec{z}}{||b\vec{y} + c\vec{z}||} \right) =$$

$$= \left(||a\vec{x} + b\vec{y} + c\vec{z}||, \frac{a\vec{x} + b\vec{y} + c\vec{z}}{||a\vec{x} + b\vec{y} + c\vec{z}||} \right) = [(a, \vec{x}) + (b, \vec{y})] + (c, \vec{z}) \quad [76]$$

c) existe elemento neutro, que es el [74]; *d*) existe el elemento inverso del $(a, \vec{x})$, que es el $(a, -\vec{x})$, porque:

$$(a, \vec{x}) + (a, -\vec{x}) = \left(0, \frac{\vec{\theta}}{||\vec{\theta}||} \right) = (0, \vec{\theta}) \quad [77]$$

De [73] se sigue que el grupo es abeliano. Obsérvese que todos los elementos de *G* tienen una representación única, salvo el neutro.

G define un espacio vectorial sobre el cuerpo de los números complejos *C*, cuando se define el producto de un elemento $(a, \vec{x})$ de *G* por un número complejo λ de la siguiente manera:

$$\lambda \cdot (a, \vec{x}) = \left(|\lambda| a, \frac{\lambda x}{|\lambda|} \right) \quad [78]$$

Se cumplen las cuatro leyes del espacio vectorial: el producto de la suma de vectores por un escalar es distributivo:

$$\lambda \cdot [(a, \vec{x}) + (b, \vec{y})] = \lambda \cdot \left(||a\vec{x} + b\vec{y}||, \frac{a\vec{x} + b\vec{y}}{||a\vec{x} + b\vec{y}||} \right) =$$

$$= \left(|\lambda| ||a\vec{x} + b\vec{y}||, \frac{\lambda(a\vec{x} + b\vec{y})}{|\lambda| ||a\vec{x} + b\vec{y}||} \right) = \lambda \cdot (a, \vec{x}) + \lambda \cdot (b, \vec{y}) =$$

$$= \left(|\lambda| a, \frac{\lambda\vec{x}}{|\lambda|} \right) + \left(|\lambda| b, \frac{\lambda\vec{y}}{|\lambda|} \right) \quad [79]$$

El producto de la suma de dos escalares por un vector también es distributivo:

$$(\lambda + \mu) \cdot (a, \vec{x}) = \left(|\lambda + \mu| a, \frac{(\lambda + \mu)\vec{x}}{|\lambda + \mu|} \right) =$$

$$= \lambda \cdot (a, \vec{x}) + \mu \cdot (a, \vec{x}) = \left(|\lambda| a, \frac{\lambda\vec{x}}{|\lambda|} \right) + \left(|\mu| a, \frac{\mu\vec{x}}{|\mu|} \right) =$$

$$= \left(|\lambda + \mu| a, \frac{(\lambda + \mu)\vec{x}}{|\lambda + \mu|} \right) \quad [80]$$

El producto de dos escalares porun vector es asociatvo:

$$\lambda \cdot [\mu \cdot (a, \vec{x})] = \lambda \cdot \left(|\mu| a, \frac{\mu\vec{x}}{|\mu|}\right) = \left(|\lambda||\mu| a, \frac{\lambda\mu\vec{x}}{|\mu||\lambda|}\right)$$
$$(\lambda\mu) \cdot (a, \vec{x}) = \left(|\lambda\mu| a, \frac{\lambda\mu\vec{x}}{|\lambda\mu|}\right) \qquad [81]$$

El producto de un vector por la unidad es el mismo vector:

$$1 \cdot (a, \vec{x}) = (a, \vec{x}) \qquad [82]$$

La estructura matemática anterior, al no utilizar el producto interno de vectores del espacio de Hilbert, hace que se pueda extender sin más a cualquier espacio vectorial normado y, por tanto, a cualquier espacio de Banach.

