

Tesis de Posgrado

Sobre los fundamentos de la estadística cuántica

Varsavsky, Oscar Alberto

1949

Tesis presentada para obtener el grado de Doctor en Química de la Universidad de Buenos Aires

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Varsavsky, Oscar Alberto. (1949). Sobre los fundamentos de la estadística cuántica. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.
http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_0588_Varsavsky.pdf

Cita tipo Chicago:

Varsavsky, Oscar Alberto. "Sobre los fundamentos de la estadística cuántica". Tesis de Doctor. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 1949.
http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_0588_Varsavsky.pdf

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires

SOBRE LOS FUNDAMENTOS DE LA ESTADÍSTICA CUÁNTICA

TESIS

OSCAR ALBERTO VARGASKY/

Fests: 588

FOFNBA

INTRODUCCION Y RESUMEN.

El objetivo general de este trabajo es reformular la fundamentación estadística de la Mecánica Cuántica tomando como elementos las "propiedades físicas" en lugar de magnitudes en general. Procuramos demostrar cuánto más sencillos de demostrar y directamente interpretables aparecen muchos resultados de la mayor importancia, sin que se pierda generalidad.

Por supuesto tomamos como punto de referencia el formidable libro de J. von Neumann: "Mathematische Grundlagen der Quantenmechanik", tan lleno de ideas aún no suficientemente explotadas y que es el cimiento obligado de cuanto edificio teórico se ha pretendido levantar en la Mecánica Cuántica.

Además de la reformulación antes mencionada, hemos conseguido varios resultados particulares nuevos; entre otros la demostración de que el método de Koopman (7) de tratar sistemas dinámicos puede aplicarse a la estadística cuántica, y por lo tanto valen también para ella ciertas demostraciones del teorema ergódico. Hacemos también una verificación del mismo mediante la comparación de una sucesión real de mediciones con una "cadena de Markov" y damos la formulación cuántica del concepto de "transitividad métrica". Hemos también simplificado la demostración de algunos teoremas y en especial generalizado para espectros continuos el referente a la ley de distribución de Maxwell.

Exponemos también, aunque sin detalles, dos ideas que nos parecen de gran importancia; la relación topológica entre espacios "complementarios" (coordenadas e impulsos por ejemplo), que aunque implícita en numerosos trabajos no hemos visto nunca enunciada claramente, y la distinción entre el tiempo del sistema y el del observador, que permite recobrar la simetría relativista.

Nueva es también la comparación entre la lógica propia de las proyecciones, la lógica usual y la de tres valores, y además el concepto de "conjunción estadística" que exponemos en un apéndice.

Teníamos también intención de utilizar el concepto estadístico de función característica de una distribución, pero encontramos que ese tema había ya sido tratado, justamente en otro trabajo de tesis, por Arnous (1), en 1947. De todos modos hemos agregado un pequeño apéndice comentando su contenido.

En cambio hemos renunciado a la inclusión de un resumen de la parte matemática, pues su volumen estaría fuera de proporción con el resto del trabajo y porque existen obras sumamente claras y apropiadas para referencia, como el capítulo correspondiente del libro de v. Neumann y otras obras que se citan en una bibliografía especial.

Aunque por comodidad nuestras fórmulas se refieren concretamente al espacio de Hilbert de las funciones de la clase L^2 (cuadrado del módulo sumable), valen siempre, con obvias modificaciones, para otros espacios de uso común.

Los teoremas que daremos por conocidos pueden encontrarse en su mayor parte en el Capítulo II del libro de v. Neumann, pero en general trataremos de dar referencias explícitas.

NOTACION.

Los "puntos" de un espacio de Hilbert se indicarán con minúsculas griegas y a veces latinas.

(φ, ψ) significa producto escalar de φ y ψ .

$\|\varphi\|$ es la norma de φ : $\|\varphi\|^2 = (\varphi, \varphi)$

Los operadores serán representados por mayúsculas de imprenta y los observables correspondientes por la misma letra en cursiva.

Los operadores proyección se indicarán con las letras E, F, G, ... pero en caso de corresponder a una sola dimensión se escribirá P_φ indicando φ el rayo sobre el cual se proyecta.

La letra \mathcal{M} indicará una multiplicidad lineal cerrada. Con un subíndice, \mathcal{M}_E , se expresa a qué proyección corresponde. El espacio total, o sea la multiplicidad del operador I (operador unidad), se designará con la letra J.

$E_\lambda, F_\lambda, \dots$, λ parámetro real, denotarán familias espectrales, o descomposiciones de la unidad. Recuérdese que a todo operador autoadjunto A corresponde una familia espectral E_λ tal que

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda \quad (\text{representación canónica; ver (8),$$

Capítulo VIII).

A v B significa que A y B conmutan.

SpA significa Spur o traza del operador A (ver (9), Cap. II, párrafo 11).

Las integrales triples con respecto a coordenadas se abreviarán así:

$$\iiint f(q_1, q_2, q_3, \dots) dq_1 dq_2 dq_3 \equiv \int_{\mathcal{D}} f(\vec{q}, \dots) d\vec{q}$$

Los números entre paréntesis se refieren a la bibliografía.

I. POSTULADOS USUALES DE LA MECÁNICA CUÁNTICA.

Como es costumbre, partiremos de los conceptos "sistema físico", "observable o magnitud física medible", "medición", "estado de un sistema", suponiendo que todos los físicos los entienden y aplican del mismo modo. El estudio detallado de estas cuestiones sólo lo ha emprendido desde un punto de vista científico la escuela filosófica del empirismo lógico (ver, por ejemplo, Frank, (5)), pero aquí nos desentenderemos de ellas.

Con esta salvedad, los postulados usuales de la Mecánica Cuántica pueden enunciarse así:

POSTULADO I. Existe una correspondencia biunívoca entre todos los estados imaginables de un sistema físico y los rayos de un espacio de Hilbert.

Comentarios. Nótese que se habla de rayos y no de puntos, es decir que ϱ y $a\varrho$ (ϱ punto del espacio de Hilbert y a número complejo) tienen el mismo significado físico. Para evitar confusiones se elige sobre cada rayo el punto a distancia unidad del origen, es decir, se normaliza.

No es necesario en rigor especificar cuál es el espacio de Hilbert en cuestión, pues son todos isomorfos, pero en la práctica se utiliza el de las funciones cuyos argumentos son coordenadas espaciales en número igual al de grados de libertad del sistema y que contienen al tiempo como parámetro (más una condición de integrabilidad). Para la teoría del electrón de Dirac se utilizan conjuntos de cuatro funciones como las antedichas.

Como pocas veces es posible conocer el estado de un sistema exactamente, el principal valor práctico de este postulado es que permite prever todas las posibilidades.

Por último, diversos autores han señalado la insuficiencia del espacio de Hilbert como representante de todos los estados posibles. El ejemplo más conocido es el de la utilísima función delta de Dirac, que no tiene cabida en ningún espacio de Hilbert. En general se ha tratado

de resolver la cuestión independizándose de la noción de estado. Véase en especial v. Neumann (13) y Segal (15).

POSTULADO II. Existe una correspondencia biunívoca entre los observables (magnitudes físicas) y los operadores autoadjuntos del espacio de Hilbert considerado.

Comentarios. En este postulado se hallan implícitas todas las características "nuevas" de la Mecánica Cuántica, pero por supuesto es poco menos que vacío si no se dan las leyes de correspondencia. Las más generales son:

- a) Sea $F(x)$ una función cualquiera, real para x real; si al observable A corresponde el operador A , al observable $F(A)$ corresponde el operador $F(A)$. (Se podría aceptar que $F(x)$ tomara también valores complejos, y entonces le correspondería un operador normal, en vez de autoadjunto, pero no se obtiene mayor generalidad.)
- b) Si los operadores A y B corresponden a los observables A y B el operador autoadjunto $A+B$ corresponde al observable $A+B$ y el operador AB , sólo en caso de ser autoadjunto, representa la magnitud $A \times B$, que de otro modo no tiene sentido físico.
- c) Entre los observables coordenada (q) e impulso canónicamente conjugado (p) existe una total simetría, a saber:
Así como las coordenadas, los impulsos pueden utilizarse como argumentos de las funciones que forman el espacio de Hilbert del sistema; pero si los valores que toma q forman un grupo topológico, los de p pertenecen al correspondiente grupo de caracteres, y viceversa.
Entre ambos espacios de Hilbert así originados hay por supuesto un isomorfismo engendrado por la transformación de Fourier adecuada, de modo que operadores correspondientes son conjugados según dicha transformación. En el espacio de las coordenadas, al observable Q (valor de la coordenada q) corresponde el operador multiplicación por q . Simétricamente, en el espacio de los impulsos, al observable

P (valor del impulso p) corresponde el operador multiplicación por p . Por eso es que en el espacio de las coordenadas al observable P corresponde, a menos de un factor constante, la derivación con respecto a q (y viceversa).

Este modo de expresar la relación entre los operadores q y p nos parece sumamente sugestivo, además de matemáticamente simple, y no tenemos conocimiento de que haya sido enunciado explícitamente. En el siguiente cuadro resumimos sus principales consecuencias:

	Espacio de las coordenadas,	Espacio de los impulsos.
Campo de variabilidad.	<ul style="list-style-type: none"> a) Eje real b) Segmento (a,b), o sea N°s. reales módulo $(a-b)$. c) Múltiplos enteros de $h/(a-b)$. 	<ul style="list-style-type: none"> Eje real Múltiplos enteros de $h/(a-b)$. Segmento (a,b), o sea N°s. reales módulo $a-b$.
Punto del esp. Hilbert.	$\varphi(q)$ (de longitud finita)	$\psi(p)$ (de longitud finita)
Transformación de Fourier.	<ul style="list-style-type: none"> a) $F\psi = \frac{1}{\sqrt{h}} \int_{-\infty}^{\infty} \psi(p) e^{\frac{2\pi i}{h} p q} dp = \varphi$ b) $F\psi = \frac{1}{\sqrt{h}} \sum_{-\infty}^{\infty} \psi(p_k) e^{\frac{2\pi i}{h} p_k q} = \varphi$ c) $F\psi = \frac{1}{\sqrt{h}} \int_a^b \psi(p) e^{\frac{2\pi i}{h} p q} dp = \varphi$ 	$\psi = F^{-1} \varphi$
Observable Q	$q \varphi(q)$	$-\frac{h}{2\pi i} \frac{\partial \psi(p)}{\partial p}$ (salvo caso b)
Observable P	$\frac{h}{2\pi i} \frac{\partial \varphi(q)}{\partial q}$ (salvo en caso c)	$p \psi(p)$
Condiciones cuánticas.	$\frac{\partial M(q)}{\partial q} = -[p, M(q)]$	$\frac{\partial N(p)}{\partial p} = [q, N(p)]$

Como ejemplo de la utilidad de este punto de vista tomemos una fórmula de la teoría de las colisiones. No es fácil interpretar una expresión como $\int_{-\infty}^{\infty} \delta_+(q-x) \varphi(x) dx$ (siendo $\delta_+(x)$ la delta positiva de Heisenberg: $\delta_+(x) = \delta(x) - \frac{1}{i\pi x}$). Sin embargo si pasamos al espacio p encontramos que el operador correspondiente es la anulación de los valores negativos de p (véase (17)); es decir, dicha fórmula representa partículas que se alejan del origen de coordenadas (ver Dirac, (3), pág. 198).

Esta dualidad matemática es la clave de la dualidad física de todo sistema. El espacio q es apropiado para partículas y el p para ondas, y el principio de incertidumbre es inmediato.

- d) Cuando se trata de hallar el operador correspondiente a la energía, en los casos en que exista un "análogo clásico", se construye el Hamiltoniano clásico y en él se sustituyen coordenadas e impulsos por sus operadores respectivos. En caso que el orden de los factores sea dudoso se aplican los criterios sugeridos por la teoría de los grupos (ver, por ejemplo, Weyl, (18)).

POSTULADO III. Si un sistema se halla en el estado φ (normalizado: $\|\varphi\| = 1$), y en él se mide la magnitud A , el valor medio del resultado es:

$$m(A; \varphi) = (A\varphi, \varphi)$$

Comentarios. El valor medio no se refiere a mediciones sucesivas sobre el mismo sistema, sino a mediciones sobre un gran número de sistemas en el mismo estado (pero véase el parágrafo sobre teorema ergódico).

Nótese que por ser A autoadjunto, los valores medios son siempre números reales.

Está implícito en el enunciado que φ debe pertenecer al dominio de A , pero estas cuestiones no se presentan en la práctica.

POSTULADO IV. Dos mediciones consecutivas (muy próximas entre sí) del mismo observable sobre el mismo sistema dan el mismo resultado.

Comentarios. Esto debe interpretarse como que, cuanto más distanciadadas en el tiempo se hallen las mediciones, tanto más probable es que den igual resultado.

Este postulado es el que permite calcular el estado de un sistema a partir del resultado de una medición, pero sólo en casos especiales, como se verá más adelante.

El Operador Tiempo.

Consideraciones relativistas elementales sugieren la ampliación de las relaciones entre coordenadas e impulsos al par tiempo-energía. Debe imaginarse entonces un espacio de los tiempos y un espacio conjugado en que la variable independiente es la energía. En este espacio el operador energía actúa simplemente multiplicando a la función de onda por su argumento, y al pasar a otro espacio queda representado por la derivación con respecto al tiempo dando automáticamente la ecuación de Schrödinger. (Debe hacerse notar que las diferencias entre espacio y tiempo exigidas por la relatividad especial se reflejan en que se pasa de t a H mediante la transformación inversa a la que permite pasar de q a p .)

