Miscelánea Matemática **75** (2022) 55-70 SMM

DOI: https://doi.org/10.47234/mm.7505

Recurrencia: análisis funcional, ergodicidad y algunas reflexiones acerca del tiempo

Rodrigo Iñigo Vargas
Sorbonne Université
Paris, Francia
rodinigo@ciencias.unam.mx

1. Introducción

Existe una frase que reza así: «las matemáticas son el arte de hacer complicado lo sencillo»[1]. Creemos, sin embargo, que aunque algunas veces es cierta, no lo es siempre. Para mostrar este hecho tomaremos como inspiración un problema de la filosofía y veremos que las matemáticas pueden transformar un problema bastante engorroso en algo más sencillo.

A finales del siglo XIX el filósofo alemán F. Nietzsche en *Así Habló Zaratustra*[2] afirmaba haber descubierto una verdad profunda: el eterno retorno. Este concepto filosófico afirma, *grosso modo*, que todas las cosas tal y como han sucedido volverán a suceder no solo una sino una infinidad de veces, de hecho ya han sucedido una infinidad de veces. ¿Es posible probar de manera rigurosa esta afirmación? Unos años más tarde el matemático francés H. Poincaré discutía un resultado que lleva hoy día su nombre y que sería probado más tarde por el matemático griego C. Carathéodory: la recurrencia de un sistema que evoluciona a lo largo del tiempo. Con dicho resultado nace una de las ramas que más desarrollo ha tenido hoy en día: la teoría ergódica. ¿Son estos resultados una prueba matemática de lo que afirmaba Nietzsche? Para poder responder procedemos a realizar un análisis más riguroso del asunto. El

Palabras clave: teoría ergódica, análisis funcional, espacios de Hilbert, proyecciones, teorema de recurrencia de Poincaré.

[1]El autor ya no recuerda de quién tomó prestada la frase pero lo que sí es seguro es que no es producto suyo.

[2]La primera vez que menciona este concepto es en *La Gaia Ciencia* (*Die fröhliche Wissenschaft*, título en alemán), publicado algunos años antes que *Así Hablo Zaratustra* (*Also sprach Zarathustra. Ein Buch für Alle und Keinen*, título en alemán), pero es en este último que lo desarrolla de manera más extensa.

objetivo es probar que, bajo ciertas hipótesis, todo sistema que evoluciona en el tiempo eventualmente se acerca tanto como queramos a su estado inicial. La herramienta principal para hacerlo es un teorema del análisis funcional. Comenzamos por definir el marco matemático en el cual se plantea el problema.

2. Espacios de Hilbert

Se supone cierta familiaridad del lector con los espacios de Banach y de Hilbert. No obstante para su comodidad recordamos algunas definiciones y resultados. Los cuales pueden ser consultados en [10].

Definición 2.1 (Producto interno)**.** Sean V un espacio vectorial, $\mathbb{K}$ el campo de los números reales o el campo de los números complejos. Un producto interno es una función $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ que satisface las siguientes propiedades:

(1) $\langle u, v \rangle = \overline{\langle v, u \rangle}$, $\forall u, v \in V$.

(2) $\langle \cdot, \cdot \rangle$ es lineal en la primera entrada.

(3) $\langle u, u \rangle > 0$, $\forall u \in V$.

Trabajaremos únicamente con espacios vectoriales sobre el campo de los reales; por lo que $\mathbb{K} = \mathbb{R}$.

Definición 2.2 (Norma)**.** Una norma es una función $|| \cdot || : V \to \mathbb{R}$ que satisface las siguientes propiedades:

(1) $||u + v|| \leq ||u|| + ||v||$, $\forall u, v \in V$.

(2) $||\lambda u|| = |\lambda| \; ||u||$, $\forall \lambda \in \mathbb{R}$, $\forall u \in V$. En donde $| \cdot |$ denota el valor absoluto.

(3) Para todo $u \in V$, si $||u|| = 0$ entonces $u = 0$.

Definición 2.3 (Espacio de Hilbert)**.** Un espacio de Hilbert H es un espacio vectorial normado, cuya norma es inducida por un producto interno y que es completo bajo esta norma. Es decir, es un espacio vectorial normado tal que para todo $x \in H$, $\|x\| = \sqrt{\langle x, x \rangle}$ y tal que cualquier sucesión de Cauchy converge en H.

En lo subsiguiente trabajaremos en un espacio de Hilbert con la topología inducida por la norma asociada al producto interno.

2.1 Ortogonalidad y proyección

Recordamos algunos resultados acerca de la ortogonalidad y proyecciones en los espacios de Hilbert que nos serán de mucha utilidad. Todos ellos son bastante conocidos; las demostraciones que no se incluyen aquí pueden ser consultadas ya sea en [10] en la sección dedicada a los espacios de Hilbert o en la misma sección sobre espacios de Hilbert en [8]. En particular, el teorema de proyección sobre un conjunto cerrado y convexo (teorema 2.6) puede ser consultado en [8, p. 189].

Definición 2.4. Sea H un espacio de Hilbert. Se dice que dos vectores $x, y \in H$ son ortogonales si $\langle x, y \rangle = 0$. Denotamos el hecho de ser ortogonales como $x \perp y$. Si $A \subset H$, definimos entonces el conjunto ortogonal a A como

$$A^{\perp} = \{x \in H | \ \forall a \in A, \ \ x \perp a\}.$$

No es difícil ver que el ortogonal de un conjunto es un subespacio vectorial cerrado.

Proposición 2.5. *$A^{\perp}$ es un subespacio vectorial cerrado de H.*

Teorema 2.6 (Proyección sobre un subespacio vectorial cerrado)**.** *Sea C un subespacio vectorial cerrado de H. Entonces para todo $x \in H$ existe un único punto dentro de C, el cual denotamos $p_C(x)$, tal que*

$$\inf_{y \in C} \|x - y\| = \|x - p_C(x)\|.$$

Observación 2.7. La función $p_C : H \to H \ x \mapsto p_C(x)$ es lineal.

