Arbor

El teorema ergódico

Juan Carlos Ferrando

Arbor CLXXV, 692 (Agosto 2003), 1485-1496 pp.

1. El teorema ergódico en media para sistemas discretos

La teoría ergódica estudia las acciones de grupos de transformaciones sobre espacios medibles. B.O. Koopman había observado ya en 1931 que si (Ω, Σ, μ) es un espacio de probabilidad y $T: \Omega \to \Omega$ una transformación μ-medible que conserva la medida (es decir, tal que

$$(i)\ T^{-1}(E) \in \Sigma \text{ y } (ii)\ \mu(T^{-1}(\mathrm{E})) = \mu(E)$$

para cada $E \in \Sigma$), el hecho de que

$$\int_\Omega |f(T(x))^2|\, d\mu(x) = \int_\Omega |f(y)^2|\, d\mu\, T^{-1}(y) = \int_\Omega |f(y)^2|\, d\mu(y)$$

garantiza que el operador lineal U: $L_2(\mu)$ definido por

$$(Uf)(x) = f(T(x))$$

para cada $x \in \Omega$ es isométrico, esto es, tal que $\|Uf\|_2 = \|f\|_2$ para cada $f \in L_2(\mu)$. Si además T es biyectiva y T^{-1} conserva la medida (en cuyo caso se dice T es una transformación *invertible que conserva la medida*), el operador U es unitario. Este hecho relaciona las transformaciones que conservan la medida con los operadores unitarios en el espacio de Hilbert $L_2(\mu)$. Incidentalmente, un resultado de Krylov y Bogoliubov garantiza la existencia de medidas que se conservan bajo muchas clases de transformaciones medibles T. Este resultado afirma que si X es un espacio métrico compacto y $T: X \to X$ una aplicación continua, existe una probabilidad regular de Borel μ tal que T conserva μ. En virtud de la observación de Koopman, cabe suponer que para obtener información sobre las propiedades de la transformación T debemos considerar las propiedades de los operadores unitarios en espacios de Hilbert. En esta línea, von Neumann y Koopman obtuvieron en 1932 algunos resultados que ligaban propiedades geométricas de la transformación T con propiedades topológicas del espectro del operador U. No obstante, la consecuencia más importante de este planteamiento general es el llamado teorema ergódico, que vamos a discutir seguidamente.

Dada una transformación medible $T: \Omega \to \Omega$ que conserva la medida, los teoremas ergódicos describen el comportamiento, en casi todo punto $x \in \Omega$ o en norma, de la media aritmética de una función medible $f: \Omega \to R$ a lo largo de la órbita $orb(x) = \{T^n (x): n \in N_0\}$ de x. En un artículo publicado en 1932 von Neumann demostró el teorema ergódico en media, que afirma que si U es un operador unitario en $L_2(\mu)$ y $f \in L_2(\mu)$, la sucesión de medias aritméticas

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k f$$

donde $U^0 f := f$, converge bajo la norma de $L_2(\mu)$ a una función $f^* \in L_2(\mu)$ tal que $Uf^* := f^*$. Expresado de otro modo, si T es una transformación invertible que conserva la medida, la sucesión de funciones

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k f \tag{1.1}$$

donde $T^0 x:=x$, converge en media cuadrática a cierta función f^*de módulo al cuadrado integrable tal que $f^* o T:=f^*$. Una prueba de este resultado, debida a Frigyes Riesz, es la siguiente. Consideremos el operador lineal acotado $S = I - U$ en $L_2(\mu)$ y los subespacios cerrados $E=\overline{\mathrm{Im} S}$ y $F=\ker S$. Como $\langle Sg, h\rangle = 0$ si y sólo si $\langle g, h-U^*h\rangle = \langle g, S^*h\rangle = 0$, se deduce que $h \perp E$ si y sólo si $U^*h = h$. Pero como $U^* = U^{-1}$ por ser U unitario, concluimos que $h \perp E$ si y sólo si $h \in F$. De modo que $E \perp F$ y, en consecuencia, $L_2(\mu) = E \oplus_t F$. De aquí que si $f \in L_2(\mu)$, existen $g \in E$ y $h \in F$ con $f = g + h$. Dado $\varepsilon>0$, escogemos $\varphi \in L_2(\mu)$ tal que $\|g-S\varphi\|_2 < \varepsilon$ y ponemos $\psi = g - S\varphi$. De este modo

