



e-ISSN  
2665-2471

# PAS XIN MATEMÁTICO

 **KONRAD  
LORENZ**  
FUNDACIÓN UNIVERSITARIA

Vol. 3

No. 1

$f(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



# PASKÍN MATEMÁTICO

<https://editorial.konradlorenz.edu.co/paskin-matematico.html>

e-mail: [paskin@konradlorenz.edu.co](mailto:paskin@konradlorenz.edu.co)

## Editor

John A. Arredondo

Fundación Universitaria Konrad Lorenz  
[alexander.arredondo@konradlorenz.edu.co](mailto:alexander.arredondo@konradlorenz.edu.co)

## Comité Editorial

Carlos Alberto Díez

Fundación Universitaria Konrad Lorenz  
[carlosa.diezf@konradlorenz.edu.co](mailto:carlosa.diezf@konradlorenz.edu.co)

Julían Jiménez Cárdenas

Universidad de los Andes  
[jo.jimenezc1@uniandes.edu.co](mailto:jo.jimenezc1@uniandes.edu.co)

Camilo Ramírez Maluendas

Universidad Nacional de Colombia, Sede Manizales  
[camramirezma@unal.edu.co](mailto:camramirezma@unal.edu.co)

Andrés Mauricio Rivera

Pontificia Universidad Javeriana, Sede Cali  
[amrivera@javerianacali.edu.co](mailto:amrivera@javerianacali.edu.co)

---

Esta publicación puede ser difundida y reproducida con fines académicos y científicos por todos aquellos que tengan a bien hacer un correcto uso de su contenido.

---

ISSN 2665-2471

**Fundación Universitaria Konrad Lorenz:** Tel: (57 1) 347 23 11, Carrera 9 Bis No. 62-43 Bogotá – Colombia, email: [info@konradlorenz.edu.co](mailto:info@konradlorenz.edu.co). **Carácter académico:** Institución Universitaria. Personería Jurídica por Resolución 18537 del 4 de noviembre de 1981 del Ministerio de Educación Nacional. Institución de Educación Superior sujeta a inspección y vigilancia por el Ministerio de Educación Nacional (Art. 2.5.3.2.10.2, Decreto 1075 de 2015).



# EL GRUPO JACOBIANO DE UN GRAFO: UN AHORRO SOLIDARIO

Ruben A. Hidalgo \*

ruben.hidalgo@ufrontera.cl

## 1. Un problema de ahorro voluntario en Purgatorio.

Hace algún tiempo, en un pueblo llamado “Purgatorio”, sus habitantes habían comenzado una iniciativa de ahorro para la vejez. Esta iniciativa consistía en que cada uno de ellos tuviese una cuenta de ahorro, donde mes a mes depositaran algún porcentaje de su sueldo.

Por desgracia, José (un habitante del pueblo) un día decidió hacer una propaganda agresiva para vender un producto a los habitantes del pueblo. Muchos de estos habitantes empezaron a comprar este producto usando dineros de sus ahorros. Después de varios años, se tenía que varios habitantes tenían un saldo negativo o cero en sus cuentas. Al ver este problema, los habitantes, en una reunión, decidieron resolver esta situación de manera solidaria (con el compromiso de no volver a caer en el mismo error en el futuro). Este proceso solidario consistía en que cada habitante podía dar o solicitar a cada uno de sus amigos una misma cantidad de dinero. Ellos pensaban que con este proceso se podría lograr que cada habitante dejase de tener saldo negativo. ¿Cómo podrían estar seguros que esto es posible?

Es claro que la problemática anterior sólo se da cuando al menos una de los habitantes de Purgatorio tiene saldo negativo en su cuenta. En lo que sigue, haremos esta asunción.

## 2. La modelación matemática del problema.

Para explicar este proceso, supondremos que para cada par de habitantes hay una secuencia de amigos consecutivos. De esta manera, a cada habitante le podría llegar dinero que pertenecía a cualquier otro habitante (aunque no fuesen amigos).

Denotemos, por simplicidad, que los habitantes se llaman  $p_1, \dots, p_r$  (aquí  $r$  es el número de habitantes de Purgatorio).

### 2.1. Interpretación matemática del problema.

Formemos la matriz  $J = (a_{ij})$ , de tamaño  $r \times r$ , donde

$a_{ii}$  denota el número de amigos de  $p_i$

y, para  $i \neq j$ ,

$$a_{ij} := \begin{cases} 0, & \text{si } p_i \text{ y } p_j \text{ no son amigos} \\ -1, & \text{si } p_i \text{ y } p_j \text{ son amigos.} \end{cases}$$

Observemos que los coeficientes de la matriz  $J$  son números enteros (fuera de su diagonal son igual a 0 ó  $-1$ ) y, por la forma en que los hemos definido, la suma de cada fila y de cada columna es cero. Más aún, se tiene que  $a_{ij} = a_{ji}$  (uno dice que la matriz  $J$  es simétrica).

Supongamos que el saldo de la cuenta de la persona  $p_j$  es el entero  $s_j \in \mathbb{Z}$  (medido en las unidades monetarias de Purgatorio). Supondremos que existe al menos una persona  $p_j$  con saldo negativo, es decir,  $s_j < 0$ .

Así, podemos pensar que los saldos de las cuentas de estos habitantes están descritos por una matriz columna (de largo  $r$ )

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_r \end{pmatrix}.$$

Denotemos por  $\mathbb{Z}^r$  a la colección de todas las posibles matrices columnas de largo  $r$  con coeficientes enteros. En  $\mathbb{Z}^r$  podemos sumar usando la siguiente regla:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_r + b_r \end{pmatrix},$$

donde la suma  $a_j + b_j$  es la suma usual de números enteros. De esta manera, la suma anterior es conmutativa (es decir, no importa el orden como la hagamos). En esta suma, el neutro aditivo es

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

y el inverso aditivo de

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix}$$

\*Departamento de Matemática y Estadística, Universidad de La Frontera. Casilla 54-D, Temuco, Chile



es dado por

$$\begin{pmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_r \end{pmatrix}.$$

En la jerga matemática, estamos diciendo que  $\mathbb{Z}^r$ , con la suma antes definida, es un grupo abeliano (infinito).

Si  $x \in \mathbb{Z}^r$ , entonces denotemos por el símbolo  $\deg(x)$  la suma de sus coeficientes; luego,  $\deg(x) \in \mathbb{Z}$ . De esta manera tenemos la función grado

$$\deg : \mathbb{Z}^r \rightarrow \mathbb{Z}$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} \mapsto a_1 + \cdots + a_r.$$

Observar que  $\deg(z + w) = \deg(z) + \deg(w)$ , para cada  $z, w \in \mathbb{Z}^r$ .

Consideremos el subconjunto

$$\mathbb{Z}_0^r = \{x \in \mathbb{Z}^r : \deg(x) = 0\} \subset \mathbb{Z}^r.$$

Se puede verificar que

1. la suma de dos elementos de  $\mathbb{Z}_0^r$  sigue perteneciendo a  $\mathbb{Z}_0^r$ ,
2. el neutro aditivo también pertenece a  $\mathbb{Z}_0^r$ , y
3. los inversos aditivos de elementos de  $\mathbb{Z}_0^r$  también están allí.

Lo anterior quiere decir que  $\mathbb{Z}_0^r$  es un subgrupo de  $\mathbb{Z}^r$ .

Para cada matriz columna  $z \in \mathbb{Z}^r$  se puede verificar que  $Jz \in \mathbb{Z}_0^r$ . Luego, podemos considerar la función

$$J : \mathbb{Z}^r \rightarrow \mathbb{Z}_0^r : z \mapsto Jz.$$

Notemos que  $J(z + w) = Jz + Jw$  (esto dice que  $J$  es un homomorfismo de grupos).

Se puede verificar que la imagen  $J(\mathbb{Z}^r)$  es un subgrupo de  $\mathbb{Z}_0^r$  de índice finito; luego el grupo cociente

$$\text{Jac} = \mathbb{Z}_0^r / J(\mathbb{Z}^r)$$

es un grupo abeliano finito; este grupo se llama el grupo jacobiano. Denotaremos por el símbolo  $[x] \in \text{Jac}$  al elemento determinado por  $x \in \mathbb{Z}_0^r$ .

Para cada  $j \in \{1, 2, \dots, r\}$ , sea  $e_j \in \mathbb{Z}^r$  la matriz columna que tiene sus coeficientes igual a cero, salvo por la  $j$ -ésima coordenada que vale 1.

Si  $\alpha \in \mathbb{Z}$ , entonces  $\alpha e_j \in \mathbb{Z}^r$  denota la matriz columna donde hemos puesto en la  $j$ -ésima coordenada a  $\alpha$  (las otras coordenadas siguen siendo iguales a cero).

La matriz columna

$$s + J\alpha e_j \in \mathbb{Z}^r$$

corresponde a la matriz que entrega los saldos de las cuentas de ahorro después que la persona  $p_j$  entregó  $\alpha$  unidades monetarias a cada uno de sus amigos.

De esta manera, el problema antes mencionado corresponde a la siguiente pregunta:

**¿Existe algún  $w \in \mathbb{Z}^r$  de manera que  $s + Jw$  tenga todos sus coeficientes no-negativos?**

Como  $\deg(Jw) = 0$ , tenemos que  $\deg(s + Jw) = \deg(s)$ . Lo que estamos observando es el hecho que el proceso hace una re-distribución de los fondos de ahorro, pero la cantidad de saldo (sea negativa, cero o positiva) debe mantenerse.

La observación anterior nos asegura que si  $\deg(s) < 0$  (es decir, el sistema de ahorro está con saldo total negativo), ¡el problema planteado por el pueblo de Purgatorio no se puede resolver!

Entonces, la pregunta que queda por responder es la siguiente:

**¿Qué podemos decir si el saldo total de las cuentas de ahorro es no-negativa?**

En otras palabras, si  $\deg(s) \geq 0$ , entonces estamos buscando una matriz columna  $w \in \mathbb{Z}^r$  de manera que  $z = s + Jw$  tenga todas sus coordenadas no-negativas. Recordemos que estamos también asumiendo que alguno de los saldos  $s_j$  es negativo (de manera contraria no debemos hacer nada para obtener nuestra solución)

## 2.2. Resultados de existencia de soluciones.

Consideremos la función

$$\begin{aligned} \Theta : \mathbb{Z}^r &\longrightarrow \mathbb{Z} \times \text{Jac} \\ u &\mapsto (\deg(u), [u - \deg(u)e_1]). \end{aligned}$$

**Afirmación 2.1.** *La respuesta a nuestra pregunta original será afirmativa si y sólo si existe  $z \in \mathbb{Z}^r$ , con todas sus coordenadas no-negativas, tal que  $\Theta(s) = \Theta(z)$ .*

En efecto, si  $z \in \mathbb{Z}^r$  tiene todas sus coordenadas no-negativas y  $\Theta(s) = \Theta(z)$ , entonces

$$(\deg(s), [s - \deg(s)e_1]) = (\deg(z), [z - \deg(z)e_1]).$$

De la igualdad en las primeras coordenadas, obtenemos que

$$\deg(s) = \deg(z)$$

y de la segunda igualdad, que existe algún  $w \in \mathbb{Z}^r$  tal que

$$s - \deg(s)e_1 + Jw = z - \deg(z)e_1,$$

de donde se obtiene

$$z = s + Jw.$$

La otra dirección es fácil de verificar.

**Debemos hacer notar que, dado  $s \in \mathbb{Z}^r$ , resolver la ecuación  $\Theta(s) = \Theta(z)$  para  $z \in \mathbb{Z}^r$  (con todas sus coeficientes**

**no-negativos) no es algo fácil de realizar en general.** La siguiente observación nos dice que cuando el grupo jacobiano es trivial (es decir, cuando  $J(\mathbb{Z}^r) = \mathbb{Z}_0^r$ ), esto es fácil de ver.

**Consecuencia 2.1.** *En el caso que  $\text{Jac} = \{0\}$ , es decir,  $J : \mathbb{Z}^r \rightarrow \mathbb{Z}_0^r$  es sobreyectiva, el problema original siempre tiene solución si  $\deg(s) \geq 0$ , es decir, cuando el saldo total de ahorro no es negativo. En efecto, cualquier  $z \in \mathbb{Z}^r$  con todos sus coeficientes no-negativos y  $\deg(z) = \deg(s)$  resuelve el problema.*

En efecto, en esta situación

$$\begin{aligned} \Theta : \mathbb{Z}^r &\longrightarrow \mathbb{Z} \times \text{Jac} = \mathbb{Z} \times \{0\} \\ u &\mapsto (\deg(u), [0]). \end{aligned}$$

Luego, para todo entero  $d \geq 0$ , tenemos que si  $\deg(s) = d$ , entonces cualquier  $z \in \mathbb{Z}^r$  con coeficientes no-negativos y  $\deg(z) = d$  resuelve nuestro problema.

### Observación 2.1.

1. Si notamos bien, la situación anterior determina un grafo conexo  $\mathcal{G}$ ; cuyos vértices son las personas  $p_j$  y los ejes son determinado por relación de amistad. La función  $J$  es usualmente llamado el operador Laplaciano discreto (que en realidad se puede definir en cualquier grafo finito con una métrica sobre sus ejes).
2. El grupo jacobiano, también conocido como el grupo de Picard, grupo crítico o el grupo del dólar) había sido obtenido de manera independiente por muchas personas, por ejemplo, R. Cori, D. Rossin, M. Baker, S. Norine [BN07], N. L. Biggs [Bi93], R. Bacher, P. de la Harpe [BHN97] y T. Nagnibeda [Na97]. Una consecuencia de un Teorema debido a Kirchhoff (Kirchhoff's Matrix-Tree Theorem [BP98]) dice que el orden del grupo jacobiano coincide con el número de árboles generadores para el grafo asociado  $\mathcal{G}$ .
3. El cálculo explíto del grupo jacobiano  $\text{Jac}$  es en general no conocido, salvo para ciertos tipo de grafos; por ejemplo, para el grafo completo  $K_r$  (es decir, todas las personas son amigas entre si) se sabe que  $\text{Jac} \cong \mathbb{Z}_r^{r-2}$  [Ho14]. Recientemente I. Mednykh y A. Mednykh han calculado tal grupo jacobiano para ciertos grafos llamados circulares [Me16].

Consideremos, para cada  $d \in \mathbb{Z}$  positivo, el conjunto

$$\mathbb{Z}_d^r = \{z \in \mathbb{Z}^r : z_j \geq 0, j = 1, \dots, r, \deg(z) = d\}$$

y consideremos la función de Abel-Jacobi

$$\begin{aligned} \delta_d : \mathbb{Z}_d^r &\longrightarrow \text{Jac} \\ z &\mapsto [z - de_1] \end{aligned}$$

La **Afirmación 2.1** nos entrega la siguiente observación:

**Consecuencia 2.2.** *Si  $d = \deg(s) \geq 0$ , entonces nuestro problema original tendrá solución positiva si existe  $z \in \mathbb{Z}_d^r$  con  $\delta_d(z) = [z - de_1]$ .*

Baker y Norine obtuvieron el siguiente resultado. Sea  $e \geq 1$  el número de pares de amigos (en términos del grafo  $\mathcal{G}$  es el número de sus ejes).

**Afirmación 2.2** (Baker-Norine [BN07]). *Sea  $d \geq 0$ .*

- (i)  $\delta_d$  es sobreyectiva sí y sólo si  $d \geq g = 1 + e - r$ .
- (ii)  $\delta_d$  es inyectiva sí y sólo si  $\mathcal{G}$  es  $(d+1)$ -eje-conexo

**Consecuencia 2.3.** *Si  $s \in \mathbb{Z}^r$  y  $\deg(s) = d \geq g = 1 + e - r$ , entonces el problema original tiene solución positiva.*

## 2.3. Ejemplos: una población de tres habitantes.

### Ejemplo 1:

Consideremos una población de tres habitantes,  $p_1, p_2, p_3$ , donde  $p_2$  es amigo de todos, pero  $p_1$  y  $p_3$  no son amigos (esto representado en el grafo mostrado en la Figura 1). En este caso,

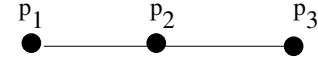
$$J_2 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$


Figura 1: El grafo del Ejemplo 1

El grupo jacobiano, en este ejemplo, es

$$\begin{aligned} \text{Jac} &= \langle x_1, x_2 : x_1 - x_2 = 0, -x_1 + 2x_2 = 0 \rangle = \\ &= \langle x_1, x_2 : x_1 = x_2, x_2 = 0 \rangle = \{0\}, \end{aligned}$$

luego, para cada ahorro  $s \in \mathbb{Z}^3$ , con  $\deg(s) \geq 0$ , tiene solución positiva a nuestro problema (ver la **Consecuencia 2.1**).

### Ejemplo 2:

Ahora, supongamos que los tres habitantes  $p_1, p_2$  y  $p_3$  son todos amigos entre si (lo que se muestra por el grafo de la figura 2). En este caso,

$$J_2 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

El grupo jacobiano ahora es

$$\text{Jac} = \langle x_1, x_2 : 2x_1 - x_2 = 0, -x_1 + 2x_2 = 0 \rangle = \langle x_1 : 3x_1 = 0 \rangle = \mathbb{Z}_3.$$

En esta situación

$$\begin{aligned} \Theta : \mathbb{Z}^3 &\longrightarrow \mathbb{Z} \times \text{Jac} = \mathbb{Z} \times \mathbb{Z}_3 \\ u &\mapsto (\deg(u), [u - \deg(u)e_1]). \end{aligned}$$



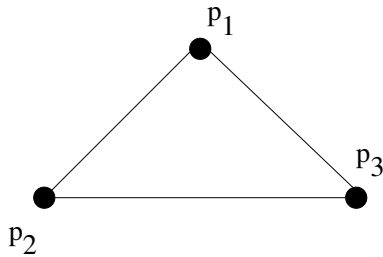


Figura 2: El grafo del Ejemplo 2

En este caso, es más difícil ver si nuestro problema tiene solución para cada

$$s = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \in \mathbb{Z}_3$$

con  $\deg(s) = s_1 + s_2 + s_3 \geq 0$ . Lo que estamos buscando es un

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{Z}^3$$

tal que

$$z_1 := s_1 + 2x_1 - x_2 - x_3 \geq 0$$

$$z_2 := s_2 - x_1 + 2x_2 - x_3 \geq 0$$

$$z_3 := s_3 - x_1 - x_2 + 2x_3 \geq 0$$

en otras palabras

$$2x_1 \geq x_2 + x_3 - s_1$$

$$2x_2 \geq x_1 + x_3 - s_2$$

$$2x_3 \geq x_1 + x_2 - s_3$$

Aquí, por simetría, podemos asumir que

$$s_1 \leq s_2 \leq s_3, \quad s_1 < 0, \quad s_3 > 0.$$

Tenemos que  $s_2$  podría ser negativo, cero o positivo, respetando las desigualdades arriba. Una bonita tarea es ver que la configuración anterior tiene solución positiva para nuestro problema.

De todas maneras, aquí vemos que ( $r = 3$ ,  $e = 3$ ), luego  $g = 1 + 3 - 3 = 1$ . Así, si  $d = s_1 + s_2 + s_3 \geq 1$ , la **Consecuencia 2.3** nos asegura que tal solución existe de manera teórica.

### Ejemplo 3:

En una sala de reuniones las sillas están ordenadas formando  $(\alpha + 1)$  filas y  $(\beta + 1)$  columnas. Supongamos que cada silla está ocupada con una persona y que cada persona puede interactuar con aquellos sentados a su derecha, su izquierda, atrás y adelante. En este caso,

$$r = (\alpha + 1)(\beta + 1), \quad e = 2\alpha\beta + \alpha + \beta, \quad g = \alpha\beta.$$

De esta manera, si el saldo total  $d \geq \alpha\beta$ , sabemos que con el proceso descrito anteriormente se puede lograr que cada persona tenga un saldo no-negativo en su cuenta de ahorros.

## Comentario Final.

Los personajes y los hechos de esta historia son ficticios y cualquier coincidencia con la realidad es sólo una mera casualidad.

## Agradecimientos.

Este artículo divulgativo ha sido parcialmente apoyado por el Proyecto Fondecyt 1190001.