Corolario 2.8 (Descomposición ortogonal)**.** *Sea F un subespacio vectorial cerrado de un espacio de Hilbert H. Entonces*

$$H = F \oplus F^{\perp} \quad y \quad x = p_F(x) + p_{F^{\perp}}(x) \quad \forall x \in H.$$

La notación $\oplus$ indica que H es la suma directa ortogonal, es decir, la suma directa de dos espacios que son ortogonales entre sí.

A continuación enunciamos un resultado técnico que nos será de utilidad en las demostraciones principales.

Proposición 2.9. *Sea A un subconjunto de un espacio de Hilbert H. Entonces $A^{\perp} = \operatorname{Span}(A)^{\perp}$ y $(A^{\perp})^{\perp} = \overline{\operatorname{Span}(A)}$.*

2.2 Operador adjunto y operador autoadjunto

Dados dos espacios de Hilbert reales $(E, \|\cdot\|_E)$ y $(F, \|\cdot\|_F)$ se considera $\mathcal{L}(E, F)$ el espacio de funciones $L : E \to F$ que son lineales y continuas. El hecho de que una función L sea continua es equivalente a que exista un número real no negativo c tal que $\|L(x)\|_F \leq c\|x\|_E$ para toda $x \in E$. Entonces se define una norma en dicho espacio de la siguiente manera:

$$\|L\|_{\mathcal{L}(E,F)} := \inf\{c \in \mathbb{R} | \ \ \|L(x)\|_F \leq c\|x\|_E\}.$$

La idea detrás de esta definición de norma es obtener el tamaño de la función L, es decir, obtener que tanto «estira» o «encoge» cada vector.

La forma más intuitiva de hacerlo es tomar el más pequeño de lo números que acotan el crecimiento de x bajo L, para toda $x \in E$; que es la definición proporcionada. Usualmente, y por comodidad, se trabaja con las siguientes definiciones equivalentes[3]

$$\|L\|_{\mathcal{L}(E,F)} := \sup_{\|x\|_E \leq 1} \|L(x)\|_F = \sup_{\|x\|_E = 1} \|L(x)\|_F.$$

A los elementos de $\mathcal{L}(E,F)$ se les llama operadores lineales y a $||\cdot||_{\mathcal{L}(E,F)}$ se le llama la norma operador. No es difícil ver que $\mathcal{L}(E,F)$ equipado con la norma $\|\cdot\|_{\mathcal{L}(E,F)}$ es un espacio vectorial normado y completo. Más aún, se tiene lo siguiente:

$$\|L(x)\|_F \leq \|L\|_{\mathcal{L}(E,F)} \|x\|_E.$$

Además si $L_1 \in \mathcal{L}(G,F)$ y $L_2 \in \mathcal{L}(E,G)$, con G un espacio de Hilbert, entonces $L_1 \circ L_2 \in \mathcal{L}(E,F)$ y

$$\|L_1 \circ L_2\|_{\mathcal{L}(E,F)} \leq \|L_1\|_{\mathcal{L}(G,F)} \ \|L_2\|_{\mathcal{L}(E,G)}.$$

En efecto, la primera desigualdad es inmediata: si $\|x\|_E = 1$ entonces por la definición de supremo:

$$\|L(x)\|_F \leq \sup_{\|y\|_E \leq 1} \|L(y)\|_F \|x\|_E = \|L\|_{\mathcal{L}(E,F)} \|x\|_E.$$

La segunda desigualdad se sigue de aplicar la primera dos veces:

$$\begin{aligned}\|(L_1 \circ L_2)(x)\|_F &\leq \|L_1\|_{\mathcal{L}(G,F)} \|L_2(x)\|_G \leq \|L_1\|_{\mathcal{L}(G,F)} \|L_2\|_{\mathcal{L}(E,G)} ||x||_E \\ &\leq \|L_1\|_{\mathcal{L}(G,F)} \|L_2\|_{\mathcal{L}(E,G)}.\end{aligned}$$

Como el supremo es la mínima cota superior se sigue el resultado.

Cuando $E = F$ se denota $\mathcal{L}(E,F)$ como $\mathcal{L}(E)$. Asimismo, para aligerar la notación, normalmente se escribe $L(x) = Lx$ para cualquier operador lineal L. En lo sucesivo se usarán las dos notaciones indistintamente. Tampoco haremos referencia al espacio sobre el cual se está tomando la norma cuando no haya motivo de confusión. Por último, notemos que se tiene dos tipos de convergencia en $\mathcal{L}(E,F)$ una es la puntual: $\{L\}_{n\in\mathbb{N}}$ converge a L puntualmente si $L_n(x)$ converge a $L(x)$ para toda $x \in E$; la otra es la convergencia en norma: $\{L\}_{n\in\mathbb{N}}$ converge a L en norma si $||L_n - L||_{\mathcal{L}(E,F)}$ converge a 0.

Observación 2.10. Una observación importante es el hecho de que $\operatorname{Ker} L$ es un subespacio vectorial cerrado de E para toda $L \in \mathcal{L}(E,F)$. En efecto, L es un operador continuo entonces $\operatorname{Ker} L = L^{-1}(\{0\})$ es cerrado.

[3]En la gran mayoría de cursos de análisis funcional la equivalencia de las normas se deja como ejercicio formativo, la prueba puede consultarse en [8, p. 92].

Teorema 2.11. *Sean H un espacio de Hilbert y $A \in \mathcal{L}(H)$. Entonces existe un único operador lineal continuo $A^* \in \mathcal{L}(H)$ tal que*

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \ \ \forall (x, y) \in H \times H.$$

Más aún, se tiene $\|A\| = \|A^*\|$.

Para la prueba referimos al lector a [8, prop. 3.4, p. 102].