Se puede dar otra estructura matemática a la superposición de vectores en espacios de Hilbert y de Banach. Sea E un espacio vectorial normado sobre el cuerpo de los números complejos C, B el subespacio de E formado por los vectores de norma la unidad, $\vec{\theta}$ el elemento neutro de E. Por f denotamos una aplicación de E en $B \cup \vec{\theta}$, definida por:

$$\forall\, \vec{x} \in E : f(\vec{x}) = \begin{cases} \vec{\theta}, \ \vec{x} = \vec{\theta} \\ \dfrac{\vec{x}}{||\vec{x}||}, \ \vec{x} \neq \vec{\theta} \end{cases} \qquad [83]$$

de la que se sigue que:

$$f(\vec{x} + \vec{y}) = \frac{\vec{x} + \vec{y}}{||\vec{x} + \vec{y}||} \neq \frac{\vec{x}}{||\vec{x}||} + \frac{\vec{y}}{||\vec{y}||} = f(\vec{x}) + f(\vec{y}) \qquad [84]$$

El espacio producto $E \times f(E)$ es un grupo abeliano para la suma definida por:

$$\vec{x}, f(\vec{x}) + \vec{y}, f(\vec{y}) = \vec{x} + \vec{y}, f(\vec{x} + \vec{y}) \qquad [85]$$

porque es cerrada, asociativa, conmutativa, tiene elemento neutro:

$$\vec{\theta}, f(\vec{\theta}) \qquad [86]$$

y todo elemento tiene inverso.

De [83] se sigue que:

$$f(\lambda\vec{x}) = \begin{cases} \dfrac{\lambda\vec{x}}{||\lambda\vec{x}||} = \dfrac{\lambda}{|\lambda|}\dfrac{\vec{x}}{||\vec{x}||} = \dfrac{\lambda}{|\lambda|} f(\vec{x}), \quad \lambda \neq 0, \ \vec{x} \neq \theta \\ f(\theta) \begin{cases} \lambda = 0, & \forall\, \vec{x} \\ \vec{x} = \theta, & \forall\, \lambda \end{cases} \end{cases} \qquad [87]$$

El espacio producto $E \times f(E)$ es un espacio vectorial sobre el cuerpo de los números complejos C, si se define el producto de un vector $\vec{x}, f(\vec{x})$ por un escalar (número complejo) λ de la siguiente manera:

$$\lambda \cdot [\vec{x}, f(\vec{x})] = \lambda\vec{x}, f(\overrightarrow{\lambda x}) \qquad [88]$$

porque se cumplen las cuatro leyes, las dos distributivas:

$$\begin{aligned} &\lambda \cdot [\vec{x}, f(\vec{x}) + \vec{y}f(\vec{y})] = \lambda \cdot [\vec{x} + \vec{y}, f(\vec{x} + \vec{y})] = \\ &= \lambda \cdot (\vec{x} + \vec{y}), f(\lambda \cdot (\vec{x} + \vec{y})) = \lambda \cdot [\vec{x}, f(\vec{x})] + \lambda \cdot [\vec{y}, f(\vec{y})] = \\ &= \lambda\vec{x}, f(\lambda\vec{x}) + \lambda\vec{y}, f(\lambda\vec{y}) = \lambda\vec{x} + \lambda\vec{y}, f(\lambda\vec{x} + \lambda\vec{y}) \end{aligned} \qquad [89]$$

y:

$$\begin{aligned} &(\lambda + \mu) \cdot [\vec{x}, f(\vec{x})] = (\lambda + \mu)\vec{x}, f((\lambda + \mu)\vec{x}) = \\ &= \lambda \cdot [\vec{x}, f(x)] + \mu \cdot [\vec{x}, f(\vec{x})] = \lambda\vec{x} + \mu\vec{x}, f(\lambda\vec{x} + \mu\vec{x}) \end{aligned} \qquad [90]$$

la asociativa:

$$\begin{aligned} &\lambda \cdot \left[\mu \cdot [\vec{x}, f(\vec{x})]\right] = \lambda \cdot [\mu\vec{x}, f(\mu\vec{x})] = \\ &= \lambda\mu\vec{x}, f(\lambda\mu\vec{x}) = (\lambda\mu) \cdot [\vec{x}, f(\vec{x})] \end{aligned} \qquad [91]$$

y la:

$$1 \cdot [\vec{x}, f(\vec{x})] = \vec{x}, f(\vec{x}) \qquad [92]$$

En los casos particulares del plano y del espacio tridimensional ordinario, se pueden visualizar y materializar geométricamente la superposición de vectores (véanse fuentes).

Las interferencias de probabilidades, surgen al considerar la superposición de dos o más vectores de onda, afectados o no de pesos de superposición, que evolucionan en el tiempo por separado unos de otros, de acuerdo con [3] ó [6] según que el tiempo sea una variable discreta o continua, que puede ser la misma para todos (igual H) o distintas.