## Referencias.

- [BHN97] R. Bacher, P. de la Harpe and T. Nagnibeda. The lattice of integral flows and the lattice of integral cuts on a finite graph. *Bull. Soc. Math. France* **125** (2) (1997), 167–198.
- [BN07] M. Baker and S. Norine. Riemann-Roch and Abel-Jacobi theory on a finite graph. *Advances in Mathematics* **215** (2) (2007), 766–788.
- [B65] Bapat, R.B. The Laplacian Matrix of a Graph. *Math. Student* **65** (1996), 214–223.
- [Be64] R. Becker and D. B. Glass. Cyclic critical groups of graphs. *Austral. Jour. of Comb.* **64** (2016), 366–375.
- [Bi93] N. Biggs. *Algebraic Graph Theory*. Cambridge Math. Lib., Cambridge Univ. Press, Cambridge, 1993. Second ed.
- [Bo98] B. Bollobás. *Modern Graph Theory*. Grad. Texts in Math. **184**, Springer-Verlag, New York, 1998.
- [BP98] F. Buekenhout and M. Parker. The Number of Nets of the Regular Convex Polytopes in Dimension  $\leq 4$ . *Disc. Math.* **186** (1998), 69–94.
- [Ve97] Colin de Verdiere, Y. Spectre d'opérateurs différentiels sur les graphes. In Random walks and discrete potential theory, Cortona, June 22–28 (1997), 1–26.
- [Ve98] Colin de Verdiere, Y. Spectres de graphes. Société Mathématique de France (1998).
- [Hi08] R. A. Hidalgo. Zeros of semilinear systems with applications to nonlinear partial difference equations on graphs. *Journal of Difference Equations and Applications* **14** No. 9 (2008), 953–969.
- [Ho14] T. Hoppenfeld. *The Structure of the Jacobian Group of a Graph*. Bachelor of Arts Thesis, Reed College, May 2014.
- [Lo91] D.J. Lorenzini. A finite group attached to the Laplacian of a graph. *Discrete Math.* **91** (3) (1991), 277–282.
- [Me16] A. D. Mednykh. Jacobians of circulant graphs. III International Conference “Quantum Topology”, June 23, 2016. Steklov Mathematical Institute of Russian Academy of Sciences.
- [Mo91] Mohar, B. The laplacian spectrum of graphs. In Graph Theory, Combinatorics, and Applications **2**. Ed. Y. Alavi, G. Chartrand, O. R. Oellermann, A. J. Schwenk. Wiley (1991), 871–898.
- [Na97] T. Nagnibeda. The Jacobian of a finite graph. In: Harmonic Functions on Trees and Buildings, New York, 1995. *Contemp. Math.* **206**, Amer. Math. Soc., Providence, RI, 1997, pp. 149–151.
- [Sm61] H. J. Smith. On systems of linear indeterminate equations and congruences. *Phil. Trans. R. Soc. Lond.* **151** (1) (1861), 293–326.

**Acerca del autor:** Rubén A. Hidalgo es actualmente profesor de la Universidad de La Frontera (desde 2015), universidad localizada en la Región de La Araucanía, en el sur de Chile. Su área de trabajo son las superficies de Riemann y dinámica compleja. Además, tiene como hobby la música folklórica andina y practica algunos instrumentos musicales como el charango, la guitarra y quena.

# SOLUCIONANDO SISTEMAS POLINOMIALES USANDO BASES DE GRÖBNER

Juan Sebastián Martínez Conejo<sup>\*</sup>  
juans.martinezc@konradlorenz.edu.co

## 1. Introducción.

Las ecuaciones algebraicas son un eje central de estudio en las matemáticas, y para estas existen diversos algoritmos como la eliminación Gaussiana para un tipo particular, las ecuaciones lineales, pero cuando se habla de otro tipo de ecuaciones, tal vez ecuaciones más complejas como lo son las polinomiales, puede ser muy complicado realizar algunos procesos como lo es el cálculo de sus raíces<sup>1</sup>, y a veces pueden ser incluso insolubles por medio de radicales. Una de las herramientas que ayuda a resolver este tipo de sistemas polinomiales son las bases de Gröbner. Este concepto fue introducido en el año 1965, junto con un algoritmo para calcular las raíces de un polinomio (el algoritmo de Buchberger), por Bruno Buchberger en su tesis de doctorado. El cómputo de bases de Gröbner generaliza la eliminación gaussiana, para el caso de sistemas de ecuaciones no lineales de varias variables. Al igual que la eliminación gaussiana en el caso lineal, las bases de Gröbner permiten encontrar una representación única y reducida de las soluciones de un sistema no lineal, haciendo posible listar todas sus soluciones. Este escrito se centrará en el estudio de las ecuaciones de tipo polinomial y se mostrará como las bases de Gröbner son una herramienta fundamental para abordar estos problemas.

## 2. De Gauss-Jordan a Gröbner.

Considere el siguiente sistema de ecuaciones lineal

$$\begin{cases} 3x_3 = 9, \\ x_1 + 5x_2 - 2x_3 = 2, \\ \frac{1}{3}x_1 + 2x_2 = 3. \end{cases} \quad (1)$$

Este sistema se transforma en reiteradas ocasiones para que sea sencillo de resolver, como se muestra a continuación.

Si se intercambia la fila uno con la fila tres quedaría

$$\begin{cases} \frac{1}{3}x_1 + 2x_2 = 3, \\ x_1 + 5x_2 - 2x_3 = 2, \\ 3x_3 = 9. \end{cases} \quad (2)$$

<sup>\*</sup>Egresado Departamento de Matemáticas, Fundación Universitaria Konrad Lorenz, Bogotá-Colombia

<sup>1</sup>Las raíces de un polinomio son números tales que hacen que un polinomio valga cero

Ahora, al multiplicar la fila uno por tres,

$$\begin{cases} x_1 + 6x_2 = 9, \\ x_1 + 5x_2 - 2x_3 = 2, \\ 3x_3 = 9. \end{cases} \quad (3)$$

Por ultimo, se suma menos un veces la fila uno a la fila dos,

$$\begin{cases} x_1 + 6x_2 = 9, \\ x_2 + 2x_3 = 7, \\ 3x_3 = 9. \end{cases} \quad (4)$$

Ahora se puede encontrar el valor de cada variable. La ecuación inferior muestra que  $x_3 = 3$ . Sustituir 3 por  $x_3$  en la ecuación central muestra que  $x_2 = 1$ . Sustituir esos dos en la ecuación superior da que  $x_1 = 3$  y así el sistema tiene una solución única, el conjunto de soluciones es  $\{(3, 1, 3)\}$ . Pero estas transformaciones que se realizaron se pueden hacer con total confianza por el método de Gauss [JH17], teorema que garantiza que al hacer estas transformaciones el conjunto de soluciones va a ser el mismo.

**Teorema 2.1** (Método de Gauss-Jordan). *Si un sistema lineal se transforma en otro aplicando una de estas operaciones*

1. Una ecuación o fila es cambiada por otra.
2. Una ecuación se multiplica en ambos lados por una constante distinta de cero.
3. Una ecuación es reemplazada por la suma de sí misma y un múltiplo de otra.

Entonces los dos sistemas tienen el mismo conjunto de soluciones.

¿Será posible extender este proceso cuando las ecuaciones no son lineales sino polinomiales? Para responder esto, primero se necesitan algunos conceptos que se introducen a continuación. Para un tratamiento detallado de los conceptos expuestos a continuación consultar [Bec93], [JH17] y [WDaGP13].

**Definición 2.1.** Se denomina **monomio** a un producto de una colección de variables  $x_1, x_2, \dots, x_n$  de la forma

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

donde  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}$ .

**Definición 2.2.** Sea  $K$  un cuerpo<sup>2</sup>. Un **polinomio**  $f$  en las variables  $x_1, x_2, \dots, x_n$  es una combinación lineal de un número finito de monomios con coeficientes en  $K$  y se representa mediante

$$f = \sum_{\alpha} c_{\alpha} x^{\alpha}, \quad c_{\alpha} \in K. \quad (5)$$

El conjunto de polinomios con coeficientes en  $K$  es un anillo conmutativo con elemento identidad y se denota por

$$K[x_1, x_2, \dots, x_n]. \quad (6)$$

Consideremos el sistema de ecuaciones homogéneo

$$\begin{cases} x^2 - y^2 - 3 = 0, \\ 2x^2 + 3y^2 - 11 = 0. \end{cases} \quad (7)$$

El método de Gauss-Jordan solo aplica para sistemas de ecuaciones lineales, entonces debemos convertir el sistema 7 en un sistema lineal. Esto lo hacemos con un cambio de variable, es decir, hagamos  $x^2 = s$  y  $y^2 = t$

$$\begin{cases} s - t - 3 = 0, \\ 2s + 3t - 11 = 0, \end{cases} \quad (8)$$

con lo cual este sistema tiene la forma matricial

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 3 \\ 11 \end{pmatrix}. \quad (9)$$

Aplicando Gauss-Jordan se obtiene

$$\begin{cases} s - t - 3 = 0, \\ -5t + 5 = 0. \end{cases} \rightarrow \begin{cases} x^2 - y^2 - 3 = 0, \\ -5y^2 + 5 = 0. \end{cases} \quad (10)$$

Por lo que se puede concluir que  $y = \pm 1$  y  $x = \pm 2$ , es decir, la solución son los puntos  $(2, 1), (2, -1), (-2, 1), (-2, -1)$  que es la solución del sistema (7). En la Figura 1 se representa esta situación.

De lo anterior surge la pregunta: ¿siempre se podrá convertir un sistema de ecuaciones polinomiales en uno lineal, para usar el método de Gauss-Jordan? Y la respuesta es negativa en general. Este hecho se puede ver en el siguiente ejemplo:

Considere el sistema de ecuaciones

$$\begin{cases} x^2 - y^3 - 3 = 0, \\ 2x^4 + 3y^9 - 11 = 0. \end{cases} \quad (11)$$

Suponga que  $s = x^2$  y  $t = y^3$ , por lo cual volviendo a expresar la ecuación (11) quedaría de la siguiente manera

$$\begin{cases} s - t - 3 = 0, \\ 2s^2 + 3t^3 - 11 = 0. \end{cases} \quad (12)$$

Note que no queda un sistema lineal, por lo cual no se puede hacer uso del método de Gauss, incluso, en algunas ocasiones no será posible hacer ningún tipo de sustitución, como en el

<sup>2</sup>En matemáticas, particularmente en álgebra moderna, un grupo es una estructura algebraica formada por un conjunto  $A$  no vacío y una ley de composición interna.

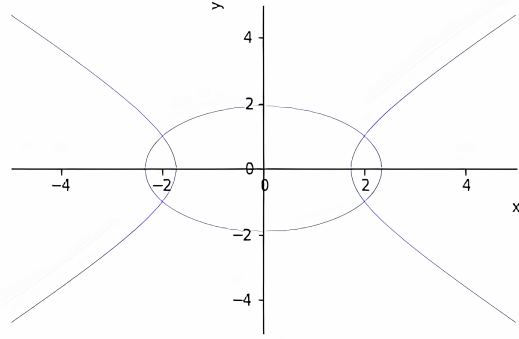


Figura 1: Note que la intersección de las cónicas es la solución del sistema (7).

siguiente caso:

Considere el sistema de ecuaciones no lineal

$$\begin{cases} x^5 - y^2 - 3 = 0, \\ 2x^7 + 3y^3 - 11 = 0. \end{cases} \quad (13)$$

Para el cual no es posible hacer ninguna sustitución, aun más, si se intentara resolver por los métodos tradicionales, quedaría una potencia alta<sup>3</sup>, por lo que en principio no es posible resolverlo analíticamente, esto se sabe gracias al teorema de Galois [Ste90] que da un criterio para saber cuándo una ecuación polinómica en una variable es resoluble o no por radicales.

Esto deja la duda de como se deben manejar los casos donde no se puede hacer una sustitución que linealice el sistema.

¿Cómo se debe afrontar estos problemas? ¿Que método se debe usar?

### 3. Bases de Gröbner.

El cálculo de bases de Gröbner [Her11] es una de las principales herramientas para resolver sistemas de ecuaciones polinomiales. Este concepto se puede ver como una generalización multivariable, no lineal, del algoritmo de Euclides para calcular los mayores divisores comunes polinomiales, y la eliminación de Gauss para sistemas lineales.

Para poder resolver este tipo de problemas, se deben formalizar algunos conceptos e introducir unos nuevos para tener más herramientas de cómo abordar estos problemas.

**Definición 3.1.** Sea  $K$  un cuerpo y  $f_1, f_2, f_3, \dots, f_s$  polinomios en

$K[x_1, x_2, \dots, x_n]$ , se define

$$V(f_1, f_2, \dots, f_s) = \{(a_1, \dots, a_n) \in K^n : f_i(a_1, \dots, a_n) = 0\}$$

<sup>3</sup>El trabajo de Galois y Abel indica que en general las ecuaciones polinomiales en una sola variable de grado mayor o igual que 5 no se pueden resolver de forma general en términos de radicales.



$$\forall i, 1 \leq i \leq s\}$$

y a  $V(f_1, f_2, \dots, f_s)$  se le conoce como *variedad afín*.

En otras palabras, una variedad afín [WDaGP13] es el conjunto de soluciones del sistema de ecuaciones:

$$f_1(x_1, x_2, \dots, x_n) = f_2(x_1, x_2, \dots, x_n) = \dots = f_s(x_1, x_2, \dots, x_n) = 0.$$

**Ejemplo 3.1.** Considere la variedad afín de  $V(2x^2 + 3y^2 - 11, x^2 - y^2 - 3)$ . Como ya se mostró anteriormente el conjunto de soluciones es  $\{(\pm 2, \pm 1)\}$ .

**Ejemplo 3.2.** Considere la variedad afín de  $V(x^2 + y^2 + z^2 - 25, x^2 + y^2 - 2x)$ . Gráficamente esto representa la intersección de una esfera con un cilindro, como se muestra en la Figura 2, se interceptan en una curva conocida como *curva de Viviani*.

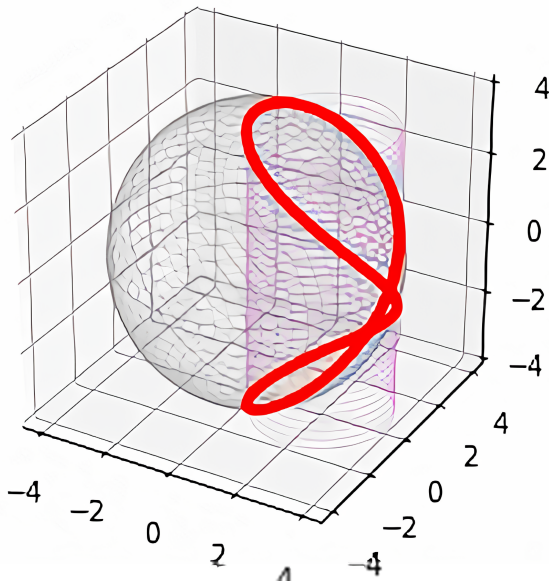


Figura 2: Note que la intersección de dos superficies puede ser una curva.

**Observación 3.1.** Una variedad afín puede ser el vacío, un conjunto de puntos discretos, una curva, un conjunto de curvas, entre otras.

**Definición 3.2.** Sea  $K[x_1, \dots, x_n]$  cualquier anillo de polinomios, un subconjunto  $I \subset K[x_1, \dots, x_n]$  es un ideal de  $K[x_1, \dots, x_n]$  si cumple que:

1.  $0 \in I$
2. Si  $f, g \in I$  entonces  $f + g \in I$
3. Si  $f \in I$  y  $h \in K[x_1, \dots, x_n]$  entonces  $hf \in I$

**Definición 3.3.** Sean  $f_1, \dots, f_s$  polinomios en  $K[x_1, \dots, x_n]$ . Entonces definimos:

$$\langle f_1, \dots, f_n \rangle = \left\{ \sum_{i=1}^s h_i f_i \mid h_1, \dots, h_s \in K[x_1, \dots, x_n] \right\}.$$

Se demostrará que en efecto  $\langle f_1, \dots, f_n \rangle$  es un ideal.

**Lema 3.1.** Si  $f_1, \dots, f_s \in K[x_1, \dots, x_n]$ , entonces  $\langle f_1, \dots, f_n \rangle$  es un ideal de  $K[x_1, \dots, x_n]$ .

A  $\langle f_1, \dots, f_n \rangle$  se conoce como el ideal generado por  $f_1, \dots, f_s$ .

**Demostración.** 1.  $0 \in \langle f_1, \dots, f_n \rangle$  ya que  $0 = \sum_{i=1}^s 0 \cdot f_i$

2. Considere  $f = \sum_{i=1}^s p_i f_i$  y  $g = \sum_{i=1}^s q_i f_i$

$$f + g = \sum_{i=1}^s p_i f_i + \sum_{i=1}^s q_i f_i = \sum_{i=1}^s (p_i f_i) + (q_i f_i) = \sum_{i=1}^s (p_i + q_i) f_i$$

Y como  $p_i + q_i$  es un elemento del anillo, entonces  $f + g$  está en el ideal.

3. Considere  $f = \sum_{i=1}^s p_i f_i$  y  $h \in K[x_1, \dots, x_n]$ , entonces

$$hf = h \sum_{i=1}^s p_i f_i = \sum_{i=1}^s (hp_i) f_i$$

Y como este es un anillo de polinomios,  $hp_i$  está en el anillo y  $hf$  está en el ideal

□

Del álgebra clásica se entiende que los polinomios usualmente son ordenados de forma descendente o ascendente respecto a la potencia de las variables. Con las siguientes definiciones se estandariza esta noción para álgebra multivariada.

**Definición 3.4.** Un orden monomial en  $K[x_1, \dots, x_n]$  es una relación ( $>$ ) sobre  $\mathbb{N}^n$ , o equivalentemente una relación sobre el conjunto de monomios  $x^\alpha$ , con  $\alpha \in \mathbb{N}^n$ , que satisface:

1.  $>$  es un orden total sobre  $\mathbb{N}^n$ .
2.  $>$  es un buen orden.
3. Si  $x^\alpha \leq x^\beta \Rightarrow x^\alpha x^\gamma \leq x^\beta x^\gamma, \forall \gamma \in \mathbb{N}^n$ .

**Definición 3.5.** Una relación de orden  $>$  sobre  $\mathbb{N}^n$ , es un buen orden si y solo si cada sucesión es estrictamente decreciente en  $\mathbb{N}^n$

$$\alpha_1 > \alpha_2 > \alpha_3 > \dots$$

eventualmente termina.

**Definición 3.6** (Orden lexicográfico). Dados dos monomios distintos  $x^\alpha, x^\beta \in K[x_1, \dots, x_n]$  con  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  y  $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$ . Diremos que  $x^\alpha <_{\text{lex}} x^\beta$  si el primer elemento no nulo (por izquierda) del vector  $\alpha - \beta \in \mathbb{Z}^n$  es negativo.

**Definición 3.7.** Sea  $f = \sum a_\gamma x^\gamma$  un polinomio en el cuerpo  $K$ , tal que  $f$  no sea el polinomio nulo. Se dice que:

- El monomio líder de  $f$  (aunque no sea necesariamente mónico) se denota  $ml(f) = x^n$  se llama monomio líder porque es el mayor según la relación de orden entre todos los monomios.
- El coeficiente líder de  $f$  se denota  $cl(f) = a_n$ .
- El término líder de  $f$  se denota  $tl(f) = a_n x^n$ .

**Definición 3.8.** Sea  $M \subset K[x_1, \dots, x_n]$  (no necesariamente un ideal). Se define el ideal del término líder de  $M$  como

$$TL(M) = \langle tl(f) : f \in M \rangle.$$

El algoritmo de Euclides es un teorema que asegura que el proceso habitual de división entre números enteros puede llevarse a cabo y que se obtiene un cociente y un residuo únicos. Este mismo algoritmo se puede extender para polinomios como se mostrará a continuación:

**Teorema 3.1** (Algoritmo de la división). Sea  $K$  un cuerpo. Entonces, dados polinomios  $P, D \in K[x]$  con  $D \neq 0$ , existen únicos polinomios  $Q$  (cociente) y  $R$  (resto) de la división de polinomios de  $P$  por  $D$ , tales que

$$P = QD + R \quad (14)$$

y  $R = 0$  (el polinomio nulo) o sino  $\deg(R) < \deg(D)$ . Donde  $\deg$  es el grado del polinomio.

Ahora se describe el algoritmo de la división en  $K[x_1, \dots, x_n]$ . Considere  $f, f_1, \dots, f_s \in K[x_1, \dots, x_n]$  con  $f_i \neq 0$ ,  $1 \leq i \leq s$  y un orden monomial  $\leq$  en  $k[x_1, \dots, x_n]$ ; se quiere dividir  $f$  entre la  $s$ -tupla de polinomios  $F = (f_1, \dots, f_s)$ . Es decir, expresar  $f$  de la forma:

$$f = h_1 f_1 + \dots + h_s f_s + r$$

con  $h_i, r \in K[x_1, \dots, x_n]$  cumpliendo las siguientes propiedades

$$r = 0 \quad \text{o} \quad r \neq 0 \quad (15)$$

y  $ml(f_i)$  no divide a ningún monomio de  $r$ , donde  $i = 1, \dots, s$ .

$$ml(f) = \max\{\max_{1 \leq i \leq s}\{ml(h_i), ml(f_i)\}, ml(r)\}. \quad (16)$$

Dado que ya se aclararon estos conceptos básicos se puede definir el concepto de base de Gröbner.

**Definición 3.9** (Bases de Gröbner). Sea  $I \subset K[x_1, \dots, x_n]$  un ideal no nulo. Un subconjunto de polinomios no nulos  $G = \{g_1, \dots, g_t\} \subset I$  es una **Base de Gröbner** del ideal  $I$  respecto al orden monomial  $\leq$  si

$$\forall f \in I, f \neq 0, \quad \exists i \in \{1, \dots, t\} \quad \text{tal que} \quad ml(g_i) | ml(f).$$

**Ejemplo 3.3.** En  $K[x, y]$  con el orden lexicográfico graduado, sean  $f_1 = x^3 - 2xy$  y  $f_2 = x^2y - 2y^2 + x$ , considere el ideal de términos líderes  $I = \langle f_1, f_2 \rangle$ . Observe que

$$x \cdot (x^2y - 2y^2 + x) - y \cdot (x^3 - 2xy) = x^2, \quad (17)$$

luego  $x^2 \in I$ . Sin embargo  $x^2$  no es divisible por  $tl(f_1) = x^3$  ni  $tl(f_2) = x^2y$ , luego  $x^2 \notin \langle tl(f_1), tl(f_2) \rangle$ . Por tanto  $\{f_1, f_2\}$  no es una base de Gröbner para  $I$ .