Definición 2.12. Al operador A^* se le llama operador adjunto de A. Un operador $A \in \mathcal{L}(H)$ se llama autoadjunto si $A = A^*$.

Un ejemplo de operador autoadjunto es el operador identidad I. En efecto, $\langle Ix, y\rangle = \langle x, y\rangle = \langle x, Iy\rangle$ para toda x y para toda y en H. Denotaremos al operador adjunto del operador adjunto $(A^*)^*$ simplemente como A^{**}

Observación 2.13. Es fácil ver que la operación $^* : \mathcal{L}(H) \to \mathcal{L}(H)$, $A \mapsto A^*$ es aditiva, es decir, $(A + B)^* = A^* + B^*$: si $x \in H$

$$\begin{aligned}\langle x, (A^* + B^*)y\rangle &= \langle x, A^*y\rangle + \langle x, B^*y\rangle = \langle Ax, y\rangle + \langle Bx, y\rangle \\ &= \langle (A + B)x, y\rangle = \langle x, (A + B)^*y\rangle.\end{aligned}$$

Proposición 2.14. *Sea $A \in \mathcal{L}(H)$. Entonces*

$$(\operatorname{Ker} A)^{\perp} = \overline{ImA^*} \ \ y \ \ \operatorname{Ker} A = (ImA^*)^{\perp}.$$

2.3 Proyección sobre el espacio invariante de un operador

El objetivo de esta subsección es demostrar el resultado principal del cual se desprenden los resultados de la teoría ergódica.

Consideremos un espacio de Hilbert H y $A \in \mathcal{L}(H)$ tal que $\|A\|_{\mathcal{L}(H)} \leq 1$. Definimos el espacio invariante de A como $\operatorname{Ker}(I - A)$ en donde I es el operador identidad. Lo llamamos espacio invariante pues consiste en todos los vectores $x \in H$ tales que $Ax = x$.

Teorema 2.15. *Sea $A \in \mathcal{L}(H)$ tal que $\|A\|_{\mathcal{L}(H)} \leq 1$. Denotamos por $p(x)$ a la proyección ortogonal*[4] *de $x \in H$ sobre $\operatorname{Ker}(I - A)$. Entonces*

$$\forall x \in H \qquad \lim_{n\to\infty} \left(\frac{1}{n+1} \sum_{k=0}^{n} A^k \right)(x) = p(x). \tag{1}$$

Antes de comenzar la demostración observemos lo que dice el teorema: nos dice que si A es un operador de norma menor o igual a 1 entonces el promedio de las iteraciones de A en un punto x converge a la proyección ortogonal de x en el espacio invariante. Lo cual puede ser pensado en analogía con el teorema del punto fijo de Banach, aquí

[4] Existe y es única pues $\operatorname{Ker}(I - A)$ es un subespacio vectorial cerrado, por lo tanto podemos aplicar el teorema 2.6. Para no hacer más pesada la notación omitimos en este caso hacer referencia al subconjunto sobre el cual se está proyectando pero tendremos en mente que $p(x) = p_{\operatorname{Ker}(I-A)}(x)$.

el punto fijo es en realidad un subespacio fijo y A es una contracción, cuando $||A||_{\mathcal{L}(H)} < 1$ entonces se dice que A es una contracción estricta.

Demostración. La demostración procede en varios pasos.

1) Mostramos que $x \in \mathrm{Ker}(I - A)$ si y solo si $\langle Ax, x\rangle = \|x\|^2$:

$\Rightarrow$] Si $x \in \mathrm{Ker}(I - A)$ entonces $Ax = x$, así $\langle Ax, x\rangle = \langle x, x\rangle = \|x\|^2$.

$\Leftarrow$] Si $x \neq 0$, $\langle Ax, x\rangle = \|x\|^2$ utilizamos la desigualdad de Cauchy-Schwarz y el hecho de que $\|A\| \leq 1$

$$\|x\|^2 = |\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|\|x\| \leq \|x\|^2.$$

Entonces $\|Ax\|\|x\| = \|x\|^2$, *i. e.*, $\|Ax\| = \|x\|$ y como la igualdad en la desigualdad de Cauchy-Schwarz se da cuando un vector es linealmente dependiente del otro, esto es, $Ax = \lambda x$ se sigue que $\|Ax\| = |\lambda|\|x\|$, por lo tanto $\lambda = \pm 1$. Ya que $\|x\|^2 = \langle Ax, x\rangle = \lambda\|x\|^2$ entonces necesariamente $\lambda = 1$, así $Ax = x$, *i. e.*, $x \in \mathrm{Ker}(I - A)$.

2) Mostramos ahora que el espacio de invariantes de A es igual al espacio de invariantes de A^*, *i. e.* $\mathrm{Ker}(I - A^*) = \mathrm{Ker}(I - A)$. Por 1) (y ya que $\|A^*\| = \|A\| \leq 1$)

$$x \in \mathrm{Ker}(I - A) \Leftrightarrow \langle Ax, x\rangle = \|x\|^2$$
$$\Leftrightarrow \langle x, A^*x\rangle = \|x\|^2 \Leftrightarrow \langle A^*x, x\rangle = \|x\|^2 \Leftrightarrow x \in \mathrm{Ker}(I - A^*).$$

3) Mostramos que $H = \mathrm{Ker}(I - A) \oplus \overline{Im(I - A)}$. En efecto, como A es un operador continuo, $I - A$ lo es también, entonces $\mathrm{Ker}(I - A)$ es un subespacio vectorial cerrado de H, podemos aplicar el corolario 2.8 y descomponer a H como

$$H = \mathrm{Ker}(I - A) \oplus \mathrm{Ker}(I - A)^{\perp}. \tag{2}$$

Por otro lado, observamos que $(I - A)^* = I^* - A^* = I - A^*$. Por 2) se tiene $\mathrm{Ker}(I - A)^{\perp} = \mathrm{Ker}(I - A^*)^{\perp} = \mathrm{Ker}([I - A]^*)^{\perp}$. La proposición 2.14 implica que

$$\mathrm{Ker}(I - A)^{\perp} = \mathrm{Ker}([I - A]^*)^{\perp} = \overline{Im\{[(I - A)^*]^*\}} = \overline{Im(I - A)}.$$

Por lo tanto $H = \mathrm{Ker}(I - A) \oplus \overline{Im(I - A)}$.