$$f = S\varphi + h + \psi = \varphi - U\varphi + h + \psi$$

con $\|\varphi\|_2 < \varepsilon$ y $h \in F$. Por tanto, si definimos $f_n = \frac{1}{n}\sum_{k=0}^{n-1} U_k f$ para cada $n \in N$ y sustituimos la expresión anterior de f, obtenemos

$$f_n = h + \frac{1}{n}\sum_{k=0}^{n-1} (U^k \varphi - U^{k+1} \varphi) + \frac{1}{n}\sum_{k=0}^{n-1} U^k \psi = h + \frac{\varphi - U^n \varphi}{n} + \frac{1}{n}\sum_{k=0}^{n-1} U^k \psi,$$

de donde resulta

$$\|f_n - h\|_2 \leq \frac{1}{n}\|\varphi - U^n\varphi\|_2 + \|\psi\|_2 < \frac{2}{n}\|\varphi\|_2 + \varepsilon$$

En consecuencia $f_n \to h$ bajo la norma de $L_2(\mu)$ y $h \in F$ verifica que $Uh=h$.

2. El teorema ergódico en casi todo punto

George David Birkhoff estableció casi simultáneamente (de hecho, su artículo apareció antes que el de von Neumann, en 1931) el llamado *teorema ergódico en casi todo punto*, que afirma que si $f \in L_1(\mu)$ y T es una transformación medible que conserva la medida, la sucesión (1.1) converge en casi todo punto $x \in \Omega$ a una función $f^* \in L_1(\mu)$ tal que $f^* o T \approx f^*$ y se verifica que

$$\int_\Omega f^* \, d\mu = \int_\Omega f d\mu.$$

El límite puntual

$$f^*(x) = \lim_{n\to\infty} \frac{1}{n} \{f(x) + f(T(x)) + \ldots + f(T^{n-1}(x))\}$$

se llama *media temporal* de f en x. Aunque los teoremas de Birkhoff y von Neumann fueron obtenidos de manera independiente, el resultado de von Neumann es una sencilla consecuencia del teorema de Birkhoff.

Si (Ω,Σ,μ) es un espacio de probabilidad, $T: \Omega \to \Omega$ una transformación medible que conserva la medida y $E \in \Sigma$, es obvio que $T^k(x) \in E$ si y sólo si $\chi_E(T^k(x))=1$. Puesto que

$$|\{x, T(x), \ldots, T^{n-1}(x)\} \cap E| = \sum_{k=0}^{n-1} \chi_E(T^k(x)),$$

se deduce que la probabilidad de que alguno de los n primeros elementos de la órbita de x esté contenido en E es igual a $\frac{1}{n}\sum_{i=0}^{n-1} \chi_E(T^i(x))$. Por consiguiente, si existe, $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_E(T^i(x))$ representa la *probabilidad* de que $orb(x) \cap E \neq \phi$, y la media temporal $f^*(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_E(T^i(x))$ es una función medible que en cada punto x determina la probabilidad de que la órbita del punto x intersecte a E. Para $f=\chi_E$, el teorema ergódico de Birchoff se limita a afirmar que f^* es realmente una función de $L_1(\mu)$ tal que $f^* o T \approx f^*$ y que $\int_\Omega f^* \, d\mu = \mu(E)$. Este resultado puede extenderse fácilmente para funciones simples y, por densidad, para cualquier función $f \in L_1(\mu)$. Para probar el teorema de Birkhoff en el caso particular de que $f = \chi_E$ se considera la sucesión de funciones simples $\{f_n\}$, donde