**Ejemplo 3.4.** Considere ahora los polinomios  $g_1 = x + z$  y  $g_2 = y - z$  de  $\mathbb{R}[x, y, z]$ . Sea  $J$  el ideal generado por estos dos polinomios,  $J = \langle g_1, g_2 \rangle$ . El conjunto  $\{g_1, g_2\}$  es una base de Gröbner de  $J \subset \mathbb{R}[x, y, z]$  con el orden lexicográfico.

**Teorema 3.2.** Considere  $I \subset K[x_1, \dots, x_n]$  un ideal distinto de cero y  $G = \{g_1, \dots, g_t\} \subset I$ . Las siguientes proposiciones son equivalentes:

1.  $G$  es una base de Gröbner de  $I$ .
2.  $\forall f \in K[x_1, \dots, x_n], f \in I$  si y solo si  $\bar{f}^G = 0$ .<sup>4</sup>
3.  $\forall f \in K[x_1, \dots, x_n], f \in I$  si y solo si  $\exists h_i \in K[x_1, \dots, x_n], i = 1, \dots, t$  tal que  $f = \sum_{i=1}^t h_i g_i$  con  $ml(f) = \max_{1 \leq i \leq t} \{ml(h_i)ml(g_i)\}$ .
4.  $TL(G) = TL(I)$ .

**Lema 3.2.** Sea  $I$  un ideal generado por un conjunto  $S$  de términos no-nulos, y sea  $f \in K[x_1, \dots, x_n]$ . Entonces  $f$  está en  $I$  si y sólo para todo término  $X$  que aparece en  $f$  existe un  $Y \in S$  tal que  $Y$  divide a  $X$ . Además, existe un sub-conjunto finito  $S_0$  de  $S$  tal que  $I = \langle S_0 \rangle$ .

*Demostración.* Si  $f \in I$ , entonces

$$f = \sum_{i=1}^l h_i X_i,$$

donde  $h_i \in K[x_1, \dots, x_n]$  y  $X_i \in S$ , para  $i = 1, \dots, l$ . Expandiendo el lado derecho se ve que cada término que aparece será divisible por algún  $X_i$  y entonces también cada término en  $f$  será divisible por algún  $X_i$ . Si cada término  $X_i$  de  $f$  es divisible por algún  $Y \in S$ , entonces claramente  $f \in \langle S \rangle = I$ .

Para probar la última afirmación, desde que  $K[x_1, \dots, x_n]$  es noetheriano<sup>5</sup>,  $I$  tiene un conjunto finito que lo genera. Luego por la primera parte del lema cada término de cada uno de los miembros de este conjunto finito es divisible por algún elemento de  $S$ . Sea  $S_0$  el conjunto finito de estos divisores, entonces  $I = \langle S_0 \rangle$ .  $\square$

**Teorema 3.3.** Todo ideal  $I \subset K[x_1, \dots, x_n]$  distinto de  $\{0\}$  tiene una base de Gröbner.

*Demostración.* Por el lema anterior, el ideal de términos  $TL(I)$  principales tiene un conjunto finito que lo genera, escriba  $tl(g_1), \dots, tl(g_t)$  con  $g_1, \dots, g_t \in I$ . Sea  $G = \{g_1, \dots, g_t\}$ , entonces  $TL(G) = TL(I)$  por lo que se llega al inciso (4) del Teorema anterior y por este teorema sabemos que  $TL(G) = TL(I)$  define una base de Gröbner.  $\square$

**Definición 3.10.** Sean  $f, g \in K[x_1, \dots, x_n]$  no nulos. Sea

$$X = \text{mcm}(ml(f), ml(g)),$$

<sup>4</sup>La notación  $\bar{f}^G$  hace referencia al residuo de la división de  $f$  entre  $G$

<sup>5</sup>Un anillo  $A$  es noetheriano si todos sus ideales son finitamente generados.

donde  $mcm$  representa el mínimo común múltiplo. Se define el  $S$ -polinomio de  $f$  y  $g$  como:

$$S(f, g) = \frac{X}{tl(f)} f - \frac{X}{tl(g)} g.$$

Donde la noción del mínimo común múltiplo de un par de monomios se puede definir del siguiente modo:

**Definición 3.11.** Considere monomios  $x^\alpha$  y  $x^\beta$ , donde  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$  y  $x^\beta = x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$ . Note que el mínimo común múltiplo (MCM) de estos se puede encontrar al tomar el máximo de cada índice

$$mcm(x^\alpha, x^\beta) = x_1^{\max(\alpha_1, \beta_1)} x_2^{\max(\alpha_2, \beta_2)} \dots x_n^{\max(\alpha_n, \beta_n)} \quad (18)$$

**Ejemplo 3.5.** Calcular el  $S$ -polinomio de  $f(x, y) = 2x^2y - 3x^2$ ,  $g(x, y) = xy^3 + 2xy^2$ , considerando  $<_{lex}$ ,

$$f(x, y) = 2x^2y - 3x^2,$$

$$g(x, y) = xy^3 + 2xy^2,$$

$$ml(f) = x^2y,$$

$$ml(g) = xy^3,$$

$$tl(f) = 2x^2y,$$

$$tl(g) = xy^3.$$

$$X = mcm(x^2y, xy^3) = x^2y^3, \quad (19)$$

$$\begin{aligned} S(f, g) &= \frac{X}{tl(f)} f - \frac{X}{tl(g)} g, \\ &= \frac{x^2y^3}{2x^2y} (2x^2y - 3x^2) - \frac{x^2y^3}{xy^3} (xy^3 + 2xy^2), \\ &= \frac{y^2}{2} (2x^2y - 3x^2) - x(xy^3 + 2xy^2), \\ &= x^2y^3 - \frac{3}{2}x^2y^2 - x^2y^3 - 2x^2y^2, \\ &= -\frac{3}{2}x^2y^2 - 2x^2y^2, \\ &= -\frac{7}{2}x^2y^2. \end{aligned}$$

Por lo tanto el  $S$ -polinomio de  $2x^2y - 3x^2$  y  $xy^3 + 2xy^2$  es  $-\frac{7}{2}x^2y^2$ .

La razón por la cual se introduce el concepto de  $S$ -polinomio es porque el algoritmo que se presentará para calcular una base de Gröbner necesita de este concepto. El algoritmo se llama algoritmo de Buchberger y se puede comprobar que este genera una base de Gröbner.

**Teorema 3.4** (de Buchberger). Sea  $G = \{g_1, \dots, g_t\} \subset K[x_1, \dots, x_n]$  con  $g_i$  no nulo para  $i \in \{1, \dots, t\}$ . Entonces,  $G$  es una base de Gröbner del ideal  $I = \langle g_1, \dots, g_t \rangle$  si y solo si  $\overline{S(g_i, g_j)}^G = 0, \forall 1 \leq i < j \leq t$ .

**Demostración.** La demostración formal se puede encontrar en [GP15] a modo de prueba se muestra el pseudocódigo.

Sea  $I = \langle f_1, \dots, f_s \rangle \subset K[x_1, \dots, x_n]$  ideal con  $f_i \neq 0, \forall i = 1, \dots, s$ . Podemos construir una base de Gröbner  $G = \{g_1, \dots, g_t\}$  para  $I$  con el siguiente algoritmo.

**Entrada:** Un conjunto de polinomios  $F$  que generan un ideal  $I$ .

**Salida:** Una base de Gröbner  $G$  para  $I$  respecto a un orden monomial dado.

1.  $G := F$ .
2. Para cada pareja  $F_i, F_j \in G$ , denotar por  $g_i$  el término dominante de  $F_i$  con respecto a un orden monomial dado, y por  $a_{ij}$  el mínimo común múltiplo de  $g_i$  y  $g_j$ .
3. Para cada pareja  $F_i, F_j \in G$  tomar  $S_{ij} = \left(\frac{a_{ij}}{g_i}\right) F_i - \left(\frac{a_{ij}}{g_j}\right) F_j$ .
4. Reducir  $S_{ij}$  usando el algoritmo de división para polinomios en varias variables respecto al conjunto  $G$  hasta que el resultado ya no se pueda reducir más.
5. Si se obtienen restos no nulos de estas divisiones, se añaden a  $G$ .
6. Repetir de 1 a 4 hasta que se hayan considerado todas las parejas posibles, incluyendo los nuevos polinomios añadidos.
7. Devolver  $G$ .

□

**Ejemplo 3.6.** Considere el sistema de ecuaciones homogéneo

$$\begin{cases} x^2 - y^2 - 3 = 0, \\ 2x^2 + 3y^2 - 11 = 0. \end{cases} \quad (20)$$

Cuya base de Gröbner es

$$G = \left[ x^2 - y^2 - 3, 2x^2 + 3y^2 - 11, -\frac{5y^2}{2} + \frac{5}{2} \right]$$

$$\begin{cases} x^2 - y^2 - 3 = 0, \\ 2x^2 + 3y^2 - 11 = 0, \\ -\frac{5y^2}{2} + \frac{5}{2} = 0. \end{cases}$$

De la última ecuación se deduce que  $y = \pm 1$  y reemplazando esto en las ecuaciones del sistema se tiene que  $x = \pm 2$ .

Ya con estos conceptos claros se puede dar una generalización para un sistema de ecuaciones cuadráticas.

**Ejemplo 3.7.** Considere el sistema de ecuaciones homogéneo

$$\begin{cases} ax^2 + by^2 + c = 0, \\ dx^2 + ey^2 + f = 0. \end{cases} \quad (21)$$

Cuya base de Gröbner es

$$G = [ax^2 + by^2 + c, dx^2 + ey^2 + f, -aey^2 - af + bdy^2 + cd].$$

Igualando  $-aey^2 - af + bdy^2 + cd$  a cero se encuentran los valores para la variable  $y$ , luego reemplazando esto en las otras dos ecuaciones se calculan los valores del eje  $x$ .



Observe un ejemplo con un grado más alto.

**Ejemplo 3.8.** Considere el sistema de ecuaciones

$$\begin{cases} ax^3 + by^3 + c = 0, \\ dx^3 + ey^3 + f = 0. \end{cases} \quad (22)$$

Cuya base de Gröbner es

$$G = [ax^3 + by^3 + c, dx^3 + ey^3 + f, -aey^3 - af + bdy^3 + cd].$$

Es decir, igualando  $-aey^3 - af + bdy^3 + cd$  a cero se encuentran los valores para la variable  $y$ , luego reemplazando esto en las otras dos ecuaciones se encuentran los valores del eje  $x$ .

Esto se puede generalizar para cualquier potencia.

**Ejemplo 3.9.** Considere el sistema de ecuaciones

$$\begin{cases} ax^n + by^n + c = 0, \\ dx^n + ey^n + f = 0. \end{cases} \quad (23)$$

Cuya base de Gröbner es

$$G = [ax^n + by^n + c, dx^n + ey^n + f, -aey^n - af + bdy^n + cd].$$

Así observamos que este proceso es la generalización del método de Gauss- Jordan para sistemas lineales.

## 4. Conclusión.

Existen muchos métodos y teoremas que establecen alternativas para resolver sistemas de ecuaciones lineales, como el método de sustitución, que consiste en aislar en una ecuación una de las incógnitas para sustituirla en las otras ecuaciones o, el método de reducción, que consiste en sumar (o restar) las ecuaciones del sistema para eliminar una de las incógnitas o, el método de igualación, que consiste en aislar una incógnita en las dos ecuaciones para igualarlas, entre muchos otros métodos. Sin embargo, cuando pasamos de sistemas lineales a un sistema más complejo, como uno polinomial no lineal por dar un caso, estas herramientas no son de utilidad. En este sentido, las bases de Gröbner son una herramienta fundamental para abordar este tipo de problemas, dado que en cierto sentido permiten extender el álgebra de un contexto lineal a uno polinomial, como se vio en este documento.

## Referencias.

- [JH17] Jim Hefferon, *Linear Algebra*, 2017.
- [WDaGP13] Wolfram Decker and Gerhard Pster, *A First Course in Computational Algebraic Geometry*, CAMBRIDGE, 2013.
- [Bec93] Thomas and Weispfenning Becker Volker, *Gröbner bases*, Springer, 1993.
- [Ste90] Ian Stewart, *Galois theory*, Chapman and Hall/CRC, 1990.
- [Her11] Marcelo Escudeiro Hernandes, *Um primeiro contato com bases de Gröbner*, Publicações Matemáticas do IMPA. [IMPA Mathematical Publications], Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2011. 28o Colóquio Brasileiro de Matemática. [28th Brazilian Mathematics Colloquium].

[GP15] Alba González Parra, *Bases de Gröbner: Eliminación y programación lineal entera* (2015).

**Acerca del autor:** Juan Sebastián Martínez Conejo es egresado del Programa de Matemáticas de la Fundación Universitaria Konrad Lorenz. Apasionado por la analítica de datos y coleccionador de Cubos de Rubik. Texto dedicado a familiares y amigos perdidos por el Covid-19.

# DEL ÁLGEBRA LINEAL Y EL ANÁLISIS A LAS MÉTRICAS RIEMANNIANAS EN $\mathbb{R}^n$

Camilo Ramírez Maluendas<sup>\*</sup>  
camramirezma@unal.edu.co

## 1. Resumen.

Iniciaremos dando un breve paseo por el producto interno y la norma en los espacios vectoriales. Luego, definiremos curva y, a partir del producto punto usual de  $\mathbb{R}^n$  y con ayuda del análisis obtendremos una fórmula para calcular la longitud de una curva de clase  $\mathcal{C}^1$  en  $\mathbb{R}^n$ . Esta longitud nos permitirá obtener una métrica (función distancia), la cual coincidirá con la métrica usual de  $\mathbb{R}^n$ . Finalmente, todas estas ideas serán extendidas y, como resultado introduciremos la noción de métrica riemanniana junto con algunas de sus propiedades en  $\mathbb{R}^n$ .

## 2. El producto interno y la norma.

Uno de los objetos clásicos en el estudio del álgebra lineal es el producto punto. Desde el punto de vista histórico, el producto interno apareció en 1840. Durante este año, el matemático alemán Hermann Günter Grassmann (1809-1877) publicó el artículo titulado *Theorie der Ebbe und Flut*, en el cual introdujo la base teórica que ha permitido desarrollar el área de los espacios vectoriales. En su manuscrito, Grassmann, definió la sustracción y adición de unos objetos matemáticos los cuales llamó *strecken*<sup>1</sup>, una de estas operaciones es el equivalente al conocido producto cruz moderno. Además, en su escrito, el alemán, presentó la operación producto lineal de dos *strecken*, la cual corresponde al producto interno, punto o escalar (véase [Cro67]). Es momento de tratar el concepto de producto interno.

Si  $V$  es un espacio vectorial sobre  $\mathbb{R}$ , entonces un **producto interno**<sup>2</sup> sobre  $V$  es una función

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}, \quad (1)$$

la cual satisface las siguientes propiedades:

1. **Linealidad por izquierda.** Para cualesquiera vectores  $u, v, w \in V$  y reales  $\lambda, \mu \in \mathbb{R}$  se cumple  $\langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle$ .
2. **Simetría.** Para cualesquiera vectores  $u, v \in V$  se cumple  $\langle u, v \rangle = \langle v, u \rangle$ .

3. **Definida positiva.** Para cualesquiera vectores  $u, v \in V$  se cumple  $\langle u, v \rangle \geq 0$ , la igualdad se tiene si y solo si  $u = \mathbf{0}$ .

De las dos primeras propiedades enunciadas anteriormente es posible verificar fácilmente la linealidad por derecha, es decir, para cualesquiera vectores  $u, v, w \in V$  y reales  $\lambda, \mu \in \mathbb{R}$  es cierta la igualdad

$$\langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle.$$

Una las ventajas de contar con un producto interno es que nos proporciona una manera de medir vectores, en otras palabras, si  $\langle \cdot, \cdot \rangle$  es un producto interno sobre al espacio vectorial  $V$ , entonces la **norma** o **magnitud** del vector  $v \in V$  es el valor real

$$\|v\| := \sqrt{\langle v, v \rangle}. \quad (2)$$

La norma de un vector cumple propiedades interesantes, mencionaremos y probaremos algunas de ellas.

**Teorema 2.1.** Si  $\langle \cdot, \cdot \rangle$  es un producto interno sobre el espacio vectorial  $V$  entonces:

1. Para todo vector  $v \in V$  se cumple que  $\|v\| \geq 0$ . La igualdad  $\|v\| = 0$  se tiene si y solo si  $v = \mathbf{0}$ .
2. Para todo real  $\lambda \in \mathbb{R}$  y todo vector  $v \in V$  se cumple que  $\|\lambda v\| = |\lambda| \|v\|$ .
3. **Desigualdad de Cauchy-Schwartz.** Para cualesquiera vectores  $u, v \in V$  se cumple

$$|\langle u, v \rangle| \leq \|u\| \|v\|. \quad (3)$$

4. **Desigualdad triangular** Para cualesquiera vectores  $u, v \in V$  se cumple

$$\|u + v\| \leq \|u\| + \|v\|. \quad (4)$$

**Demostración.** La demostración de los incisos (1) y (2) es consecuencia directa de la definición de norma. Veamos que se satisface el inciso (3). Consideremos el vector  $v \neq \mathbf{0}$  en  $V$  y definamos el vector  $w = \frac{\langle u, v \rangle}{\|v\|^2} v$ . Entonces

<sup>\*</sup>Departamento de Matemática y Estadística, Universidad Nacional de Colombia, Sede Manizales.

<sup>1</sup>Tramo, traducción del alemán al castellano.

<sup>2</sup>También se le conoce como producto escalar o producto punto.

$$\begin{aligned}
0 &\leq \langle u - w, u - w \rangle = \langle u, u \rangle - \langle u, w \rangle - \langle w, u \rangle + \langle w, w \rangle, \\
&\leq \langle u, u \rangle - 2\langle u, w \rangle + \langle w, w \rangle, \\
&\leq \langle u, u \rangle - 2 \left\langle u, \frac{\langle u, v \rangle}{\|v\|^2} v \right\rangle + \left\langle \frac{\langle u, v \rangle}{\|v\|^2} v, \frac{\langle u, v \rangle}{\|v\|^2} v \right\rangle, \\
&\leq \langle u, u \rangle - \frac{2}{\|v\|^2} \langle u, v \rangle \langle u, v \rangle + \frac{1}{\|v\|^2 \|v\|^2} \langle u, v \rangle \langle u, v \rangle \langle v, v \rangle, \\
&\leq \langle u, u \rangle - \frac{2}{\|v\|^2} \langle u, v \rangle^2 + \frac{1}{\|v\|^2} \langle u, v \rangle^2, \\
&\leq \|u\|^2 - \frac{1}{\|v\|^2} \langle u, v \rangle^2.
\end{aligned}$$

De esta última expresión despejamos el término  $\langle u, v \rangle^2$  y luego tomamos la raíz cuadrada

$$\begin{aligned}
\langle u, v \rangle^2 &\leq \|u\|^2 \|v\|^2, \\
|\langle u, v \rangle| &\leq \|u\| \|v\|.
\end{aligned}$$

Finalmente, veamos que se satisface el inciso (4). Observemos que se tiene

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v \rangle, \\
&= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle, \\
&= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2.
\end{aligned}$$

A esta última igualdad le aplicamos la desigualdad de Cauchy-Schwarz y obtenemos

$$\begin{aligned}
\|u + v\|^2 &\leq \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \leq \|u\|^2 + 2\|u\| \|v\| + \|v\|^2, \\
&\leq (\|u\| + \|v\|)^2.
\end{aligned}$$

□

La norma induce una función distancia sobre el espacio vectorial  $V$ , dicho de otra manera, si  $V$  es un espacio vectorial con un producto interno  $\langle \cdot, \cdot \rangle$ , entonces la función  $d : V \times V \rightarrow \mathbb{R}$ , definida mediante

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}, \quad (5)$$

es una métrica sobre el espacio vectorial  $V$ , para cualesquiera  $u, v \in V$ . La prueba de esta afirmación se puede verificar fácilmente usando las propiedades descritas en el Teorema 2.1.

Si el lector está interesado en continuar el abordaje del producto interno y la norma, le sugerimos echar un vistazo a e.g., [Ser86], [FILE89]. Ahora, presentaremos algunos ejemplos clásicos.

**Ejemplo 2.1.** En  $\mathbb{R}^n$  definimos el **producto interno usual**  $\langle \cdot, \cdot \rangle_e$  tal que para cualesquiera dos elementos  $\mathbf{x} = (x_1, \dots, x_n)$ ,  $\mathbf{y} = (y_1, \dots, y_n)$  en  $\mathbb{R}^n$  se tiene

$$\langle \mathbf{x}, \mathbf{y} \rangle_e := \sum_{i=1}^n x_i y_i.$$

Los vectores  $e_1, \dots, e_n$  corresponden a la base canónica o estándar para  $\mathbb{R}^n$ , entonces el producto interno usual

$$\langle e_i, e_j \rangle_e = \delta_{i,j}$$

es el **delta de Kroneécker**, donde  $\delta_{i,j} := 1$  si  $i = j$ , con  $j, i \in \{1, \dots, n\}$ . Contrariamente, si  $i \neq j$ , entonces  $\delta_{i,j} = 0$ . Adicionalmente, la norma del vector  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  es

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Por otro lado, la distancia entre los vectores  $\mathbf{x} = (x_1, \dots, x_n)$  y  $\mathbf{y} = (y_1, \dots, y_n)$  está dada por

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Además, se le suele llamar **la métrica euclidiana usual**.