4) Mostramos que si $x \in Im(I - A)$ entonces

$$\lim_{n\to\infty} \left(\frac{1}{n+1} \sum_{k=0}^{n} A^k \right)(x) = 0.$$

En efecto, si $x \in Im(I - A)$ existe $y \in H$ tal que $(I - A)(y) = x$ entonces $x = y - Ay$, así $A^k x = A^k y - A^{k+1} y$, de esta manera obtenemos

$$\sum_{k=0}^{n} A^k x = \sum_{k=0}^{n} [A^k y - A^{k+1} y] = A^0 y - A^{n+1} y = y - A^{n+1} y,$$

entonces

$$\left(\frac{1}{n+1}\sum_{k=0}^{n} A^k\right)(x) = \frac{1}{n+1}\left[y - A^{n+1}y\right]$$

Al tomar la norma tenemos

$$\left\|\left(\frac{1}{n+1}\sum_{k=0}^{n} A^k\right)(x)\right\| = \frac{1}{n+1}\left\|y - A^{n+1}y\right\| \leq \frac{1}{n+1}\Big(\|y\| + \|A^{n+1}y\|\Big)$$

$$\leq \frac{1}{n+1}\Big(\|y\| + \|A\|^{n+1}\|y\|\Big) \leq \frac{1}{n+1}\Big(\|y\| + 1^{n+1}\cdot\|y\|\Big)$$

$$= \frac{2\|y\|}{n+1} \to 0 \quad \text{cuando } n \to \infty.$$

Por lo tanto,

$$\lim_{n\to\infty}\left(\frac{1}{n+1}\sum_{k=0}^{n} A^k\right)(x) = 0 \quad \text{si } x \in Im(I-A).$$

5) Mostramos por último que si $x \in \overline{Im(I-A)}$ entonces

$$\lim_{n\to\infty}\left(\frac{1}{n+1}\sum_{k=0}^{n} A^k\right)(x) = 0.$$

Si $x \in \overline{Im(I-A)}$, por la definición de cerradura se tiene que para todo $\epsilon > 0$ existe $y \in Im(I-A)$ tal que $\|x-y\| < \epsilon$. Por otro lado, denotamos $S_n = \frac{1}{n+1}\sum_{k=0}^{n} A^k$, entonces

$$\|S_n\| \leq \frac{1}{n+1}\sum_{k=0}^{n}\|A^k\| \leq \frac{1}{n+1}\sum_{k=0}^{n}\|A\|^k \leq \frac{1}{n+1}\sum_{k=0}^{n} 1^k = 1.$$

Entonces

$$\begin{aligned}\|S_n(x)\| &\leq \|S_n(x-y)\| + \|S_n(y)\| \\ &\leq \|S_n\|\|(x-y)\| + \|S_n(y)\| < \epsilon + \|S_n(y)\|,\end{aligned}$$

así,

$$\limsup_{n\to\infty}\|S_n(x)\| \leq \epsilon + \limsup_{n\to\infty}\|S_n(y)\| = \epsilon + 0.$$

Por lo tanto $S_n(x) \to 0$ cuando $n \to \infty$.

Finalmente, gracias a la descomposición mostrada en 3) se tiene que cualquier x en H se puede descomponer como la suma de su proyección sobre $\operatorname{Ker}(I-A)$ y su proyección sobre $\overline{Im(I-A)}$, que denotamos $p(x)$ y $p'(x)$ respectivamente. Como $S_n(y) = y$ para todo $y \in \operatorname{Ker}(I-A)$, pues

$$S_n(y) = \frac{1}{n+1}\sum_{k=0}^{n} A^k y = \frac{1}{n+1}(n+1)y = y,$$

$p(x) \in \mathrm{Ker}(I - A)$ y $p'(x) \in \overline{Im(I - A)}$, entonces por 5.

$$S_n(x) = S_n(p(x) + p'(x)) = S_n(p(x)) + S_n(p'(x)) = p(x) + S_n(p'(x)).$$

Al tomar el límite cuando n tiende a infinito obtenemos

$$\lim_{n\to\infty} \left(\frac{1}{n+1} \sum_{k=0}^{n} A^k \right)(x) = p(x) + 0 = p(x). \qquad \square$$

3. Teoría ergódica y sistemas dinámicos

Un sistema que evoluciona con el tiempo (en este caso un tiempo discretizado) puede ser modelado mediante un espacio (los estados del sistema) y una transformación (el cambio del sistema) del espacio en sí mismo que al iterarla nos da la nueva disposición del sistema. Debido al gran número de sistemas en los que se conserva una cantidad (la energía, el volumen, etc.) resulta natural introducir alguna noción de medida para poder medir dicha cantidad. Es importante notar que se debe pensar el sistema no necesariamente como una representación exacta de lo que sucede en el espacio y en el tiempo sino una representación abstracta del fenómeno estudiado, por ejemplo: el espacio de estados de un péndulo no tiene que ser necesariamente la posición del péndulo en el espacio, puede ser los estados de energía del péndulo, por ejemplo. Los teoremas 3.3 y 3.10 y la proposición 3.9 son tomados de [3].

Definición 3.1. Sea $(X, \mathcal{F}, \mu)$ un espacio de medida, en donde $\mathcal{F}$ es una σ-álgebra de subconjuntos de X y μ una medida sobre la σ-álgebra $\mathcal{F}$. Una transformación que preserva la medida es una función medible $T : X \to X$ tal que para todo conjunto medible B, $\mu(T^{-1}(B)) = \mu(B)$. También se dice que la medida μ es invariante bajo T. Un subconjunto $B \in \mathcal{F}$ se dice que es invariante bajo T si $T^{-1}(B) = B$ y una función medible f se dice que es invariante bajo T si $f \circ T = f$.