$$f_n(x) := \frac{1}{n}\sum_{k=0}^{n-1} \chi_E(T^k(x)) = \frac{1}{n}\sum_{k=0}^{n-1} \chi_{T^{-k}(E)}(x)$$

y las funciones $f^*(x) = \overline{\lim_{n\to\infty}} f_n(x)$ y $f_*(x) = \underline{\lim_{n\to\infty}} f_n(x)$. Puesto que

$$(f^* oT)(x) = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=0}^{n-1} \chi_E(T^{k+1}(x)) = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=1}^{n} \chi_E(T^k(x)) =$$

$$= \overline{\lim_{n\to\infty}} \left(\frac{n+1}{n} f_{n+1}(x) - \frac{1}{n}\chi_E(x)\right) = \overline{\lim_{n\to\infty}} f_{n+1}(x) = f^*(x)$$

en cada punto $x \in \Omega$, resulta que $f^* oT = f^*$. Se obtiene análogamente que $f_* oT = f_*$.

Para establecer que $f^* \simeq f_*$ se consideran los conjuntos medibles

$$E_{\alpha\beta} = \{x \in \Omega: f_*(x) < \alpha,\ \beta < f^*(x)\},$$

$\alpha, \beta \in R$, y se observa que

$$\{x \in \Omega: f_*(x) < f^*(x)\} = \cup \{E_{\alpha\beta}: \alpha < \beta,\ \alpha, \beta \in Q\}$$

por lo que basta probar que $\mu(E_{\alpha\beta}) = 0$ si $\alpha < \beta$. Gracias a las ecuaciones $f^* oT = f^*$ y $f_* oT = f_*$, se tiene

$$T^{-1}(E_{\alpha\beta}) = \{x \in \Omega: f_*(Tx) < \alpha,\ \beta < f^*(Tx)\} = E_{\alpha\beta}.$$

Por otro lado, dado que

$$E_{\alpha\beta} \subseteq \{x \in \Omega : \inf_k \sup_{n<k} f_n(x) > \beta\} \subseteq \{x \in \Omega : \sup_{n>1} f_n(x) > \beta\}$$

si $x \in E_{\alpha\beta}$ debe existir un $N \in \mathbf{N}$ tal que $\sum_{k=0}^{N-1} \chi_E(T^k(x)) \geq N\beta$, lo que implica

$$\sum_{k=0}^{N-1} \int_{E_{\alpha\beta}} \chi_E(T^k(x))\, d\mu(x) \geq N\beta \int_{E_{\alpha\beta}} d\mu$$

es decir,

$$\sum_{k=0}^{N-1}\int_{E_{\alpha\beta}} \chi_{T^{-k}(E)}\, d\mu\,(x) > N\,\beta \int_{E_{\alpha\beta}} d\,\mu \tag{2.1}$$

Pero como $T^{-k}(E_{\alpha\beta})=E_{\alpha\beta}$ para cada $k\in N$, resulta

$$\mu(E_{\alpha\beta} I T^{-k}(E))=\mu(T^{-k}(E_{\alpha\beta} IE))=\mu(E_{\alpha\beta} IE)$$

y, en consecuencia,

$$\sum_{k=0}^{N-1}\int_{E_{\alpha\beta}} \chi_{T^{-k}(E)}\, d\mu = \sum_{k=0}^{n-1}\int_{E_{\alpha\beta}} \chi_{E}\, d\mu = N\int_{E_{\alpha\beta}} \chi_{E}\, d\mu$$

Luego, en virtud de (2.1), se obtiene

$$\int_{E_{\alpha\beta}} \chi_{E}\, d\mu > \beta\int_{E_{\alpha\beta}} d\mu = \beta\mu(E_{\alpha\beta})$$