**Ejemplo 2.2.** Dado un real positivo  $c > 0$ , entonces en  $\mathbb{R}^n$  se define el producto interno  $\langle \cdot, \cdot \rangle$  tal que si  $\mathbf{x}, \mathbf{y}$  son elementos en  $\mathbb{R}^n$  entonces

$$\langle \mathbf{x}, \mathbf{y} \rangle := \frac{\langle \mathbf{x}, \mathbf{y} \rangle_e}{c}.$$

La norma del vector  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  está dada por

$$\|\mathbf{x}\| = \frac{\sqrt{c}}{c} \sqrt{\sum_{i=1}^n x_i^2}.$$

Por otro lado, la distancia entre los vectores  $\mathbf{x} = (x_1, \dots, x_n)$  y  $\mathbf{y} = (y_1, \dots, y_n)$  es

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \frac{\sqrt{c}}{c} \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

**Ejemplo 2.3.** Tomemos  $V$  el espacio vectorial de las funciones continuas real valuadas en el intervalo cerrado  $[-1, 1]$ . Si  $f$  y  $g$  son vectores de  $V$ , entonces definimos

$$\langle f, g \rangle := \int_{-1}^1 f(x)g(x)dx.$$

De las propiedades básicas de la integral (véase e.g., [Spi70]) se sigue que la igualdad anterior es un producto punto. Adicionalmente, la norma de la función  $f \in V$  está dada por

$$\|f\| = \sqrt{\int_{-1}^1 [f(x)]^2 dx}.$$

Por otro lado, la distancia entre las funciones  $f$  y  $g$  de  $V$  es

$$d(f, g) = \|f - g\| = \sqrt{\int_{-1}^1 [f(x) - g(x)]^2 dx}.$$



### 3. La métrica euclidiana desde un punto de vista geométrico “usando las curvas”.

El matemático Adolf Hurwitz (1859-1919) propuso en 1879 durante el *I Congreso Internacional de Matemáticos* dar una respuesta razonable a la siguiente pregunta: ¿Qué es una curva? (véase [TF98]). En 1921-22 el matemático ruso Pável Uryson (1898-1924) definió una **curva**, desde el punto de vista topológico, como un *continuo*<sup>3</sup> de dimensión 1 (véase e.g., [Kur66], [Kur68]). Parte del trabajo de Uryson originó la teoría de la dimensión [Eng89, p. 392]. Dado que nuestro propósito consiste en definir la métrica euclidiana (función distancia) en  $\mathbb{R}^n$  con ayuda de la longitud de una curva, definiremos curva usando funciones continuas.

Una **curva**  $\gamma$  en  $\mathbb{R}^n$  es una función continua desde un intervalo  $I^4 \subset \mathbb{R}$  hasta  $\mathbb{R}^n$ . Si  $x_1, \dots, x_n$  son las coordenadas de  $\mathbb{R}^n$ , entonces la regla de asignación de  $\gamma$  está descrita mediante

$$\gamma(t) = (x_1(t), \dots, x_n(t)).$$

La imagen directa  $\gamma(I) \subset \mathbb{R}^n$  es llamada **la traza de la curva**  $\gamma$ , algunos autores también le llaman **trayectoria de la curva**  $\gamma$  (véase Figura 1). Cuando la curva  $\gamma$  está definida en el intervalo cerrado  $I = [a, b]$ , decimos que las  $n$ -uplas  $\gamma(a)$  y  $\gamma(b)$  de  $\mathbb{R}^n$  son los **extremos o puntos finales de la curva**  $\gamma$ .

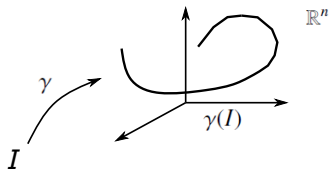


Figura 1: Traza de una curva.

La curva  $\gamma$  es **derivable** en el punto  $t_0 \in I$ , si para cada  $i \in \{1, \dots, n\}$  existe la derivada parcial

$$x'_i(t_0) = \frac{\partial x_i}{\partial t}(t_0) = \lim_{t \rightarrow t_0} \frac{x_i(t) - x_i(t_0)}{t - t_0},$$

entonces definimos la  $n$ -upla

$$\gamma'(t_0) = D\gamma(t_0) = (x'_1(t_0), \dots, x'_n(t_0)), \quad (6)$$

la cual es conocida como la **derivada de la curva  $\gamma$  en el punto  $t_0$** . Desde el punto de vista geométrico, la derivada  $\gamma'(t_0)$  representa el vector tangente de  $\gamma$  en el punto  $\gamma(t_0)$ . Por otro lado, desde el punto de vista físico, la derivada  $\gamma'(t_0)$  representa el vector velocidad de  $\gamma$  en el punto  $\gamma(t_0)$ . Si para  $k \in \mathbb{N}$  y cada  $i \in \{1, \dots, n\}$  existe la derivada parcial

$$x_i^k(t_0) = \frac{\partial^k x_i}{\partial t^k}(t_0),$$

entonces definimos la  **$k$ -ésima derivada de la curva  $\gamma$  en el punto  $t_0$**  como la  $n$ -upla dada por

$$\gamma^{(k)}(t_0) = (x_1^k(t_0), \dots, x_n^k(t_0)).$$

La curva  $\gamma$  se dice que es de **clase  $\mathcal{C}^k$**  si sus primeras  $k$  derivadas existen y son continuas en el intervalo  $I$ . La curva  $\gamma$  se dice que es **suave** o de **clase  $\mathcal{C}^\infty$**  si es de clase  $\mathcal{C}^k$ , para cada  $k \in \mathbb{N}$ .

**Ejemplo 3.1.** La función  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  cuya regla de asignación está dada por

$$\gamma(t) = (t, t^2)$$

es una curva suave o de clase  $\mathcal{C}^\infty$ , cuya traza es  $\gamma(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : x^2 = y\}$ <sup>5</sup> (véase Figura 2). La derivada de  $\gamma$  en el punto  $t \in \mathbb{R}$  es  $\gamma'(t) = (1, 2t)$ .

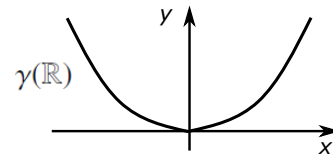


Figura 2: Traza de la curva  $\gamma(t) = (t, t^2)$ .

**Ejemplo 3.2.** Dadas las  $n$ -uplas  $\mathbf{a} = (a_1, \dots, a_n)$  y  $\mathbf{b} = (b_1, \dots, b_n)$  en  $\mathbb{R}^n$ , definimos la función  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$  tal que

$$\gamma(t) = \mathbf{a}t + \mathbf{b},$$

la cual es curva suave o de clase  $\mathcal{C}^\infty$ , cuya traza  $\gamma(\mathbb{R})$  es la línea recta con dirección el vector  $(a_1, \dots, a_n)$  y que atraviesa por el punto  $(b_1, \dots, b_n)$  (véase Figura 3). La derivada de  $\gamma$  en el punto  $t \in \mathbb{R}$  es  $\gamma'(t) = \mathbf{a}$ .

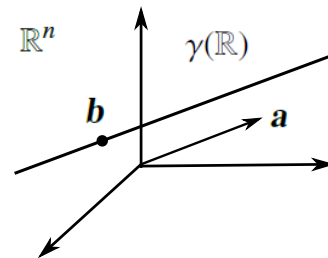


Figura 3: Traza de la curva  $\gamma(t) = \mathbf{a}t + \mathbf{b}$ .

Si el lector está interesado en conocer un zoológico de curvas, le sugerimos revisar [dC16], [Spi79]. A continuación, enunciaremos algunas de las propiedades de la derivada e introduciremos la definición de longitud de una curva.

<sup>3</sup>Un continuo es un espacio topológico conexo, compacto y Hausdorff.

<sup>4</sup>El intervalo  $I$  puede ser abierto  $(a, b)$ , cerrado  $[a, b]$  o todo  $\mathbb{R}$ .

<sup>5</sup>Las curvas cuyas trazas están en el plano euclidiano  $\mathbb{R}^2$  son llamadas **curvas planas**.

**Teorema 3.1.** Dadas las curvas  $\gamma_1, \gamma_2 : I \rightarrow \mathbb{R}^n$  de clase  $\mathcal{C}^1$  con reglas de asignación  $\gamma_1(t) = (x_1(t), \dots, x_n(t))$  y  $\gamma_2(t) = (y_1(t), \dots, y_n(t))$ , entonces se satisfacen las siguientes propiedades<sup>6</sup>:

1. **Linealidad.** Si  $\lambda$  y  $\mu$  son números reales, entonces consideramos la curva  $\lambda\gamma_1 + \mu\gamma_2 : I \rightarrow \mathbb{R}^n$ , definida mediante

$$(\lambda\gamma_1 + \mu\gamma_2)(t) = (\lambda x_1(t) + \mu y_1(t), \dots, \lambda x_n(t) + \mu y_n(t)).$$

Entonces para cada valor  $t \in I$  la derivada

$$(\lambda\gamma_1 + \mu\gamma_2)'(t) = \lambda\gamma_1'(t) + \mu\gamma_2'(t).$$

2. **Fórmula de Leibniz.** Dada la curva  $f : I \rightarrow \mathbb{R}$  de clase  $\mathcal{C}^1$ , entonces consideramos la curva  $f\gamma : I \rightarrow \mathbb{R}^n$ , definida mediante

$$f\gamma(t) = (f(t)x_1(t), \dots, f(t)x_n(t))$$

Entonces para cada valor  $t \in I$  la derivada

$$(f\gamma)'(t) = \gamma(t)f'(t) + f(t)\gamma'(t).$$

3. **Fórmula de Leibniz para el producto interno usual.** Definimos la curva  $\langle \gamma_1, \gamma_2 \rangle : I \rightarrow \mathbb{R}^n$  definida mediante

$$\langle \gamma_1, \gamma_2 \rangle(t) = \sum_{i=1}^n x_i(t)y_i(t).$$

Entonces para cada valor  $t \in I$  la derivada cumple

$$\langle \gamma_1, \gamma_2 \rangle'(t) = \langle \gamma_1', \gamma_2 \rangle(t) + \langle \gamma_1, \gamma_2' \rangle(t).$$

**Demostración.** Iniciaremos probando la propiedad 1. De las propiedades de linealidad de la derivada (cálculo diferencial) se tiene que para cada  $i \in \{1, \dots, n\}$ , la derivada

$$(\lambda x_i + \mu y_i)'(t) = \lambda x_i'(t) + \mu y_i'(t), \text{ para cada } t \in I.$$

De esta relación se sigue que la derivada de la curva  $\lambda\gamma_1 + \mu\gamma_2$  en el punto  $t \in I$  está dada por

$$\begin{aligned} (\lambda\gamma_1 + \mu\gamma_2)'(t) &= (\lambda x_1'(t) + \mu y_1'(t), \dots, \lambda x_n'(t) + \mu y_n'(t)), \\ &= (\lambda x_1'(t), \dots, \lambda x_n'(t)) + (\mu y_1'(t), \dots, \mu y_n'(t)), \\ &= \lambda(x_1'(t), \dots, x_n'(t)) + \mu(y_1'(t), \dots, y_n'(t)), \\ &= \lambda\gamma_1'(t) + \mu\gamma_2'(t). \end{aligned}$$

Ahora verificaremos la propiedad 2. Dado que las curvas  $\gamma_1$  y  $f$  son de clase  $\mathcal{C}^1$ , entonces para cada  $i \in \{1, \dots, n\}$ , la derivada

$$(fx_i)'(t) = f(t)x_i'(t) + x_i(t)f'(t), \text{ para cada } t \in I.$$

<sup>6</sup>El lector puede referirse a [PVRV08] para más propiedades de la derivada.

De esta última relación se sigue que la derivada de la curva  $f\gamma_1$  en el punto  $t \in I$  está dada por

$$\begin{aligned} (f \cdot \gamma_1)'(t) &= (f(t)x_1'(t) + x_1(t)f'(t), \dots, f(t)x_n'(t) + x_n(t)f'(t)), \\ &= (f(t)x_1'(t), \dots, f(t)x_n'(t)) + (x_1(t)f'(t), \dots, x_n(t)f'(t)), \\ &= f(t)(x_1'(t), \dots, x_n'(t)) + f'(t)(x_1(t), \dots, x_n(t)), \\ &= f(t)\gamma_1'(t) + f'(t)\gamma_1(t). \end{aligned}$$

Finalmente, demostraremos que se satisface la propiedad 3. Dado que las curvas  $\gamma_1$  y  $\gamma_2$  son de clase  $\mathcal{C}^1$ , entonces, de las reglas de la derivada del cálculo diferencial para cada  $t \in I$ , la derivada de la curva  $\langle \gamma_1(t), \gamma_2(t) \rangle$  está dada por

$$\begin{aligned} \langle \gamma_1, \gamma_2 \rangle'(t) &= \left( \sum_{i=1}^n x_i(t)y_i(t) \right)', \\ &= x_1'(t)y_1(t) + y_1'(t)x_1(t) + \dots + x_n'(t)y_n(t) + y_n'(t)x_n(t), \\ &= \sum_{i=1}^n x_i'(t)y_i(t) + \sum_{i=1}^n y_i'(t)x_n(t), \\ &= \langle \gamma_1', \gamma_2 \rangle(t) + \langle \gamma_2', \gamma_1 \rangle(t). \end{aligned}$$

□

Dada la curva  $\gamma : [a, b] \rightarrow \mathbb{R}^n$ , tomamos una partición  $P$  del intervalo cerrado  $[a, b]$  conformada por los reales

$$a = t_0 < t_1 < \dots < t_{n-1} < t_n = b, \text{ para algún } m \in \mathbb{Z}^+.$$

A partir de los puntos  $\gamma(t_i)$  en  $\mathbb{R}^n$  con  $i \in \{0, \dots, m\}$  definimos **la curva poligonal**  $\gamma_P$ , que tiene como traza los segmentos de línea recta con extremos  $\gamma(t_{i-1})$  y  $\gamma(t_i)$ , con  $i \in \{1, \dots, m\}$  (véase Figura 4).

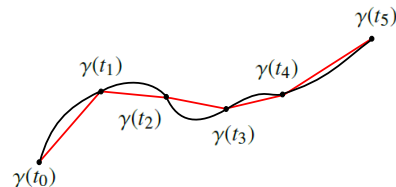


Figura 4: Curva poligonal.

Naturalmente, **la longitud de la curva**  $\gamma_P$  es real

$$\ell(\gamma, P) =: \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\|.$$

Nótese que de la desigualdad triangular (4) se sigue

$$\|\gamma(b) - \gamma(a)\| \leq \ell(\gamma, P) = \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\|. \quad (7)$$

El valor  $\ell(\gamma, P)$  lo podemos entender como lo que podría ser una aproximación a **la longitud de la curva**  $\gamma$ . Entre más puntos tenga la partición  $P$ , mejor será la aproximación a la longitud de la curva  $\gamma$ . Esta idea sugiere la definir la longitud de una curva como sigue.

Dada la curva  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  de clase  $\mathcal{C}^1$ , definimos la **longitud de  $\gamma$**  mediante

$$\ell(\gamma) := \sup\{\ell(\gamma, P) : P \text{ es una partición del intervalo } [a, b]\}. \quad (8)$$

Si  $\ell(\gamma) < \infty$ , entonces la curva  $\gamma$  es llamada **rectificable**.

De la desigualdad (7) se sigue

$$\|\gamma(a) - \gamma(b)\| \leq \ell(\gamma). \quad (9)$$

Consideremos la curva  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  tal que su respectiva  $i$ -ésima componente es la función con valores reales  $x_i(t)$ , la cual es integrable sobre el intervalo cerrado  $[a, b]$ , para cada  $i \in \{1, \dots, n\}$ . Entonces definimos la **integral de la curva  $\gamma$  sobre el intervalo cerrado  $[a, b]$**  como la  $n$ -upla

$$\int_a^b \gamma(t) dt = \left( \int_a^b x_1(t) dt, \dots, \int_a^b x_n(t) dt \right). \quad (10)$$

Si  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  es una curva continua y  $\beta: [a, b] \rightarrow \mathbb{R}^n$  es una curva de clase  $\mathcal{C}^1$  tal que  $\gamma(t) = \beta'(t)$  para cada  $t \in [a, b]$ , es inmediato del Teorema Fundamental del Cálculo que

$$\int_a^b \gamma(t) dt = \beta(b) - \beta(a). \quad (11)$$

**Teorema 3.2.** Consideremos la curva  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  tal que  $\int_a^b \gamma(t) dt$  existe. Entonces satisfacen las siguientes propiedades

1. La función  $\|\gamma\|: [a, b] \rightarrow \mathbb{R}$ , definida mediante

$$\|\gamma(t)\| = \sqrt{\sum_{i=1}^n (x_i(t))^2},$$

es integrable sobre el intervalo  $[a, b]$ .

$$2. \left\| \int_a^b \gamma(t) dt \right\| \leq \int_a^b \|\gamma(t)\| dt.$$

*Demostración.* La prueba del inciso 1 se deja para que el lector la verifique y, puede ser consultada en [Rud76]. Comprobemos el inciso 2. Denotemos mediante  $y_i := \int_a^b x_i(t) dt$  para cada  $i \in \{1, \dots, n\}$ , entonces

$$\int_a^b \gamma(t) dt = (y_1, \dots, y_n) := \mathbf{y}.$$

De las anteriores relaciones se sigue que la norma al cuadrado de la  $n$ -upla  $\mathbf{y}$  es

$$\|\mathbf{y}\|^2 = \sum_{i=1}^n y_i^2 = \sum_{i=1}^n y_i y_i = \sum_{i=1}^n \left( y_i \int_a^b x_i(t) dt \right). \quad (12)$$

Dado que  $y_i$  es una constante para cada  $i \in \{1, \dots, n\}$ , entonces de las propiedades de linealidad de la integral la ecuación (12) se reescribe como sigue

$$\|\mathbf{y}\|^2 = \sum_{i=1}^n \left( \int_a^b y_i x_i(t) dt \right) = \int_a^b \left( \sum_{i=1}^n y_i x_i(t) \right) dt. \quad (13)$$

Por otro lado la Desigualdad de Cauchy-Schwartz (3) nos dice que para cualesquier par de vectores  $u, v$  en  $\mathbb{R}^n$  se cumple  $|\langle u, v \rangle| \leq \|u\| \|v\|$ . Usando dicha propiedad se sigue que para cada  $t \in [a, b]$

$$\begin{aligned} \sum_{i=1}^n y_i x_i(t) &= (y_1, \dots, y_n) \cdot (x_1(t), \dots, x_n(t)), \\ &= \mathbf{y} \cdot \gamma(t) \leq \|\mathbf{y}\| \|\gamma(t)\| \end{aligned}$$

Como la función  $\|\gamma\|$  es integrable, de los teoremas de comparación de la integral se sigue

$$\begin{aligned} \int_a^b \left( \sum_{i=1}^n y_i x_i(t) \right) dt &= \int_a^b (\mathbf{y} \cdot \gamma(t)) dt, \\ &\leq \int_a^b \|\mathbf{y}\| \|\gamma(t)\| dt = \|\mathbf{y}\| \int_a^b \|\gamma(t)\| dt. \end{aligned} \quad (14)$$

De las ecuaciones (13) y (14) se sigue

$$\|\mathbf{y}\|^2 \leq \|\mathbf{y}\| \int_a^b \|\gamma(t)\| dt. \quad (15)$$

Ahora bien, si  $\mathbf{y} = \mathbf{0}$  se verifica el inciso 2 del teorema. Contrariamente, si  $\mathbf{y} \neq \mathbf{0}$  se sigue de la ecuación (15)

$$\left\| \int_a^b \gamma(t) dt \right\| = \|\mathbf{y}\| \leq \int_a^b \|\gamma(t)\| dt.$$

□

**Teorema 3.3.** Si  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  es una curva de clase  $\mathcal{C}^1$  tal que  $\gamma'$  es continua en  $[a, b]$ , entonces

$$\ell(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

La prueba para este teorema fue tomada de [Shi13], las ideas se parecen a las presentadas por M. Spivak [Spi70] en la verificación del Teorema Fundamental del Cálculo.

*Demostración.* Consideremos la partición  $P$  del intervalo cerrado  $[a, b]$  tal que  $P = \{a = t_0 < t_1 < \dots < t_n = b\}$ . De la ecuación (11) se sigue

$$\ell(\gamma, P) = \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| = \sum_{i=1}^n \left\| \int_a^{t_i} \gamma'(t) dt \right\|.$$

Al aplicar el teorema 3.2 a la relación anterior junto con las propiedades de la integral obtenemos

$$\ell(\gamma, P) = \sum_{i=1}^n \left\| \int_a^b \gamma'(t) dt \right\| \leq \sum_{i=1}^n \int_a^b \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt.$$

Como hemos considerado una partición arbitraria  $P$  del intervalo cerrado  $[a, b]$ , entonces la anterior desigualdad es válida para cualquier partición del intervalo cerrado  $[a, b]$ . Así, concluimos que

$$\ell(\gamma) \leq \int_a^b \|\gamma'(t)\| dt. \quad (16)$$

Para cada  $t \in [a, b]$  tomamos la curva restricción  $\gamma_t = \gamma|_{[a, t]}$ . Entonces definimos la función  $s : [a, b] \rightarrow \mathbb{R}$ , que a cada valor  $t \in [a, b]$  determina la longitud de la curva  $\gamma_t$ , es decir

$$s(t) = \ell(\gamma_t).$$

Probaremos que  $s(t) = \ell(\gamma_t) = \int_a^t \|\gamma'(t)\| dt$ , para cada  $t \in [a, b]$ .