Todo esto es muy satisfactorio, incluso desde el punto de vista estético, pero deja de serlo en cuanto se pasa a considerar el operador conjugado de la energía H , o sea el tiempo. Según los mismos razonamientos de paralelismo con q , en el espacio t debe existir un operador "multiplicación por t " que represente el valor de la coordenada temporal del sistema, y también ~~que~~ en el producto escalar figure una integración con respecto al tiempo. Así entonces $|\psi(\vec{q}, t)|^2$ sería la densidad de probabilidad en cuatro dimensiones. Esta densidad no podría ser absoluta, pues como $\int |\psi(\vec{q}, t)|^2 d\vec{q} = 1$ para todo t , una nueva integración con respecto a t diverge y no es posible normalizar.

Por otra parte, a primera vista no tiene mucho sentido la cuestión: "¿Cuál es el tiempo del sistema?". Cuando se pregunta cuáles son sus coordenadas espaciales se sobreentiende "al tiempo t ", pero no parece muy claro preguntar cuál es el tiempo al tiempo t .

Sin embargo creemos que existe una posibilidad de aclaración total de esta cuestión, si se observa que la frase "al tiempo t " significa en realidad "en el momento de la medición".

Nos parece que deben considerarse dos clases de tiempo.

- 1). El tiempo del sistema, representado por un operador en estricto paralelo con las coordenadas espaciales.
- 2). El tiempo del observador, que es un parámetro macroscópico, que sólo

~~se refiere a un sistema de referencia~~

sirve para ordenar las mediciones e indicar los lapsos transcurridos entre ellas, y que no interfiere con la medición de la energía del sistema.

Como ilustración de esto veamos dos ejemplos.

Si se quiere medir la posición de un sistema, puede imaginarse un reticulado en el espacio y se observa que en cierto momento coinciden el sistema y cierto retículo (sistema y retículo están al mismo tiempo en el mismo sitio). En este caso no hay diferencia entre el tiempo del observador (retículo) y el del sistema, así como tampoco la hay entre sus coordenadas espaciales.

Pero supongamos que se mide la energía de un sistema. El observador verifica que en cierto instante, que puede determinar con toda precisión, la aguja de un instrumento se detiene en cierta división de la escala. A partir de este dato, empleando fórmulas basadas en leyes físicas, deduce un cierto valor para la energía del sistema. Pero ahora el tiempo de observación no es el del sistema. Aquél es leído directamente y con toda exactitud; éste sólo puede calcularse con la ayuda de nuevas leyes físicas, y de tal modo que se satisfaga el principio de Heisenberg. Entonces es compatible un conocimiento muy exacto de la energía del sistema y del instante de observación, pero en tal caso poco podrá decirse sobre el momento en que el sistema poseía esa energía. Es lo mismo que ocurre con el par p - q , como lo hizo notar el primero v. Neumann (11) al señalar que las observaciones macroscópicas de posición e impulso son computables, pero de ellas sólo pueden deducirse los valores para el sistema microscópico con las limitaciones del principio de incertidumbre. Es decir, se puede conocer exactamente la posición del observador que mide el impulso del sistema, pero de ella no se puede deducir exactamente la posición del sistema.

De modo entonces que una función de onda debería llevar indicación de dos tiempos; uno como argumento (el del sistema) y otro como parámetro.

Así $|\varphi_{\theta}(\vec{q}, t)|^2$ indica con qué probabilidad una propiedad verificada al tiempo θ era poseída por el sistema al tiempo t y en la posición q . Con esta interpretación, la integral triple $\int |\varphi_{\theta}(\vec{q}, t)|^2 d\vec{q}$ no es constante ni igual a uno, y una nueva integración con respecto a t da la probabilidad de que la propiedad fuera poseída por el sistema en algún momento y algún lugar y por hipótesis deba ser igual a uno.

Más sencilla resulta la interpretación pasando al espacio complementario. Sea $\Psi_{\theta}(\vec{p}, H)$ la función correspondiente a $\varphi_{\theta}(\vec{q}, t)$ (nótese que la correspondencia exige valores iguales del parámetro). Ahora $|\Psi_{\theta}(\vec{p}, H)|^2$ indica la probabilidad de que, medidos al tiempo θ , el impulso tenga un valor en el entorno de p y la energía un valor en el entorno de H .

La suposición implícita en las formulaciones usuales es que la función de onda es de la forma: $|\varphi_{\theta}(\vec{q}, t)|^2 = \delta(\theta - t) |\varphi(\vec{q})|^2$ lo que ocurre en realidad sólo si el sistema se observa directamente. Es natural entonces que se obtenga el mismo resultado promediando con respecto al tiempo del sistema (o sea integrando con respecto a t) o ignorándolo y reemplazándolo directamente por el parámetro θ . Y nótese que esto vale incluso cuando actúa sobre φ_{θ} un operador que contenga explícitamente al tiempo. En efecto, sea por ejemplo $H(\vec{q}, \vec{p}, t)$ el hamiltoniano de un sistema en un campo de potencial variable. El valor medio de la energía en el estado $\varphi_{\theta}(\vec{q}, t)$ es:

$$\begin{aligned} \text{vm}(H; \varphi_{\theta}) &= \iint H(\vec{q}, \vec{p}, t) \varphi_{\theta}(\vec{q}, t) \cdot \overline{\varphi_{\theta}(\vec{q}, t)} d\vec{q} dt = \\ &= \iint \delta(\theta - t) \cdot H(\vec{q}, \vec{p}, t) \varphi(\vec{q}) \cdot \overline{\varphi(\vec{q})} d\vec{q} dt = \int H(\vec{q}, \vec{p}, \theta) \varphi(\vec{q}) \cdot \overline{\varphi(\vec{q})} d\vec{q} \end{aligned}$$

que es la fórmula usual.

Especialmente interesante es la posibilidad de ampliar esta idea al caso de más de un sistema, puesto que ahora no se presenta ninguna dificultad conceptual. Una función de onda para dos sistemas sería de la forma $\varphi_{\theta}(\vec{q}_1, \vec{q}_2, t_1, t_2)$, donde se ve que las quejas contra la

existencia de un solo tiempo de los que quieren relativizar la Cuántica no tienen base sólida (ver Dirac (3), pág. 252). El tiempo común a que ellos se refieren es nuestro Θ , que expresa que hay un solo parámetro de observación. En cambio cada sistema tiene su coordenada temporal y es que con respecto a estas que deben exigirse propiedades de covariancia. Ahora $\|\varphi_{\Theta}(\vec{q}, t)\| = 1$ (integración cuádruple) es un escalar con todo derecho, así como lo son la masa o el tiempo propio, y no hay que buscar un cuadrivector del que sea la componente temporal (ver Dirac (4)). Además la conservación de la probabilidad total deja de tener el sentido usual. Este puede adaptarse a una coordenada espacial en vez del tiempo: definamos el producto como una integral triple con respecto a q_2, q_3 y t ; es evidente que para la norma correspondiente hay conservación con respecto a q_1 :

$$\frac{d}{dq_1} \iiint |\varphi_{\Theta}(\vec{q}, t)|^2 dq_2 dq_3 dt = 0$$

La nueva "ley de conservación" exige simplemente que $\|\varphi_{\Theta}(\vec{q}, t)\| = 1$ para todo Θ . En aproximación no relativista esto sólo exige aceptar la razonable hipótesis $\varphi_{\Theta+\tau}(\vec{q}, t) = \varphi_{\Theta}(\vec{q}, t-\tau)$. En el caso relativista sirve como una condición adicional para las ecuaciones diferenciales. Por lo tanto, si las velocidades no son muy grandes la ecuación de Schrödinger vale también derivando con respecto al parámetro Θ .

No es objeto de este trabajo desarrollar estas nuevas ideas. Las exponemos sólo porque no es posible dejar de hacerlo al enunciar, por brevemente que sea, los fundamentos de la Mecánica Cuántica, y no se nos oculta que hasta no ser aplicadas a casos concretos no puede estimarse su valor.

Condiciones Cuánticas.

Sea $M(q)$ un operador función del operador q . Su acción sobre la función de onda $\varphi(q)$ será multiplicativa, por el postulado II, a). Se puede entonces derivar como un producto de funciones:

$$\frac{\partial}{\partial q} [M(q) \cdot \varphi(q)] = \frac{\partial M}{\partial q} \varphi + M \frac{\partial \varphi}{\partial q}$$

$\frac{\partial M}{\partial q}$ es una función de la variable q que multiplica a $\varphi(q)$ y por lo tanto puede considerarse como un operador función del operador q . En cuanto a la derivación con respecto a q es la aplicación del operador p salvo un factor numérico:

$$p M \varphi = \frac{h}{2\pi i} \frac{\partial M}{\partial q} \varphi + M p \varphi$$

como esto vale para toda φ , se deduce la igualdad entre operadores:

$$\frac{2\pi i}{h} (p M - M p) \equiv -[p, M] = \frac{\partial M}{\partial q}$$

En completa analogía, si $N(t)$ es función del operador t , se tendrá:

$$[H, N] = \frac{\partial N}{\partial t}$$

pero debe recordarse que esta fórmula sólo vale para producto escalar definido con una integración con respecto al tiempo.

Si en cambio se considera a t como parámetro y en el producto escalar se integra sólo con respecto a las coordenadas espaciales, puede deducirse, como es sabido:

$$\frac{d}{dt} (N \varphi, \psi) = \left(\frac{\partial N}{\partial t} \varphi, \psi \right) + ([N, H] \varphi, \psi)$$

relación funcional que permite definir (siempre en tres dimensiones) el operador

$$\frac{dN}{dt} = [N, H] + \frac{\partial N}{\partial t} \quad \left(\frac{\partial N}{\partial t} = \lim_{\Delta t \rightarrow 0} \frac{N(t+\Delta t) - N(t)}{\Delta t} \right)$$

pero ahora $\frac{\partial N}{\partial t} \neq [H, N]$, salvo en casos particulares, pues t no es aquí un operador.

Por supuesto puede obtenerse un resultado análogo eliminando la integración con respecto a cualquier otro argumento en la definición de producto escalar.

Aplicaciones de la Representación Canónica.

Sea A una magnitud cuyo operador correspondiente A tiene la familia espectral E_λ . Entonces:

- a). La probabilidad de obtener un resultado comprendido entre λ y $\lambda + \Delta\lambda$ al medir A es $\|E_{\Delta\lambda} \varphi\|^2$ si el sistema se hallaba en el estado φ (normalizado). ($E_{\Delta\lambda} \equiv E_{\lambda+\Delta\lambda} - E_\lambda$)
- Pues si $c(\lambda)$ es la función característica del intervalo $(\lambda; \lambda + \Delta\lambda)$ (igual a uno dentro y a cero fuera de él), esa probabilidad es

igual al valor medio de la magnitud \mathcal{A} , y por lo tanto:

$$p_{\text{rob}}(\mathcal{A}; \Delta\lambda; \varphi) = \text{vm}(c(\mathcal{A}); \varphi) = (c(\mathcal{A})\varphi, \varphi) = \int_{-\infty}^{\infty} c(\lambda) d\|E_{\lambda}\varphi\|^2 = \int_{\lambda}^{\lambda+\Delta\lambda} d\|E_{\lambda}\varphi\|^2 = \|E_{\Delta\lambda}\varphi\|^2$$

b). $\|E_{\lambda}\varphi\|^2$ es la función de distribución de la magnitud \mathcal{A} , y su función característica define un grupo continuo y abeliano de operadores unitarios.

En efecto, $\|E_{\lambda}\varphi\|^2$ según a) representa la probabilidad de que el valor de \mathcal{A} sea menor que λ . Es una función no decreciente de λ ; $\|E_{-\infty}\varphi\|^2 = 0$; $\|E_{\infty}\varphi\|^2 = 1$ y es continua por la derecha. Por lo tanto es la función de distribución de \mathcal{A} . La función característica es:

$$f(s) = \int_{-\infty}^{\infty} e^{i\lambda s} d\|E_{\lambda}\varphi\|^2 = (e^{i s \mathcal{A}} \varphi, \varphi)$$

o sea es el valor medio del operador unitario $e^{i s \mathcal{A}}$, siendo s un parámetro real. La operación de grupo es: $e^{i s \mathcal{A}} \cdot e^{i s' \mathcal{A}} = e^{i(s+s')\mathcal{A}}$ obteniéndose la unidad para $s=0$, y la inversa cambiando s por $-s$.

c). La probabilidad de obtener un resultado determinado r sólo es distinta de cero si r pertenece al espectro puntual del observable. Pues sólo en los puntos del espectro puntual puede haber discontinuidad de la función de distribución:

$$\lim_{\Delta\lambda \rightarrow 0} \|E_{\lambda, \lambda+\Delta\lambda}\varphi\|^2 = 0 \text{ si } \Delta\lambda \text{ es entorno de un punto del espectro continuo.}$$

d). La probabilidad de obtener un resultado r es uno cuando y sólo cuando el estado es autofunción del observable que se mide y r el autovalor correspondiente.

Puesto que son equivalentes las condiciones:

$$E_{\lambda}\varphi = 0 \quad \text{para } \lambda < r; \quad E_{\lambda}\varphi = \varphi \quad \text{para } \lambda \gg r$$

$$\text{y} \quad A\varphi = \int_{-\infty}^{\infty} \lambda dE_{\lambda}\varphi = r\varphi.$$

II. LAS PROPIEDADES FÍSICAS.

Se llama "propiedad física" a toda magnitud que se mide con dos resultados posibles: "sí" o "no", y tal que no varía por conjunción consigo misma.