A veces, como sucede con la ecuación de Schrödinger, al vector de onda se le puede asociar una distribución de probabilidad. En el caso más general, esto puede hacerse de la siguiente manera: Sea E un espacio vectorial, $\mathscr{C}$ un clan unitario de subconjuntos de un conjunto F ($F \in \mathscr{C}$), $\mathscr{F}$ un conjunto de normas de E que cumplen las siguientes condiciones:

1.º $\mathscr{C}$ y $\mathscr{F}$ son biyectivos.

2.º Para todo elemento $\vec{x}$ de E las normas de $\mathscr{F}$ definen una medida sobre F, $\mathscr{C}$.

Entonces los elementos de E para los que la norma asociada a F es la unidad gozan de la propiedad de que para cada uno de ellos las normas de $\mathscr{F}$ definen una probabilidad sobre F, $\mathscr{C}$.

En este caso la superposición de vectores de E, da origen a interferencias de probabilidades.

3. UNA INTERPRETACION PROBABILISTICA DE LA POLARIZACION DE LA LUZ

Cuando la luz atraviesa un polarizador, un fotón cualesquiera tiene una probabilidad de atravesar el polarizador igual a $\cos^2\theta$, siendo θ el ángulo que forma el vector de polarización del fotón, con la dirección de transmisión del polarizador. Si la luz incidente está formada por n_0 fotones, y está linealmente polarizada, el número n de fotones que atraviesan el polarizador es una v.a. que sigue la distribución binomial de función característica:

$$(\mathrm{sen}^2\theta + e^{iz}\cos^2\theta)^{n_0} \qquad [93]$$

siendo su valor medio $\bar{n}$ y su varianza σ_n^2:

$$\bar{n} = n_0\cos^2\theta \quad ; \quad \sigma_n^2 = n_0\cos^2\theta\,\mathrm{sen}^2\theta \qquad [94]$$

como n_0 es muy grande, la v.a.:

$$\frac{n-\bar{n}}{\sigma_n} \qquad [95]$$

sigue la ley normal (es una v.a. gaussiana) de valor medio nulo y varianza la unidad.

Pero el polarizador no actúa solamente como un filtro, que obstaculiza el paso de los fotones, sino que modifica el estado físico de los fotones que lo atraviesan, cambiando su vector de polarización, que pasa a ser paralelo a la dirección de transmisión del polarizador. Si la luz es monocromática, de frecuencia ν, la energía asociada a los n fotones que atraviesan el polarizador es:

$$E = nh\nu \qquad [96]$$

que es una v.a. de valor medio y varianza:

$$\bar{E} = E_0\cos^2\theta \quad ; \quad \sigma_E^2 = n_0(h\nu)^2\cos^2\theta = E_0h\nu\cos^2\theta\,\mathrm{sen}^2\theta \qquad [97]$$

siendo E_0 la energía de la luz incidente. Por tanto, la varianza de la energía transmitida por el polarizador es proporcional a la frecuencia de la luz ν; h es la constante de Planck.

Para cualquier magnitud física proporcional a la energía, como es la intensidad $I = kE$, se cumple que:

$$I = I_0\cos^2\theta \quad ; \quad \sigma_I^2 = \sigma_E^2k^2 = I_0kh\nu\cos^2\theta\,\mathrm{sen}^2\theta \qquad [98]$$

Y la medida de las fluctuaciones en la luz transmitida da un medio de calcular la constante de proporcionalidad con el número de fotones.

Esta interpretación probabilística redescubre la ley de Malus, primera [98], pero abre un campo nuevo, que es una teoría de fluctuaciones.

Obsérvese que la probabilidad depaso de los fotones es similar a la distribución de probabilidad asociada a un punto sobre una circunferencia, como describimos en el parágrafo 1, y que la acción del polarizador es similar a la contracción del paquete de ondas, como describimos también en el mismo parágrafo.

Si la luz no está linealmente polarizada, hay que cambiar las fórmulas anteriores. Si el vector de polarización está repartido uniformemente al azar, la función característica de la v.a. de que un fotón cualquiera atraviese o no el polarizador es:

$$\frac{1}{2\pi}\int_0^{2\pi}(\operatorname{sen}^2\theta + e^{iz}\cos^2\theta)\,d\theta = \frac{1}{2} + \frac{e^{iz}}{2} \qquad [99]$$

y como son v.a. independientes para todos los fotones, el valor medio y la varianza para la suma de los n_0 fotones valen, respectivamente, $n_0/2$ y $n_0/4$. Por tanto, en las fórmulas [94], [97] y [98] hay que efectuar los cambios:

$$\cos^2\theta \longrightarrow \frac{1}{2} \quad ; \quad \operatorname{sen}^2\theta\cos^2\theta \longrightarrow \frac{1}{4} \qquad [100]$$