Para  $h > 0$ , la diferencia  $s(t+h) - s(t)$  es la longitud de la curva  $\gamma|_{[t, t+h]}$  y satisface (véase figura 5)

$$\|\gamma(t+h) - \gamma(t)\| \leq s(t+h) - s(t) = \ell(\gamma|_{[t, t+h]}). \quad (17)$$

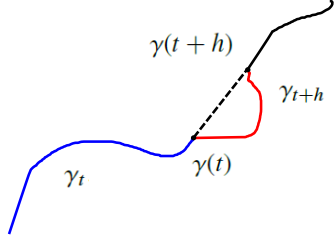


Figura 5: Curvas  $\gamma_t$  y  $\gamma_{t+h}$ .

De la ecuación (16) se sigue

$$\begin{aligned} \|\gamma(t+h) - \gamma(t)\| &\leq s(t+h) - s(t), \\ &\leq \int_a^{t+h} \|\gamma'(t)\| dt - \int_a^t \|\gamma'(t)\| dt. \end{aligned}$$

Nótese que

$$\int_a^{t+h} \|\gamma'(t)\| dt - \int_a^t \|\gamma'(t)\| dt = \int_t^{t+h} \|\gamma'(t)\| dt.$$

Ahora, en esta última desigualdad tomamos el cociente

$$\begin{aligned} \frac{\|\gamma(t+h) - \gamma(t)\|}{h} &\leq \frac{s(t+h) - s(t)}{h}, \\ &\leq \frac{1}{h} \left( \int_a^{t+h} \|\gamma'(t)\| dt - \int_a^t \|\gamma'(t)\| dt \right). \end{aligned} \quad (18)$$

Nótese que

$$\frac{1}{h} \left( \int_a^{t+h} \|\gamma'(t)\| dt - \int_a^t \|\gamma'(t)\| dt \right) = \frac{1}{h} \left( \int_t^{t+h} \|\gamma'(t)\| dt \right).$$

Por otro lado, tenemos los siguientes límites

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{\|\gamma(t+h) - \gamma(t)\|}{h} &= \|\gamma'(t)\|, \\ &= \lim_{h \rightarrow 0^+} \frac{1}{h} \left( \int_a^{t+h} \|\gamma'(t)\| dt - \int_a^t \|\gamma'(t)\| dt \right). \end{aligned}$$

Ahora, del teorema del Emparedado aplicado a la ecuación (18) y del anterior límite se sigue

$$\lim_{h \rightarrow 0^+} \frac{s(t+h) - s(t)}{h} = \|\gamma'(t)\|.$$

Con argumentos similares es posible probar que para  $h < 0$  se cumple

$$\lim_{h \rightarrow 0^-} \frac{s(t+h) - s(t)}{h} = \|\gamma'(t)\|.$$

De las anteriores ecuaciones tenemos que para cada  $t \in [a, b]$  la derivada

$$s'(t) = \|\gamma'(t)\|.$$

Del Teorema Fundamental de Cálculo se sigue que para cada  $t \in [a, b]$  se tiene

$$s(t) = \int_a^t \|\gamma'(t)\| dt.$$

Finalmente, si consideramos  $t = b$  obtenemos

$$s(b) = \ell(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

□

Nos encontramos preparados para proponer una función distancia en  $\mathbb{R}^n$  con ayuda de la longitud de curva. Definamos **la distancia  $\rho$  entre los puntos  $\mathbf{x}$  y  $\mathbf{y}$  en  $\mathbb{R}^n$**  como: el ínfimo de las longitudes de las curvas de clase  $\mathcal{C}^1$  en  $\mathbb{R}^n$  con extremos los puntos  $\mathbf{x}$  y  $\mathbf{y}$ , es decir,  $\rho(\mathbf{x}, \mathbf{y})$  es igual a

$$\inf \{ \ell(\gamma) : \gamma \text{ es una curva de clase } \mathcal{C}^1 \text{ con extremos } \mathbf{x} \text{ y } \mathbf{y} \}.$$

Con los resultados que hemos probado hasta el momento, podemos encontraremos el valor exacto para  $\rho(\mathbf{x}, \mathbf{y})$ . Primero debemos notar que de la desigualdad (9) se sigue que

$$\|\mathbf{x} - \mathbf{y}\| \leq \rho(\mathbf{x}, \mathbf{y}).$$

Nuestro siguiente paso es mostrar que existe una curva  $\gamma$  de clase  $\mathcal{C}^1$  con extremos  $\mathbf{x}$  y  $\mathbf{y}$  tal que  $\|\mathbf{x} - \mathbf{y}\| = \ell(\gamma)$ , así podremos concluir que

$$\rho(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$



Consideremos la curva  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$  de clase  $\mathcal{C}^1$ , definida mediante  $\gamma(t) = t\mathbf{x} + (1-t)\mathbf{y}$ , la cual tiene puntos extremos  $\mathbf{x}$  y  $\mathbf{y}$ . Además, la derivada  $\gamma'(t) = \mathbf{x} - \mathbf{y}$ , para cada  $t \in [0, 1]$ . Entonces

$$\begin{aligned}\ell(\gamma) &= \int_0^1 \|\gamma'(t)\| dt = \int_0^1 \|\mathbf{x} - \mathbf{y}\| dt, \\ &= \|\mathbf{x} - \mathbf{y}\| \int_0^1 dt = \|\mathbf{x} - \mathbf{y}\|.\end{aligned}$$

Hemos probado que la función distancia  $\rho: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  tal que  $\rho(\mathbf{x}, \mathbf{y})$  es igual a

$$\inf\{\ell(\gamma) : \gamma \text{ es una curva de clase } \mathcal{C}^1 \text{ con extremos } \mathbf{x} \text{ y } \mathbf{y}\}$$

es la métrica usual euclidiana. Nótese que hemos encontrado una curva  $\gamma$  en  $\mathbb{R}^n$  de clase  $\mathcal{C}^1$  cuya longitud realiza la distancia entre los puntos  $\mathbf{x}$  y  $\mathbf{y}$ , es decir,  $\ell(\gamma) = \rho(\mathbf{x}, \mathbf{y})$ , a este tipo de curvas se les llama **geodésicas**.

## 4. Métrica riemanniana.

Vamos a extender varias de las ideas descritas en la sección anterior con el objetivo de introducir el concepto de métrica riemanniana sobre subconjuntos abiertos y conexos de  $\mathbb{R}^n$ . Comenzaremos definiendo el espacio ambiente (espacio vectorial) en el cual vive la derivada de una curva y, seguidamente introduciremos la definición de métrica riemanniana junto con algunos ejemplos y propiedades.

Dado el subconjunto abierto y conexo  $U$  de  $\mathbb{R}^n$ , entonces el **plano tangente** asociado al punto  $\mathbf{x} = (x_1, \dots, x_n) \in U$  es el conjunto

$$T_{\mathbf{x}}U := \{\mathbf{x}\} \times \mathbb{R}^n.$$

Nótese que el plano tangente es isomorfo al espacio vectorial  $\mathbb{R}^n$ , es decir, existe una transformación lineal biyectiva  $L: T_{\mathbf{x}}U \rightarrow \mathbb{R}^n$ , cuya inversa  $L^{-1}$  también es una transformación lineal. Si  $\gamma: I \rightarrow U$  es una curva de clase  $\mathcal{C}^1$ , entonces para cada  $t \in I$ , la derivada  $\gamma'(t)$  es un vector en el espacio  $T_{\gamma(t)}U$ .

Una **métrica riemanniana**  $g$  en el subconjunto abierto y conexo  $U \subset \mathbb{R}^n$  es una correspondencia (función), la cual asocia a cada punto  $\mathbf{x} \in U$  un producto interno  $g(\mathbf{x}, \cdot) := \langle \cdot, \cdot \rangle_{\mathbf{x}}$  sobre el espacio tangente  $T_{\mathbf{x}}U$ ,

$$g(\mathbf{x}, \cdot) = \langle \cdot, \cdot \rangle_{\mathbf{x}}: T_{\mathbf{x}}U \times T_{\mathbf{x}}U \rightarrow \mathbb{R}. \quad (19)$$

Además, esta correspondencia varía continua y diferenciablemente sobre los puntos de  $U$ . Dicho de otra manera, la función  $g_{i,j}: U \rightarrow \mathbb{R}$ , definida mediante

$$g_{i,j}(\mathbf{x}) := g(\mathbf{x}, e_i, e_j) = \langle e_i, e_j \rangle_{\mathbf{x}},$$

es diferenciable<sup>7</sup>, para cada  $i, j \in \{1, \dots, n\}$ . Un abierto  $U$  de  $\mathbb{R}^n$  dotado con una métrica riemanniana se le conoce con el nombre de **variedad riemanniana**.

**Ejemplo 4.1.** La **métrica riemanniana usual**  $g_e$  en  $\mathbb{R}^n$  asocia a cada punto  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  el producto interno  $\langle \cdot, \cdot \rangle_{\mathbf{x}}$  en el plano tangente  $T_{\mathbf{x}}\mathbb{R}^n$ , tal que  $g(\mathbf{x}, \cdot)$  es el producto interno usual, es decir

$$g_e(\mathbf{x}, \cdot) := \langle \cdot, \cdot \rangle_e. \quad (20)$$

La función  $g_{i,j}: \mathbb{R}^n \rightarrow \mathbb{R}$ , está definida mediante

$$g_{i,j}(\mathbf{x}) = g_e(\mathbf{x}, e_i, e_j) = \langle e_i, e_j \rangle_e = \delta_{i,j},$$

para cada  $i, j \in \{1, \dots, n\}$ , entonces  $g_{i,j}$  es diferenciable porque es una función constante.

**Ejemplo 4.2.** Consideremos el abierto

$$U = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : x_n > 0\} \subset \mathbb{R}^n,$$

entonces la correspondencia  $g$ , la cual asocia a cada punto  $\mathbf{x} \in U$  el producto interno  $g(\mathbf{x}, \cdot)$  en el plano tangente  $T_{\mathbf{x}}U$ , tal que

$$g(\mathbf{x}, \cdot) := \frac{\langle \cdot, \cdot \rangle_e}{x_n^2},$$

es una métrica riemanniana. Nótese que la función  $g_{i,j}: U \rightarrow \mathbb{R}$  está definida mediante

$$g_{i,j}(\mathbf{x}) = g(\mathbf{x}, e_i, e_j) = \frac{\langle e_i, e_j \rangle_e}{x_n^2} = \frac{\delta_{i,j}}{x_n^2},$$

para cada  $i, j \in \{1, \dots, n\}$ . Entonces la función  $g_{i,j}$  es diferenciable porque sus derivadas parciales existen y son continuas. El abierto  $U$  con la métrica  $g$  se le llama el **el espacio hiperbólico  $n$  dimensional, modelo del semi espacio**.

**Ejemplo 4.3.** Consideremos el abierto

$$U = \left\{ \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : \|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2} < 1 \right\} \subset \mathbb{R}^n,$$

entonces la correspondencia  $g$ , la cual asocia a cada punto  $\mathbf{x} \in U$  el producto interno  $g(\mathbf{x}, \cdot)$  en el plano tangente  $T_{\mathbf{x}}U$ , tal que

$$g(\mathbf{x}, \cdot) = \frac{4\langle \cdot, \cdot \rangle_e}{(1 - \|\mathbf{x}\|^2)^2},$$

es una métrica riemanniana. Nótese que la función  $g_{i,j}: U \rightarrow \mathbb{R}$  está definida mediante

$$g_{i,j}(\mathbf{x}) = g(\mathbf{x}, e_i, e_j) = \frac{4\langle e_i, e_j \rangle_e}{(1 - \|\mathbf{x}\|^2)^2} = \frac{4\delta_{i,j}}{(1 - \|\mathbf{x}\|^2)^2},$$

para cada  $i, j \in \{1, \dots, n\}$ . Entonces la función  $g_{i,j}$  es diferenciable porque sus derivadas parciales existen y son continuas. El abierto  $U$  con la métrica  $g$  se le llama el **el espacio hiperbólico  $n$  dimensional, modelo de la bola**.

Dado que la métrica riemanniana  $g$  asocia a cada punto  $\mathbf{x}$  del subconjunto abierto y conexo  $U \subset \mathbb{R}^n$  el producto interior  $g(\mathbf{x}, \cdot)$  sobre el espacio tangente  $T_{\mathbf{x}}U$ , entonces definimos la **norma** o **magnitud** del vector  $\mathbf{v} \in T_{\mathbf{x}}U$  mediante

$$\sqrt{g(\mathbf{x}, \mathbf{v}, \mathbf{v})}.$$

<sup>7</sup>Si el lector aún no está familiarizado con el concepto de diferenciabilidad le sugerimos revisar [MT98].

Si  $\gamma: [a, b] \rightarrow U$  es una curva de clase  $\mathcal{C}^1$ , entonces para cada  $t \in I$  la derivada  $\gamma'(t)$  es un vector en el espacio vectorial  $T_{\gamma(t)}U$  cuya norma está dada mediante

$$\sqrt{g(\gamma'(t), \gamma'(t), \gamma'(t))}.$$

Adicionalmente, la longitud de  $\gamma$  con respecto a la métrica riemanniana  $g$  está definida mediante

$$\ell_g(\gamma) := \int_a^b \sqrt{g(\gamma'(t), \gamma'(t), \gamma'(t))} dt. \quad (21)$$

Con ayuda de la longitud de una curva definimos la **distancia  $\rho$  entre dos puntos  $\mathbf{x}$  y  $\mathbf{y}$  en  $U$**  como: el ínfimo de las longitudes con respecto a la métrica riemanniana  $g$  de las curvas de clase  $\mathcal{C}^1$  con puntos finales  $\mathbf{x}$  y  $\mathbf{y}$ , es decir,  $\rho(\mathbf{x}, \mathbf{y})$  es igual a

$$\inf\{\ell_g(\gamma) : \gamma \text{ es una curva de clase } \mathcal{C}^1 \text{ con extremos } \mathbf{x} \text{ y } \mathbf{y}\}. \quad (22)$$

Vamos a enunciar dos resultados que se siguen de la función  $\rho$ . No los vamos a probar. Estas afirmaciones y otras más se extienden a otros objetos llamados **variedades diferenciales** (véase e.g., [dC92], [Lee18]).

**Teorema 4.1.** La función  $\rho: U \times U \rightarrow \mathbb{R}$ , tal que  $\rho(\mathbf{x}, \mathbf{y}) =$

$$\inf\{\ell_g(\gamma) : \gamma \text{ es una curva de clase } \mathcal{C}^1 \text{ con extremos } \mathbf{x} \text{ y } \mathbf{y}\},$$

es una métrica, es decir, la función  $\rho$  satisface las siguientes propiedades:

- i. **Definida positiva.** Para cada  $\mathbf{x}, \mathbf{y} \in U$  se cumple  $0 \leq \rho(\mathbf{x}, \mathbf{y}) < \infty$ . La igualdad se cumple si y solo si  $\mathbf{x} = \mathbf{y}$ .
- ii. **Simetría.** Para cada  $\mathbf{x}, \mathbf{y} \in U$  se cumple  $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{y}, \mathbf{x})$ .
- iii. **Desigualdad triangular.** Para cada  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in U$  se cumple  $\rho(\mathbf{x}, \mathbf{y}) \leq \rho(\mathbf{x}, \mathbf{z}) + \rho(\mathbf{z}, \mathbf{y})$ .

**Teorema 4.2.** La topología  $\tau(\rho)$  definida por la métrica  $\rho$  en  $U$  coincide con la topología (subespacio) en  $U$  heredada por topología usual de  $\mathbb{R}^n$ .

## Agradecimientos.

El autor fue apoyado por la Universidad Nacional de Colombia, Sede Manizales. Además, agradece a los miembros del Semillero de Geometría y Dinámica (código HERMES 50953) por sus valiosos comentarios y apreciaciones: Martín Carmona Díaz, Omar Steven Gutiérrez Rojas, Julio César Ospina Bustos y Andrés Steven Puertas Londoño. Camilo Ramírez Maluendas ha dedicado este trabajo a su hermosa familia: Marbella y Emilio, en apreciación por su amor y cariño.

## Referencias.

- [Cro67] Michael J. Crow, *A History of Vector Analysis, The Evolution of the Idea of a Vectorial System.*, Notre Dame, Indiana: University of Notre Dame Press, 1967.
- [dC16] Manfredo do Carmo, *Differential Geometry of Curves & Surface*, Second Edition, Dover Publications, Inc., 2016.
- [dC92] Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.
- [Eng89] Ryszard Engelking, *General Topology*, Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag Berlin, 1989.
- [FILE89] Stephen H. Friedberg, Arnold J. Insel, and Spence Lawrence E., *Linear Algebra*, Second Edition, Prentice-Hall, 1989.
- [Kur66] K. Kuratowski, *Topology*, Second Edition, Vol. 1, Academic Press, Inc., 1966.
- [Kur68] ———, *Topology*, Second Edition, Vol. 2, Academic Press, Inc., 1968.
- [Lee18] John M. Lee, *Introduction to Riemannian Manifolds*, Graduate Text in Mathematics, vol. 176, Springer, 2018.
- [MT98] Jerrold E. Marsden and Anthony J. Tromba, *Cálculo Vectorial*, Cuarta Edición, Addison Wesley Longman, 1998.
- [PVRV08] Óscar A. Palmas Velasco and J. Guadalupe Reyes Victoria, *Curso de Geometría Diferencial. Parte 1. Curvas y superficies*, Facultad de Ciencias, UNAM, Las prensas de ciencias, 2008.
- [Rud76] Walter Rudin, *Principles of mathematical analysis*, International series in pure and applied mathematics, McGraw-Hill, Inc., 1976.
- [Ser86] Lang Serge, *Introduction to Linear Algebra*, Second Edition, Undergraduate Texts in Mathematics, Springer-Verlag, 1986.
- [Shi13] Theodore Shifrin, *Differential Geometry: A First Course in Curves and Surfaces*, Preliminary Version, Preprint, 2013.
- [Spi70] Michael Spivak, *Calculus, Cálculo Infinitesimal*, Vol. 1, Editorial Reverté, S.A., 1970.
- [Spi79] ———, *A Comprehensive Introduction to Differential Geometry*, Second Edition, Vol. II, Publish or Perish, Inc., 1979.
- [TF98] Juan. Tarrés Freixenet, *Sobre la historia del concepto topológico de curva*, Gaceta de la Real Sociedad Matemática Española **1** (1998), 60–7.

**Acerca del autor:** El profesor Camilo estudia el problema de realización de grupos como grupo de Veech de superficies de translación de tipo infinito. También investiga las posibles estructuras holomorfas sobre las superficies de tipo infinito. Ha despertado una curiosidad por el café, lo disfruta en varias ocasiones del día y, goza el café con notas cítricas chocolatadas.

# INTEGRACIÓN NUMÉRICA AL RESCATE: DOS ALTERNATIVAS PARA INTEGRAR $f(x) = e^{-x^2}$ CON PYTHON

Marco Blanco Ariza \*

marco.blanco01@correo.usa.edu.co

## 1. Introducción.

En algunos casos, encontrar el valor de una integral definida de una función  $f : [b, c] \rightarrow \mathbb{R}$ , integrable sobre el intervalo compacto  $[b, c]$ , con  $b, c \in \mathbb{R}$ , puede no resultar fácil. Esto puede en parte atribuirse a la *imposibilidad* de expresar sus antiderivadas (primitivas) en términos de *funciones elementales*<sup>1</sup>. El hecho anterior fue establecido por el matemático francés Joseph Liouville, quien demostró que bajo ciertas condiciones especiales, la antiderivada de una función puede expresarse en términos de funciones elementales (ver sección 4 de [Con05]). En esta situación, el cálculo de la integral definida se puede tratar (es decir, aproximar dicho valor) por diversos métodos numéricos, entre ellos: la *regla de los trapecios*, el *método de Simpson* o los *métodos de Monte Carlo*. En el presente artículo, se exponen dos métodos para aproximar numéricamente el resultado de integrar  $f(x) = e^{-x^2}$  en  $[0, 1/3]$ ; función cuyas antiderivadas no se pueden expresar en términos elementales<sup>2</sup>.

Por último, como parte importante de este texto, se presenta una aplicación para cada método a partir de un algoritmo en Python, con el fin de mostrar cómo la programación funcional facilita llevar a la práctica los métodos numéricos expuestos.

## 2. Teorema Fundamental del Cálculo.

Uno de los objetivos más importantes de todo curso de cálculo integral es el estudio del *Teorema Fundamental del Cálculo*. Es usual que este importante resultado se presente en dos enunciados; para nuestro interés, recordamos aquel que dice: Sea  $f : I \rightarrow \mathbb{R}$ , una función continua en el intervalo abierto  $I \subset \mathbb{R}$ . Si  $b, c \in I$  y  $F : [b, c] \rightarrow \mathbb{R}$  es una función primitiva (cualquiera) de  $f$  en  $[b, c]$ , entonces se verifica que:

$$\int_b^c f(x) dx = F(c) - F(b). \quad (1)$$

La expresión (1) se conoce como *Regla de Barrow*.

Como se observa, aplicar la regla de Barrow para calcular la integral definida de  $f$  sobre el compacto  $[b, c]$ , implica

\*Estudiante de Matemáticas, Universidad Sergio Arboleda.

<sup>1</sup>En [Mor15] se presentan ejemplos al respecto.

