Anal. Acad. Nac. Cs. Ex. Fís. Nat., Buenos Aires, tomo 45, 1993

EL TEOREMA ERGODICO MULTIPARAMETRICO

por Norberto A. Fava

Varias circunstancias me mueven hoy a expresar mi agradecimiento. En primer lugar deseo manifestarlo a los Señores Académicos que han debido extremar su benevolencia a la hora de decidir sobre mi incorporación a esta Academia. También quiero agradecer que de la habitual presentación haya aceptado hacerse cargo uno de los maestros más admirados de nuestro ambiente científico: al agradecer las palabras del Dr. Santaló debo confesar que nada me gustaría tanto como poder creer yo mismo que las merezco.

Muy especialmente quiero agradecer la presencia de mi amigo el Dr. Manuel Balanzat y su familia.

Vaya también mi agradecimiento a todos los colegas, amigos, parientes y familiares que han llegado hasta este lugar por razones de afecto o cortesía. Su presencia me reconforta y me alienta.

La teoría ergódica tiene su origen en una hipótesis surgida durante los comienzos de la Mecánica Estadística a fines del siglo pasado, según la cual el promedio temporal de una magnitud ligada al estado de un sistema mecánico tiende al promedio de dicha magnitud sobre todos los estados del sistema compatibles con el nivel de energía en el que se encuentra.

Como ha ocurrido muchas veces la busca de una demostración para la hipótesis física dió lugar a unos desarrollos matemáticos y aplicaciones que a primera vista no tienen una relación directa con el objeto inicial de la teoría.

Conferencia pronunciada durante su incorporación como Académico Titular, el día 18 de septiembre de 1992.

El estado de un sistema mecánico con s grados de libertad se caracteriza por un punto $x = (q, p)$ del espacio euclidiano R^{2s} llamado el *espacio de las fases*, donde $q = (q_1, q_2, \dots, q_s)$ son las coordenadas generalizadas y $p = (p_1, p_2, \dots, p_s)$ los correspondientes impulsos. La evolución del sistema en el tiempo está regida las ecuaciones de Hamilton:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (i = 1,2,\dots,s) \tag{1}$$

donde $H(q,p) = H(q_1,\dots,q_s,p_1,\dots,p_s)$ es la función de Hamilton del sistema considerado.

Una trayectoria del sistema en el espacio de las fases es una solución $x(t) = (q(t),p(t))$ del sistema de ecuaciones (1) y queda determinada por el estado inicial $x_0 = x(0)$.

A todo lo largo de cualquier trayectoria se cumple $H(x(t)) = H(q(t),p(t)) = c =$ *constante*, de manera que cualquier trayectoria real del sistema en el espacio de las fases tiene lugar sobre una superficie Σ_c caracterizada por la ecuación $H(x) = c$.

Si para cada punto $x_0 \in \Sigma_c$ ponemos $\theta_t(x_0) = x(t)$, tendremos $\theta_{t+s} = \theta_t \circ \theta_s$, con lo cual se obtiene un grupo de transformaciones $(\theta_t, t \in R)$ de la superficie Σ_c en sí misma. Por otra parte, como consecuencia del teorema de Liouville, el grupo (θ_t) preserva la medida μ definida sobre Σ_c por medio de la fórmula

$$\mu(E) = \int_E \frac{1}{\| grad\, H \|} d\sigma_c, \tag{2}$$

donde $d\sigma_c$ representa el elemento de área de la superficie Σ_c. Lo cual significa que para cualquier conjunto $E \subset \Sigma_c$ y cualquier número real t se cumple $\mu(\theta_t E) = \mu(E)$.

En estas condiciones la *hipótesis ergódica* introducida por Boltzmann puede expresarse matemáticamente diciendo que para cualquier función $f(x)$ integrable con respcto a μ sobre Σ_c se cumple

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(\theta_t x)\,dt = \frac{1}{\mu(\Sigma_c)}\int_{\Sigma_c} f\,d\mu \tag{3}$$

El primer paso hacia una demostración de esta hipótesis fue el célebre teorema de recurrencia de Poincare que data del año 1890, aún cuando las ideas intuitivas usadas en aquel trabajo debieron esperar una docena de años antes de verse legitimadas por la teoría de Lebesgue (cf. [7]).

Pero el avance hacia una solución satisfactoria debió esperar al desarrollo de la teoría general de la medida, llevado a cabo durante las primeras décadas de este siglo. Recién en 1931 el matemático norteamericano Geoge Birkhoff probó que el miembro izquierdo de (3) existe para casi todo punto x de Σ_c bajo la condición de que f sea integrable con respecto a μ.