Si no hay direcciones privilegiadas para el vector de polarización, pero éste no es aleatorio, el valor medio y la varianza de la v.a. de atravesar o no el polarizador, cualquier fotón, cuyo vector de polarización forme un ángulo θ con la dirección de transmisión del polarizador es:

$$\cos^2\theta \quad ; \quad \operatorname{sen}^2\theta\cos^2\theta \qquad [101]$$

y como son v.a. independientes, el valor medio y la varianza para los n_0 fotones son:

$$\frac{n_0}{2\pi}\int_0^{2\pi}\cos^2\theta\,d\theta = \frac{n_0}{2} \quad ; \quad \frac{n_0}{2\pi}\int_0^{2\pi}\operatorname{sen}^2\theta\cos^2\theta\,d\theta = \frac{n_0}{8} \qquad [102]$$

Por tanto, en las fórmulas [94], [97] y [98] hay que efectuar los cambios:

$$\cos^2\theta \longrightarrow \frac{1}{2} \quad ; \quad \operatorname{sen}^2\theta\cos^2\theta \longrightarrow \frac{1}{8} \qquad [103]$$

En ambos casos el valor medio es el mismo, pero el de la varianza no.

Si la luz no es monocromática, y $n_0(\nu)$ es la función de densidad del número de fotones incidentes, para el número de fotones emergentes las anteriores fórmulas subsisten, porque:

$$n_0 = \int_0^{\infty} n_0(\nu)\,d\nu \qquad [104]$$

Pero para la energía, por las fórmulas [96] y [97] y por ser v.a.

independientes, el hecho de que elfotón atraviese o no el polarizador, hay que efectuar en ellas los cambios:

$$\begin{aligned} hn_0\nu &\longrightarrow \int_0^\infty h\nu n_0(\nu)\, d\nu = \int_0^\infty E_0(\nu)\, d\nu = E_0 \\ n_0 h^2 \nu^2 &\longrightarrow \int_0^\infty n_0(\nu) h^2 \nu^2\, d\nu = h \int_0^\infty \nu E_0(\nu)\, d\nu \end{aligned} \qquad [105]$$

La fórmula del valor medio subsiste, pero cambia la de la varianza (*).

4. CADENAS DE MARKOV ASOCIADAS A LA MECANICA CUANTICA. DOS TEOREMAS ERGODICOS

Vamos a adoptar para los operadores hermíticos la notación:

$$p_i(A) = \text{Prob}\,(A = \lambda_i) \quad ; \quad p_j(B) = \text{Prob}\,(B = \eta_j) \qquad [106]$$

con lo que [21] se puede escribir:

$$\vec{p}(A) \| |\langle \vec{u}_i, \vec{v}_j \rangle|^2 \| = \vec{p}(A)(A, B) = \vec{p}(B) \qquad [107]$$

donde la matriz (A, B) es biestocástica, por ser los $\vec{u}_i$, $\vec{v}_j$ dos bases ortonormales de C^m. Se cumplen las:

$$(A, A) = \mathrm{I} \quad ; \quad (A, B) = (B, A)^t \quad ; \quad (A, C) \neq (A, B)(B, C) \qquad [108]$$

en la que el supraíndice t significa la traspuesta. La última [108] significa que la matriz de las probabilidades de transición de A a C, es distinta si se pasa directamente de A a C, o por intermedio de B. La segunda [108] significa que la probabilidad de que al aplicar el operador B se obtenga su valor propio η_j (véase el parágrafo 1), después de que la aplicación del operador A haya dado su valor propio λ_i, es igual a la probabilidad de que al aplicar el operador A se obtenga su valor propio λ_i, después de que la aplicación del operador B haya dado su valor propio η_j, porque estas probabilidades valen, respectivamente:

$$|\langle \vec{u}_i, \vec{v}_j \rangle|^2 \quad ; \quad |\langle \vec{v}_j, \vec{u}_i \rangle|^2 \qquad [109]$$

Como dijimos en el parágrafo 1, las aplicaciones de estos operadores son sucesivas, pero instantáneas, es decir, sin dejar evolucionar elvector de onda entre una y otra aplicación.