Podemos definir, entonces, un sistema dinámico que conserva la medida (measure-preserving dynamical system).

Definición 3.2. Sean $(X, \mathcal{F}, \mu)$ un espacio de medida y $T : X \to X$ una transformación que preserva la medida μ. Se dice que $(X, \mathcal{F}, \mu, T)$ es un sistema dinámico discreto que preserva la medida.

El siguiente teorema es el teorema ergódico de von Neumann también conocido como «mean ergodic theorem». Para poder enunciarlo recordamos la definición de los espacios L^2. Consideremos un espacio de medida $(X, \mathcal{F}, \mu)$ y el conjunto $\mathbb{L}^2(X, \mu) := \{f : X \to \mathbb{R} | \int_X |f|^2 d\mu < \infty\}$. Se define la siguiente función, $|| \cdot ||_{\mathbb{L}^2(x,\mu)} : \mathbb{L}^2(X, \mu) \to [0, \infty)$,

$f \mapsto \int_X |f|^2 d\mu$. Esta función cumple todas las propiedades de una norma excepto que $||f||_{\mathbb{L}^2(X,\mu)} = 0$ no necesariamente implica que $f = 0$. Por ello se define el nuevo conjunto $L^2(X,\mu)$ como el conjunto de clases de equivalencia módulo μ, es decir, se considera que f y g pertenecen a la misma clase de equivalencia si $f = g$ salvo en un conjunto de medida cero. En este caso se define $||f||_{L^2(X,\mu)} = \int_X |f|^2 d\mu$ para toda $f \in L^2(X,\mu)$. Entonces $(L^2(X,\mu), ||\cdot||_{L^2(X,\mu)})$ es un espacio de Banach. Más aún, este es un espacio de Hilbert con el producto interno definido por $\langle f, g\rangle = \int_X fg\ d\mu$. Cuando se está trabajando con una medida μ fija, este espacio se denota simplemente como $L^2(X)$. Recordemos también que se tiene dos tipos de convergencia en $L^2(X)$; una es la convergencia puntual —llamada también convergencia casi donde sea—, es decir $\{f\}_{n\in\mathbb{N}} \subset L^2(X)$ converge casi donde sea hacia f si $f_n(x)$ converge a $f(x)$ para toda $x \in X$, salvo en un conjunto de medida cero; la otra, es la convergencia en norma $||\cdot||_{L^2(X)}$, es decir $\{f\}_{n\in\mathbb{N}} \subset L^2(X)$ converge en norma $||\cdot||_{L^2(X)}$ hacia f si $||f_n - f||_{L^2(X)}$ converge a 0. No es nuestro cometido recordar todas las implicaciones entre los distintos tipos de convergencia; para fines de este trabajo basta recordar que la convergencia en norma implica que existe una subsucesión que converge casi donde sea.

Teorema 3.3 (von Neumann (1932))**.** *Sea $(X, \mathcal{F}, \mu)$ un espacio de medida y $T : X \to X$ una transformación que preserva la medida μ. Sean $f \in L^2(X)$ y $\Gamma = \{g \in L^2(X)|\ g \circ T = g\}$ el espacio de funciones invariantes bajo T. Entonces*

(a) Γ es un subespacio vectorial cerrado de L^2, y por lo tanto $p_\Gamma(f)$ existe.

(b) Se tiene

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} f \circ T^k = p_\Gamma(f) \quad \textit{en norma } \|\cdot\|_{L^2}. \tag{3}$$

Demostración. Definimos el operador lineal continuo $A_T : L^2(X) \to L^2(X)$ mediante $A_T f = f \circ T$. La linealidad se deduce de la definición de la suma y la multiplicación por escalar en $L^2(X)$, *i. e.* si $f, g \in L^2(X)$ y $\lambda \in \mathbb{R}$

$$A_T(\lambda f + g) = (\lambda f + g) \circ T = (\lambda f) \circ T + g \circ T = \lambda A_T f + A_T g.$$

Más aún, A_T es una isometría, es decir, $\|A_T f\| = \|f\|$. En efecto, es simplemente una aplicación del teorema de cambio de variable para medidas: como T preserva la medida μ se tiene $\mu(B) = \mu(T^{-1}(B)) = T_*\mu(B)$ para todo $B \in \mathcal{F}$ en donde $T_*\mu$ es el *push-forward*[5] de μ a través

[5] También llamada medida imagen de μ bajo T, cf. [8, p. 432].

de T y como la norma en $L^2(X)$ está dada por la integral entonces

$$\begin{aligned}\|A_T f\|^2 &= \int_X |f \circ T|^2 \ d\mu = \int_{T^{-1}(X)} |f \circ T|^2 \ d\mu \\ &= \int_X |f|^2 \ dT_* \mu = \int_X |f|^2 \ d\mu = \|f\|^2,\end{aligned}$$

en donde en la segunda igualdad simplemente reescribimos X como $T^{-1}(X)$.

(a) La demostración de que es un subespacio vectorial es el mismo argumento para ver que A_T es lineal. En efecto, claramente la función constante cero es invariante bajo T; si $f, g \in \Gamma$ y $\lambda \in \mathbb{R}$ entonces

$$(\lambda f + g) \circ T = \lambda f \circ T + g \circ T = \lambda f + g,$$

i. e. $\lambda f + g$ es invariante bajo T. Ahora bien, observemos que $\Gamma = \operatorname{Ker}(I - A_T)$, como $I - A_T$ es un operador lineal continuo entonces $\Gamma = (I - A_T)^{-1}(\{0\})$ es cerrado.