Las "propiedades" corresponden entonces al proceso de verificar la existencia o inexistencia de propiedades (ahora en el sentido usual) en un sistema físico. La magnitud es: "¿Existe la propiedad tal en el sistema en cierto estado?" La conjunción mencionada se expresaría: "¿Posee el sistema la propiedad tal y esa misma propiedad?" Es obvio que no hay diferencia.

Un ejemplo típico sería: "¿La medición de cierta magnitud dará un resultado comprendido entre a y b?"

1) A las propiedades corresponden los operadores proyección y la correspondencia es biunívoca.

Pues por ser magnitudes les corresponden operadores autoadjuntos, y los únicos idempotentes entre estos son, por definición, las proyecciones. Al autovalor uno se le atribuye el significado de "sí", y al cero el de "no".

2) La probabilidad de que un sistema en el estado φ (normalizado) posea la propiedad \mathcal{E} de proyección E es $\|E\varphi\|^2$

Pues como las propiedades sólo pueden ^{dar} como resultado uno o cero, su probabilidad es igual a su valor medio y

$$v_m(E; \varphi) = (E\varphi, \varphi) = \|E\varphi\|^2$$

3) Sea \mathcal{M}_E la multiplicidad lineal cerrada de E . La propiedad existe entonces con certeza en el estado φ si $\varphi \in \mathcal{M}_E$, y no existe, con certeza, si $\varphi \in \mathcal{J} - \mathcal{M}_E$. De modo que a la negación lógica de la propiedad \mathcal{E} , o sea a la propiedad "no \mathcal{E} ", corresponde la proyección $I - E$. Si $0 \neq E\varphi \neq \varphi$ no se puede afirmar nada sobre el resultado de la verificación de la existencia de esa propiedad en el sistema.

Sean \mathcal{E}^1 y \mathcal{E}^2 dos propiedades, y $\mathbb{E}^1, \mathbb{E}^2$ sus proyecciones. Su conjunción lógica expresa que el sistema posee ambas propiedades, o puede decirse, la propiedad " \mathcal{E}^1 y \mathcal{E}^2 ", de modo que debe esperarse un resultado uno cuando en efecto existen ambas y cero cuando no existe al menos una de ellas. Le corresponde entonces el operador $\mathbb{E}^1 \mathbb{E}^2$. Pero este operador sólo es proyección cuando $\mathbb{E}^1 \vee \mathbb{E}^2$, de modo que no siempre puede efectuarse la conjunción de propiedades (pero véase el apéndice E). Lo mismo ocurre con la disyunción " \mathcal{E}^1 o \mathcal{E}^2 ", que equivale lógicamente a " $\text{no} [(\text{no } \mathcal{E}^1) \text{ y } (\text{no } \mathcal{E}^2)]$ ". Es también necesario y suficiente que $\mathbb{E}^1 \vee \mathbb{E}^2$ para que en efecto sea una propiedad, y entonces su proyección es $\mathbb{E}^1 + \mathbb{E}^2 - \mathbb{E}^1 \mathbb{E}^2$, y no $\mathbb{E}^1 + \mathbb{E}^2$ como podría parecer a primera vista. Lo que se suma son las multiplicidades lineales correspondientes, de modo que la de " \mathcal{E}^1 o \mathcal{E}^2 " es $\mathcal{M}_{\mathbb{E}^1} \cup \mathcal{M}_{\mathbb{E}^2}$, así como la de " \mathcal{E}^1 y \mathcal{E}^2 " es $\mathcal{M}_{\mathbb{E}^1} \cap \mathcal{M}_{\mathbb{E}^2}$.

Resulta así evidente la íntima relación entre las operaciones de la lógica o el álgebra de clases y las de los operadores proyección, pero con una significativa diferencia: aquí hay operaciones "prohibidas".

Los autores que escriben sobre filosofía de la Mecánica Cuántica reconocen que el lenguaje usado en ella no puede seguir las mismas reglas que el común, debido a las expresiones de valor lógico indeterminado que aparecen (como por ejemplo, el sistema tiene tal posición y tal impulso). Hay dos salidas para esta situación: o guiarse por una lógica de tres valores (ver por ejemplo Reichenbach (14), parágrafo 30) o restringir el significado del lenguaje, de modo que existan frases sin sentido.

Más parece que el álgebra de las proyecciones incluye hasta cierto punto ambas posibilidades, pues por una parte añade un tercer valor a los enunciados (sí, no, indeterminado) y por otra restringe la aplicación de ciertas operaciones (conjunción, disyunción) quitándoles todo

significado cuando los prediandos no satisfacen cierta condición (conmutabilidad). Por supuesto la lógica de tres valores llega también a algo equivalente, pero el camino parece ser distinto.

Para fijar ideas he aquí un cuadro comparativo con la lógica usual:

<u>Lógica (sintaxis)</u>	<u>Algebra de Boole</u>	<u>Proyecciones.</u>
Sujeto	a (punto)	φ (punto de un espacio de Hilbert)
Prediando	X (conjunto)	E (operador proyección)
Enunciado, con dos valores de verdad.	$a \in X$ $a \notin X$	$E\varphi = \varphi$ (sí) $E\varphi = 0$ (no) $0 \neq E\varphi \neq \varphi$ (indeterm.)
Negación	$I-X \equiv X'$	$I - E$ (negación diametral)
Conjunción	$X \cap Y$	$EF ; E \vee F$
Disyunción	$X \cup Y$	$E + F - EF ; E \vee F$
Inclusión	$X \subset Y$	$E \leq F$
Implicación	$X \rightarrow Y (\equiv X' \cup Y)$	$I - E + EF ; E \vee F$
Tautología	$f(X, Y) \equiv I$	$f(E, F) \equiv I ; E \vee F$
absurdo	$f(X, Y) \equiv 0$	$f(E, F) \equiv 0 ; E \vee F.$

La negación es diametral porque transforma cierto en falso, falso en cierto e indeterminado en indeterminado. Hagamos notar que no es posible aquí definir otras negaciones, típicas de la lógica de tres valores (como la cíclica y la completa), pues todo lo que transforme lo indeterminado totalmente en cierto o falso coincide con el operador I y entonces los otros dos valores no sufrirían modificación. Como esas otras negaciones son necesarias para definir el concepto de complementaridad (que es el equivalente a la no-conmutabilidad) deducimos que las proyecciones no constituyen un caso típico de lógica de tres valores, a pesar de incluir un tercer valor de verdad.

Relaciones entre Propiedades y Magnitudes.

Es evidente que las propiedades son magnitudes de manejo simple, que por estar representadas mediante operadores acotados no dan lugar a complicaciones matemáticas, más aún recordando que poseen siempre un espectro discreto y con sólo dos autovalores.

Por ello sería muy de desear que pudieran reemplazarse las magnitudes en general por propiedades, y como se verá esto es perfectamente factible y aún ventajoso desde el punto de vista empirista.

En efecto, medir una magnitud A es hacer la pregunta "¿cuánto vale A ?" y es evidente que esto equivale a la sucesión de preguntas: "¿Está el valor de A comprendido entre a_1 y a_2 ?" "¿...entre a_2 y a_3 ?"..., que no son otra cosa que propiedades. Y advirtamos que esta formulación posee dos importantes ventajas: 1°), son más "reales", en el sentido que expresan más fielmente lo que ocurre en la experiencia, pues toda medición tiene su límite de error, de modo que nunca se da el valor de una magnitud a secas sino el intervalo en que se halla contenida. Por supuesto en el caso de convenirnos dar un valor exacto a cierta magnitud, ello también equivaldría a una propiedad: "¿Es a el valor exacto de A ?", y 2°) este método de ir "buscando por secciones" el valor de una magnitud elimina los infinitos o les da un sentido directo, por ejemplo: "¿Es el valor de A mayor que a ?", que es en rigor el único modo experimental de tratar el asunto.

¿Cómo se reconstruye una magnitud a partir de sus propiedades componentes? Supongamos que a la propiedad de que el valor de A esté comprendido entre a_n y a_{n+1} corresponde la proyección E_n . Para un estado φ cualquiera, la probabilidad de que esa propiedad se verifique es $\|E_n \varphi\|^2 = (E_n \varphi, \varphi)$. Como las propiedades de estar en intervalos diferentes se excluyen, se cumple $E_j E_k = E_j \delta_{jk}$ o sea las proyecciones distintas son ortogonales. Por lo tanto pueden sumarse y su suma será otra proyección. Definimos entonces las proyecciones $E_\lambda = \sum_{a_k \in \lambda} E_k$

que forman una familia espectral, como puede comprobarse de inmediato ($E_{-\infty} = 0$, $E_{\infty} = I$, $E_{\lambda} < E_{\mu}$ para $\lambda < \mu$, continuidad por la derecha). Ahora un E_{λ_n} puede escribirse $\Delta E_{\lambda_n} = E_{\lambda_n}$. El valor medio de \mathcal{A} en el estado φ será entonces, siendo λ_n un punto cualquiera del intervalo correspondiente: $vm(\mathcal{A}; \varphi) \cong \sum_n \lambda_n (\Delta E_{\lambda_n} \varphi, \varphi)$ y esto tiende a $\int_{-\infty}^{\infty} \lambda d(E_{\lambda} \varphi, \varphi) = (\mathcal{A} \varphi, \varphi)$ a medida que tomamos más pequeños los intervalos (el límite existe por hipótesis). Así hemos recobrado la expresión usual para el valor medio de \mathcal{A} . Si no pasamos al límite y nos quedamos con la suma en lugar de la integral, habremos representado un observable que "aproxima" a \mathcal{A} y puede reemplazarlo para todo uso experimental. En cambio en general habremos perdido una expresión matemática más sencilla.

Tomemos como ejemplo el operador coordenada q . La propiedad de que el sistema se halle en el intervalo (a_k, a_{k+1}) tiene como proyección a E_k , definida por: $E_k \varphi(q) = 0$ para $q < a_k$; $E_k \varphi = \varphi$ para $a_k < q < a_{k+1}$ y $E_k \varphi = 0$ para $q > a_{k+1}$. Es evidente que las distintas E_k son ortogonales entre sí y sumadas dan una familia espectral E_{λ_n} que en el límite coincide con la de q : $E_{\lambda} \varphi = \varphi$ para $q < \lambda$ y $E_{\lambda} \varphi = 0$ para $q > \lambda$.

Entonces el operador $q' = \sum_k q_k E_k$ (siendo q_k un punto cualquiera de ese intervalo) es una aproximación al operador q , tanto mejor cuanto más pequeños son los intervalos k .

III. ESTADÍSTICA CUÁNTICA.

El postulado III permite predecir resultados de observaciones, siempre que se conozca el estado ψ del sistema, pero esto es un caso de excepción. En efecto, sólo se tiene como arma el postulado IV, que permite eliminar los estados no compatibles con el resultado de la medición previa. Pero esto da un resultado único sólo en caso que la magnitud medida tenga espectro discreto y el valor observado sea un autovalor no degenerado. Sólo entonces hay correspondencia biunívoca entre autovalor y autofunción y puede deducirse que el estado del sistema está representado por esa autofunción. En todos los demás casos el valor medido corresponde una multiplicidad lineal que puede tener cualquier número de dimensiones y no hay regla alguna para elegir uno de sus puntos como representante del estado.

Es para estos casos en que los datos no alcanzan a determinar el estado del sistema que se introducen los métodos estadísticos, tanto clásicos como cuánticamente.

Para ello se estudia el comportamiento medio de un agregado formado por un gran número de sistemas iguales al de interés y luego, mediante alguna hipótesis de tipo ergódico, se extraen conclusiones con respecto al comportamiento individual de sus componentes.

En primerísimo lugar hay que aclarar el significado de "sistemas iguales". En general se dice que dos sistemas son iguales cuando dan resultados iguales para mediciones de todas las magnitudes imaginables efectuadas en las mismas condiciones. Sistemas iguales satisfacen el principio de causalidad; esto es una tautología.

Esta definición de igualdad no es útil pues es imposible de aplicar. La definición que implícitamente se usa es la siguiente: sistemas iguales son aquellos que poseen en común un cierto número de propiedades fijadas de común acuerdo por todos los observadores y que en general se resumen mediante la adjudicación de un nombre al sistema (electrón, molécula de hidrógeno). Además se fija también un cierto número de parámetros que por hipótesis bastan para determinar el estado del sistema.

Entonces sistemas iguales y en iguales condiciones significa que tienen igual nombre e iguales valores de esos parámetros. El principio de causalidad se convierte así en el problema de calcular cuantos y cuales parámetros son necesarios para definir la igualdad. Si parece no cumplirse eso sólo significará que existen otros parámetros "ocultos" cuya influencia no se había tomado en cuenta y que en adelante habrá también que medir.

Para la Mecánica Cuántica el problema no está así bien planteado, pues bastaría que entre los parámetros necesarios hubiese algún par no conmutable para que fuese imposible verificar la igualdad de condiciones según esa definición. Existe si un cierto número de propiedades constantes en el tiempo y conmutables con gran número de magnitudes, y ellas se usan para dar nombre a los sistemas del mismo modo que antes. Luego se consideran todas aquellas propiedades que se han verificado y se sabe que coexisten; por ser simultáneas sus proyecciones conmutan y pueden combinarse por conjunción; se tiene en definitiva una sola propiedad E que las resume a todas. Se dirá entonces que dos sistemas son equivalentes si tienen el mismo nombre y poseen la misma propiedad E . El término igualdad es preferible no utilizarlo. Es claro que esta definición es igualmente útil en el caso clásico, pero aquí se postula que existe una propiedad tan "completa" que permite sustituir la palabra equivalencia por igualdad. Un agregado estadístico formado por sistemas iguales ~~mutuamente~~ tendría las dos características siguientes: 1°) sería puro, lo que significa que el valor medio de cualquier magnitud tomado sobre el agregado total es igual al tomado sobre una parte cualquiera de él, de modo que por división en partes se obtienen agregados estadísticos idénticos. 2°) sería exacto, o sea que todas las magnitudes tendrían dispersión nula.