Haciendo $\chi_E \to -\chi^E$, $\alpha \to -\beta$, $\beta \to -a$, es obvio que $E_{\alpha\beta} \to E_{\alpha\beta}$, de donde

$$\int_{E_{\alpha\beta}} \chi_{E}\, d\mu < \alpha\mu\,(E_{\alpha\beta}).$$

De aquí que $\beta\mu(E_{\alpha\beta}) < \int_{E_{\alpha\beta}} \chi_E d\mu \le \alpha\mu(E_{\alpha\beta})$, es decir, $(\alpha-\beta)\mu(E_{\alpha\beta})\geq 0$. Luego si $\alpha<\beta$, entonces $\mu(E_{\alpha\beta})=0$. De manera que $f^*\approx f_*$, o bien

$$f^*(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{N-1}\chi_E\,(T^k(x)) = \lim_{n\to\infty} f_n\,(x)$$

en casi todo punto $x\in\Omega$. Finalmente, para probar que $f^* \in L_1(\mu)$ y $\int_\Omega f^*\, d\mu = \int_\Omega \chi_E\, d\mu$ basta notar que

$$\int_\Omega f_n\, d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\mu\,(T^{-k}(E)) = \mu\,(E)$$

debido a que $\chi_E\, oT^k = \chi_{T^{-k}(E)}$ y $\mu(T^{-k}(E))=\mu(E)$. Como $f_n\leq 1$ para cada $n\in N$, una aplicación del teorema de la convergencia dominada lleva a que

$$\int_\Omega f^*(X)\, d\mu\,(x) = \lim_{n\to\infty}\int_\Omega f_n\, d\mu = \mu\,(E) < \infty.$$

Respecto a la prueba del teorema ergódico de von Neumann, dada $f\in L_2(\mu)$ y $\varepsilon>0$ se elige $g\in L_\infty(\mu)$ tal que $\|f-g\varepsilon\|_2<\frac{\varepsilon}{4}$. Poniendo

$S_n(f)(x) := \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x)$, es fácil notar que el hecho de que $\|S_n(f)\|_2 < \|f\|_2$ implica

$$\|S_m(f)-S_n(f)\|_2 < \|S_m(f)-S_m(g_\varepsilon)\|_2+\|S_n(g_\varepsilon)-S_n(f)\|_2+\|S_m(g_\varepsilon)-S_n(g_\varepsilon)\|_2 \leq$$
$$\leq \|f-g_\varepsilon\|_2+\|S_m(g_\varepsilon)-S_n(g_\varepsilon)\|_2 \leq \frac{\varepsilon}{2} + \|S_m(g_\varepsilon)-S_n(g_\varepsilon)\|_2$$

por lo que basta demostrar que $\{S_n(g_\varepsilon)\}$ es una sucesión de Cauchy en $L_2(\mu)$ para que $\{S_n(f)\}$ sea también una sucesión de Cauchy en $L_2(\mu)$ y que, por consiguiente, exista una función $f^* \in L_2(\mu)$ tal que $S_n(f) \to f^*$ bajo la norma de $L_2(\mu)$. El hecho de que $\{S_n(g_\varepsilon)\}$ es una sucesión de Cauchy $L_2(\mu)$ en es una sencilla consecuencia del teorema de Birkhoff. En efecto, por el teorema de Birkhoff existe una función medible g^*_ε tal que $S_n(g_\varepsilon)(x) \to g^*_\varepsilon(x)$ en casi todo punto $x \in \Omega$.