(b) Como A_T es una isometría ($\|A_T\| = 1$), entonces el resultado se sigue directamente del teorema 2.15. En efecto, el lado de izquierdo de (3) se reescribe en términos del operador A_T como

$$\lim_{n \to \infty} \left(\frac{1}{n+1} \sum_{k=0}^{n} A_T^k \right) (f),$$

por el teorema 2.15 este último límite es igual a $p_\Gamma(f)$. □

Observación 3.4. El lector interesado en la teoría de la probabilidad puede notar que si μ es un medida de probabilidad entonces $p_\Gamma(f)$ es una esperanza condicional. En efecto, es posible mostrar que el conjunto de subconjuntos de X que son medibles e invariantes bajo T forman una σ-álgebra $\mathcal{I}$ (se deduce del «buen comportamiento» de la imagen inversa respecto a la intersección y unión de conjuntos). Más aún, una función g es invariante bajo T si y solo si g es medible respecto a $\mathcal{I}$, (no es difícil mostrarlo, de hecho es un buen ejercicio para el lector; la prueba puede consultarse en [3, p. 6]). Entonces $p_\Gamma(f) = E(f|\mathcal{I})$, en donde esta última denota la esperanza condicional[6] de f respecto a $\mathcal{I}$, y en este caso el resultado puede ser escrito como

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} f \circ T^k = E(f|\mathcal{I}). \tag{4}$$

Usualmente se trabaja con transformaciones denominadas ergódicas. Intuitivamente una transformación que preserva la medida es ergódica si los únicos subconjuntos medibles que deja intactos son el espacio total

[6] El concepto de esperanza condicional puede ser consultado en cualquier libro de probabilidad intermedia y/o avanzada. Sugerimos al lector [4].

y el vacío, es decir, la transformación cambia localmente el espacio pero globalmente lo conserva. La definición formal es la siguiente:

Definición 3.5. Sea $(X, \mathcal{F}, \mu)$ un espacio de probabilidad. Una transformación $T : X \to X$ que preserva la medida es ergódica si para cualquier conjunto $B \in \mathcal{F}$ que es invariante bajo T se tiene $\mu(B) = 1$ o bien $\mu(B) = 0$.

Notemos que si T es ergódica entonces los únicos subconjuntos invariantes bajo T son X y $\emptyset$, salvo conjuntos de medida cero. Así, $\mathcal{I}$ es la σ-álgebra trivial y entonces $E(f|\mathcal{I}) = E(f)$, por lo que (4) se escribe como

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} f \circ T^k = E(f) = \int_X f \ d\mu \quad \text{en norma } \|\cdot\|_{L^2}. \tag{5}$$

Aquí mencionamos de manera explícita que la convergencia es en norma para que no exista motivo de confusión: en (5) el término de la izquierda es una función y el término de la derecha es la función constante $E(f)$. Lo que estamos diciendo es que la proyección de f sobre el espacio de invariantes de T es la función constante definida por la integral de f. Es posible mostrar este hecho sin hacer referencia a la teoría de la probabilidad, lo hicimos de esta manera con fines estéticos para mostrar una interesante conexión con el lenguaje probabilístico. Este último resultado es conocido como el teorema de Birkhoff e informalmente nos dice que para tiempos muy largos el comportamiento promedio en el tiempo es igual al comportamiento promedio en el espacio.

Probamos por último el teorema de recurrencia de Poincaré. Para ello mostramos primero que p es un operador autoadjunto, y segundo que si $f \in L^2(X)$ es positiva entonces su proyección ortogonal en el conjunto de invariantes de T es también positiva.

Definición 3.6. Sea V un espacio vectorial. Se dice que una función lineal y continua $\pi : V \to V$ es una proyección si $\pi^2 = \pi$.

Observación 3.7. El operador proyección ortogonal sobre un subespacio vectorial cerrado de un espacio de Hilbert H es una proyección en el sentido de la definición anterior.

Sea $\pi : H \to H$ una proyección sobre un subespacio vectorial cerrado de H. Notemos que $H = \mathrm{Ker}(\pi) + Im(\pi)$ y $\mathrm{Ker}(\pi) \cap Im(\pi) = \{0\}$. Así, se tiene el siguiente lema.

Lema 3.8. *π es autoadjunto si y solo si la suma directa es ortogonal, es decir,* $\mathrm{Ker}(\pi)^\perp = Im(\pi)$.

Proposición 3.9. *Sea $f \in L^2(\mu)$ tal que $\mu(X) < \infty$, $T : X \to X$ una función que preserva la medida y $p_\Gamma(f)$ la proyección de f en Γ, el espacio de invariantes de T. Entonces se tiene lo siguiente:*

(a) Para toda $f \in L^2(X)$, $\int_X p_\Gamma(f)\, d\mu = \int_X f\, d\mu$.

(b) Para toda $f \in L^2(X)$ tal que $f(x) > 0$ casi en todas partes, se tiene $p_\Gamma(f)(x) > 0$ casi en todas partes.

Demostración. Primero observemos que como p_Γ es la proyección ortogonal, por el lema 3.8 se tiene $p_\Gamma = p_\Gamma^*$. Entonces, si $f, g \in L^2(X)$ y si g es invariante bajo T se sigue

$$\int_X p_\Gamma(f)\, g \;\, d\mu = \langle p_\Gamma f, g\rangle = \langle f, p_\Gamma g\rangle = \langle f, g\rangle = \int_X f\, g \;\, d\mu. \tag{6}$$

Así, como $\mu(X) < \infty$, la función constante 1 es integrable e invariante bajo T, al usar $g = 1$ obtenemos (a). Observemos además, que para cualquier conjunto medible B invariante bajo T, la función $g = 1_B$ es invariante bajo T; entonces la identidad (6) nos proporciona

$$\int_B p_\Gamma(f)\, d\mu = \int_B f\, d\mu.$$

La prueba de (b) es ligeramente más técnica: sean $N > 0$ y

$$C_N := \{x \in X | p(f)(x) \leq -1/N\}.$$

el conjunto C_N es medible respecto a $\mathcal{F}$ pues $p_\Gamma(f)$ lo es. Entonces

$$-\frac{1}{N}\mu(C_N) = \int_{C_N} -\frac{1}{N}\, d\mu \geq \int_{C_N} p(f)\, d\mu = \int_{C_N} f\, d\mu \geq 0.$$

La única manera de que esto suceda es si $\mu(C_N) = 0$ para toda $N \in \mathbb{N}$, *i. e.*, $\mu(\{x | p(f)(x) \leq 0\}) = 0$. □

Probamos ahora el teorema de recurrencia de Poincaré.