Casi simultáneamente el matemático ruso A. I. Khinchin [6] probó que la hipótesis ergódica puede demostrarse si se admite que el grupo de transformaciones (θ_t) posee una propiedad que se conoce bajo el nombre de "ergodicidad", en tanto que B. O. Koopman y J. von Neumann demostraban la convergencia de los promedios temporales en el sentido de la norma del espacio L^2.

Desde el comienzo se fue advirtiendo que el desarrollo de esta teoría, en la que se toman promedios sobre intervalos cada vez más grandes, mostraba un asombroso paralelismo con la teoría de la derivación, en la que se consideran promedios sobre conjuntos cada vez más pequeños.

Un indicio para la explicación de este interesante fenómeno está contenido en el trabajo de F. Riesz [7], en el que muestra que el teorema ergódico de Birkhoff es consecuencia de su famoso "lema del sol naciente" usado por él mismo para estudiar la derivabilidad de funciones monótonas.

Sin embargo, la clave de la semejanza entre ambas teorías no fue revelada hasta la aparición del trabajo de Calderón [2]. En él se demuestra que el teorema fundamental de la teoría ergódica, llamado "teorema ergódico maximal" es una consecuencia de la desigualdad maximal de Hardy y Littlewood, piedra angular en la teoría de la derivación. El interés por el tema fue creciendo con el tiempo (como lo prueba la extensión de la bibliografía acumulada) por sus aplicaciones al estudio de los procesos estocásticos, los sistemas dinámicos, la teoría de la información y la geometría estocástica, en la que surge la necesidad de considerar grupos de transformaciones que dependen de varios parámetros.

Grupos multiparamétricos

Por lo visto se comprende que el marco general adecuado para el desarrollo de la teoría está formado por un espacio de medida $(X, \mathcal{M}, \mu)$ en el que actúa un grupo de transformaciones $(q_t, t \in R^k)$ dependiente de un multiparámetro $t = (t_1, t_2, ..., t_k)$. Se supone que el grupo de transformaciones *preserva* la medida μ, en el sentido de que para cualquier conjunto $E \in \mathcal{M}$ y para cualquier t se cumplen

$$\theta_t E \in \mathcal{M}, \qquad \mu(\theta_t E) = \mu(E). \tag{4}$$

En estas condiciones, el matemático norteamericano Norbert Wiener [8] probó que los promedios

$$A_\alpha f(x) = \frac{1}{|B_\alpha|}\int_{B_\alpha} f(\theta_t x)dt, \tag{5}$$

donde B_α representa la bola de radio a con centro en el origen de R^k y las barras verticales denotan la medida de Lebesgue, convergen en casi todo punto x con respecto a la medida μ cuando α tiende a infinito. La función límite

$$\hat{f}(x) = \lim_{\alpha\to\infty} A_\alpha f(x)$$

es *invariante*, lo cual significa que para cada t, $\hat{f}(\theta_t x) = \hat{f}(x)$ para casi todo x, y es integrable; en realidad, $\|\hat{f}\|_1 \le \|f\|_1$. Además, si se agrega la hipótesis de que la medida de X sea finita, entonces los promedios $A_\alpha f$

convergen hacia f también en el espacio $L^1(\mu)$. Es decir,

$$\int |A_\alpha f(x) - f(x)|\, d\mu(x) \to 0 \text{ cuando } \alpha \to \infty.$$

Este es el enunciado del teorema ergódico multiparamétrico de Norbert Wiener que está encontrando muchas aplicaciones en el estudio de la Geometría Estocástica, una de las ramas más recientes de la Matemática contemporánea.

Para que el teorema resulte útil en las aplicaciones es necesario identificar el límite $\hat{f}(x)$, al menos en el caso en que la medida $\mu(X)$ sea finita, cosa que supondremos ahora. Un conjunto $E \in \mathcal{M}$ se llama *invariante* si para cada t, $\mu(E \,\Delta\, \theta_t E) = 0$. Los conjuntos invariantes forman una σ-álgebra $\mathcal{Y} \subset \mathcal{M}$ y una función $f(x)$ es invariante si es medible con respecto a $\mathcal{Y}$.

Resulta fácilmente que para cualquier conjunto invariante E, se cumple

$$\int_E \hat{f}\, d\mu = \int_E f\, d\mu \quad (E \in \mathcal{Y}). \tag{6}$$

Lo cual muestra que en el caso de que μ sea una medida de probabilidad, $\hat{f}$ representa la esperanza condicional de f con respecto a la σ-álgebra $\mathcal{Y}$.

Un caso particularmente importante se presenta cuando el grupo (θ_t) es *ergódico*, lo cual significa que todo conjunto invariante o bien es de medida nula o bien es el complemento de un conjunto de medida nula; es decir, $\mathcal{Y}$ está incluida en la σ-álgebra generada por los conjuntos de medida nula.