Como $|S_n(g_\varepsilon)(x)-g^*_\varepsilon(x)|^2 < \|S_n(g_\varepsilon)-g^*_\varepsilon\|_\infty$ para casi todo punto $x \in \Omega$, se deduce que $g^*_\varepsilon \in L_\infty(\mu)$ y que, por tanto,

$$|S_n(g_\varepsilon)(x)-g^*_\varepsilon(x)|^2 \leq \|S_n(g_\varepsilon)-g^*_\varepsilon\|^2_\infty$$

en casi todo punto. Puesto que (Ω,Σ,μ) es un espacio de medida finita, el teorema de la convergencia dominada garantiza que $S_n(g_\varepsilon) \to g^*_\varepsilon$ en $L_2(\mu)$, de modo que $\{S_n(g_\varepsilon)\}$ es una sucesión de Cauchy en $L_2(\mu)$.

Los sistemas en los que coinciden las medias temporales $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$ y espaciales $\int_\Omega f(y)\, d\mu(y)$ se llaman ergódicos. Puede probarse que un sistema dinámico (Ω,Σ,μ,T), donde (Ω,Σ,μ) es un espacio de probabilidad y T una transformación medible que conserva la medida, es ergódico si los únicos elementos A de Σ invariantes bajo T (es decir, tales que $T^{-1}(A) = A$) son los conjuntos de medida uno o de medida cero. De hecho, si el sistema (Ω,Σ,μ,T) es ergódico, es posible demostrar que la función f^* del teorema de Birkhoff es constante en casi todo punto, por lo que

$$f^*(x) = \int_\Omega f^*(y)\, d\mu(y) = \int_\Omega f(y) d\mu(y)$$

en casi todo $x \in \Omega$. De modo que si el sistema (Ω,Σ,μ,T) es ergódico, entonces

$$lím_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^{k)}) = \int_{\Omega} f(y)\, d\,\mu\,(y) \qquad (2.2)$$

en casi todo punto $x\in\Omega$. Dicho de otro modo, si el sistema (Ω,Σ,μ,T) es ergódico entonces

$$\mu\left(\left\{x\in\Omega : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T_k(x)) \neq \int_{\Omega} f(y)\, d\,\mu\,(y)\right\}\right)= 0.$$

Notemos que, en el caso particular en que $f=\chi_E$, la ecuación (2.2) afirma que, para casi todo punto $x\in\Omega$, la probabilidad $lím_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_E (T^k(x))$ de que la órbita de x intersecte a E coincide con la medida

$$\int_{\Omega} \chi_E (y)\, d\,\mu\,(y) = \mu\,(E)$$

de E.

Si $(\Delta,2^{\Delta},\upsilon)$ denota el espacio de probabilidad $\Delta=\{0,1\}$, $\nu(\phi)=0$, $\nu(\Delta)=1$ y $\nu(\{i\})={}^1/_2$, $0\leq i\leq 1$ y $\Omega := \Delta^{N0}$ es el conjunto de Cantor, Σ la σ-álgebra sobre Ω generada por los n-cilindros

$$C_{i_1,\ldots,i_n}[\varepsilon_1,\ldots,\varepsilon_n] = \{x \in \Omega : x_{i_1} = \varepsilon_1, \ldots, x_{i_n} = \varepsilon_n\} = C_{i_1}[\varepsilon_1] \cap \ldots \cap C_{i_n}[\varepsilon_n]$$

para cada $n\in N$ y $\mu:=\otimes_{i=0}^{\infty} \upsilon$, la transformación $T : X\to X$ definida por

$$T\,(x_0, x_1, \ldots x_n, \ldots) = (x_1, x_2, \ldots\, x_{n+1}, \ldots)$$

se llama desplazamiento de Bernoulli y es un ejemplo de transformación ergódica.