Cualquiera de estas dos características implica la otra.

En la estadística cuántica en cambio puede demostrarse que tal suposición es incorrecta y que aunque hay agregados puros no los hay exactos.

Representación Matemática de Agregados.

Un agregado estadístico queda determinado si se dan las correspondientes funciones de distribución para todas las magnitudes. Así $f_Q(\lambda)$ indica la probabilidad de obtener un resultado menor que λ al medir la magnitud Q , y está dada clásicamente por el cociente entre el número de sistemas que tienen un valor de Q menor que λ y el número total de sistemas del agregado.

Conocidas las $f_Q(\lambda)$ se pueden calcular los valores medios de las Q : $vm(Q) = \int_{-\infty}^{\infty} \lambda d f_Q(\lambda)$ y los de cualquier función de Q
 $vm[F(Q)] = \int_{-\infty}^{\infty} F(\lambda) d f_Q(\lambda)$.

Pero para introducir la noción de operador estadístico, debida a v. Neumann (ver (11) y también (10)) es preferible caracterizar al agregado mediante el dato de los valores medios de todas las magnitudes y todas sus funciones. Que esto es equivalente a dar las funciones de distribución se verifica fácilmente: sea $F_\lambda(x)$ la función característica del intervalo $(-\infty, \lambda)$ (o sea igual a uno para $x \leq \lambda$ y a cero para $x > \lambda$); entonces el valor medio de $F_\lambda(Q)$ es por definición la función de distribución $f_Q(\lambda)$

Dar un agregado estadístico significa entonces dar una función $vm(Q)$, cuyo argumento es cualquier magnitud imaginable.

Pero para que a la inversa una función $vm(Q)$ represente a un agregado estadístico debe cumplir ciertas condiciones:

- a). Ser siempre real, y si Q es una magnitud que por su naturaleza no puede dar resultados negativos, debe ser $vm(Q) \geq 0$.
- b). Si $Q = aA + bB$ siendo A y B otras dos magnitudes y a y b números reales, debe ser $vm(Q) = a.vm(A) + b.vm(B)$ independientemente de que A y B sean o no medibles simultáneamente en un mismo sistema.

Sin embargo nosotros introduciremos aún otra modificación que consiste en reemplazar las magnitudes por propiedades, y se verá como

ello simplifica conceptos y demostraciones.

Un agregado estadístico estará dado por la función $p(\mathcal{E})$, igual a la probabilidad de que se verifique afirmativamente la propiedad \mathcal{E} que esto es equivalente a la caracterización anterior es inmediato, pues la función $v_n(\mathcal{R})$ utilizada en la página anterior es una propiedad. A la inversa, como toda propiedad es a la vez una magnitud, resulta que la función $v_n(\mathcal{R})$ es sólo una prolongación de la $p(\mathcal{E})$, o sea $v_n(\mathcal{E}) = p(\mathcal{E})$, pero $p(\mathcal{R})$ no está definida si \mathcal{R} no es una propiedad. Esto explica que el trabajo con la $p(\mathcal{E})$ sea más sencillo, pues $v_n(\mathcal{R})$ contiene datos superabundantes que sólo sirven para complicar las formulaciones.

Las condiciones para que una función $p(\mathcal{E})$ caracterice a un agregado serán ahora:

- a) $p(\mathcal{E})$ real y mayor o igual a cero.
- b) Si \mathcal{E} y \mathcal{F} son dos propiedades excluyentes, o sea $\mathcal{E} \cap \mathcal{F} = 0$ debe cumplirse $p(\mathcal{E} \cup \mathcal{F}) = p(\mathcal{E}) + p(\mathcal{F})$
De aquí se deduce que si $\mathcal{E} \subset \mathcal{G}$, $p(\mathcal{E}) \leq p(\mathcal{G})$, pues \mathcal{G} puede descomponerse en $\mathcal{E} \cup \mathcal{F}$ siendo $\mathcal{E} \cap \mathcal{F} = 0$ y el teorema resulta del postulado a).

$p(\mathcal{E})$ será llamada la "función estadística" del agregado.

- a) Si se tienen dos agregados de funciones estadísticas $p_1(\mathcal{E})$ y $p_2(\mathcal{E})$ y con ellos se forma uno solo, su función estadística será $p(\mathcal{E}) = \alpha p_1(\mathcal{E}) + \beta p_2(\mathcal{E})$ siendo α y β la proporción numérica de cada agregado en el agregado total.
- d) $p(\mathcal{J}) \geq 0$. \mathcal{J} es una propiedad que poseen todos los sistemas, sin excepción, pues su único resultado posible es uno. Entonces equivale a afirmar que el sistema, que existía por hipótesis, en efecto existe, u otra tautología cualquiera. Pero lo más conveniente es considerarla como la propiedad de existir. Lo lógico entonces sería exigir $p(\mathcal{J}) = 1$ y entonces todas las probabilidades serían absolutas, normalizadas. Sin embargo a veces es conveniente

trabajar con probabilidades relativas, poniendo $p(\gamma) = \infty$ y renunciando a normalizar. Pero como nuestro objeto es justamente no introducir infinitos en las formulaciones físicas, supondremos siempre $p(\gamma)$ finita, y salvo advertencia explícita, normalizada.

e). Un agregado de función estadística $p(\mathcal{E})$ será puro cuando cualquier otro agregado formado con parte de sus componentes tiene también como función estadística a $p(\mathcal{E})$.

f). Un agregado $p(\mathcal{E})$ es exacto cuando los únicos valores que puede tomar $p(\mathcal{E})$ son cero y uno.

Pues en un agregado exacto todas las magnitudes tienen dispersión nula, por definición, o sea: $vm(Q^2) = [vm(Q)]^2$ para todo Q

Pero si Q es una propiedad \mathcal{E} , como $\mathcal{E}^2 = \mathcal{E}$, resulta

$p(\mathcal{E}) = [p(\mathcal{E})]^2$ y esto sólo se cumple para los números uno y cero.

g). Todo agregado exacto es puro. Pues si lo dividimos en dos partes en proporción α y β ($\alpha + \beta = 1$) y las funciones estadísticas de esas partes son $p_1(\mathcal{E})$ y $p_2(\mathcal{E})$, vale

$p(\mathcal{E}) = \alpha p_1(\mathcal{E}) + \beta p_2(\mathcal{E})$ y como son todas cantidades no negativas, si $p(\mathcal{E}) = 0$ tiene que ser $p_1(\mathcal{E}) = p_2(\mathcal{E}) = 0$; y si $p(\mathcal{E}) = 1$, como $p(\gamma - \mathcal{E}) = 0$ (por b).), resulta

$p_1(\gamma - \mathcal{E}) = p_2(\gamma - \mathcal{E}) = 0$, o sea $p_1(\mathcal{E}) = p_2(\mathcal{E}) = 1$. Y como uno y cero son los únicos valores que puede tomar el agregado original, queda demostrado el teorema.

El teorema recíproco no es cierto, como se verá más adelante.

El Operador Estadístico.

La elección de la forma de la función $p(\mathcal{E})$ o $vm(Q)$ implica en realidad un nuevo postulado que sustituye y generaliza al postulado III. El siguiente razonamiento sugiere cual ha de ser esa forma.

Sea $\{\varphi_n\}$ un sistema ortonormal cualquiera y Q una magnitud. Si

un sistema se halla en el estado φ_k , el valor medio de Q es $(R \varphi_k, \varphi_k)$. Pero si en lugar de un sistema se tiene un agregado que es mezcla de sistemas en los estados $\{\varphi_n\}$ en la proporción u_n para el estado φ_n , el valor medio de Q sobre todo el agregado será $vm(Q) = \sum_n u_n (R \varphi_n, \varphi_n)$, y siendo U un operador cuyos autovalores son los u_n y cuyas autofunciones son las φ_n , esa suma es el Spur de UR , pues

$$\sum u_n (R \varphi_n, \varphi_n) = \sum (R \varphi_n, U \varphi_n) = \sum (UR \varphi_n, \varphi_n) = Sp UR$$

Entonces se adoptará como nuevo postulado:

POSTULADO III'. El valor medio de cualquier magnitud Q está dado por la función $vm(Q) = SpUR$, o también, la probabilidad de cualquier propiedad \mathcal{E} está dada por la función $p(\mathcal{E}) = SpUR$; siendo U un operador que caracteriza al agregado estadístico representativo del sistema.

U se llama "operador estadístico" del agregado o del sistema que ese agregado representa. Pero hay que verificar que la función $SpUR$ es una función estadística. Y en efecto:

a). Esta condición se satisface exigiendo que U sea autoadjunto y definido positivo, pues el producto de dos operadores definido positivos tiene Spur mayor o igual a cero (ver (9), Cap. II, párrafo 11), y las proyecciones también lo son.

b). $SpU(E+F) = SpUE + SpUF$ pues el Spur es una función lineal.

Esto vale también si $EF \neq 0$, pero entonces $E+F$ no es una propiedad y se trata más bien de la función $vm(Q)$ que de la $p(\mathcal{E})$.

c). Si U_1 y U_2 son los operadores estadísticos de dos agregados y con estos se forma uno solo en el que entran en la proporción $\alpha:\beta$, $\alpha+\beta=1$, el op.est. del agregado suma es

$$U = \alpha U_1 + \beta U_2$$

d). Si se exige que el operador U tenga la probabilidad uno (normalización), ello equivale a pedir: $\text{Sp}U = \text{Sp}U = 1$, de modo que U , por tener Spur finito es un operador "vollstetig" y tiene espectro puntual puro.

e). Agregados puros. Si se divide un agregado de op.est. U en dos partes en proporción numérica $\alpha : \beta$, $\alpha + \beta = 1$, y cuyos op.est. son U_1 y U_2 , el agregado inicial se llamará puro si, cualquiera sea el modo de partición, se cumple $U_1 = U_2 = U$.

Teorema. La condición necesaria y suficiente para que un agregado sea puro es que su operador estadístico sea una proyección sobre una sola dimensión $U = P_\varphi$

La condición es suficiente. Pues sea $U = P_\varphi$ y f una función cualquiera ortogonal a φ . Como U, U_1, U_2 son definido positivos, vale: $0 \leq \alpha (U_1 f, f) \leq \alpha (U_1 f, f) + \beta (U_2 f, f) = (U f, f) = (P_\varphi f, f) = 0$. Pero de $(U_1 f, f) = 0$ se deduce, por la positividad, $U_1 f = 0$. Si ahora g es una función cualquiera, $(f, U_1 g) = (U_1 f, g) = 0$, de modo que toda función f ortogonal a φ es ortogonal a $U_1 g$, o sea $U_1 g = a_g \varphi$ para todo g , y considerando que $\text{Sp}U_1 = 1$ resulta $U_1 = P_\varphi$ y por lo tanto también $U_2 = P_\varphi$.

Para demostrar la necesidad de la condición, v. Neumann (ver (9), pág. 169) procede de la siguiente manera: como de la igualdad $U = \alpha U_1 + \beta U_2$ se deduce ahora por hipótesis, $\underbrace{U = U_1 = U_2}$ bastará demostrar que cualquier operador U puede descomponerse en suma de una proyección unidimensional y algo más (que también sea un op.est.), es decir: $U = \alpha P_\varphi + \beta U_2$, y de aquí entonces resulta $U = P_\varphi$.

Pero para efectuar dicha descomposición nosotros seguiremos un camino distinto, que nos parece más ilustrativo. Por ser $\text{Sp}U = 1$, U posee espectro puntual puro; sean $\{\lambda_n\}$ sus autovalores ($\sum \lambda_n = 1$) y P_n las multiplicidades lineales características correspondientes. Entonces U puede escribirse $U = \sum \lambda_n P_n$ (que correspondería a una mez-

ela de agregados de op.est. P_n con peso λ_n respectivamente).
Elijamos un P_k tal que $\lambda_k \neq 0$ y supongamos que $\sum'_k \lambda_n \neq 0$ (\sum'_k indica que la suma se extiende a todo $n \neq k$). Entonces:

$$U \equiv \lambda_k P_k + \sum' \lambda_n P_n = r \lambda_k \frac{P_k}{r} + \sum' \lambda_n \frac{\sum' \lambda_n P_n}{\sum' \lambda_n}$$

siendo r el número de dimensiones de P_k . Si ahora identificamos:

$$r \lambda_k = \alpha \quad ; \quad \sum' \lambda_n = \beta \quad ; \quad \frac{P_k}{r} = U_1 \quad , \quad \frac{\sum' \lambda_n P_n}{\sum' \lambda_n} = U_2 \quad ,$$

tenemos una descomposición de U físicamente significativa y que contradice la hipótesis, pues evidentemente $U_1 \neq U_2$. De aquí se deduce que $\sum' \lambda_n$ no puede ser distinto de cero, y por lo tanto U es múltiplo de una proyección $U = P_k / r$. Pero si r es mayor que uno, es evidente que P_k puede descomponerse en suma de dos proyecciones ortogonales de modo que nuevamente habría contradicción con la hipótesis. Entonces debe ser $r=1$ y $U = P_\varphi$.

f). No hay agregados exactos. Pues en caso de haberlos, $SpUE$ debe tomar sólo los valores uno y cero, pero si K es proyección sobre una sola dimensión, $K = P_\varphi$, $SpUP_\varphi$ es función continua de φ , de modo que, o es constantemente nula y entonces $U=0$, o es constantemente igual a uno y entonces $U=I$, pero ambos casos están excluidos pues no tienen Spur uno, y además por el párrafo anterior, U tendría que ser una proyección unidimensional, ya que todo agregado exacto es puro.

g). Como se acaba de ver, un agregado puro no es exacto, a diferencia de lo que ocurre en la estadística clásica. Pero decir que un agregado es puro es lo más cerca que podemos llegar del concepto clásico de un agregado cuyos componentes son iguales, pues ello indica una completa homogeneidad de propiedades en los componentes. Sin embargo aún así habrá magnitudes de dispersión no nula. Como esto no puede deberse a la existencia de "parámetros ocultos", pues las deducciones son independientes del número y clase de parámetros, esto significa

que no se cumple el principio de causalidad. Es decir, no es posible definir la igualdad de dos estados mediante la igualdad de los resultados de mediciones efectuadas en ellos.