Teorema 3.10 (Recurrencia de Poincaré)**.** *Sea $(X, \mathcal{F}, \mu, T)$ un sistema dinámico discreto que preserva la medida y tal que $\mu(X) < \infty$. Sea $B \subset X$ un conjunto de medida positiva. Entonces para casi todo $x \in B$ existe una cantidad infinita de $n \in \mathbb{N}$ tales que $T^n(x) \in B$.*

Demostración. El teorema 3.3 aplicado a 1_B, que pertenece a $L^2(X)$ pues la medida de X se supone finita, nos proporciona

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} 1_B \circ T^k = p_\Gamma(1_B) \quad \text{en norma } \|\cdot\|_{L^2},$$

en donde $p_\Gamma(1_B)$ denota la proyección de 1_B en Γ, el conjunto de funciones invariantes bajo T. La convergencia en la norma L^2 implica que existe una subsucesión $\{n_j\}_{j\in\mathbb{N}}$ tal que para casi todo $x \in X$

$$\lim_{j\to\infty} \frac{1}{n_j+1} \sum_{k=0}^{n_j} (1_B \circ T^k)(x) = p_\Gamma(1_B)(x).$$

El inciso (b) de la proposición 3.9 implica que $p_\Gamma(1_B)(x) > 0$ para casi toda $x \in B$. Supongamos que existe solo una cantidad finita de $n \in \mathbb{N}$ tales que $T^n(x) \in B$, en este caso

$$0 < p_\Gamma(1_B)(x) = \lim_{j\to\infty} \frac{1}{n_j+1} \sum_{k=0}^{n_j} (1_B \circ T^k)(x) = 0.$$

lo cual es una contradicción. □

Es posible demostrar el teorema 3.10 mediante la teoría de la medida, sin utilizar el teorema 3.3; lo hicimos de esta manera con fines ilustrativos. En efecto, la prueba del teorema 3.10 usando la teoría de la medida es más corta; no obstante, se pierde en riqueza de conceptos e ideas, sobre todo del análisis funcional.

La ergodicidad juega un papel importante dentro de las matemáticas, pues proporciona una idea poderosa y sencilla al mismo tiempo, informalmente nos dice que, bajo ciertas hipótesis, un sistema termina por aproximarse a todos los estados para un tiempo largo. Las conexiones con otras áreas de las matemáticas han resultado muy fructíferas, por ejemplo con la geometría, con la teoría de números[7], entre muchas otras. Una introducción a todos estos temas puede ser consultada en [3] y en [5].

4. Algunas reflexiones acerca de los resultados

La idea principal de la recurrencia puede compararse con el principio de la caja de Dirichlet comúnmente conocido en teoría de conjuntos como el principio del palomar. Informalmente el principio dice que si hay poco espacio y mucho que albergar siempre habrá un traslapamiento. La idea es retomada en teoría de la probabilidad con el comúnmente llamado teorema del mono infinito[8]: cuando se considera un tiempo muy largo el mono logra escribir cualquier obra de Shakespeare. Un ejemplo literario

[7]Por ejemplo: el teorema de Birkhoff es usado para probar de manera muy sencilla el teorema de Borel sobre los número normales que dice en pocas palabras que casi todos los números en $[0,1)$ son normales en base 2, *i. e.* que para casi todo número en $[0,1)$ la frecuencia de 1's en su expansión binaria es 1/2. Es importante destacar que probar que un número es normal o su negación es altamente no trivial, de hecho hoy en día no se sabe si muchos números importantes como π lo son, véase [11, p. 35].

[8]Luis Rincón tiene un excelente artículo también en *Miscelanea Matemática* llamado *Sobre el problema del mono que escribe caracteres al azar* en el cual calcula el tiempo que tardaría un mono eterno en escribir las obras enteras de Shakespeare; dicha cifra es, por supuesto, más que astronómica. cf. [9]

de esta misma idea es el cuento de Jorge Luis Borges, *La Biblioteca de Babel*, aparecido en su libro *Ficciones*; en él el tiempo es representado como una biblioteca que se especula es infinita pero que nadie lo sabe de manera certera; los libros son las posibles situaciones, o las combinaciones de la materia, si se quiere pensar. Dado que hay un número finito de símbolos —un número finito de átomos— entonces en algún punto de la biblioteca los libros, que suponemos con un numero finito de páginas y un número finito de caracteres por página, se terminan y se comienzan a repetir, y así sucesivamente. En otro cuento llamado *El Inmortal*, Borges reflexiona acerca de la inmortalidad y concluye que una persona inmortal terminaría por desear solamente la muerte —aquello que no puede tener— pues vive todas las vidas posibles, termina siendo todas las personas posibles y por lo tanto la identidad individual se diluye; la vida pierde todo significado pues se cometen todos los actos loables pero también todos los actos deleznables. En resumen, un inmortal se aproximaría a todos los estados posibles, no una sino una infinidad de veces. Podemos pensar estas ideas como una misma idea: una aproximación a todos lo estados del sistema cuando se considera una cantidad extremadamente grande, ya sea de tiempo o espacio. Claro está que todos estos ejemplos (con excepción del principio del palomar, que, de hecho, puede generalizarse a conjuntos infinitos) suponen un hipotético infinito que, hasta donde se sabe, no es realizable. ¿Además qué significaría tener un tiempo sin fin?[9].