3. El teorema ergódico para sistemas dinámicos continuos

El teorema ergódico tiene su origen en la mecánica estadística. Si consideramos un sistema de N puntos materiales moviéndose en una región del espacio euclídeo tridimensional bajo la acción de fuerzas conocidas, el estado x del sistema está determinado por las posiciones q_i e impulsos p_i de cada una de las N partículas en el instante t_0. A medida que el tiempo avanza, el sistema evoluciona de acuerdo con las ecuaciones diferenciales que gobiernan el movimiento, es decir, las ecuaciones de Hamilton

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i},\qquad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}$$

con condiciones iniciales $q_i(t_0)=q_{i,0}$ y $p_i(t_0)=p_{i,0}$, $1<i<3N$. Si el hamiltoniano H no depende explícitamente del tiempo, entonces H coincide con la energía total E del sistema y es, por tanto, una constante del movimiento, es decir, $H(q_i,p_i) = E$ para cada $t\geq t_0$. Si, dadas las condiciones iniciales $q_i(t_0)$, $p_i(t_0)$, $i=1,...,3N$, las ecuaciones de Hamilton tienen una solución única, dicha solución corresponde a una curva (parametrizada por el tiempo) en el llamado *espacio de las fases* $\Phi\subseteq R^{6N}$, curva que representa la historia del sistema mecánico. Si $x=x(0)$ es un punto del espacio de las fases que corresponde al estado de nuestro sistema en el instante $s = 0$ y $T_t(x)$ es el punto correspondiente a la evolución del estado del sistema en el instante t, entonces T es una transformación del espacio de las fases en sí mismo definida por $T_t(x) := x(t)$ que tiene las propiedades (i) $T_0=id_\Phi$ y (ii) $T_{t+s}=T_t \, o T_s$. Así pues, podemos considerar $\{T_t : t\geq t_0\}$ como un semigrupo continuo de transformaciones uniparamétricas en Φ. Además, puesto que el hamiltoniano es constante a lo largo de la historia del sistema, se asume que cada hipersuperficie Ω de energía constante, digamos $\Omega := H^{-1}(E)$, es globalmente invariante bajo la acción de T_t, en el sentido de que $T_t(\Omega)=(\Omega)$ para cada $t\geq 0$.

En mecánica clásica tanto T_t como los observables f del sistema (la energía, el impulso y el momento angular), son funciones diferenciables en la variedad Ω. La mecánica estadística estudia el comportamiento asintótico de los observables f de un sistema (estados de equilibrio), es decir, evalúa medias temporales tales como

$$\lim_{t\to\infty}\int_0^t f(T_s(x))\, ds$$

Dado un sistema de muchas partículas, es objetivo de la mecánica estadística mostrar que este límite, llamado límite termodinámico, existe y poder calcularlo. Ahora bien, para determinar la función $t\to T_t(x)$ hay que resolver un sistema de $6N$ ecuaciones diferenciales (las ecuaciones de Hamilton), donde N vale típicamente 10^{23}. Para evitar esto Boltzmann supuso que, a medida que $t \to +\infty$, cada punto x del espacio de las fases recorre toda la región del mismo energéticamente accesible. Esto es, que la trayectoria $x(t) = T_t(x)$, $t>0$, descrita por cada estado inicial $x\in\Phi$ de energía E pasa por todos los puntos de la superficie $\Omega := H^{-1}(E)$ (hipótesis ergódica). De esta hipótesis dedujo Bolzmann el teorema ergódico que lleva su nombre, a saber: que el valor medio temporal de la función $t \to f(x(t))$ coincide con el valor esperado (media espacial) de f en la hipersuperficie Ω, es decir,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(T_s(x))\, ds = \int_\Omega f(y)\, d\mu_\Omega(y),$$

donde μ_Ω es la medida (normalizada) de superficie correspondiente a la hipersuperficie (compacta) Ω de R^{6N}. Esta medida se llama medida de Liouville. Si r es un número real positivo fijo y A un subconjunto μ_Ω–medible de Ω, la hipótesis ergódica junto con el hecho de que $T_r(\Omega) = \Omega$ implican que $\mu_\Omega(T^{-1}{}_r(A)) = \mu_\Omega(A)$. En efecto, por el teorema ergódico de Bolzmann se cumple