En rigor hagamos notar que el derrumbe no es total ni mucho menos. En efecto, si conocemos "exactamente" el estado de un sistema (o sea la función de onda φ), podemos en muchos casos "aplicar la ley de causalidad", predecir con exactitud muchos resultados, a saber los de aquellas magnitudes de cuyos operadores sea φ autofunción (todas las conmutables con la que permitió determinar φ). En cuanto a las otras, si no podemos predecir exactamente sus valores, podemos predecir la ley de distribución de estos, lo cual no es poco.

Conocer "exactamente" el estado de un sistema es el caso límite de la estadística (tanto clásica como cuánticamente), es el caso en que la estadística no es necesaria. El agregado sería puro y su operador estadístico la proyección sobre la función de onda de ese estado $U = P_{\varphi}$. Y en efecto, tal caso el valor medio de una magnitud cualquiera es $SpUA = SpP_{\varphi}A = (A\varphi, \varphi)$, como es inmediato si para calcular el Spur se toma a φ como uno de los ejes de coordenadas. En este caso se dice que el agregado es un estado, en los demás es una mezcla de estados.

Cálculo del Operador Estadístico.

Para que el tratamiento del problema físico sea completo es necesario aún indicar cómo se calcula el op.est. de un agregado. Esto debe ser siempre posible, pues para eso se introducen los métodos estadísticos.

Como dato se tiene el resultado de ciertas mediciones (conmutables), que como se ha visto, pueden resumirse todas en una cierta propiedad \mathcal{E} . Decir que el sistema tiene la propiedad \mathcal{E} significa afirmar que los resultados de tales mediciones son tales.

Como hipótesis se supone que todo el conocimiento adquirido sobre el sistema está resumido en \mathcal{E} . Nada puede afirmarse sobre una propiedad no contenida en \mathcal{E} ; ni que es cierta ni que es falsa ni que es indeterminada.

El agregado estadístico que representará al sistema de interés tiene entonces que cumplir dos condiciones.

1°). Todos sus componentes poseen la propiedad \mathcal{E} .

2°). Ningún otro factor debe ser tomado en cuenta.

Llamando U al operador estadístico buscado, la primera condición dice: $\text{Sp}UX = 1$. La segunda se expresa usualmente diciendo "hay que elegir iguales probabilidades a priori", y seguidamente se verá su formulación matemática.

Para satisfacer la primera condición el procedimiento es el siguiente: se van tomando distintos sistemas (que puede suponerse formaban parte de un cierto agregado U_0) y en ellos se mide la propiedad \mathcal{E} . Si la poseen se incluyen en el agregado que se está construyendo; si no, se dejan de lado.

Pero ¿de cuántos modos puede medirse una propiedad \mathcal{E} ? Sea \mathcal{M}_E su multiplicidad (que se supone de más de una dimensión, pues si no todo es trivial) y $\{\varphi_n\}$ cualquier sistema de coordenadas en ella. Se completa este sistema en $\mathcal{J}-\mathcal{M}_E$ con otro, sea $\{\psi_n\}$, y se establece una correspondencia biunívoca $\varphi_n \leftrightarrow \lambda_n, \psi_n \leftrightarrow \mu_n$ siendo los λ_n y μ_n números reales cualesquiera pero todos distintos entre sí. Se tiene así definido un operador autoadjunto $\overset{R}{U}$ cuyos autovalores son los $\{\lambda_n\}, \{\mu_n\}$ y autofunciones $\{\varphi_n\}, \{\psi_n\}$. Por el postulado II le corresponde una magnitud física medible. Fórmese ahora la función $F(R)$, siendo $F(x) = 1$ para $x \in \{\lambda_n\}$ y $F(x) = 0$ para $x \in \{\mu_n\}$ (y fuera de esos puntos arbitraria, pero integrable). El operador $F(R)$ tiene como únicos autovalores uno y cero. Uno para las funciones de \mathcal{M}_E y cero para las ortogonales a ellas. Por lo tanto $F(R) = E$ y la medi-

ción de R equivale a la verificación de \mathcal{E} . Pero hay infinitas R admisibles, pues hay infinitos modos de elegir los sistemas $\{\varphi_n\}$ y $\{\psi_n\}$ y los números $\{\lambda_n\}$ y $\{\mu_n\}$.

Hay entonces infinitas maneras de verificar una propiedad.

Por otra parte, si el agregado original tenía op.est. U_0 , el agregado obtenido con aquellos de sus componentes que dan $F(R) \neq 1$ tendrá como operador estadístico:

$$U = \sum_{\{\varphi_n\}} (U_0 \varphi_n, \varphi_n) \frac{P_{\varphi_n}}{\text{Sp} U_0 E}$$

pues al medir R todos los sistemas quedan en alguno de sus autoestados, $\{\varphi_n\}$ o $\{\psi_n\}$, y eso con probabilidad $\text{Sp} U_0 P_{\varphi_n} = (U_0 \varphi_n, \varphi_n)$ etc. y se utilizan sólo los $\{\varphi_n\}$. U es una mezcla de los estados P_{φ_n} .

El factor $1/\text{Sp} U_0 E$ es para normalizar.

Como $R \vee P_{\varphi_n}$ para todo n , resulta $U \vee R$, o sea, U conmuta con todo operador R que se utilice para la formación de su agregado estadístico.

Es ahora necesario satisfacer la segunda condición. Se ha visto que hay infinitos modos de elegir el operador R . No hay ninguna razón a priori para preferir algunos de ellos o rechazar otros. Si nada se sabe del sistema fuera de tener la propiedad \mathcal{E} , todos los procedimientos de verificación deben ser admisibles. Esto no es más que el postulado de equiprobabilidad a priori, aunque aquí parece mucho más "natural" que en el espacio de las fases. En realidad sigue siendo un postulado que dice: "no hay ninguna ley física que dé normas a priori sobre selección de métodos de verificación de propiedades". v. Neumann (9), que fué el primero en dar la expresión cuántica precedente del postulado, parece dar a entender que ahora, más que postulado, es una necesidad lógica. A nosotros nos parece que el problema ha sido llevado a un terreno mucho más concreto al traducirse a la Cuántica, pero que no es inconcebible que un día se enuncie otro postulado que prohíba por ejemplo la utilización de operadores

R si no son acotados, o positivos, o algo así. Pero aceptando que no se rechaza ningún R pueden demostrarse los siguientes teoremas:

Teorema. El operador estadístico de un agregado formado para representar a un sistema que posee la propiedad \mathcal{E} (que resume todo lo que se sabe de él) es $U = \frac{1}{r} E$ (siendo $r = SpK = K^0$ de dimensiones de \mathcal{M}_E).

Pues sea $\varphi \in \mathcal{M}_E$. Hay algún R del que es autofunción :

$R\varphi = \lambda\varphi$, y entonces, como $R \vee U$: $RU\varphi = UR\varphi = U\lambda\varphi = \lambda U\varphi$ o sea $U\varphi$ también es autofunción de R y para el mismo autovalor, pero como estos son todos simples vale $U\varphi = a\varphi$.

Como esto vale para todo $\varphi \in \mathcal{M}_E$, el factor a debe ser independiente de φ , ~~maximamente~~ Recordando que $U\psi = 0$ para todo $\psi \in \mathcal{J} - \mathcal{M}_E$ (pues no hay ningún sistema del agregado que tenga la propiedad no \mathcal{E}), resulta entonces $U = aE$, y de $SpU = 1$ se deduce $a = 1/SpE = 1/r$.

Teorema. Condición suficiente para que pueda resultar $U = E/r$ es que, siendo U_0 el op.est. del agregado original, se cumpla $U_0 E = E$.

Pues como ya se ha visto es

$$U = 1/SpU_0 E \sum_{\{\varphi_n\}} (U_0 \varphi_n \cdot \varphi_n) P_{\varphi_n} = 1/SpE \sum_{\{\varphi_n\}} P_{\varphi_n} = E/SpE .$$

Un agregado original que sirviera para formar el agregado estadístico correspondiente a cualquier propiedad tendría entonces que satisfacer a la ecuación $U_0 E = E$ para toda E , o sea $U_0 = I$.

Sin embargo I no es normalizable ($SpI = \infty$), de modo que no cabe en nuestra exposición. Es que en realidad resulta de una generalización demasiado amplia. En efecto, las propiedades de la realidad no son infinitas, no cubren todo el espacio de Hilbert. Bastará entonces considerar un operador I' que las contenga a todas y del cual en cierto modo I sería el límite al tomarse en cuenta todas las propiedades teóricamente imaginables. I' tiene una multiplicidad lineal de enorme, pero no infinite, número de dimensiones.

Para admitir esto tenemos que investigar más a fondo el cará-

ter de ciertas propiedades que ya por sí solas no son normalizables. Tomemos como ejemplo el siguiente: el sistema tiene la propiedad de que su coordenada x se halla en el intervalo Δx . Por pequeño que sea Δx , la multiplicidad correspondiente tiene infinitas dimensiones, pues es bien sabido que hay infinitas funciones de L^2 linealmente independientes y nulas fuera de un intervalo dado. Pero esto ocurre porque se adjudica a x un espectro continuo, de modo que el intervalo Δx puede decrecer sin término y siempre corresponderle una probabilidad mayor que cero (relativa al intervalo unidad).

Nos parece que eso no es correcto. En efecto, existen para el impulso de un sistema límites dados por la finitud de la energía total del universo. Ningún sistema puede tener impulso infinito. Pero según lo dicho en página 6, eso significa que la coordenada espacial sólo puede tomar como valores los múltiplos enteros de cierta cantidad que equivaldría a un "cuanto de posición", de modo que el pasaje de una posición a otra sería discontinuo. Como ahora en lugar de funciones hay vectores, al autovalor λ_k corresponde el autovector $\{v_n\}$ siendo $v_n = \delta_{nk}$, y por lo tanto es simple. Entonces a todo intervalo finito corresponde una multiplicidad lineal de número finito de dimensiones.

Hay otra dificultad, pero menos grave: si \mathcal{E} tiene dimensión finita, "no \mathcal{E} " la tendría infinita, pero es que ahora no le hacemos corresponder el operador $I - K$ sino el $I' - K$.

Podemos entonces suponer que todas las propiedades experimentales corresponden a proyecciones de Spur finito.

En el límite $I' \rightarrow I$ puede considerarse como la mezcla en iguales proporciones de todos los estados P_φ mutuamente excluyentes. Es fácil demostrar que la medición de cualquier propiedad en él no lo altera (ver v. Neumann, (9), pág. 183), es decir que corresponde al más alto grado de equilibrio imaginable.

IV. APLICACIONES.

Postulación definitiva.

Después de todas estas consideraciones estamos en condiciones de resumir los métodos de la estadística cuántica en dos postulados, en apariencia sencillos:

POSTULADO I'. Hay correspondencia biunívoca entre propiedades físicas y operadores proyección en un espacio adecuado.

POSTULADO II'. Si la propiedad de proyección E resume todo lo que se sabe de un sistema físico, la probabilidad de que una verificación de otra propiedad F dé resultado afirmativo es: $SpEF/SpE$.

El postulado I' equivale exactamente al II (pág. 5) y merece los mismos comentarios.

En cuanto al II', se supone que la propiedad E existe en el momento en que el observador procede a la verificación de F . Si E se había medido con una diferencia apreciable de tiempo de observación es necesario dar alguna regla que permita sustituirla por otra más adecuada. Como se sabe dicha regla es: sustituir la proyección E por la proyección $T_t E T_t^{-1}$, donde t es el tiempo transcurrido entre las dos mediciones y T_t es un grupo abeliano continuo de operadores unitarios cuya forma debe ser dada por nuevas reglas, la más sencilla de las cuales se expresa: $T_t = e^{iHt}$, siendo H el hamiltoniano del sistema. (Véase por ejemplo Dirac (3) Cap.V).

Variación del Operador Estadístico con el Transcurso del Tiempo.

En primer lugar se verá la influencia de las mediciones sobre el operador estadístico. Esta puede considerarse desde dos puntos de vista: 1°). Si interesa seguir la evolución de un solo sistema, de modo que

al agregado estadístico es sólo un medio de estudiarlo, el proceso es sencillo: se sabe que el sistema tiene la propiedad \mathcal{E} ; entonces el op.est. del agregado que lo representa es $U = E/SpE$; se verifica luego la propiedad \mathcal{F} , que existe con probabilidad $SpEF/SpE$. Si el resultado es afirmativo, se sabe ya otra cosa que antes: el sistema tiene la propiedad \mathcal{F} , de modo que el nuevo agregado que debe representar estadísticamente al sistema tiene como op.est. $U = F/SpF$ (pero véase el Apéndice I). Si el resultado fué negativo sólo hay que reemplazar F por $I' - F$.