Por ser biestocásticas las matrices (A, B) se sigue que si el vector $\vec{p}(A)$ tiene todas sus componentes iguales, también las tiene el $\vec{p}(B)$ y cualquier otro que le siga. Si después de aplicar el operador A, aplicamos sucesiva e instantánea-

(*) Véase mi memoria publicada en esta Revista en 1984: «Distribuciones binomiales y de Bernoulli de probabilidad aleatoria. Teoremas del límite central».

mente, de modo alternativo n veces el B y $n - 1$ veces el A, se obtiene, finalmente:

$$\vec{p}(B) = [(A, B)(B, A)]^{n-1}(A, B) \qquad [110]$$

y como la matriz $(A, B) \cdot (B, A)$ es también biestocástica, cuando n tiende a infinito, $\vec{p}(B)$ de [110] tiende a tener todas sus componentes iguales, cualquiera que sea el valor inicial de $\vec{p}(a)$, y por lo dicho anteriormente, cualquier otro vector de probabilidad de cualquier otro operador después de la operación anterior, tiende a tener todas sus componentes iguales, lo que constituye un teorema ergódico.

Si $\vec{x}'$ y $\vec{x}''$ son dos vectores de onda que evolucionan de acuerdo con [13] se conserva en el tiempo su producto escalar, porque:

$$\langle \vec{x}'_n, \vec{x}''_n \rangle = \langle M_n \vec{x}'_{n-1}, M_n \vec{x}''_{n-1} \rangle = \langle \vec{x}'_{n-1}, \vec{x}''_{n-1} \rangle = \cdots = \langle \vec{x}'_0, \vec{x}''_0 \rangle \qquad [111]$$

al ser las matrices M unitarias. En el caso en que el tiempo es una variable continua, si evolucionan de acuerdo con [6], también se conserva su producto escalar, porque:

$$\frac{d}{dt} \langle \vec{x}'(t), \vec{x}''(t) \rangle = \left\langle \frac{d\vec{x}'}{dt}, \vec{x}'' \right\rangle + \left\langle \vec{x}', \frac{d\vec{x}''}{dt} \right\rangle =$$
$$= \langle -iH\vec{x}', \vec{x}'' \rangle + \langle \vec{x}', -iH\vec{x}'' \rangle = \langle \vec{x}', (iH - iH)\vec{x}'' \rangle = 0 \qquad [112]$$

por ser H hermítico. Por tanto, si $\vec{x}'$ y $\vec{x}''$ para $t = 0$, son ortogonales, lo son a lo largo del tiempo, y se sigue también que si inicialmente un conjunto de vectores de onda formauna base ortogonal del espacio de Hilbert, lo siguen formando en todo otro instante.

La anterior propiedad nos permite desarrollar la teoría de las aplicaciones sucesivas de operadores no instantáneamente. Si en los instantes $t_1, t_2, \ldots, t_n$ aplicamos los operadores $A_1, \ldots, A_n$, se obtiene:

$$|||\langle \vec{x}(t_1), \vec{u}_{1i}(t_1) \rangle|^2|| = \vec{p}(A_1) \quad ; \quad \vec{p}(A_1)|||\langle \vec{x}_{1i}(t_2), \vec{u}_{2j}(t_2) \rangle|^2|| = \vec{p}(A_2) \qquad [113]$$

donde $\vec{x}_{1i}(t)$ es la solución de [6], tal que $\vec{x}_{1i}(t_1) = \vec{u}_{1i}(t_1)$. Y así hasta:

$$\vec{p}(A_{n-1})|||\langle \vec{x}_{n-1i}(t_n), \vec{u}_{nj}(t_n) \rangle|^2|| = \vec{p}(A_n) \qquad [114]$$

donde $\vec{x}_{n-1i}(t)$ es la solución de [6] tal que $\vec{x}_{n-1i}(t_{n-1}) = \vec{u}_{n-1i}(t_{n-1})$. De las anteriores se sigue que:

$$\vec{p}(A_1) \prod_{k=1}^{n-1} |||\langle x_{ki}(t_{k+1}), u_{k+1j}(t_{k+1}) \rangle|^2|| = \vec{p}(A_n) \qquad [115]$$

Como todas las matrices anteriores son biestocásticas, por lo que demostramos anteriormente de que todos los vectores $\vec{x}_{ki}$, $\vec{u}_{ki}$ (para un mismo k e i variable) forman bases ortonormales del espacio de Hilbert, se sigue que si a partir de un instante cualquiera t_k es $\vec{p}(A_k)$ un vector de probabilidad con todas sus componentes iguales, todos los restantes vectores de probabilidad de subíndice superior a k, tienen todas sus componentes iguales, cualesquiera que hayan sido los vectores de probabilidad de índice inferior a k. Ello constituye un segundo teorema ergódico.