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \chi_A(T_s T_r(x))\, ds = \int_\Omega \chi_A(T_r(x))\, d\mu_\Omega(T_r(x)) = \int_\Omega \chi_A(y)\, d\mu_\Omega(y)$$

ya que $T_t(\Omega) = \Omega$. Pero, por otro lado, como

$$\frac{1}{t}\int_0^t \chi_A(T_s T_r(x))\, ds = \frac{1}{t}\int_0^t \chi_A(T_r T_s(x))\, ds = \frac{1}{t}\int_0^t \chi_{T_r^{-1}(A)}(T_s x)\, ds$$

una nueva aplicación del teorema ergódico lleva a que

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \chi_A(T_s T_r(x))\, ds = \lim_{t\to\infty}\frac{1}{t}\int_0^t \chi_{T_r^{-1}(A)}(T_s(x))\, ds = \int_\Omega \chi_{T_r^{-1}(A)}(x)\, d\mu_\Omega(x)$$

Por consiguiente las transformaciones T_r conservan la medida de Liouville. Esta afirmación constituye la tesis del *Teorema de Liouville*. Vemos pues que las transformaciones que conservan la medida surgen de manera natural en mecánica estadística como consecuencia de la hipótesis ergódica.

Si Φ denota el espacio de las fases de un sistema mecánico holónomo, Ω es una hipersuperficie de energía constante y f es un observable del sistema, como se ha dicho antes fue a Koopman a quien se le ocurrió en 1931 considerar, en vez del proceso continuo $x(s) \to x(s+t)$, el operador $(U_t f)(x(s))=(f\, oT_t)x(s)=f(x(s+t))$ definido sobre funciones reales continuas de cuadrado integrable respecto a la medida de Liouville μ_Ω de Ω. Como las funciones $s \to |f(x(s))|^2$ y $s \to |f(x(s+t))|^2$ tienen la misma integral sobre Ω, el operador U_t es isométrico en términos de la norma de $L_2(\mu_\Omega)$ y admite un inverso U_{-t}. El operador U_t puede extenderse entonces a un operador unitario en todo $L_2(\mu_\Omega)$. Puesto que (*i*) $T_0 = id_\Omega$ y (*ii*) $T_{t+s} = T_t o T_s$, se deduce que $U_0 = id_{L^2(\mu\Omega)}$ y $U_{t+s} = U_t\, oU_s$, lo que nos permite aplicar el teorema de Stone y expresar cada operador unitario U_t, $-\infty<t<\infty$, en la forma (integral de Lebesgue-Stieltjes)

$$U_t = \int_{-\infty}^{+\infty} e^{it\lambda}\, dE_{\lambda.}$$

donde $\{E_\lambda : \lambda \in R\}$ es una cierta familia espectral sobre la recta real. Si ponemos

$$M(t,\lambda) := \frac{1}{t}\int_0^t e^{is\lambda}\, ds$$

$$\langle M(t)f, g\rangle := \frac{1}{t}\int_0^t \langle U_s f, g\rangle ds = \int_{-\infty}^{+\infty} M(t,\lambda) d\langle E^\lambda f, g\rangle$$

para cada $f, g \in L_2(\mu\Omega)$, entonces $M(t,\lambda) \to \delta(\lambda)$, donde $\delta(\lambda) = 0$ si $\lambda \neq 0$ y $\delta(\lambda)=1$ si $\lambda=0$. Si ponemos $E(0) = E_0 - E_0 - = E_0 - \lim_{\lambda \to 0^-} E_\lambda$ y tenemos en cuenta la propiedad fundamental de la integral de Stieltjes

$$\int_{[a,a]} f d\alpha = f(a)\,[\alpha(a^+) - \alpha(a^-)],$$

se obtiene

$$\langle E(0)f, g\rangle = \int_{-\infty}^{+\infty} \delta(\lambda)\, d\langle E(0)f, g\rangle$$

para cada $f, g \in L_2(\mu_\Omega)$. Por tanto, usando la relación

$$d\langle E_\lambda f, M(t)f - E(0)f\rangle = (M(t,\lambda) - \delta(\lambda))\, d\langle E_\lambda f, f\rangle$$

y observando que $\langle E_\lambda f, f\rangle = \langle E_\lambda^2 f, f\rangle = \|E_\lambda f\|_2^2$, resulta la ecuación

$$\| M(t)f - E(0)f\|_2^2 = \int_{-\infty}^{+\infty} | M(t,\lambda) - \delta(\lambda)|^2\, d\| E_\lambda f\|_2^2.$$