En tal caso las únicas funciones invariantes son las esencialmente constantes, de donde resulta fácilmente que

$$\lim_{\alpha \to \infty} \frac{1}{|B_\alpha|} \int_{B_\alpha} f(\theta_t x)\, dt = \frac{1}{\mu(X)} \int f\, d\mu \tag{7}$$

para casi todo punto x con respecto a la medida μ.

Es natural preguntarse qué ocurre con el teorema de Wiener si se reemplaza la familia de bolas con centro en el origen y radio α por otra familia de regiones más generales.

Analicemos solamente dos ejemplos en los que las respuestas presentan aspectos muy diferentes:

1°) Consideramos una familia de regiones (D_α) dependiente de un parámetro real positivo α de modo tal que si $\alpha < \beta$, entonces $D_\alpha \subset D_\beta$ y además todo punto de R^k pertenece a alguna de dichas regiones. Entonces el teorema de Wiener subsiste para esta familia, siempre que se cumplan las siguientes condiciones:

(i) $\quad |D_\alpha - D_\alpha + D_\alpha| \leq \text{const. } |D_\alpha|$

(ii) para cada t, $\displaystyle \lim_{\alpha \to \infty} \frac{|D_\alpha \,\Delta\, (t + D_\alpha)|}{|D_\alpha|} = 0.$

La familia formada por todas las bolas B_α satisface a estas condiciones; más aún: ha sido posible probar que cualquier familia de regiones convexas también las satisface, obteniendo así una generalización natural del teorema de N. Wiener ([3] y [4]).

Por este camino que dejó abierto el trabajo de Calderón [2] resultó posible no sólo generalizar el teorema de Wiener sino también lograr una apreciable simplificación de la demostración original, de lo cual el mejor exponente es el trabajo de Becker [1].

En la demostración de estos resultados juega un papel muy importante el operador M definido por la fórmula

$$Mf(x) = \sup_{\alpha > 0} \frac{1}{|D_\alpha|} \int_{D_\alpha} |f(\theta_t x)|\; dt. \tag{8}$$

Este operador satisface las propiedades siguientes:

(M1) $\quad M(f+g) \leq Mf + Mg, \quad M(\lambda f) = |\lambda|\, Mf,$

(M2) $\quad \|Mf\|_\infty \leq \|f\|_\infty$

(M3) existe una constante C tal que para cualquier f y cualquier número $\lambda > 0$,

$$\mu\,\{x\colon Mf(x) > \lambda\} \leq \frac{C}{\lambda}\, \|f\|_1.$$

La tercera propiedad es la que se conoce bajo el nombre de "teorema ergódico maximal" y representa la clave fundamental para la demostración del teorema ergódico.

2°) Supongamos ahora que $\alpha = (\alpha_1, \alpha_2, ..., \alpha_k)$ es una k-upla de números positivos y que D_α es el intervalo de R^k definido por las relaciones

$$|t_1| \leq \alpha_1, |t_2| \leq \alpha_2, ..., |t_k| \leq \alpha_k,$$

de modo que $|D_\alpha| = 2^k \alpha_1\alpha_2...\alpha_k$.

En esas condiciones se puede probar que si $f \varepsilon L^p$ para algún $p > 1$, entonces los promedios

$$A_\alpha f(x) = \frac{1}{|D_\alpha|} \int_{D_\alpha} f(\theta_t x)\, dt$$

convergen en casi todo punto x cuando los números α_i tienden a infinito independientemente. La función límite $f(x)$ pertenece al espacio L^p y $A_\alpha f$ tiende hacia f en la norma del espacio L^p. Sin embargo, ninguna de estas afirmaciones subsiste cuando $p = 1$, es decir cuando se supone simplemente que f es integrable, como lo habíamos hecho en el caso anterior.

¿Cuál es entonces el espacio adecuado para asegurar la validez del teorema para estas regiones D_α?

La mejor respuesta que hemos podido dar a esta pregunta es que *los promedios $A_\alpha f(x)$ convergen en casi todo punto x hacia una función límite f(x) cuando los números α_i tienden a infinito independientemente, siempre que para cualquier número $\lambda > 0$ se cumpla la condición*

$$\int_{|f|>\lambda} \frac{|f|}{\lambda} \left(\log \frac{|f|}{\lambda}\right)^{k-1} d\mu < \infty.$$

Las funciones que la satisfacen forman un subespacio cerrado del espacio de Orlicz $L(log^+L)^{k-1}$ que coincide con éste cuando la medida $\mu(X)$ es finita [3].