2°). Si interesa seguir la evolución del agregado estadístico en sí, verificando la propiedad \mathcal{F} en todos sus componentes pero sin luego separar ninguno ni agregar otros nuevos, el razonamiento es como sigue: el agregado tenía el op.est. $U = E/SpE$. En la fracción $SpEF/SpE$ de sus componentes existe la propiedad \mathcal{F} y en la fracción $Sp(I'-F)E/SpE = 1 - SpEF/SpE$ no existe, o sea existe $I'-F$.

Entonces el nuevo operador estadístico es

$$U' = \frac{SpEF}{SpE} \frac{F}{SpF} + \frac{SpE(I'-F)}{SpE} \frac{(I'-F)}{Sp(I'-F)} \quad (\text{pero véase el Apéndice I}).$$

Hagamos observar que el hecho de haber trabajado con propiedades en lugar de magnitudes cualesquiera nos ahorra una cantidad de complicaciones formales.

3°). En el intervalo entre dos mediciones U también se modifica. Si se considera el tiempo del sistema como parámetro, es decir, si no se incluye la cuarta integración en el producto escalar, puede obtenerse la variación de U definiendo la derivada como límite del cociente de incrementos:

$$\frac{\partial U}{\partial t} = \lim_{\Delta t \rightarrow 0} \frac{U_{t+\Delta t} - U_t}{\Delta t}$$

Supóngase en primer lugar que U_t es proyección unidimensional sobre

\mathcal{C}_t , $U_t = P \varphi_t$, y sea ψ una función cualquiera:

$$\frac{\partial U}{\partial t} \psi = \lim_{\Delta t} \frac{U_{t+\Delta t} - U_t}{\Delta t} \psi = \lim_{\Delta t} \frac{(\psi, \varphi_{t+\Delta t}) \varphi_{t+\Delta t} - (\psi, \varphi_t) \varphi_t}{\Delta t}$$

y sumando y restando $(\psi, \varphi_t) \varphi_{t+\Delta t}$ en el numerador y pasando al límite

$$\frac{\partial U}{\partial t} \psi = (\psi, \frac{\partial \varphi}{\partial t}) \varphi + (\psi, \varphi) \frac{\partial U}{\partial t} = \frac{2\pi i}{h} [(H\psi, \varphi) \varphi - H(\psi, \varphi) \varphi] = \frac{2\pi i}{h} [UH\psi - HU\psi]$$

y como ψ era cualquiera: $\frac{\partial U}{\partial t} = [H, U]$

Si U es combinación lineal de proyecciones unidimensionales vale la misma conclusión puesto que todas las operaciones efectuadas son lineales. Como ese es siempre nuestro caso, esa igualdad vale para todos nuestros operadores estadísticos.

Resalquemos que aun cuando formalmente esta fórmula es igual a la de pág. 12, su significado es distinto. Aquí el tiempo figura como parámetro (aunque es el tiempo del sistema, no el de observación) y por eso es necesaria una demostración especial que no vale en general.

Pues una propiedad, a menos que en su enunciación intervenga explícitamente el tiempo, es siempre la misma (su enunciado no cambia) y su "derivado parcial" es nula. Pero si a esa misma propiedad se la toma como operador estadístico, debe acompañar en su variación al sistema que representa, o sea su multiplicidad lineal debe transformarse como si se aplicara la ecuación de Schrödinger a cada uno de sus rayos.

Identificando, como se hace siempre, los tiempos de observación y del sistema, se deduce que la variación de U puede describirse también mediante el grupo de operadores unitarios T_t , del siguiente modo: $U_t = T_t U T_t^{-1}$.

El Teorema Ergódico.

Sin detenernos en la importancia y significado de la hipótesis ergódica, haremos en primer lugar un breve resumen de su planteamiento clásico.

En el espacio de las fases de un sistema, su punto representativo q_t describe una trayectoria, a la que en general se le exige no salir de cierto dominio prefijado, por ejemplo una hipersuperficie \mathcal{H} de energía constante. El teorema dice entonces: "el tiempo medio de

permanencia de q_t en un dominio cualquiera \mathcal{D} de volumen finito tiende, para $t \rightarrow \infty$, al cociente $\text{vol}(\mathcal{D})/\text{vol}(\mathcal{H})$ si \mathcal{H} tiene volumen finito, y a cero en caso contrario".

Este teorema permaneció sin demostración hasta que en 1931 Koopman (7) encontró que los movimientos en el espacio de las fases podían representarse mediante un grupo de operadores unitarios en un espacio de Hilbert, del siguiente modo: el teorema de Liouville permite establecer la invariancia del volumen ocupado por un número fijo de puntos representativos. Como medida invariante ese volumen permite definir un producto escalar apropiado: sea $f(q)$ una función cuyo argumento q pertenece al espacio de las fases; diremos que pertenece al espacio de Hilbert "ergódico" si $\int_{\mathcal{H}} |f(q)|^2 dv(q) < \infty$ siendo $dv(q)$ el volumen infinitesimal de un entorno de q , y si $g(\cdot)$ también pertenece a ese espacio, definiremos

$$(f, g) = \int_{\mathcal{H}} f(q) \cdot \bar{g}(q) dv(q)$$

En este espacio de Hilbert se definen ahora los operadores unitarios V_t mediante: $V_t f(q) = f(Q_t)$, siendo Q_t la posición que ocuparía el punto q al cabo del tiempo t . Que son unitarios resulta de la invariancia de $v(\cdot)$, pues:

$$(V_t f, V_t g) = \int f(Q_t) \cdot \bar{g}(Q_t) dv(Q) = \int f(Q) \cdot \bar{g}(Q) dv(Q) = (f, g)$$

Como por supuesto $V_t \cdot V_{t'} = V_{t+t'}$ y $V_0 = I$ y $V_t^{-1} = V_{-t}$, los operadores V_t forman un grupo abeliano, continuo.

Un dominio \mathcal{D} del espacio de las fases estará representado en el espacio ergódico por su función característica $c_{\mathcal{D}}(q)$, igual a uno si $q \in \mathcal{D}$ y a cero en caso contrario.

Ahora el teorema ergódico se expresa así:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T V_t c_{\mathcal{D}}(q) dt = \int_{\mathcal{H}} |c_{\mathcal{D}}(q)|^2 dv(q) = \|c_{\mathcal{D}}\|^2$$

donde hemos supuesto que el volumen total del espacio de las fases que interviene está normalizado a uno ($\int_{\mathcal{H}} dv(\cdot) = 1$), cosa que seguiremos aceptando en lo sucesivo.

Con esta formulación el teorema ergódico fué rápidamente demostra-

do por v. Neumann (12) para los casos de convergencia en promedio y por Birkhoff (2) para la convergencia puntual, aunque usando métodos totalmente distintos y que no vamos a resumir. La parte matemáticamente importante del teorema es la demostración de la existencia del límite temporal; luego es fácil ver que si existe debe ser igual al promedio espacial, con una condición: que haya "transitividad métrica", lo que significa que en la parte del espacio de las fases de la que por hipótesis no puede salir el sistema (\mathcal{H}) no deben existir dominios invariantes en el tiempo, o sea que $\forall t, \phi_t(Q) \neq Q$ para casi todo t , a menos que $\mathcal{D} = \mathcal{H}$ o $\mathcal{D} = 0$.

Pasemos ahora a la Mecánica Cuántica y tomemos al tiempo del sistema como parámetro. En primer lugar no puede utilizarse el espacio de las fases, pues está prohibido señalar en él puntos o incluso dominios limitados, de modo que una trayectoria carece de sentido en cualquier aproximación. A primera vista esto parece una catástrofe, pues todos los conceptos utilizados en el teorema ergódico: trayectorias, celdas, superficies de energía constante, necesitan del espacio de las fases para ser expresados. Sin embargo ya en 1929, v. Neumann (11) al tratar de demostrar este teorema, reconoció que el espacio de las fases puede y debe reemplazarse por el espacio de Hilbert del sistema, y en realidad ya hemos visto como todos los resultados de la estadística de Gibbs se expresan y demuestran en ese espacio.

Procedamos pues a reformular el problema.

- a). En lugar de espacio de las fases del sistema de interés tendremos su espacio de Hilbert.
- b). El agregado estadístico representativo, que se describía mediante una densidad de probabilidad función de las coordenadas e impulsos, está caracterizado ahora por un operador en dicho espacio de Hilbert: el operador estadístico U , que podemos suponer proporcional a una proyección: $U = L/S_p L$ ($S_p L$ finito).

c). Un punto del espacio de las fases es un rayo en el espacio de Hilbert. Al desplazamiento $Q \rightarrow Q_t$ corresponde la transformación

$$\varphi \rightarrow T_t \varphi$$

d). Decir que la trayectoria de un punto no puede salir de una cierta hipersuperficie (no necesariamente la que corresponde a energía constante) equivale a afirmar que $\|U\varphi\| = \|UT_t\varphi\|$ para todo t y toda φ , o sea $U \nabla T_t$, de modo que en todos los casos en que T_t es función de H solamente se tiene $U \nabla H$ y U es una constante del movimiento.

e). En lugar de celdas o dominios en el espacio de las fases tendremos multiplicidades lineales o sus proyecciones. La celda ocupada por varios puntos se transforma en la multiplicidad determinada por los rayos correspondientes. L se subdivide entonces en multiplicidades lineales disjuntas, que son las celdas.

f). El volumen de una celda es sustituido por el Spur de la proyección correspondiente (o sea el N° de dimensiones de la multiplicidad lineal). *SpL celda*

g). La fracción $\delta N/N$ del número total de sistema que se halla en la celda $dv(Q)$, era: $\delta N/N \approx \rho(Q)dv(Q)$. Ahora diremos, la fracción $\delta N/N$ que tiene la propiedad \mathcal{E} es: $\delta N/N = SpUE = SpLE/SpL$ y como E es una celda de L , es $E \subset L$, o sea $\delta N/N = SpE/SpL$.

Es ahora invariante el volumen? Si, pues el transcurso del tiempo t equivale a la transformación $E \rightarrow T_t E T_t^{-1}$, y el Spur es un invariante unitario. Pero si se quiere mayor paralelismo con el caso clásico, puede demostrarse también en el espacio de Hilbert un teorema de Liouville, pues $\frac{dU}{dt} = [U, H] + \frac{\partial U}{\partial t}$, pero por ser U un operador estadístico es $\frac{\partial U}{\partial t} = [H, U]$ de donde $\frac{dU}{dt} = 0$. Entonces si seguimos a una fracción fija de sistemas en el transcurso del tiempo tendremos formalmente:

$$0 = \frac{d(\delta N/N)}{dt} = \frac{d}{dt} SpUE = Sp\left(\frac{dU}{dt} E + U \frac{dE}{dt}\right) = Sp U \frac{dE}{dt}$$

grat

y si E por hipótesis no puede "salirse" de L , se deduce $\frac{dE}{dt} = 0$ o sea $\frac{d}{dt} \text{Sp}E = 0$. (Nótese que para esta demostración suponemos que son conmutables la derivación y el cálculo del Spur).

b). Pero si ha sido posible definir una medida invariante podemos aprovechar para la Mecánica Cuántica la misma demostración del teorema ergódico que en el caso clásico. Para ello definimos un nuevo espacio de Hilbert "ergódico". Cada punto de este espacio es una función compleja cuyo argumento es una celda genérica del espacio de Hilbert del sistema. La "hipersuperficie" L se subdivide de un modo cualquiera en celdas disjuntas y a cada una de éstas le corresponde un número. Sea $\{E_n\}$ una tal subdivisión y $f(E_n)$ una función en ella definida. (que las celdas sean disjuntas significa que $E_j E_k = 0$ para $j \neq k$.) Llamaremos integral de f a $\int f \cdot dv = \sum_n f(E_n) \text{Sp}UE_n$ (que como es evidente no varía si las E_n se subdividen a su vez).

Como $\text{Sp}UE_n = \text{Sp}E_n / \text{Sp}L$, a medida que L aumenta de dimensiones esa suma tiende a una integral.

Consideraremos en primer lugar que todas las subdivisiones de L son conmutables, es decir, puede darse un sistema de coordenadas común a todas ellas. Entonces si $\{F_n\}$ es otra subdivisión y $g(F_n)$ una función definida en ella, y $\{\varphi_n\}$ el sistema ortonormal común, la suma $f+g$ se formará sumando los valores de f y g para cada P_{φ_n} , poniendo naturalmente $f(P_{\varphi_n}) = f(E_n)$ para $P_{\varphi_n} \subset E_n$ y lo análogo para g . Asimismo, el producto escalar será:

$$(f, g) = \sum_n f(P_{\varphi_n}) \cdot \bar{g}(P_{\varphi_n}) \cdot \text{Sp}UP_{\varphi_n} = \sum_n f(P_{\varphi_n}) \cdot \bar{g}(P_{\varphi_n}) \Big|_{\text{Sp}L} = \int f \cdot \bar{g} \cdot dv$$

y por lo tanto la norma: $\|f\|^2 = \sum_n |f(E_n)|^2 \text{Sp}UE_n = \int |f|^2 dv$

Generalicemos esto para el caso de no conmutabilidad. A la función $f(E_n)$ le hacemos corresponder el operador $A = \sum_n f(E_n) E_n$ y a $g(F_n)$ el operador $B = \sum_n g(F_n) F_n$; A y B en general no son autoadjuntos. Definiremos entonces: $(f, g) = \text{Sp}AB^* / \text{Sp}L$. (B^* es el adjunto de B .) Esto será en general un número complejo, pero es fácil ver que la definición de norma coincide con la precedente y lo mismo ocurre en el

caso especial de conmutabilidad. La suma de f y g se define mediante la suma de sus operadores; $A+B$ es un operador normal y de Spur finito (pues así son A y B) y por lo tanto tiene espectro puntual puro. $f+g$ será la función que en cada multiplicidad característica de $A+B$ tiene como valor el autovalor correspondiente. Con esto ya pueden deducirse fácilmente las propiedades de este espacio de Hilbert ergódico.

i). En el espacio de Hilbert ergódico definimos ahora los operadores V_t mediante: $V_t f(E_n) = f(T_t E_n T_t^{-1})$ o también $V_t A = T_t A T_t^{-1}$; siendo A el operador definido en el parágrafo h) y T_t los varias veces mencionados operadores unitarios del espacio del sistema.