Como $|M(t,\lambda) - \delta(\lambda)| < 1$ para cada $\lambda \in R$ y $\int_{-\infty}^{+\infty} d\| E_\lambda f\|_2^2 = \langle U_0 f, f\rangle = \| f\|^2 < \infty$, el teorema de la convergencia dominada garantiza que

$$\| M(t)f - E(0)f\|_2^2 = \int_{-\infty}^{+\infty} |M(t,\lambda) - \delta(\lambda)^2|\, d\| E_\lambda f\|_2^2 \to 0$$

cuando $t \to \infty$ para cada $f \in L_2(\mu\Omega)$. Dado que $(M(t)f)(x) = t^{-1}\int_0^t (foT_{s)}(x) ds$, se deduce

$$\lim_{t\to\infty} \| \frac{1}{t}\int_0^t f o T_s ds - f^* \|_2 \to 0,$$

donde $f^* = E(0)f$. Como además $U_t f^* = U_t E(0)f = E(0)f = f^*$ gracias a que $E_\lambda E_0 = E_\lambda$ para $-\infty < \lambda \leq 0$, concluimos que $foT_t = f$. Esta es la versión «continua» del teorema ergódico en media, que fue establecida por von Neumann en su artículo de 1932.

Respecto a la versión «continua» del teorema de Birkhoff, notemos que si consideramos un semigrupo continuo de transformaciones $\{T_t: t \geq 0\}$ en vez del semigrupo discreto $\{T^n: n \in N_0\}$, el análogo «continuo» de la ecuación (2.2) es la ecuación

$$\lim_{t \to +\infty} \int_0^t f(T_s(x))ds = \int_\Omega f(y)d\mu_\Omega(y), \tag{3.1}$$

verificándose la igualdad en μ_Ω–casi todo punto $x \in \Omega$. Este es precisamente el teorema ergódico (en casi todo punto) de Boltzmann. Este resultado fue establecido de hecho por Birkhoff en su trabajo de 1931, demostrando que si (Ω, Σ, μ) es un espacio de probabilidad, $\{T_t: t \in R\}$ es una familia de transformaciones μ-medibles sobre Ω que conservan la medida y verifican que $T_t o T_s = T_{t+s}$ para cada $t, s \in R$ y $f \in L_1(R_+)$, el límite del lado izquierdo de la ecuación anterior existe para μ–casi todo punto $x \in \Omega$. Para que se verifique exactamente la ecuación (3.1) es condición necesaria y suficiente que el sistema sea *métricamente transitivo*, es decir, que los únicos conjuntos invariantes A frente a $\{T_t: t \in R_+\}$ (esto es, tales que $T_s^{-1}(A) = A$ para cada $s \geq 0$) sean los que tienen medida cero o uno. Esta consideración suscita la interesante cuestión de si los sistemas físicos son métricamente transitivos. La respuesta es negativa para la mayoría de los sistemas de la mecánica clásica, por lo que en tales sistemas la no-ergodicidad parece ser la regla.

Agradecimiento. Este trabajo sirvió de base para la conferencia del mismo título impartida por el autor en la Real Academia de Ciencias Exáctas, Físicas y Naturales de Madrid el 5 de marzo de 2003 con motivo del centenario del nacimiento de von Neumann. Deseo expresar aquí mi gratitud a la Academia por haber tenido la gentileza de invitarme.