Los resultados anteriores subsisten para el caso en que el tiempo es una variable discreta.

Supongamos que inicialmente, para $t = 0$, el vector de onda en vez de ser cierto es aleatorio, siendo $\vec{p}_0$ el vector de probabilidad de que $\vec{x}(0)$ sea, respectivamente, $\vec{x}_{01}(0)$, ..., $\vec{x}_{0m}(0)$ (se trata de C^m, si es el espacio de Hilbert clásico es lo mismo, con m infinito numerable), entonces es:

$$\vec{p}_0 \| |\langle \vec{x}_{0i}(t_1), \vec{u}_j(t_1) \rangle|^2 \| = \vec{p}(A_1) \qquad [116]$$

siendo $\vec{x}_{0i}(t)$ la solución de [6], que para $t = 0$, vale el dado $\vec{x}_{0i}(0)$.

Apliquemos el resultado anterior al caso en que los $\vec{x}_{0i}(0)$ son iguales a los respectivos vectores propios de H para $t = 0$, que denotamos por $\vec{h}_i(0)$.

Sea entonces una población de vectores de onda, formada por n_1 que para $t = 0$ la aplicación de H dé su valor propio γ_1, n_2 que dé el valor propio γ_2, y así sucesivamente hasta n_m que dé el valor propio γ_m. Si tomamos uno de dichos vectores de onda al azar, hay que hacer igual en [116] $\vec{p}_0$ a:

$$\vec{p}_0 = \left\| \frac{n_1}{N}, \frac{n_2}{N}, \frac{n_m}{N} \right\| \quad ; \quad N = \sum_{i=1}^{m} n_i \qquad [117]$$

y [116] junto a [115] permiten calcular cualquier $\vec{p}(A)$.

Otro problema que podemos resolver es el de hallar la función generatriz de la variable aleatoria que en el instante t_n da la repartición de esta población de vectores de onda entre los vectores propios de A_n, o lo que es lo mismo, de la repartición de los valores de A_n entre sus valores propios. Se tiene, siendo las z reales, que:

$$z_1^{n_1} \cdots z_m^{n_m} \qquad [118]$$

es la función generatriz de la repartición inicial (en este casocierta) de los vectores de onda, entre los vectores propios de H, y:

$$\langle \vec{p}_1(A_n), \vec{z} \rangle^{n_1} \cdots \langle \vec{p}_m(A_n), \vec{z} \rangle^{n_m} \qquad [119]$$

la función generatriz de la repartición de los vectores de onda entre los vectores propios de A_n en el instante t_n. En [119] las $\vec{p}_i(a_n)$ son las soluciones de [116] y [115] con los valores iniciales de $\vec{p}_0$ iguales respectivamente a los vectores $\vec{e}_i$ [35].

Si la población inicial de vectores de onda no es cierta, sino aleatoria, y su función generatriz en vez de [118] es:

$$g(z_1, ..., z_m) \qquad [120]$$

la solución del problema para $t = t_n$, se obtiene efectuando en [120] la sustitución:

$$z_1 \rightarrow \langle \vec{p}_1(A_n), \vec{z} \rangle, \ldots, z_m \rightarrow \langle \vec{p}_m(A_n), \vec{z} \rangle \qquad [121]$$

El teorema ergódico que hemos demostrado para los vectores de probabilidad, se traduce para las funciones generatrices en que en ellas hay que sustituir cada z por la media aritmética de todas ellas, con lo que se transforman en funciones simétricas. En [118] y [119] hay que efectuar los cambios:

$$z_1 \rightarrow \frac{z_1 + \cdots + z_m}{m}; \ldots ; z_m \rightarrow \frac{z_1 + \cdots + z_m}{m} \qquad [122]$$

Véase mi memoria publicada en la *Revista de la Real Academia de Ciencias* en 1981 titulada: «Cadenas de Markov en poblaciones aleatorias y probabilidades en cadena generalizadas».

Los procesos estocásticos aquí definidos son irreductibles a los del movimiento browniano, porque en esto últimos lo que evoluciona en el tiempo es la función de frecuencia de una variable aleatoria, mientras que en los aquí tratados lo que evoluciona en el tiempo es la raíz cuadrada de una función de frecuencia.