Es interesante señalar que este enunciado amplía la validez del teorema ergódico aún en el caso clásico de un grupo de transformaciones uniparamétrico, pues en tal caso $k = 1$ y los promedios $A_\alpha f(x)$ convergen en casi todo punto siempre que para cualquier $\lambda > 0$ la función f sea integrable sobre el conjunto donde $|f| > \lambda$; propiedad que define una clase de funciones mucho más amplia que el espacio L^1.

También en este ejemplo juega un papel decisivo el operador M definido por la fórmula

$$Mf(x) = \sup_\alpha |A_\alpha f(x)|$$

cuyas propiedades son ahora esencialmente distintas, pues ya no goza de la propiedad M3 del ejemplo anterior.

Sin embargo, es posible probar la existencia de una constante C, independiente de f, tal que para cualquier $\lambda > 0$ se cumple

$$\mu\,\{x: Mf(x) > \lambda\} \leq C \int_{|f|>\lambda} \frac{|f|}{\lambda}\left(\log \frac{|f|}{\lambda}\right)^{k-1} d\mu$$

y esta desigualdad es ahora la clave de la demostración.

Uso del teorema ergódico en la geometría estocástica

El teorema ergódico multiparamétrico ha encontrado su aplicación más importante en la fundamentación de la geometría estocástica.

Imaginemos una colección infinita de figuras geométricas distribuidas sobre el plano R^2 de acuerdo con una ley de probabilidad P. Esto es en forma muy sucinta lo que llamamos un *proceso aleatorio geométrico* en el plano.

Cada disposición ω de la colección de figuras se llama una *realización* del proceso y el conjunto Ω formado por todas las disposiciones posibles es el *espacio muestral* de dicho proceso. Un evento básico E está formado por todas las realizaciones ω que verifican una propiedad determinada en una región acotada del plano.

Cada translación $x \rightarrow x + t$ del plano R^2 induce una transformación θ_t del espacio muestral Ω en sí mismo y es claro que se cumple la propiedad de grupo $\theta_{t+s} = \theta_t \circ \theta_s$.

El proceso se llama *estacionario* si las probabilidades de eventos que se refieren a regiones que resultan de transladar una misma región G son iguales. Esto se traduce en el hecho de que para cualquier evento E se cumple $P(\theta_t E) = P(E)$; es decir, el grupo (θ_t) preserva la probabilidad P. En muchos procesos importantes ocurre además que dos eventos que se refieren a regiones cada vez más alejadas son asintóticamente independientes, lo cual se traduce en el hecho de que para dos eventos cualesquiera E y F se cumple la relación llamada *condición "mixing"*:

$$\lim_{t \to \infty} P(\theta_t E \cap F) = P(E)\, P(F)$$

A su vez esta condición implica que el grupo (θ_t) es ergódico, pues si para cada t, $P(\theta_t E \,\Delta\, E) = 0$, entonces $P(E) = P(\theta_t E \cap E)$ y en consecuencia, $P(E) = P(E)^2$, de donde se sigue que $P(E) = 0$ ó $P(E) = 1$.

De lo anterior resulta, por ejemplo, que si $X(\omega)$ es una variable aleatoria con valor medio finito, entonces el promedio

$$\frac{1}{|D_r|} \int_{D_r} X(\theta_t \omega)\, dt,$$

donde D_r representa al disco de radio r con centro en el origen, converge casi seguramente al valor medio de X cuando r tiende a infinito. Y en esta forma el teorema ergódico multiparamétrico se convierte en un recurso técnico importante para el estudio de la geometría aleatoria.

REFERENCIAS

1. BECKER, M. E., Multiparameter groups of measure preserving transformations; a simple proof of Wiener's ergodic theorem, Annals of Probability, Vol. 9, No. 3 (1981), 504-509.
2. CALDERON, A. P., Ergodic theory and translation invariant operators, Proc. Nat. Acad. Sci. (U.S.A.), Vol. 59, (1968), 349-353.
3. FAVA, N. A., Weak type inequalities for product operators, Studia Math. vol. 42 (1972), 271-288.
4. -- k-parameter semigroups of measure preserving transformacions, Trans. Amer. Math. Soc., vol. 177 (1973), 345-352.
5. FAVA, N. A. y NANCLARES, J. H., Norbert Wiener's ergodic theorem for convex regions, Trans. Amer. Math. Soc. vol. 235 (1978), 403-406.
6. KHINCHIN, A. I., Mathematical Foundations of Statistical Mechanics, Dover, New York, 1949.
7. RIESZ, M., Sur la théorie ergodique, Comment. Math. Helvetici. 17 (1945), 221-239.
8. WIENER, N., The ergodic theorem, Duke Math. Journal, vol. 5 (1939), 1-18.

Manuscrito recibido en marzo de 1993