<sup>2</sup>Usando coordenadas polares, es posible calcular el valor de la integral de esta función sobre el eje real  $(-\infty, \infty)$ . Es decir, existe la posibilidad de realizar el cálculo preciso sobre toda la recta real pero no en un caso finito.

conocer necesariamente al menos una primitiva de  $f$ . Sin embargo, a pesar de que  $f$  admita primitivas, no siempre éstas se pueden expresar en términos de funciones elementales, por ejemplo, la función  $f(x) = e^{-x^2}$  no es una función elemental<sup>3</sup>. La proposición anterior se demuestra usando el *Teorema de Liouville*, en el cual se establecen condiciones para determinar si la primitiva de una función se puede expresar o no en términos de funciones elementales (ver [Con05]).

Por lo anterior, y dado que en muchas ocasiones interesa más conocer el resultado numérico, se presenta a continuación una aproximación de la solución de  $\int_b^c e^{-x^2} dx$  por el método de Monte Carlo y aproximando  $f(x) = e^{-x^2}$  por Polinomios de Taylor<sup>4</sup>.

## 3. El método Monte Carlo.

La simulación de Monte Carlo es un método no determinista que permite aproximarse numéricamente a la solución de ciertas expresiones matemáticas complejas (por ejemplo: ecuaciones diferenciales parciales con dominio de solución infinito), a partir de la generación de grandes cantidades de números *pseudoaleatorios*<sup>5</sup>. Gracias a su utilidad, este método es ampliamente utilizado en diversos campos de investigación, por ejemplo: en finanzas, facilita la valoración de *opciones*; en física, ayuda a simular el transporte de la radiación en la materia, a estudiar sistemas cuánticos y agregados atómicos; en física médica, ayuda a solucionar problemas de dosimetría; en administración, permite la optimización de modelos de administración de inventarios, entre otras.

Para implementar el método de Monte Carlo, es necesario realizar dos supuestos importantes (ver [Saa08]):

1.  $X$  es una variable aleatoria que se distribuye uniforme en el intervalo  $[0, 1]$ . Por lo tanto, su función de densidad de probabilidad está dada por:

$$f_X(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1, \\ 0 & \text{en otro caso.} \end{cases} \quad (2)$$

<sup>3</sup>Aunque la demostración de este importante hecho se encuentra fuera del alcance de este artículo, el lector la puede consultar en [Con05] ejemplo 4.6.

<sup>4</sup>En [Tho06] sección 8.7, se presentan la *regla de los trapecios* y el *método de Simpson*.

<sup>5</sup>Hace referencia a una sucesión de números generados a partir de un algoritmo, los cuales se aproximan de buena manera a un conjunto de números aleatorios.

2. Si  $Y = G(X)$ , entonces:

$$E[G(X)] = \int_0^1 G(X) f_X(x) dx = \int_0^1 G(X) dx \approx \frac{1}{n} \sum_{i=1}^n G(x_i). \quad (3)$$

La aproximación de  $E[G(X)]$  por el promedio de los  $G(X_i)$  se da gracias a la *ley de los grandes números*, para un número  $n$  muy grande de valores.

A partir de la expresión (3) se puede integrar cualquier función siempre y cuando esté definida en el intervalo  $[0, 1]$ . Sin embargo, esto sigue siendo poco útil, ya que no siempre se requiere integrar en dicho intervalo. Para poder generalizar el resultado y que se pueda integrar cualquier función en cualquier intervalo, se siguen los siguientes pasos:

- Se parte de:

$$\int_b^c G(x) dx. \quad (4)$$

- Para poder transformar los límites de integración de (4), se crea la siguiente variable:

$$u = \frac{x-b}{c-b}. \quad (5)$$

Si  $b = 0$  y  $c = 1$  entonces  $u = x$ .

- Despejando  $x$  en (5) y diferenciando a ambos lados, se llega a:

$$x = b + (c-b)u. \quad (6)$$

$$dx = (c-b) du. \quad (7)$$

- Sustituyendo (6) y (7) en (4) y empleando la *ley de los grandes números*, se obtiene:

$$= \int_0^1 G(b + (c-b)u) (c-b) du \quad (8)$$

$$= (c-b) \int_0^1 G(b + (c-b)u) du \approx \frac{c-b}{n} \sum_{i=1}^n G(b + (c-b)u_i). \quad (9)$$

Donde  $u_i$  denota números pseudoaleatorios generados con una distribución uniforme en el intervalo  $[0, 1]$ . A partir de (9) podemos integrar cualquier función en cualquier intervalo.

### 3.1. El algoritmo.

A continuación, se presenta la implementación de (9) en Python<sup>6</sup>. Se aclara que se crean dos funciones: la primera, define la función a integrar; la segunda, define la función que realiza la integración. Asimismo, en algunas líneas se

incluyen comentarios para describir lo que el código está realizando:

Paso 1: importar las librerías necesarias

```
import numpy as np
```

Paso 2: definir función a integrar

```
def funcion_a_integrar(x):
    return np.exp(-x**2)
```

Paso 3: definir la función que integra

```
def integracion_monte_carlo(funcion_a_integrar,
    ,limiteinf,limitesup,n):
    suma = 0
    for i in range(n): ##Particiona, evalúa y acumula.
        suma += funcion_a_integrar(limiteinf+
            (limitesup-limiteinf)*np.random.uniform(0,1,1))
    return ((limitesup-limiteinf)/n)*suma[0]
```

Figura 1: Integración Monte Carlo con Python.

Para probar el algoritmo, definimos como función a integrar  $f(x) = e^{-x^2}$  en el intervalo  $[0, \frac{1}{3}]$  y variamos  $n = 100$ ,  $n = 1.000$ ,  $n = 10.000$ ,  $n = 100.000$  y  $n = 1.000.000$ . A continuación, se presentan los resultados:

| $n$       | $\int_0^{\frac{1}{3}} e^{-x^2} dx$ |
|-----------|------------------------------------|
| 100       | 0.320895                           |
| 1.000     | 0.321610                           |
| 10.000    | 0.321296                           |
| 100.000   | 0.321373                           |
| 1.000.000 | 0.321379                           |

Tabla 1: Aproximación numérica para diferentes tamaños de muestra ( $n$ ).

Si se calcula la integral por medio del software *Mathematica*, el valor que se obtiene es 0.321389, aproximadamente. Como se observa, a partir de  $n = 100.000$  se pudo obtener un valor cercano al obtenido con *Mathematica*; en ese sentido, se puede concluir que la aproximación numérica por Monte Carlo mejora conforme  $n \rightarrow \infty$ .

### 3.2. Polinomios de Taylor.

Los polinomios de Taylor (ver [Apo67] pág. 333), como una forma de aproximar funciones, resultan ser una herramienta de trabajo muy cómoda, ya que evaluar un valor implica realizar una cantidad finita de adiciones y multiplicaciones, lo cual permite concluir que es más fácil evaluar un polinomio que una función exponencial. En ese sentido, el objetivo de esta sección es hallar el polinomio de Taylor asociado a la función exponencial  $f(x) = e^{-x^2}$ . Con este propósito, a continuación se introduce el concepto de *polinomios de Taylor*.

<sup>6</sup>En [Arm19] se presenta otra opción en Python y R.



Si  $f: U \rightarrow \mathbb{R}$  es una función  $n-1$  veces derivable en entorno  $U$  de un punto  $a \in \mathbb{R}$ , y  $n$  veces derivable en  $a$ , entonces,  $f$  admite un desarrollo de Taylor de orden  $n \in \mathbb{N}$  en  $a$ , si para todo  $x \in U$  se tiene:

$$f(x) = p_n(x) + r_n(x) \quad \text{con} \quad r_n(x) = o[(x-a)^n].$$

Siendo  $r_n(x)$  el resto  $n$ -ésimo (notación de Landau) y  $p_n(x)$  dado por:

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Al polinomio  $p_n(x)$  se le llama Polinomio de Taylor de grado  $n$  de  $f$  en  $a$ . La expresión anterior se puede reescribir en la forma (ver [Apo67] pág. 336):

$$T_n(f) = p_n(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k. \quad (10)$$

El polinomio de Taylor generado por  $f$  de grado  $n$  se notará como  $T_n(f)$ . Para hallar  $T_n(e^{-x^2})$  en  $a=0$ , primero se hallará  $T_n(e^y)$  en  $a=0$ , así:

1. Evaluamos la función y sus derivadas en  $a=0$ :

$$\begin{aligned} f(y) &= e^y \longrightarrow f(0) = 1, \\ f'(y) &= e^y \longrightarrow f'(0) = 1, \\ &\vdots \\ f^{(k)}(y) &= e^y \longrightarrow f^{(k)}(0) = 1. \end{aligned}$$

2. Con los resultados anteriores, determinamos una expresión para los coeficientes del polinomio, así:

$$C_k = \frac{f^{(k)}(0)}{k!} = \frac{1}{k!}.$$

3. Por lo tanto, el polinomio de Taylor de grado  $n$  generado por  $e^y$  en el punto  $a=0$ , está dado por:

$$T_n(e^y) = \sum_{k=0}^n \frac{y^k}{k!}. \quad (11)$$

Una de las propiedades más útiles en los polinomios de Taylor es la de sustitución (ver [Apo67] pág. 338). En este caso, estamos buscando  $T_n(e^{-x^2})$ , por lo tanto, sustituyendo  $y = -x^2$  en (11), se llega a:

$$T_n(e^{-x^2}) = \sum_{k=0}^n \frac{(-x^2)^k}{k!} = \sum_{k=0}^n \frac{(-1)^k (x^{2k})}{k!}. \quad (12)$$

Es importante resaltar que  $f(y) = e^y$  tiene como entorno  $U$  a toda la recta real, es decir,  $U = (-\infty, \infty)$  para todo punto

$a \in \mathbb{R}$ , lo cual garantiza que al realizar el cambio de variable  $y = -x^2$  el entorno  $U$  sea el mismo, concluyendo que la sustitución se realizó correctamente.

La expresión (12) permite aproximarse a  $f(x) = e^{-x^2}$  en cada punto  $x \in \mathbb{R}$ , mediante la siguiente fórmula:

$$|e^{-x^2} - T_n(-x^2)| = |r_n(-x^2)|.$$

Donde  $|r_n(-x^2)|$  es el resto  $n$ -ésimo.

En este punto, surge la necesidad de acotar  $|r_n(-x^2)|$  con el fin de conocer cuál es el error máximo que resulta de generar el polinomio de Taylor de grado  $n$ ; la utilidad de lo anterior, radica en que, dado un nivel de aproximación requerido, es posible encontrar el grado  $n$  que debe tener el polinomio de Taylor para satisfacerlo (ver [Apo67] ejemplo 1 pág. 344). Una expresión que permite realizar dicha acotación para el caso  $n=4$  es<sup>7</sup>:

$$0 < r_4(-x^2) \leq \frac{x^{10}}{5! \cdot e^c}. \quad (13)$$

Si  $0 \leq x \leq \frac{1}{3}$  (intervalo de integración):

$$\begin{aligned} 0 < r_4(-x^2) &\leq \frac{\left(\frac{1}{3}\right)^{10}}{5! \cdot e^{\left(\frac{1}{3}\right)}} \\ 0 < r_4(-x^2) &\leq 0.0000001. \end{aligned} \quad (14)$$

Con lo cual se concluye que el error que se obtiene al aproximar  $f(x) = e^{-x^2}$  mediante  $T_4(-x^2)$  en  $[0, \frac{1}{3}]$  es, como máximo, 0.0000001. Teniendo en cuenta lo anterior, se calcula  $\int_0^{\frac{1}{3}} e^{-x^2} dx$  empleando (12 y 14) con  $n=4$ , obteniendo como resultado:

$$\begin{aligned} \int_0^{\frac{1}{3}} e^{-x^2} dx &= \frac{1}{3} - \frac{1}{3 \cdot 3^3} + \frac{1}{10 \cdot 3^5} - \frac{1}{42 \cdot 3^7} + \\ &\quad \frac{1}{216 \cdot 3^9} + r_4(-x^2) = 0.321389. \end{aligned} \quad (15)$$

Se evidencia que la aproximación por polinomios de Taylor arroja el mismo valor que el calculado con el software *Mathematica*, por lo cual se concluye que este método tiene un alto nivel de precisión. A continuación, se presenta la implementación en Python.

### 3.3. El algoritmo.

Para realizar el algoritmo, se empleó la librería Sympy, la cual permite realizar cálculos simbólicos con facilidad. Asimismo, en algunas líneas se incluyen comentarios para describir lo que el código está realizando. En ese sentido, se presenta una propuesta de algoritmo que permite evaluar el polinomio de Taylor hallado en 12:

<sup>7</sup>Ver el ejemplo 3 pág. 346 y los teoremas 7.6 y 7.7 de [Apo67].

```
import sympy as sp
import math
```

Paso 2: definir las variables

```
x = sp.symbols('x')
y = sp.symbols('y')
```

Paso 3: construir la función

```
def aprox_pol_taylor(a,b,n):
    evalinf = [] #cada elemento resulta de evaluar 'a'
                  en cada término del polinomio.
    evalsup = [] #cada elemento resulta de evaluar 'b'
                  en cada término del polinomio.

    x = a
    ## Se crea cada termino del polinomio,
    se evalua en 'a' y se guarda en evalinf.
    for i in range(n):
        evalinf.append((-1)**i*x**(2*i+1)
                       /(math.factorial(i)*(2*i+1)))

    y = b
    #se crea cada termino del polinomio,
    se evalua en 'b' y se guarda en evalsup.
    for i in range(n):
        evalsup.append((-1)**i*y**(2*i+1)
                       /(math.factorial(i)*(2*i+1)))
    #resultado límite superior menos límite inferior.
    return(sum(evalsup)-sum(evalinf))
```

Figura 2: Integración mediante Polinomios de Taylor con Python.

Al ejecutar la función<sup>8</sup> para  $b = 0$ ,  $c = \frac{1}{3}$  y  $n = 5$ , se llega a 0.321389; lo cual permite concluir que a través de los polinomios de Taylor se puede realizar integración numérica con un buen nivel de aproximación.

## 4. Conclusiones.

Del desarrollo anterior se destacan dos conclusiones:

1. Las diversas técnicas de integración numérica constituyen un recurso valioso para la matemática y los campos en donde es aplicada. Se puede apreciar cómo dos métodos diferentes permiten obtener respuestas similares, lo cual reafirma que en algunos casos no hay un único camino para solucionar un problema.
2. El uso de lenguajes de programación como Python permite a los estudiantes poder aterrizar la teoría en ejercicios prácticos que, además de afianzar los conocimientos, ayudan a desarrollar un pensamiento orientado a la programación.

## Referencias.

[Apo67] T. Apostol, *Calculus - Volumen 1*, Wiley, 1967.

<sup>8</sup>Se debe tener en cuenta que el polinomio de Taylor evaluado en 15 es de grado  $n = 4$ ; sin embargo, la función  $range(k)$  genera los números desde 0 hasta  $k - 1$ , por lo tanto, para que Python genere el polinomio de grado  $n = 4$ , se debe dejar como entrada  $n = 5$ .

[Saa08] P. Saavedra, *Integración numérica por Monte-Carlo* (2008).

[Arm19] D. Armendáriz, *Resuelve cualquier integral usando el azar—integración con montecarlo—python—R*, 2019.

[Tho06] G. Thomas, *Cálculo - una variable*, Pearson/Addison Wesley, 2006.

[Con05] B. Conrad, *Impossibility theorems for elementary integration* (2005).

[Mor15] W. Mora, *¿Se puede saber si una función tiene primitiva elemental?*, Revista digital matemática, educación e internet, 2015.

## Agradecimientos.

A los profesores Oscar Eduardo Martínez y Hermes Jackson Martínez, porque sin sus enseñanzas no hubiese sido posible la elaboración de este artículo. Al profesor Andrés Rivera Acevedo, por sus valiosos comentarios en cada revisión y al profesor John Arredondo, por la oportunidad de realizar esta publicación.

**Acerca del autor:** Marco es un estudiante de cuarto semestre de Matemáticas de la Universidad Sergio Arboleda. Entre los campos de la matemática, encuentra especial interés en la teoría de la medida, la probabilidad, el cálculo y su relación con el mundo financiero. Practica ciclismo con regularidad y espera algún día enseñar matemáticas y finanzas cuantitativas a nivel universitario.

# UNA ZAMBULLIDA EN LAS ECUACIONES CON RETARDO

Pablo Amster<sup>\*</sup>  
pamster@dm.uba.ar

Imaginemos, con ánimo matemático, que queremos darnos una ducha. Todos tenemos nuestras preferencias personales respecto de la temperatura del agua (sobre duchas no hay nada escrito), así que entramos y accionamos el grifo en pos de alcanzar ese ideal. Podemos suponer que se trata de una única llave y el agua se calienta o se enfría de modo proporcional al ángulo de rotación. Eso nos brinda al menos la esperanza de que, con un poco de destreza, la temperatura del agua se acercará cada vez más a la deseada:

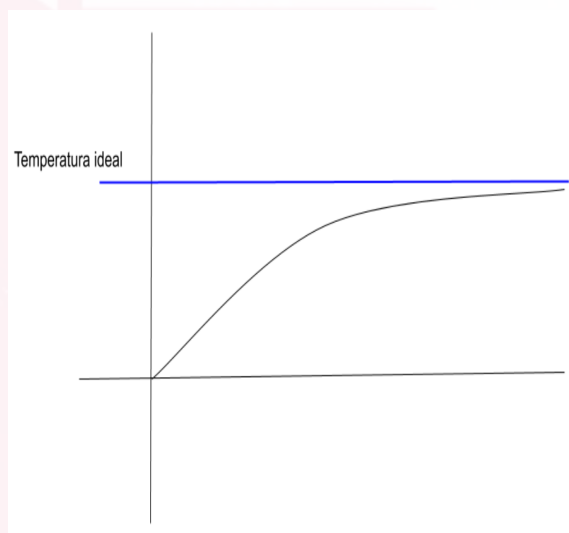


Figura 1: Una ducha sin contratiempos.

La situación anterior se puede describir mediante una ecuación diferencial, según la cual, la variación de la temperatura en cada instante  $t$  depende de la diferencia entre su valor en  $t$  y la temperatura ideal  $T_i$ . Simplificando muchísimo (la ducha no es lugar para cálculos finos), podemos pensar que la temperatura  $x$  se rige por la fórmula

$$x'(t) = a[T_i - x(t)]$$

donde  $a$  es una constante positiva. El término de la izquierda es, por supuesto, la derivada de la función  $x$ , que representa la tasa instantánea de cambio de la temperatura; el término de la derecha se puede pensar como una suerte de *control*: en cualquier momento  $t$ , si la temperatura es menor que la ideal,

entonces el término de la derecha es positivo y  $x$  aumenta, pues queremos que el agua se ponga más calentita. En cambio, cuando la temperatura  $x(t)$  es mayor que  $T_i$  movemos el grifo hacia el otro lado, para que la temperatura disminuya. Esto no es lo que se muestra en el gráfico anterior aunque podría ocurrir, por ejemplo, que nos llamen por teléfono y, al volver, el agua se haya calentado demasiado. Entonces movemos la llave para enfriarla:

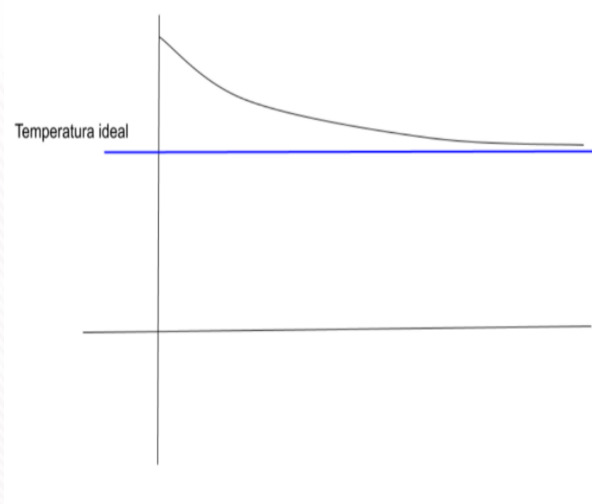


Figura 2: Un enfriado de emergencia.

Pero todo aquel que alguna vez se haya duchado sabe que la realidad no es tan idílica y, en general, está más cerca de la situación ilustrada magistralmente hace más de sesenta años el gran dibujante argentino conocido como Calé (ver Figura 3).

En efecto, lo habitual es pasarse una y otra vez de frío y calor hasta, por fin, lograr una temperatura aceptable. Las causas de tan penosa experiencia son fáciles de explicar: en general, el resultado de nuestro control sobre la mezcla de agua se manifiesta con cierto retraso, debido a la distancia entre la llave y la regadera de la ducha. La ecuación anterior se transforma entonces en

$$x'(t) = a[T_i - x(t - r)]$$

donde  $r$  expresa el tiempo que tarda nuestra acción en hacer efecto. Esto suele provocar oscilaciones en torno a la temperatura ideal (Figura 4).

No es difícil comprobar que si la distancia del grifo a la regadera es pequeña (por ejemplo, las duchas de un club en

<sup>\*</sup>Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires — IMAS-CONICET



Figura 3: Calé, *Buenos Aires en camiseta*.