Los operadores V_t son unitarios, pues

$$SpL \cdot (V_t f, V_t g) = Sp(T_t A T_t^{-1} \cdot T_t B^* T_t^{-1}) = Sp T_t A B^* T_t^{-1} = Sp A B^* = (f, g) \cdot SpL$$

gracias a la invariancia de la medida frente a T_t .

Es inmediato que los operadores V_t tienen exactamente las mismas propiedades que los homónimos definidos en el caso clásico.

j). Llamaremos función característica de la celda E a $e(E;P)$ (P es el argumento, una proyección unidimensional), igual a uno para $P \leq E$ y a cero para $EP = 0$. Podemos entonces poner $e(E;P) = SpEP$. Nótese que el operador correspondiente a $e(E;P)$ según el parágrafo h) es E ; a la inversa, sea δ la función del espacio ergódico cuyo operador sería P ; entonces podemos escribir: $e = (e, \delta) SpL$, de modo que a menos del factor SpL , P corresponde a una verdadera función delta.

Sea $P_t = T_t P T_t^{-1}$; el teorema ergódico se enuncia entonces:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T SpEP_t dt = SpE / SpL = SpUE$$

que es la traducción directa de la fórmula de pág. 35.

Como demostración vale la misma que para el caso clásico (ver Hopf, (6) para una demostración simplificada).

k). Por supuesto para que valga la fórmula es necesario que haya tran-

atividad métrica. ¿Qué significa ésta en el espacio del sistema? La invariancia de una proyección P se expresa así: $T_t P T_t^{-1} = P$ para todo t , o sea, P reduce a T_t , de modo que cuando T_t es sólo función de H eso significa $P \nabla H$ (por hipótesis es $P \nabla U$, y $U \nabla H$, pero la conmutabilidad no tiene carácter transitivo).

Y en efecto, supongamos que K es invariante, entonces:

$$SpEP_t = SpET_t P T_t^{-1} = SpT_t P T_t^{-1} = SpEP \text{ constante, de donde}$$

$$\frac{1}{T} \int_0^T SpEP_t dt = SpEP, \text{ lo cual no tiene por qué ser igual a}$$

$SpUE$ como lo exigiría el teorema.

Asimismo, si K no fuese invariante, pero se cumpliera para todo t : $T_t P T_t^{-1} < L'$, siendo L' invariante y menor que L , podemos volver a formular el teorema ergódico reemplazando en todas partes L por L' , y llegaríamos al absurdo $SpL = SpL'$.

También hay que tener en cuenta que P puede ser una multiplicidad característica de la energía H y entonces $P_t \neq P$ para todo t y otra vez $SpEP_t$ constante, pero en rigor el teorema sólo vale en promedio cuadrático o en casi todo punto, y así como la recta tiene medida nula con respecto al plano, podemos despreciar los puntos de cualquier multiplicidad lineal menor que L . De modo que podemos pedir que las proyecciones unidimensionales P no pertenezcan a ninguna multiplicidad invariante menor que L .

En resumen, el teorema ergódico sólo vale para celdas no invariantes ni contenidas en ninguna celda invariante menor que L . Entonces no vale en general, pues en L siempre hay celdas invariantes (las formadas con autofunciones de T_t o H).

1). Sin embargo, en ningún momento hemos utilizado la conmutabilidad de U y T_t , de modo que podemos librarnos de esa condición clásica y trabajar con un U o L cualquiera. Entonces sí habrá muchos casos en que no exista en L ninguna celda invariante: aquellos en que L no contiene ninguna autofunción de H o T_t .

a).

m). Acabamos de ver que si una propiedad $E < L$ no es constante del movimiento puede aplicársele al teorema ergódico. Entonces el promedio temporal puede reemplazarse por el promedio espacial: $SpUE$, que es la fórmula que hemos aceptado como postulado, basándonos en el criterio de equiprobabilidades a priori. Estas dos fundamentalísimas hipótesis son pues equivalentes para aquellas propiedades que no son constantes del movimiento. El teorema ergódico no vale en otro caso.

Pero estudiemos más de cerca el significado de "promedio temporal". En Mecánica Cuántica esto no puede tener otro sentido que el de efectuar mediciones sucesivas sobre el sistema y promediar los resultados. Lo que hace el sistema entre dos mediciones no tiene significado físico. Supongamos entonces que se trata de verificar la propiedad E .

Si E es constante del movimiento, todas las mediciones darán el mismo resultado que la primera, de modo que el promedio temporal será uno o cero, o sea en general distinto del promedio espacial $SpUE$.

Si E no es constante, las mediciones sucesivas darán en general resultados diferentes. Supongamos que se efectúan a intervalos iguales de t segundos. Inicialmente sabemos que el sistema tiene la propiedad L . Verificamos la propiedad $E < L$. La posee con probabilidad $SpUE = SpLE / SpL$. Si el resultado es afirmativo, el sistema queda representado por el operador estadístico E / SpE , que, transcurrido el tiempo t se transforma en $E_t = T_t E T_t^{-1}$ dividido por SpE , pues el $Spur$ es invariante. Si es negativo, el nuevo op. est. es $(L-E) / Sp(L-E)$ pues $E < L$, y al cabo del tiempo t : $(L_t - E_t) / Sp(L-E)$.

Así seguimos, verificando cada t segundos si E existe o no.

Sea p_k la probabilidad de que el resultado sea afirmativo en la k -ésima medición. Entonces tomaremos como promedio temporal:

$$\lim_{r \rightarrow \infty} P_r, \text{ siendo } P_r = \frac{1}{r} \sum_{k=1}^r p_k$$

Ahora bien, como el resultado de cada medición depende sólo del de la inmediata anterior y no de las otras, esta sucesión de resultados constituye una cadena de Markov simple. Sea $n = SpE = SpE_t$ y

$N = SpL = SpL_t$. Entonces, para todo k , es $p'_k = SpE_t E/n$ la probabilidad de obtener un resultado afirmativo si el anterior había sido afirmativo y $p''_k = Sp(L_t - E_t)E/(N-n)$ si el anterior había sido negativo. Las llamaremos respectivamente p' y p'' .

Entonces p_k está dada por la fórmula (ver Uspensky(16), pág.75):

$$p_k = \frac{p''}{1+p''-p'} + \left(p_1 - \frac{p''}{1+p''-p'} \right) (p' - p'')^{k-1}$$

siendo p_1 la probabilidad inicial: $p_1 = SpLE/N$.

Pongamos $\delta = p' - p''$ y $a = \frac{p''}{1-\delta}$, o sea $p_k = a + (p_1 - a)\delta^{k-1}$

y la probabilidad media P_r está dada por (ibid.):

$$P_r = \frac{1}{r} \sum p_k = a + \frac{p_1 - a}{r} \frac{1 - \delta^r}{1 - \delta}$$

Como en este caso es $p' - p'' = \delta = \frac{SpEE_t}{n} - \frac{Sp(L_t - E_t)E}{N-n}$

$$= \frac{N}{n(N-n)} SpEE_t - \frac{1}{N-n} SpE \leq \frac{N}{n(N-n)} SpE - \frac{1}{N-n} SpE = \frac{SpE}{n} = 1$$

puede escribirse:

$$\lim_{r \rightarrow \infty} P_r = a = \frac{\frac{1}{N-n} Sp(L_t - E_t)E}{1 + \frac{Sp(L_t - E_t)E}{N-n} - \frac{SpEE_t}{n}} = \frac{Sp(L_t - E_t)E}{N-n + Sp(L_t - \frac{N}{n}E_t)E} = \frac{SpE - SpEE_t}{SpL_t - \frac{N}{n}SpEE_t}$$

$$= \frac{n(1 - \frac{1}{n}SpEE_t)}{N(1 - \frac{1}{n}SpEE_t)} = \frac{n}{N} = \frac{SpE}{SpL} = \frac{SpEL}{SpL}$$

de modo que el promedio temporal verdadero tiende al promedio espacial tal como lo establece el teorema ergódico. Esta verificación es sumamente satisfactoria, pues demuestra que no hay contradicción entre los postulados cuánticos. Pero se ha visto que la postulación que hemos adoptado es más fuerte que la hipótesis ergódica.

Aplicaciones a la Termodinámica.

El agregado estadístico se presta naturalmente a consideraciones termodinámicas, pues por definición no es otra cosa que el conjunto representativo de un sistema según el método de Gibbs.

Como sería casi imposible exponer los razonamientos generales que llevan a la deducción de la fórmula fundamental con más claridad y sencillez que en el capítulo correspondiente del libro de v. Neumann (Capítulo V-2), y como tampoco tenemos nada que agregar, tomaremos

como punto de partida la expresión que por definición llamaremos entropía del agregado de op.est. U formado por N sistemas componentes:

$$S = -Nk \text{Sp}(U \ln U)$$

donde k es la constante de Boltzmann y el logaritmo no trae dificultades pues U es definido positivo. La coincidencia con la magnitud de igual nombre definida clásicamente está demostrada en op.cit. p.200. Cuando es $U = \text{Sp}E/E$, la expresión puede simplificarse:

$$S = -Nk \text{Sp} \left(\frac{E}{\text{Sp}E} \ln \frac{E}{\text{Sp}E} \right) = Nk \cdot \ln(\text{Sp}E)$$

Se ve ahora claramente que la entropía es una función positiva, igual a cero sólo en caso que E sea unidimensional, o sea represente un estado, lo cual era de esperar pues eso representa el conocimiento máximo que puede tenerse de un sistema y por lo tanto debe corresponderle entropía mínima.

a) La entropía crece con el número de dimensiones de E o sea con nuestra ignorancia de las propiedades del sistema.

b). La mezcla de dos agregados ($U_1 = E/\text{Sp}E$; $U_2 = F/\text{Sp}F$) no disminuye la entropía, y sólo la conserva cuando $E=F$.

Sean N_1 y N_2 los números de componentes de ambos agregados y $N = N_1 + N_2$

$\alpha = N_1/N$; $\beta = N_2/N$. Se trata entonces de demostrar:

$$\frac{S}{Nk} = \alpha \ln \text{Sp}E + \beta \ln \text{Sp}F \leq -\text{Sp} \left[\left(\frac{\alpha E}{\text{Sp}E} + \frac{\beta F}{\text{Sp}F} \right) \ln \left(\frac{\alpha E}{\text{Sp}E} + \frac{\beta F}{\text{Sp}F} \right) \right] = \frac{S'}{Nk}$$

o sea que la suma de las entropías de ambos agregados aislados no es mayor que la de su mezcla.

Consideremos el caso $E \vee F$. Existe siempre una proyección G tal que $E = G + E'$; $F = G + F'$ y $E'F' = 0$ (E' y F' también proyecciones)!

Entonces:

$$\frac{S'}{Nk} = -\text{Sp} \left[G \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) + \frac{\alpha E'}{\text{Sp}E} + \frac{\beta F'}{\text{Sp}F} \right] \ln \left[G \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) + \frac{\alpha E'}{\text{Sp}E} + \frac{\beta F'}{\text{Sp}F} \right]$$

y eligiendo para calcular el Sp un sistema de autofunciones de G :

$$\frac{S'}{Nk} = - \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) \text{Sp}G \cdot \ln \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) + \frac{\alpha \text{Sp}(E-G)}{\text{Sp}E} \ln \frac{\text{Sp}E}{\text{Sp}E} + \frac{\beta \text{Sp}(F-G)}{\text{Sp}F} \ln \frac{\text{Sp}F}{\text{Sp}F}$$

Tomemos a $\text{Sp}G$ como una variable continua y derivemos:

$$\frac{1}{Nk} \frac{\partial S'}{\partial (\text{Sp}G)} = - \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) \ln \left(\frac{\alpha}{\text{Sp}E} + \frac{\beta}{\text{Sp}F} \right) + \frac{\beta}{\text{Sp}F} \ln \frac{\beta}{\text{Sp}F} + \frac{\alpha}{\text{Sp}E} \ln \frac{\alpha}{\text{Sp}E} < 0$$

Por lo tanto, dados α , β , $\text{Sp}E$ y $\text{Sp}F$, o sea, dada U , el valor mí-

nimo de S' corresponde al caso $E > F$, o $F < E$ (pues SpG es máximo cuando $G = F$ o $G = E$). No es difícil, pero sí engorroso, demostrar que lo mismo vale si E y F no conmutan y G es la mayor proyección contenida en ambos.