Volvamos a la cuestión de Nietzsche. La recurrencia no sucede en todos los sistemas, de hecho solo sucede en los sistemas que se denominan conservativos, los que no lo son se llaman disipativos. Hasta el día de hoy la comunidad científica sigue discutiendo qué tipo de sistema podría ser el universo o incluso si tiene sentido hacer un modelo de él con este tipo de herramientas. Otro problema es el hecho de que el tiempo que le toma a un sistema aproximarse al estado en el que inició es inversamente proporcional a la medida de dicho estado; las cifras que se obtienen son, de nuevo, más que astronómicas[10]. Consideremos esto último y supongamos que el universo es un sistema conservativo; hoy

[9]El problema del tiempo es una cuestión tratada por casi todos los filósofos, entre ellos Immanuel Kant quién postuló en *Crítica de la Razón Pura* que el tiempo no es más que un concepto de la mente que hace posible el conocimiento, es decir, el tiempo no es una entidad por sí misma física sino, como él la llama, una *intuición pura a priori*. Por otro lado, Martin Heidegger sostuvo en *Ser y Tiempo* que la pregunta *¿qué es el tiempo?* no tiene sentido, es en realidad un error lingüístico pues solo se *es* en el tiempo y por lo tanto el tiempo no puede *ser*.

[10]En [7] D. Page hace el cálculo del tiempo que tardaría en regresar a su estado inicial un hipotético agujero negro del tamaño del sol confinado en una caja y resultó ser $10^{10^{10^{76.66}}}$, ni siquiera tiene sentido poner unidades a este número, es tan grande que no hay diferencia entre medirlo en tiempos de Planck —que es el intervalo de tiempo más pequeño que puede ser medido— o milenios.

en día lo más aceptado[11] es que el destino del universo será alcanzar el equilibrio termodinámico —el cuál será su fin—[12] en un tiempo calculado aproximadamente entre 10^{100} años y $10^{32^{44}}$ milenios, lo cual es infinitamente menor a los tiempos de regreso calculados en [7] tan solo para un sistema que describe a un agujero negro: es imposible que el universo regrese a un estado anterior antes de llegar a su fin.

Quizá el lector se vio tentado a pensar que exhibiríamos una prueba matemática de lo que afirmaba Nietszche (pues así lo insinuamos en la introducción); lamentamos desilusionarlo (sí es que así fue) pero como hemos visto hay varias dificultades y, por el contrario, esto nos ayuda a dudar de su aseveración[13]. Tal parece que el tiempo es irrecuperable y nuestro destino es el olvido; quizá Proust tenía razón y *las casas, los caminos, los paseos, desgraciadamente son tan fugitivos como los años.*

5. Comentarios finales

Creemos que a pesar de todos los problemas inherentes a estas ideas (algunas bastante especulativas) estas por sí mismas sirven para fomentar la creatividad de la mente y de paso la creatividad dentro de las matemáticas que, para nosotros, es lo importante. Sin embargo, hay que saber distinguir entre lo que es matemática y lo que no lo es. Por último, hemos comprobado también que la frase mostrada en la introducción no es del todo cierta pues hemos usado las matemáticas para esclarecer un problema filosófico que, en general, son más problemáticos. En cuanto a la veracidad e implicaciones de los temas tratados, dejamos que el lector juzgue y obtenga sus propias conclusiones. No obstante, al discutir todo esto no podemos dejar de pensar en una frase —de Borges— distorsionada por nuestra opinión: El mundo afortunadamente es real[14]. Y entonces recobramos el tiempo.

---

[11]La evolución del universo se estudia a través de las distintas geometrías que puede tener, notablemente la curvatura de este. El escenario más plausible es el de un universo abierto que sigue una expansión sin fin.

[12]Decimos «fin» ya que es el único concepto temporal con el que estamos familiarizados pero, como el tiempo está intrínsecamente relacionado con la entropía, cuando el universo alcance su equilibrio termodinámico no se sabe exactamente que sucederá pero se cree que el tiempo dejará de transcurrir tal y como lo experimentamos.

[13]No es que estemos en contra del eterno retorno, de hecho poco importa, literariamente hablando, si es verdad o no; su lugar en el sistema de Nietzsche es solo como figura retórica para establecer una ética.

[14]La frase original es con la que termina su ensayo *Nueva Refutación del Tiempo* perteneciente al libro de ensayos *Inquisiciones/Otras Inquisiciones*: «El mundo desgraciadamente es real, yo desgraciadamemte soy Borges».

6. Agradecimientos

Me gustaría agradecer a Ana Meda por motivarme a escribir un artículo para Miscelánea Matemática. Aunque pasó algo de tiempo entre su invitación y el momento en el que logré formar las ideas del presente artículo, fue gracias a ella que este vio la luz. Me gustaría agradecer también a los dos revisores anónimos por sus valiosos comentarios y correcciones que ayudaron a mejorar la presentación y la claridad de este trabajo.

Bibliografía

[1] J. Borges, *Inquisiciones/otras inquisiciones*, Debolsillo, 2011.
[2] ______, *Ficciones*, Debolsillo, 2021.
[3] Y. Coudène, *Ergodic theory and dynamical systems*, Springer-Verlag, 2016.
[4] R. Durret, *Probability, theory and examples*, 5.ª ed., Cambridge University Press, 2019.
[5] D. Kerr y H. Li, *Ergodic theory*, Springer-Verlag, 2016.
[6] F. Nietzsche, *Así Habló Zaratustra*, Alianza, 2011.
[7] D. N. Page, «Information loss in black holes and/or conscious beings?», 1994.
[8] J.-P. Penot, *Analysis*, Springer-Verlag, 2016.
[9] L. Rincón, «Sobre el problema del mono que escribe caracteres al azar», *Miscelanea Matemática*, núm. 42, 2006, 79–90.
[10] W. Rudin, *Functional analysis*, 2.ª ed., McGraw-Hill, 1991.
[11] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, 1981.