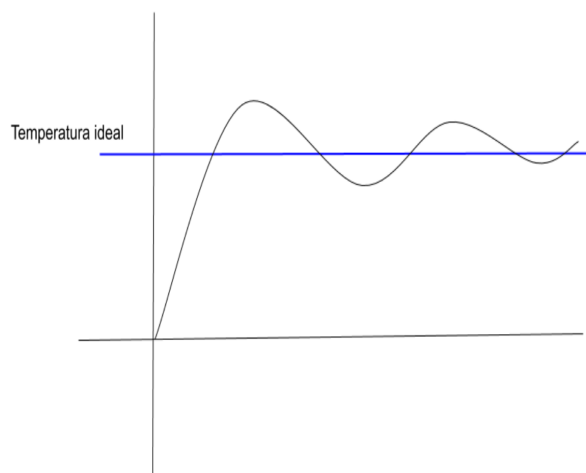


Figura 4: Dificultad para regular la temperatura.

Liliput) entonces las oscilaciones son pequeñas; en cambio, se la distancia es muy grande el valor de  $r$  aumenta y da lugar a oscilaciones cada vez mayores, incluso no acotadas (Figura 5). A modo de ejemplo, se puede suponer que  $T_i = 0$ . Esto no quiere decir que vivamos en el Ártico o nos gusten las emociones extremas sino, simplemente, al hecho de que se puede hacer el cambio de variables dado por  $y(t) = x(t) - T_i$ . Si además suponemos por simplicidad que  $a = 1$ , la ecuación resultante es muy simple:

$$y'(t) = -y(t-r).$$

Pero ahora podemos observar que para  $r = \frac{\pi}{2}$  la función  $y(t) = \sin(t)$  es precisamente una solución, ya que

$$y'(t) = \cos(t) = \sin\left(\frac{\pi}{2} - t\right) = -\sin\left(t - \frac{\pi}{2}\right) = -y(t-r).$$

La teoría de oscilaciones permite probar que este valor está justamente en el límite de que la ducha se transforme en un

acot.png

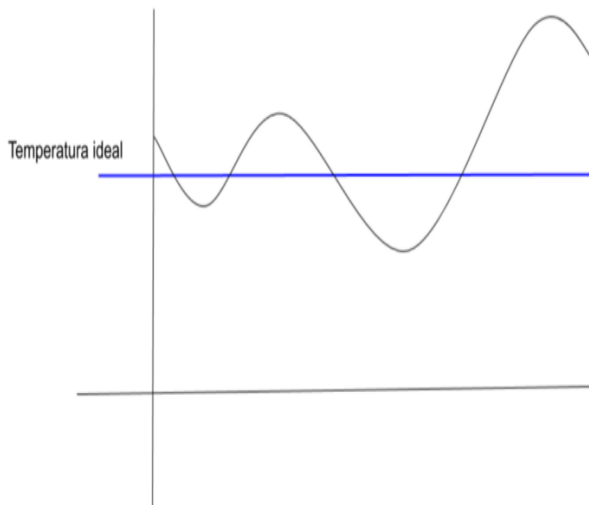


Figura 5: Oscilaciones no acotadas.

verdadero suplicio: si  $r > \frac{\pi}{2}$ , entonces hay al menos una solución con oscilaciones no acotadas. Sin llegar a tanto, se demuestra también que si el retardo es mayor que  $\frac{1}{e}$  entonces todas las soluciones oscilan. Este dato es muy útil, aunque no es fácil llevarlo a la práctica, por ejemplo, cuando uno efectúa una reserva de hotel, a menos que esté dispuesto a sostener un diálogo como el siguiente:

- Diga, ¿tiene habitación con baño privado?
- Claro que sí.
- ¿Me podría indicar la longitud de la cañería, por favor?

Las ecuaciones diferenciales con retardo sirven para modelar distintos fenómenos: entre ellos, uno de los más conocidos es el de la dinámica poblacional. La historia se remonta a los trabajos del clérigo y economista Malthus [7] quien, muy a tono con la temática de este artículo, murió en 1834 en la ciudad inglesa de *Bath*. El modelo original de Malthus es muy simple: si se supone una tasa  $b$  de nacimientos y una tasa  $d$  de muertes, entonces la población total  $P$  crece en cada instante  $t$  una cantidad proporcional  $bP(t)$  y disminuye una cantidad  $dP(t)$ , es decir:

$$P'(t) = bP(t) - dP(t).$$

Sin embargo, este modelo no es muy realista y da lugar a un crecimiento exponencial de la población, cosa que no ocurre ni en las mejores familias de conejos. O también puede ser decrecimiento, en caso de que  $b$  sea menor que  $d$ ; de todas formas, es claro que la ecuación no captura el comportamiento real de una población. Existen otros modelos, como el de Verhulst [10], que suponen una población que se autorregula, condicionada por cierto factor que limita su crecimiento: por ejemplo, la falta de algún recurso:

$$P'(t) = aP(t)[M - P(t)].$$



La ecuación de Verhulst, también llamada *ecuación logística*, tiene aplicación también en otros contextos, como las redes sociales.

Llegado este punto, podemos preguntar: ¿qué tiene que ver la ducha con la dinámica poblacional? Dejando de lado el hecho de que algunas poblaciones pueden ser más limpias que otras, lo que nos interesa considerar aquí es el retardo, que tiene, en este contexto, una interpretación muy razonable. Volvamos por un momento a la ecuación de Malthus, en la que el término  $-dP(t)$  es inobjetable, pues los individuos que mueren lo hacen una vez, con la posible excepción del señor Valdemar, aquel personaje del cuento de Poe que se mantuvo entre la vida y la muerte durante varias semanas [9]. Pero dejando de lado los casos extraordinarios, es razonable pensar que la tasa de mortandad provoca una caída instantánea de la población. En cambio, con los individuos que nacen la historia es otra, ya que transcurre un tiempo hasta que son capaces, a su vez, de reproducirse. Por eso es natural colocar allí un retardo, que permite dar cuenta del tiempo que transcurre desde que un individuo nace hasta que alcanza la madurez:

$$P'(t) = bP(t-r) - dP(t).$$

A la luz de algunos casos que conocemos, se podrá objetar que algunos individuos no alcanzan nunca tal madurez. Pero esto va más allá de un comentario jocoso ya que, en efecto, puede ocurrir que un individuo muera antes de alcanzar la edad reproductiva. Por tal motivo, en muchos modelos se incluye en el término correspondiente un factor adicional que contempla la probabilidad de llegar a la madurez. Tal es el caso del modelo de Nicholson [8], en el que se asume una distribución exponencial dependiente de la población total:

$$P'(t) = be^{-P(t-r)}P(t-r) - dP(t).$$

El comportamiento de este tipo de ecuaciones es muy diferente al caso sin retardo. Por ejemplo, la ecuación logística se resuelve por completo; en cambio, su versión con retardo propuesta por Hutchinson [4]

$$P'(t) = aP(t)[M - P(t-r)]$$

es mucho más complicada, a tal punto que una pregunta surgió a partir de esta ecuación se mantuvo sin respuesta por más de cincuenta años. Cabe mencionar que la motivación original proviene de estudiar la densidad de una población de carácter muy diferente: no se trata de personas, bacterias o moscas (como en el modelo de Nicholson) sino de números primos. Más específicamente, lo que Hoffman de Visme [3] intentó hacer, basándose en una idea previa de Cherwell, es dar un argumento heurístico para el célebre teorema de los números primos, según el cual la cantidad  $\pi(N)$  de primos menores que un cierto  $N$  es, aproximadamente,  $\frac{N}{\log N}$ . El resultado es conocido y su historia está ligada a la no menos célebre *hipótesis de Riemann*, cuya verdad o falsedad aún no ha podido demostrarse. Pero, el teorema de los números primos fue demostrado ya a fines del siglo XIX; las pruebas de Hadamard y de la Vallée Pousin no requieren toda la fuerza de la hipótesis de Riemann.

Sin embargo, se trata de resultados trabajosos y de Visme pretendió encontrar una explicación más elemental. Lo notable es que su planteo deriva en una ecuación diferencial: concretamente, si  $y(x)$  es la mejor aproximación suave de  $\pi(x)$ , entonces de Visme mostró que vale

$$y''(x) = -\frac{y'(x)y'(\sqrt{x})}{2x}.$$

Cabe decir que esta ecuación no tiene un aspecto muy amigable, así que el autor decidió consultar a un especialista en el tema, el inglés E. M. Wright, quien propuso un ingenioso cambio de variables:

$$y'(x) \ln x = z, \quad 2^t = \ln x.$$

De esta forma, la ecuación se convierte en

$$z'(t) = \ln 2 z(t)[1 - z(t-1)]$$

que no es otra que la ecuación de Hutchinson con  $a = \ln 2$  y  $M = r = 1$ . Wright logró demostrar entonces que la solución constante  $z \equiv 1$  es un atractor de todas las soluciones positivas, es decir: si  $z(t) > 0$  es solución entonces  $z(t) \rightarrow 1$  para  $t \rightarrow +\infty$ . Y esto permite ver lo que quería de Visme, usando un truco muy conocido por los estudiantes, la regla de L'Hôpital:

$$\lim_{x \rightarrow +\infty} \frac{y(x) \ln x}{x} = \lim_{L'H} \left( y'(x) \ln x + \frac{y(x)}{x} \right) = 1$$

Wright probó que su resultado no solo vale para  $a = \ln 2$  sino para cualquier  $a \leq \frac{3}{2}$ ; sin embargo, para la ecuación linealizada es fácil ver que la atractividad del equilibrio se mantiene todavía un poquito más allá, para  $a < \frac{\pi}{2}$ . La cercanía entre las dos cantidades le llamó la atención, de modo que en 1955 formuló la conjetura de que la última condición es también suficiente para la ecuación no lineal. La solución llegó, como no podía ser de otra manera, con algo de retraso: recién en 2017 J. B. van den Berg y J. Jaquette [2] lograron probar que Wright estaba en lo cierto.<sup>1</sup>

Por supuesto, el mundo de ecuaciones con retardo no acaba aquí: existen modelos para todas las situaciones y todos los gustos. Un problema especialmente importante en muchas aplicaciones es el de encontrar órbitas periódicas, para el que se han desarrollado muy diversos métodos. Entra entonces en escena uno de los más grandes matemáticos de todos los tiempos, el francés Henri Poincaré, quien propuso diversas técnicas para hallar tales órbitas. Entre ellas, un método que lleva su nombre y responde a una idea encantadora y sencilla: buscar una condición inicial cuya trayectoria, al cabo de cierto intervalo de tiempo  $T$ , vuelve al punto de partida. De esta forma, la solución de la ecuación se muerde la cola para cerrarse sobre sí misma y dar lugar a la periodicidad. La novedad respecto de las ecuaciones ordinarias es que ahora una "condición inicial" no es un número o -si se trata de un sistema- un vector en  $\mathbb{R}^n$ , sino una función: para que valga el teorema de

<sup>1</sup> En inglés suena mejor: *Wright was right*, aunque la consigna no fue usada para hablar del matemático sino del arquitecto estadounidense Frank Lloyd Wright.

existencia y unicidad de soluciones en el caso de una ecuación con un retardo  $r$ , es preciso prescribir los valores de la solución en todo un intervalo de longitud  $r$ , lo que vuelve el problema mucho más complicado.

Obtener una condición inicial a partir de la cual se generan trayectorias periódicas puede parecer tan difícil como buscar una aguja en un pajar; sin embargo, existen maneras de hacerlo o bien, cuando menos, de achicar el pajar. De lo que se trata es de encontrar un punto fijo (la condición inicial), tarea que motivó algunos de los más interesantes desarrollos de la matemática del siglo XX. Claro que esto es tema para otro trabajo: por el momento, suficientemente empapados de ecuaciones diferenciales, podemos cerrar en este punto el grifo y envolvernos en la toalla.

## Apéndice: un teorema de punto fijo en dimensión infinita.

Para el lector interesado, veamos un caso sencillo en el que se puede probar la existencia de soluciones periódicas. Consideremos el anterior modelo de Nicholson, donde ahora los coeficientes  $b$  y  $d$  son funciones positivas y periódicas, con un período fijo  $T$ :

$$P'(t) = b(t)e^{-P(t-\tau)}P(t-r) - d(t)P(t). \quad (1)$$

En el caso sin retardo, la situación es fácil de resolver, pues el problema se puede escribir de la siguiente forma:

$$P'(t) = (b(t)e^{-P(t)} - d(t))P(t).$$

Para cualquier  $t$ , es claro que si el valor de la población en ese instante es muy grande, entonces el término  $b(t)e^{-P(t)}$  se hace pequeño y, en consecuencia,  $P'(t) < 0$ . Esto quiere decir que si fijamos un valor inicial  $P(0) = M \gg 0$ , entonces la trayectoria inicialmente decrece y luego no puede volver a subir hasta  $M$ : en otras palabras,  $P(t) < M$  para todo  $t > 0$ .

Supongamos ahora que vale  $b(t) > d(t)$  para todo  $t$ , entonces cuando el valor  $P(t)$  es muy cercano a 0 resulta  $b(t)e^{-P(t)} > d(t)$  y la derivada de  $P$  es positiva. Esto significa que si  $P(0) = \varepsilon$ , un valor positivo muy pequeño, entonces la trayectoria comienza creciendo y luego se mantiene siempre mayor que  $\varepsilon$ .

¿Qué conclusiones podemos sacar de todo esto? La respuesta es muy simple: a partir de resultados clásicos de ecuaciones diferenciales ordinarias, sabemos que las trayectorias que tienen valor inicial entre  $\varepsilon$  y  $M$  se encuentran definidas para todo  $t \in [0, +\infty)$ , así que tiene sentido evaluarlas en  $t = T$ . Esta es, precisamente, la aplicación  $\mathcal{P}$  definida por Poincaré, que resulta continua en virtud de la dependencia continua respecto de las condiciones iniciales. Pero, más aún, de acuerdo con lo que vimos, ya sabemos que si  $P(0) = M$  entonces  $P(T) < M$  y si  $P(0) = \varepsilon$  entonces  $P(T) > \varepsilon$ . Como las trayectorias no se cruzan, se deduce que la aplicación de Poincaré envía el intervalo  $[\varepsilon, M]$  dentro de sí mismo. En particular,

$$\mathcal{P}(\varepsilon) - \varepsilon > 0 > \mathcal{P}(M) - M.$$

Luego, por el teorema de Bolzano, existe algún valor inicial  $P_0$  tal que la correspondiente trayectoria satisface  $P(T) = P_0 = P(0)$ , dando lugar a una órbita cerrada. La condición  $b(t) > d(t)$  es bastante precisa: a modo de ejercicio, el lector puede comprobar que si la desigualdad se invierte, entonces no puede haber soluciones periódicas positivas: más aún, si  $b(t) \leq d(t)$  para todo  $t$ , entonces cualquier trayectoria con dato inicial  $P_0 > 0$  verifica que  $P(t) \rightarrow 0$  para  $t \rightarrow +\infty$ , es decir: el equilibrio trivial es un atractor global de todas las soluciones positivas.

Veamos ahora la situación análoga para el problema con  $\tau > 0$ : ¿será cierto también ahora que si  $b(t) > d(t)$  para todo  $t$  entonces hay soluciones  $T$ -periódicas positivas? La respuesta es que sí, aunque el resultado está lejos de ser tan trivial como en el caso sin retardo.

Comencemos por la cuestión planteada algunos párrafos más arriba: en este contexto, el clásico teorema de existencia y unicidad requiere que el valor inicial de la ecuación no sea un número sino una función: por ejemplo, dada  $\varphi : [-\tau, 0] \rightarrow \mathbb{R}$  continua, se prueba que existe una única solución  $P(t)$  de (1) tal que  $P = \varphi$  en  $[-\tau, 0]$ . Se trata, entonces, de hallar  $\varphi$  positiva de modo tal que la correspondiente solución vuelva a tomar los mismos valores en el intervalo  $[T - \tau, T]$ , es decir:

$$P(T+t) = \varphi(t) \quad t \in [-\tau, 0].$$

Como antes, no es difícil probar la existencia de  $M > 0$  tal que si  $\varphi(t) \leq M$  para  $t \in [-\tau, 0]$  entonces  $P(t) < M$  para  $t > 0$ . Con un poco más de esfuerzo (que queda a cargo del lector), se prueba también, bajo la hipótesis  $b(t) > d(t)$  para todo  $t$ , que existe  $\varepsilon > 0$  tal que si  $\varepsilon \leq \varphi(t) \leq M$  para todo  $t \in [-\tau, 0]$  entonces  $P(t) > \varepsilon$  para  $t > 0$ . Sin embargo, ahora el problema es en dimensión infinita y la existencia de un punto fijo  $\varphi$  de la aplicación de Poincaré ya no queda tan clara. Una primera dificultad, por mencionar alguna, es que para  $\tau > 0$  las soluciones pueden cruzarse, sin que eso contradiga la mencionada unicidad (¿por qué?). En definitiva, en este caso no hay Bolzano que valga... o quizás sí, aunque se trata de un resultado de carácter bastante diferente. Repasemos nuestra situación: en el problema sin retardo teníamos  $\mathcal{P}([\varepsilon, M]) \subset [\varepsilon, M]$ ; ahora tenemos, en cambio,  $\mathcal{P}(C) \subset C$ , donde

$$C := \{\varphi \in C[-\tau, 0] : \varepsilon \leq \varphi(t) \leq M \quad t \in [-\tau, 0]\}.$$

Un gran teorema viene entonces en nuestra ayuda, para garantizar la existencia de al menos un punto fijo en espacios de dimensión infinita; nos referimos al teorema de Schauder, que dice lo siguiente:

**Teorema 0.1.** *Sea  $E$  un espacio de Banach y  $C \subset E$  un conjunto convexo, cerrado y acotado. Sea  $F : C \rightarrow E$  continua tal que  $F(C) \subset C$  y además  $\overline{F(C)}$  es compacto. Entonces existe  $x \in C$  tal que  $F(x) = x$ .*

En nuestro caso, el teorema viene como anillo al dedo, tomando  $E = C[-\tau, 0]$ ,  $C$  el conjunto de antes y  $F = \mathcal{P}$ . Se puede suponer (sin perder generalidad) que  $\tau < T$ ; en ese caso es fácil ver -empleando otro teorema célebre, el de Arzelà-Ascoli- que la clausura de  $\mathcal{P}(C)$  es un conjunto compacto.

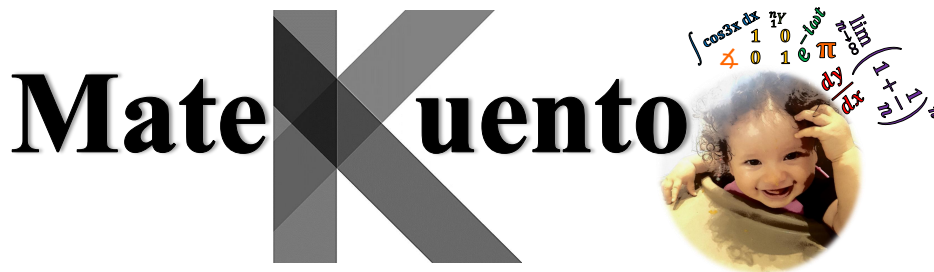
Luego, el teorema de Schauder garantiza la existencia de al menos una función  $\varphi \in C$  tal que  $\mathcal{P}(\varphi) = \varphi$ , y asunto terminado.

Tal como anunciamos, la ecuación de Nicholson es apenas un ejemplo de una gran variedad de problemas con retardo que se pueden resolver mediante métodos topológicos, que además del teorema de Schauder incluyen otros más sofisticados, así como la llamada *teoría de grado*. Existen numerosos modelos biológicos de interés por sus aplicaciones, para los cuales dichos métodos han mostrado su importancia. Pero además, la teoría de ecuaciones con retardo ofrece una amplia gama de problemas abiertos, incluso en situaciones que se encuentran completamente resueltas para el caso  $\tau = 0$ . La dinámica de las ecuaciones con retardo se revela mucho más rica y complicada: para dar apenas una idea de esto, basta mencionar que un modelo tan simple como el de Nicholson ofrece preguntas que aún esperan respuesta [1].

## Referencias.

- [1] L. Berezanski, E. Braverman and L. Idels, *Nicholson's blowflies differential equations revisited: Main results and open problems*. Applied Mathematical Modelling 34 No. 6 (2010), 1405–1417.
- [2] J. B. van den Berg, J. Jaquette, *A proof of Wright's conjecture*. Journal of Differential Equations 264 No. 12 (2018), 7412–7462.
- [3] G. Hoffman de Visme, *The density of prime numbers*. Math. Gazette 45, (1961), 13–14.
- [4] G. Hutchinson, *Circular causal systems in ecology*. Ann. N.Y. Acad. Sci. 50 (1948–1950) 221–246.
- [5] V. Kolmanovskii, A. Myshkis, *Introduction to the theory and applications of functional differential equations*. Mathematics and its Applications, Vol. 463, Kluwer Academic Publishers, 1999.
- [6] E. Liz, G. Robledo y S. Trofimchuk, *La conjetura de Wright: entre la distribución de los números primos y el crecimiento de la población*. Cubo. Matemática Educacional 3 (2001), 89–107.
- [7] T. Malthus, *An essay on the principle of population*. Oxford University Press, 1999.
- [8] A. Nicholson, *An outline of the dynamics of animal populations*. Aust. J. Zool. 2 (1954) 9–65.
- [9] E. Allan Poe, *The Facts in The Case of M. Valdemar*. American Whig Review, 1845.
- [10] P. Verhulst *Notice sur la loi que la population poursuit dans son accroissement*. Correspondance mathématique et physique 10 (1838), 113–121.