Supongamos, para concretar, $E > F$, o sea $G = F$; entonces:

$$\frac{S'}{Nk} = \frac{\alpha SpE'}{SpE} \ln \frac{SpE}{\alpha} - \left(\beta + \alpha \frac{SpF}{SpE} \right) \ln \left(\frac{\beta}{SpF} + \frac{\alpha}{SpE} \right)$$

Pero por una parte $-\beta \ln \left(\frac{\alpha}{SpE} + \frac{\beta}{SpF} \right) \geq \beta \ln \frac{SpF}{\alpha + \beta} = \beta \ln SpF$

(pues $SpE \geq SpF$. Hay igualdad sólo si $E = F$).

y por otra queda: $\frac{\alpha}{SpE} [SpE' \ln \frac{SpE}{\alpha} - SpF \ln \left(\frac{\alpha}{SpE} + \frac{\beta}{SpF} \right)]$

y como esto es función creciente de SpE' (recordemos que $E = F + E'$), su valor mínimo será ($E' = 0, E = F$): $\alpha \ln(SpE)$.

$\alpha \ln(SpE) + \beta \ln(SpF)$ es entonces el *minimum minimorum* de todos los valores posibles de S'/Nk , y se alcanza sólo para $E = F$.

e). Toda medición produce un aumento de entropía, salvo que se mida la misma propiedad que caracteriza al agregado.

Sea un agregado de op.est. $U = E/SpE$, y entropía $S = Nk \cdot \ln(SpE)$.

Si se verifica la propiedad F en sus componentes, sin separar ninguno (proceso 2°, pág. 33), el agregado así modificando tiene el op. est.

$$U' = \frac{SpEF}{SpE} \frac{F}{SpF} + \left(1 - \frac{SpEF}{SpE} \right) \frac{(I)-F}{Sp(I)-F}$$

Poniendo $\alpha = SpEF/SpE$ y $\beta = 1 - \alpha$; y procediendo como en el teorema anterior, hallamos ($G = 0$) que la entropía correspondiente S' es:

$$\frac{S'}{Nk} = \alpha \ln \frac{SpF}{\alpha} + \beta \ln \frac{Sp(I)-F}{\beta}$$

Fijos SpE y SpF , el valor de S' depende sólo de α , y así:

$$\frac{1}{Nk} \frac{\partial S'}{\partial \alpha} = \ln SpF - \ln Sp(I)-F + \ln \beta - \ln \alpha.$$

Pero $\frac{1}{Nk} \frac{\partial^2 S'}{\partial \alpha^2} = -\frac{1}{\beta} - \frac{1}{\alpha} < 0$, pues tanto α como β son positivos siempre. Entonces en el intervalo $0 \leq \alpha \leq 1$ la curva $S'(\alpha)$ tiene un máximo en $SpF/Sp(I)-F = \frac{\alpha}{\beta} = SpEF/Sp(I)-F$, y mínimos son los extremos del intervalo.

Si $\alpha = 1$, $S' = Nk \cdot \ln SpF$, pero como entonces $EF = E$, resulta

$SpF > SpE$ y $S' \gg S$. Si $\alpha = 0$, $S' = Nk \cdot \ln Sp(I'-F)$, pero entonces $EF = 0$, o sea $Sp(I'-F) \gg SpE$ y $S' \gg S$.

Queda así demostrado este teorema, que, nótese, es en realidad el teorema H de Boltzmann desde un punto de vista empirista (pues no se puede conocer la entropía si no se hace alguna medición). Tenemos además un criterio para averiguar, dados E y SpF , cuál es la propiedad F que más aumenta la entropía al ser verificada.

d). Agregado estadístico en equilibrio. Plantearemos el problema del siguiente modo: tenemos distintos agregados estadísticos caracterizados cada uno por la propiedad de que la energía está comprendida en el intervalo $(\lambda, \lambda+d\lambda)$, de modo que, siendo $H = \int_{-\infty}^{\infty} \lambda dE_{\lambda}$, las proyecciones correspondientes son las $dE_{\lambda} = E_{\lambda+d\lambda} - E_{\lambda}$.

Se trata ahora de mezclar estos agregados para formar otro de energía media dada, combinando las proporciones de tal manera que la entropía total resulte máxima. Como se puede suponer todo encerrado en un recipiente de volumen grande y constante, las condiciones exigidas (energía constante, entropía máxima) caracterizan el estado de equilibrio del agregado. Veamos su expresión matemática.

Llamemos $f(\lambda) \cdot Sp(dE_{\lambda})$ a la proporción en que interviene el agregado de op. est. $dE_{\lambda}/Sp(dE_{\lambda})$ en el total. Entonces el operador estadístico resultante será:

$$U = \int_{-\infty}^{\infty} f(\lambda) \cdot dE_{\lambda}, \text{ o sea } U = f(H)$$

sujeto a la condición $SpU = 1 = \int_{-\infty}^{\infty} f(\lambda) \cdot d(SpE_{\lambda})$.

Sea \bar{H} el valor medio de la energía en el agregado total; entonces:

$$\bar{H} = \text{const.} = SpUH = \int_{-\infty}^{\infty} \lambda f(\lambda) \cdot dSpE_{\lambda}$$

y por último la entropía es:

$$-S/Nk = Sp(U \cdot \ln U) = \int_{-\infty}^{\infty} f(\lambda) \cdot \ln f(\lambda) \cdot d(SpE_{\lambda})$$

Se trata de buscar el mínimo de $-S/Nk$ para SpU y $SpUH$ constantes. La variable es f , y siendo r y s dos multiplicadores de Lagrange, eso equivale a:

$$\frac{\partial}{\partial f} (f \cdot \ln f) + r \frac{\partial}{\partial f} f + s \frac{\partial}{\partial f} (\lambda f) = 0$$

para todo λ , o sea $f(\lambda) = a \cdot e^{-s\lambda}$ y $U = a \cdot e^{-sH}$, siendo $a = 1/\text{Sp}(e^{-sH})$, y estando s determinado por la condición

$$\text{Sp}UH = \text{Sp}(H \cdot e^{-sH}) / \text{Sp}(e^{-sH}) = \bar{H}$$

$\text{Sp}(e^{-sH}) \equiv Z(s)$ corresponde por definición a la "Zustandssumme" o "partition function".

La demostración que hemos dado, aunque sobre las mismas líneas que la de v. Neumann, la generaliza, pues vale para cualquier clase de espectro que tenga H .

No proseguiremos con aplicaciones termodinámicas, contentándonos con hacer notar que la entropía puede ponerse en función de $Z(s)$:

$$S = -Nk \cdot \text{Sp} \left[\frac{e^{-sH}}{Z(s)} \cdot \ln \frac{e^{-sH}}{Z(s)} \right] = Nk \left[\ln Z(s) - sZ'(s)/Z(s) \right]$$

siendo $Z'(s) = \frac{d}{ds} Z(s)$,

y de aquí pueden deducirse las demás funciones termodinámicas.

APENDICE I.

La Conjunción Estadística.

El concepto de conjunción de propiedades merece ser analizado con más detenimiento.

Si un sistema posee la propiedad E y luego se verifica en él la F con resultado positivo, su nuevo op.est. es, como sabemos, F/SpF . Pero supongamos $F > E$; antes de medir F ya sabemos que dará afirmativo, de modo que el sistema posee simultáneamente ambas propiedades; pero $EF = E$, de modo que nada debería cambiar. Pero es que si llevamos a cabo la medición y el procedimiento no está adecuadamente elegido, el sistema puede dejar de poseer con certeza la propiedad E. En efecto, si para verificar F utilizamos el procedimiento descrito en pág. 28-29 y elegimos las $\{\varphi_n\}$ de modo que no sean autofunciones de E, habremos llegado a esa situación, y nótese que ello significa aumentar la entropía del sistema. En este caso entonces no podemos hablar de conjunción en el sentido usual.

Más en general, incluso cuando E y F no conmutan puede encontrarse una multiplicidad lineal máxima contenida en ambos. Sea G su proyección; se cumple entonces $GE = G$; $GF = G$. Si al medir F utilizamos un sistema ortonormal que sea también característico de G, los sistemas que se hallan en esa parte de E (que tiene la probabilidad SpG/SpE) no serán modificados por la medición, y lo que se hallan en E-G pasarán a F-G o a $I'-F$, de modo que nuevo op.est. es:

$$U' = \frac{SpG}{SpE} \cdot \frac{G}{SpG} + \frac{Sp(E-G)F}{SpE} \cdot \frac{(F-G)}{Sp(F-G)} + \frac{Sp(I'-F)E}{SpE} \cdot \frac{(I'-F)}{Sp(I'-F)}$$

mientras que si no se puede escoger el sistema de medida el resultado es:

$$U' = \frac{SpEF}{SpE} \cdot \frac{F}{SpF} + \frac{Sp(I'-F)E}{SpE} \cdot \frac{(I'-F)}{Sp(I'-F)}$$

Se podría entonces definir una "conjunción más favorable" o "conjunción estadística" $E \times F$ mediante la siguiente igualdad:

$$E \times F \equiv \frac{SpG}{SpE} G + \frac{Sp(E-G)F}{SpE} (F-G)$$

$E \times F$ siempre representa un observable, aunque E y F no conmuten,

pero en cambio sólo es una proyección si $G=E$, o sea $F>E$, que es el caso que vimos al principio.

El segundo término es nulo siempre que $K \vee F$.

Supongamos $G=F$; entonces $K \times F = \frac{SpF}{SpK} F$, que es la expresión matemática del hecho que, medir primero K y luego F sólo da resultado positivo en una fracción SpF/SpK de los casos en que daría positivo si se midiera primero F y luego K .

Podemos decir entonces que el producto estadístico corresponde al siguiente observable: "verificar primero la propiedad K y luego la F ", pero en las condiciones más favorables, es decir, eligiendo adecuadamente el procedimiento de medida.

Por supuesto no pueda haber conmutatividad en este nuevo producto salvo en casos especiales ($EF=0$).

Este nuevo tipo de conjunción permite entonces definir magnitudes que se miden en dos etapas sucesivas sin que sea necesaria la conmutabilidad.

APENDICE II.

Funciones Características.

Hemos visto ya (pág. 13, b)) que la función característica de una magnitud \mathcal{A} en el estado φ , es el valor medio del operador $e^{is\mathcal{A}}$ (s parámetro real) , genérico de un grupo de operadores unitarios.

El estudio de este grupo es equivalente al del operador \mathcal{A} , y como se trata de operadores acotados, puede resultar ventajoso. Arnous (1) , se propone utilizar este hecho, aunque no menciona que el grupo U_s con que sustituye al operador \mathcal{A} se obtiene simplemente mediante $U_s = e^{is\mathcal{A}}$, y eso complica en general las demostraciones.

Para ilustrar esto tomemos el primer ejemplo de Arnous.

Sea D_s el operador unitario definido por $D_s f(x) = f(x-s)$.

El producto escalar

$$(D_s f, f) = \int f(x-s) \cdot \bar{f}(x) \cdot dx = F(s)$$

no es más que una convolución: $F(s) = \bar{f}(x) * f(-x)$, y por lo tanto representando con M la transformación de Fourier:

$$M^{-1}(F) = M^{-1}\bar{f}(x) \cdot M^{-1}f(-x) = |Mf|^2$$

que es la densidad de probabilidad del impulso conjugado de la coordenada x . Por lo tanto $F(s)$ es la función característica del problema.

Pero esto no debe extrañarnos, pues es bien conocida la relación entre el operador desplazamiento D_s y el operador impulso $P = \frac{h}{2\pi i} \frac{d}{dx}$

$$D_s = e^{-\frac{ih}{2\pi} s P}$$

La ventaja de sustituir las magnitudes por sus funciones características consiste en la sencillez de los operadores unitarios comparados con los autoadjuntos (por ejemplo su dominio es todo el espacio de Hilbert). Pero las proyecciones también son operadores acotados y además tienen un significado físico directo que en nuestra opinión las hace muy preferibles como instrumentos de trabajo.

O Vannan 4

BIBLIOGRAFIA CITADA.

- (1). ARNOUS E. Jour, Phys.et Rad. VIII - 87 - 1947.
 - (2). BIRKHOFF G.D. Proc.Nat.Acad.Sci. 17 - 656 - 1931.
 - (3). DIRAC P.A.M. "The Principles of Quantum Mechanics". Oxford
3° edición - 1947.
 - (4). " " Proc.Roy.Soc. A 180 - 1 - 1942 .
 - (5). FRANK PH. Inter.Enoye.of Unified Science. Tomo 3 .
 - (6). HOPF E. Proc.Nat.Acad.Sci. 18 - 93 - 1932.
 - (7). KOOPMAN B.O. Proc.Nat.Acad.Sci. '18 - 315 - 1931.
 - (8). NAGY v.Sz. "Spektraldarstellung linearer Transformationen des
Hilbertschen Raumes". Springer - 1942.
 - (9). NEUMANN J.v. "Mathematische Grundlagen der Quantenmechanik".
Springer - 1932.
 - (10). " " G5tt. Nachr. 11 - 245 - 1927.
 - (11). " " Zeits.f.Phys. 57 - 50 - 1929.
 - (12). " " Proc.Nat.Acad.Sci. 18 - 70 - 1932.
 - (13). " " Sbornik (Rec.Math.) 1 - 415 - 1936.
 - (14). REICHENBACH H. "Philosophic foundations of Quantum Mechanics".
U.of California - 1946.
 - (15). SEGAL I.E. Ann.o.Math. 48 - 930 - 1947.
 - (16). USPENSKY J.V. "Introduction to Mathematical Probability"
McGraw-Hill - 1937.
 - (17). VARSAVSKY O.A. A aparecer en Bol.Mat.
 - (18). WEYL H. "Gruppentheorie und Quantenmechanik" Hirzel - 1928.
- JULIA G. "Introduction mathématique aux théories quantiques" Paris
1928 (dos tomos).
- STONE M.H. "Linear transformations in Hilbert space" New York,1932.