**Acerca del autor:** Profesor Titular del Departamento de Matemática de la Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Investigador Principal del CONICET. Es autor de unos 120 trabajos de investigación científica y tres libros en el área de ecuaciones diferenciales. Colabora en diferentes proyectos de investigación en universidades argentinas y extranjeras. Por otra parte, dicta con frecuencia conferencias y seminarios de divulgación, y ha publicado diversos libros destinados a un público amplio. Ha organizado varios talleres sobre matemática y música, así como numerosas “conferencias” de espíritu tanguero. Para dar una idea de este espíritu, cabe mencionar que el presente artículo se basa en algunos trabajos previos de divulgación que llevaron por título *Poincaré junto a un calefón*, una suerte de versión matemática del célebre tango *Cambalache*.



## Matando el tiempo, una lucha por la supervivencia. Confesiones de un periodo de enclaustramiento, y otras ficciones

Recluido, solitario, aguardo a que llegue el final. Es incierto el momento exacto en que ocurrirá, lo sé desde que acepté que, *a ciencia cierta*, nadie entiende lo que ha sucedido, ni siquiera somos capaces de dimensionar la severidad de las circunstancias actuales, por demás desbordadas e incontrolables.

En cada conversación que virtualmente rompe el aislamiento, me prodigo en tímidas sentencias de esperanza, son dosis escrupulosamente administradas de colectiva fe en el futuro; en el fondo, algo así como un presentimiento o una revelación si se quiere, pone en evidencia la paupérrima actuación en la que repito como un mantra los placebos que nos permiten huir momentáneamente de la angustia y la desesperación, propias del inexorable ocaso que nos espera.

Pasos de verdugo en el cadalso, eso es para mí el constante golpeteo del segundero, el mecanismo del reloj: inoportuno, sistemático, imperturbable, se atraviesa en mi mente, ocupa mis pensamientos, lo descubro omnipresente en cada recuerdo o experiencia de este último mes de confinamiento.

Las reflexiones de Søren Kierkegaard parecen muy pertinentes en estos momentos, el existencialismo y la angustia se llevan bien con estas circunstancias, sin embargo, una lectura diferente de los argumentos de Kierkegaard me hace pensar acerca de las implicaciones de una parte de su obra en el concepto de la estructura del tiempo. Postula el filósofo danés que el hombre en medio de una infinidad de alternativas puede perderse, en una pausa indefinida en la que extravía su individualidad. Enfrentado a una infinidad de alternativas, el hombre llega a asumir su individualidad sólo a través de un *instante de elección*, y ese instante resulta ser una *decisión de eternidad* en la que el hombre se abre al futuro y a todas las posibilidades que trae como consecuencia una *existencia auténtica* basada en decisiones propias.

Indirectamente Kierkegaard propone la existencia de un instante en medio de una infinidad, a su vez dicho instante contiene a la eternidad (infinidad), todo esto en el marco de un devenir de hechos. ¿Es esto posible en la estructura del tiempo? Sin algo mejor que hacer me dispongo a abordar estas cuestiones desde un punto de vista físico-matemático.

Cada tic-tac separado en el tiempo por un segundo, entre cada tic-tac se da el lapso exacto para 9.192.631.770 oscilaciones de radiación emitida por el cesio 133 a una temperatura de... ¡pamplinas! ¡eso no es el tiempo! ¿qué es el tiempo? En

rigor tampoco lo sabemos. Tal vez lo mejor que pueda decirse al respecto es que es *lo que transcurre* en medio de dos eventos, parece que hay cierto consenso en torno de una definición semejante. Entonces si no *ocurriera algo*, si no sucedieran eventos ¿se detendría el tiempo? Desde el punto de vista de la termodinámica, tal vez sí. Es indistinguible un universo en el que no ocurre nada -ni siquiera a nivel subatómico- de un universo en el que se detuvo el tiempo.

Es común encontrar en la literatura alusiones a la entropía cuando se trata el tema del sentido en el que transcurre el tiempo; de hecho, ambos elementos: la dinámica de la entropía y el sentido del tiempo suelen tratarse de manera indistinta. La cantidad de entropía o el grado de desorden en un sistema (o del universo) aumenta con el *paso del tiempo*, así lo enuncia el segundo principio de la termodinámica, la explicación para esto puede plantearse de múltiples modos, por ejemplo, desde un enfoque mecanicista la actividad a cualquier nivel (un evento) implica el uso de energía. Si, por ejemplo, se realiza un trabajo para ordenar un sistema, por el principio de Carnot una parte de la energía empleada se pierde en forma de calor, una forma desordenada de energía. Puede demostrarse que este aumento del desorden es siempre mayor que el aumento del orden que se le *aporta* al sistema, por lo que un aumento de la entropía es consecuencia de la ocurrencia de un evento y a la vez del transcurrir del tiempo.

Otra forma de ver el aumento de la entropía en relación con el *paso del tiempo* se reduce al hecho probabilístico de que hay una mayor cantidad de “estados de desorden” en relación con los “estados de orden” para cualquier sistema -es decir, el orden es por definición un estado especial-, por lo que un cambio en el sistema generalmente aumenta su desorden.

Entonces: que transcurra el tiempo es tanto como decir que hay un aumento del desorden (o de la cantidad de entropía), lo que a su vez equivale al hecho de que ocurran eventos... ese maldito tic-tac de nuevo, parte el tiempo a su conveniencia, ahora conspira en mi contra con el silencio, siento que pierdo la cordura, ¡ya no es sistemático! En situaciones de medrosa depresión su traqueteo es lento, cada vez menos frecuente, se evade... me asalta un anhelo agobiante por que este sonido no vuelva más. En angustiosa y plena desesperación, por el contrario, crepita su frecuencia *in crescendo*, pienso que llegará a ser un sonido sostenido, aun más asfixiante, como de ruido blanco.



Si el tiempo se define -al menos parcialmente- a través de la ocurrencia de eventos, ¿cuál es la *ocurrencia mínima*? Es decir, ¿qué es lo *menos* que puede pasar?

En la física clásica el tiempo y el espacio son considerados continuos, divisibles infinitamente, también localmente homeomorfos con sus equivalentes euclídeos, por lo que proponer una unidad mínima de tiempo o de espacio es ilógico.

En la relatividad general, aun cuando la medida del tiempo sea inherente a cada observador, y sea posible que el espacio-tiempo pueda deformarse, se mantiene el tratamiento del tiempo como un continuo, atribuyéndosele apenas características que lo hacen menos “suave”.

La materialidad del mundo físico nos permite asumir, que la ocurrencia de un evento debería verse representada en el cambio de la posición o de la velocidad (desplazamiento) de una onda/partícula.

En la mecánica cuántica, a diferencia de lo que ocurre en la mecánica clásica o en la relativista, para una longitud menor a la de Planck (de 35 órdenes de magnitud inferior al metro), el espacio posee una geometría distinta a la clásica; por lo que no es posible tratar distancias inferiores a la longitud de Planck en el sentido convencional del término distancia. El tiempo de Planck (de 44 órdenes de magnitud inferior al segundo) es el lapso que tarda en recorrer un fotón la longitud de Planck (a la velocidad de la luz). Por lo anterior, cualquier cosa que *ocurre* en un sentido físico, acontece en un lapso mayor o igual al tiempo de Planck; lo que ocurra en un lapso inferior habrá generado un *desplazamiento* de la onda/partícula (o de cualquier ente) inferior a la longitud de Planck, lo que por efectos cuánticos y del principio de incertidumbre no puede verificarse, no habría evidencia alguna de que el desplazamiento *haya ocurrido*.

Así las cosas, tenemos una unidad básica de tiempo: el tiempo de Planck, nada ocurre en un lapso inferior; un intervalo menor de tiempo equivale a la *no ocurrencia* de eventos, pues dichos eventos no pueden verificarse; por lo que un lapso menor al tiempo de Planck equivale a que el tiempo esté detenido. Desde el inicio del universo con el Big Bang han transcurrido aproximadamente  $8 \times 10^{60}$  tiempos de Planck, lo que no es poca cosa, pero es una cifra finita. Si asumimos que el tiempo *no terminará*, este recubrimiento hace al tiempo infinito numerable ( $\aleph_0$ ).

Sabemos que todos los intervalos abiertos de la recta son homeomorfos con el intervalo  $(0, 1)$ , por lo que, manteniendo

las unidades empleadas, el intervalo de un tiempo de Planck es homeomorfo al intervalo de tiempo transcurrido desde el Big Bang a la actualidad. Este homeomorfismo nos permite de manera inmediata concluir que el cardinal del intervalo de un tiempo de Planck es ( $\aleph_1$ ) y que éste es el mismo cardinal del intervalo de tiempo entre el Big Bang y la actualidad.

Así las cosas, al devenir de eventos o transcurrir del tiempo que es un intervalo infinito no numerable ( $\aleph_1$ ), le pertenece un instante, un intervalo de tiempo que es un elemento entre una cantidad finita de momentos de Planck y que es un subconjunto propio de un conjunto numerable ( $\aleph_0$ ), que contiene a su vez a una infinidad no numerable, al intervalo de tiempo de Planck ( $\aleph_1$ ). De alguna forma -un poco aparatosa- se ha desarrollado un argumento compatible con la estructura del tiempo propuesta indirectamente por Kierkegaard.

Se han forzado algunas definiciones, -trayendo a Cantor por los cabellos- para dar forma al argumento indirecto de Kierkegaard, tal vez mi subrepticio afán sea el de más bien, justificar de esta manera su argumento directo en procura de una actitud constructiva: en medio de un sinfín de posibilidades del mundo, la decisión que se toma en un instante de tiempo nos conduce a la eternidad e infinitud de una existencia auténtica. Del estoicismo se ha hecho famosa la frase que dice que: no son las cosas que nos pasan las que nos hacen sufrir, sino lo que nosotros nos decimos sobre esas cosas<sup>1</sup>. Es fácil caer en la trampa de sentirnos víctimas de las circunstancias, sobre todo cuando nos enfrentamos a situaciones que ponen al límite nuestro carácter. No podemos determinar todos los eventos que nos ocurren, pero sí somos dueños de la manera en que los asumimos<sup>2</sup>. Por cierto, ahora debo lidiar con el hecho de haber descubierto que, hace más de un año, no tengo un reloj que haga “tic-tac”.

**Dayron Fabián Achury-Calderón**

[dayronf.achuryc@konradlorenz.edu.co](mailto:dayronf.achuryc@konradlorenz.edu.co)

**Acerca del autor:** Fabián Achury-Calderón es ingeniero de la Universidad Distrital Francisco José de Caldas y matemático de la Fundación Universitaria Konrad Lorenz. Se ha interesado por estudiar la interacción de las matemáticas con otras ciencias. Asesora compañías en asuntos financieros y de dirección estratégica. Le apasionan la ópera, el ballet y las ciencias sociales. Sus pasatiempos favoritos son la lectura y la colección de registros de radio internacional. Disfruta recorrer museos cuando viaja y admite que se conmueve con facilidad. Algún día espera aprender a tocar piano.

<sup>1</sup>Frase atribuida a Epicteto.

<sup>2</sup>Extraído de los Aforismos Sobre el Arte de Saber Vivir de Arthur Schopenhauer.

# K' Problema

Complejos, cuaternios, octonios y muchos más, en matemáticas se encuentran sistemas numéricos para todo y para todos. Los mortales no necesitan ir tan lejos y se quedan en los reales, para los más mortales con los racionales es suficiente, y para los mortales acérrimos (la gran mayoría), con los naturales se tiene para toda la vida. Suena a que estos números, tan usados por los mortales están muy bien estudiados y muy bien comprendidos. Nada más lejos de la realidad, tan sólo en los apetecidos naturales las preguntas sin respuesta están a la orden del día. Y si vamos más allá, a un mundo no tan conocido, los irracionales,

Un número es irracional si posee infinitas cifras decimales no periódicas, esto quiere decir que no se puede expresar en forma de una fracción, o sea, no se puede representar como  $\frac{p}{q}$  para  $p, q$  dos números enteros. Por razones históricas, aplicaciones, apariciones públicas y no tan públicas, chismes y otros devenires, el número irracional más famoso es  $\pi$ , que se define como la relación entre la longitud de la circunferencia y su diámetro [BBB04]. El famoso 3,14159265359..., que podemos encontrar en presentaciones mucho más elegantes:

$$\begin{aligned}\pi &= 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}, \\ &= 2 \prod_{p \text{ odd}} \frac{p}{p + (-1)^{(p-1)/2}}, \\ &= \lim_{n \rightarrow \infty} 2^n \sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \dots \sqrt{2}}}}}}.\end{aligned}\quad (1)$$

¿Pero como estar seguros de que es irracional? Con una demostración.

A continuación desarrollaremos lo que en diversos textos y artículos aparece como una prueba elemental de la irracionalidad de  $\pi$ .

Supongamos que  $\pi^2$  es racional, esto es, puede escribirse como

$$\pi^2 = \frac{p}{q}, \quad p, q \in \mathbb{Z}$$

y tomemos la siguiente serie

$$\begin{aligned}G(x) &= b^n [\pi^{2n} f_n(x) - \pi^{2n-2} f_n''(x) + \pi^{2n-4} f_n^{(4)}(x) - \dots \\ &\quad + (-1)^n f_n^{(2n)}(x)].\end{aligned}$$

cuyos términos satisfacen que

$$b^n \pi^{(2n-2k)} = b^n (\pi^2)^{n-k} = b^n \left(\frac{a}{b}\right)^{n-k} = a^{n-k} b^k.$$

Si además, definimos

$$f_n(x) = \frac{1}{n!} \sum_{i=0}^{2n} c_i x^i,$$

con coeficientes  $c_i$  enteros, se verifica (ánimo, usted puede hacerlo) que  $f_n^{(k)}(0)$  y  $f_n^{(k)}(1)$  son enteros, esto es,  $G(0)$  y  $G(1)$  son enteros. Si derivamos  $G(x)$  dos veces se obtiene que

$$G''(x) = b^n [\pi^{2n} f_n''(x) - \pi^{2n-2} f_n^{(4)}(x) + \dots + (-1)^n f_n^{(2n+2)}(x)].$$

donde el último término es cero (¿por qué?). De esta manera, sumando  $G(x)$  y  $G''(x)$  se obtiene que

$$G''(x) + \pi^2 G(x) = b^n \pi^{2n+2} f_n(x) = \pi^2 a^n f_n(x). \quad (2)$$

Definamos ahora una nueva serie

$$H(x) = G' \sin - \pi G(x) \cos \pi x,$$

derivandola y usando la ecuación (2) se obtiene

$$H'(x) = \pi^2 a^n f_n(x) \sin \pi x.$$

que aplicando el teorema fundamental del cálculo nos permite obtener

$$\begin{aligned}\pi^2 \int_0^1 a^n f_n(x) \sin \pi x \, dx &= H(1) - H(0) \\ &= \pi [G(1) + G(0)].\end{aligned}$$

Así, podemos asegurar que

$$\pi \int_0^1 a^n f_n(x) \sin \pi x \, dx,$$

es un entero. Pero por otra parte,  $0 < f_n(x) < \frac{1}{n!}$  para  $0 < x < 1$ , de modo que

$$0 < \pi a^n f_n(x) \sin \pi x < \frac{\pi a^n}{n!},$$

y en consecuencia se obtiene que

$$0 < \pi \int_0^1 a^n f_n(x) \sin \pi x \, dx < \frac{\pi a^n}{n!}.$$

Hasta este punto nuestro razonamiento no toma en cuenta el valor que tome  $n$ , así que si  $n$  es lo suficientemente grande (¿que tan grande será necesario?), entonces

$$0 < \pi \int_0^1 a^n f_n(x) \sin \pi x \, dx < \frac{\pi a^n}{n!} < 1,$$

pero esto es una contradicción, ya que la integral, como lo habíamos mencionado es un entero, y no existe ningún número entero entre 0 y 1. Así, la suposición original de que  $\pi^2$  es racional debe ser falsa y por lo tanto se obtiene que  $\pi^2$  es irracional y así mismo debe serlo  $\pi$  (¿sí se ve por qué?).

Que  $\pi$  es un número irracional se sabe hace mucho tiempo y se ha probado de diversas maneras, la demostración que les he presentado, tiene la peculiaridad, de que guarda una idea central que se puede aplicar para probar la irracionalidad de otros números. De hecho, este escrito no está dedicado a  $\pi$ , todo esto ha sido una antesala para presentarles a uno de los personajes más famosos del complejo mundo de la “irracionalidad”, la constante de Euler-Mascheroni o constante gama de Euler, el famoso

$$\gamma \approx 0,5772156649015328606065120900824024310421593.$$

Si este numerito nos les suena de algún sitio, lo cuál ya de por sí es extraño, veámoslo en su forma primitiva, esta constante mide la diferencia entre las sumas parciales de la serie armónica y la función logarítmica, así

$$\gamma := \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \ln n \right). \quad (3)$$

Y que por qué es importante, pues si el mismísimo Euler dedicó años de su vida a su estudio, será por algo, pero el lector interesado por aplicaciones, apariciones y milagros de esta constante, puede entre otros consultar [Lag13] (se encuentra libre en línea), donde se expone de manera divulgativa y cronológica muchos resultados que involucran a la constante de Euler. De momento algunas de sus elegantes presentaciones:

$$\begin{aligned} \gamma &= \lim_{n \rightarrow \infty} \left( \sum_{p \leq n} \ln \left( \frac{p}{p-1} \right) - \ln(\ln(n)) \right), \\ &= \lim_{n \rightarrow \infty} \left[ n - \Gamma \left( \frac{1}{n} \right) \right], \\ &= \lim_{s \rightarrow 1} \left( \zeta(s) - \frac{1}{s-1} \right). \end{aligned}$$

Finalmente y sin dar más rodeos, a lo que quería llegar ¿es  $\gamma$  un número irracional? Mi estudiante Esteban Burbano, en su tesis de pregrado abordó esta pregunta, muy a la  $\pi$ -style, veamos:

Supongamos que  $\gamma$  es racional, entonces podemos escribir

$$\gamma = \frac{p}{q}$$

donde  $p$  y  $q$  son enteros positivos y  $q \neq 0$ , de la definición de  $\gamma$  tendríamos que

$$\gamma = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} - \ln n = \frac{p}{q}$$

multiplicando a ambos lados por  $q$  se obtiene

$$\lim_{n \rightarrow \infty} \left( q \sum_{k=1}^n \frac{1}{k} - q \ln n \right) = p$$

usando propiedades de límites podemos escribir

$$\lim_{n \rightarrow \infty} q \sum_{k=1}^n \frac{1}{k} - \lim_{n \rightarrow \infty} q \ln n = p$$

dato que  $q \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} = q \sum_{k=1}^{\infty} \frac{1}{k}$  esta última ecuación toma la forma

$$q \sum_{k=1}^{\infty} \frac{1}{k} - \lim_{n \rightarrow \infty} q \ln n = p.$$

Partiendo la serie armónica en dos sumas reescribimos como

$$q \left( \sum_{k=1}^j \frac{1}{k} + \sum_{k=j+1}^{\infty} \frac{1}{k} \right) - \lim_{n \rightarrow \infty} q \ln n = p$$

que se puede ver como

$$q \sum_{k=j+1}^{\infty} \frac{1}{k} - \lim_{n \rightarrow \infty} q \ln n = p - q \sum_{k=1}^j \frac{1}{k}$$

donde podemos observar que el término de la derecha en esta igualdad resulta ser un número racional para algún  $j$ , mientras que a la izquierda, tomando el límite obtenemos

$$q \sum_{k=j+1}^{\infty} \frac{1}{k} = \infty = q \lim_{n \rightarrow \infty} \ln n.$$

Conclusión, callejón sin salida. No hemos obtenido un argumento final que nos permita abalar, ni desmentir, la hipótesis original. Será otro el camino, ¿pero cuál? Como Esteban, muchos otros desde 1734, han tratado de verificar lo que la mayoría cree, que  $\gamma$  es un número irracional, sin que hasta la fecha, alguien haya tenido éxito, por tanto, el estatus de  $\gamma$  dentro de los reales, no pasa de conjetura. Decidir si un número es o no irracional, vaya que problema...

## Referencias.

- [BBB04] L. Berggren, J. Borwein, and P. Borwein, *Pi: A source book, Third Edition*, Springer., 2004.
- [Lag13] Jeffrey, Lagarias, *Euler's work and modern developments*, Bulletin of the AMS **50** (2013), no. 7, 527-628.

**John A. Arredondo**  
alexander.arredondo@konradlorenz.edu.co

## Política Editorial

El Paskín Matemático es una producción del Programa de Matemáticas de la Fundación Universitaria Konrad Lorenz. Está abierto a todos y todas. Si tienes interés en escribir un artículo con cualquier tipo de contenido matemático, reseñar a un personaje, comentar un problema o tienes comentarios sobre nuestra publicación, envía un mensaje a [paskin@konradlorenz.edu.co](mailto:paskin@konradlorenz.edu.co)

$$\frac{\pi}{4} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$$

El Programa de Matemáticas de la Fundación Universitaria Konrad Lorenz fue creado en 1998. Conoce más acerca de nuestro programa en nuestra página institucional

<http://www.konradlorenz.edu.co/es/aspirantes/carreras-universitarias/carrera-de-matematicas/presentacion.html>

Bogotá, Colombia. 2021

# PASKÍN MATEMÁTICO

## Volumen 3, No. 1, 2021

### Contenido

- Ruben A. Hidalgo  
El grupo jacobiano de un grafo: un ahorro solidario.....1
- Juan Sebastián Martínez Conejo  
Solucionando sistemas polinomiales usando Bases de Gröbner .....5
- Camilo Ramírez Maluendas  
Del álgebra lineal y el análisis a las métricas riemannianas en  $\mathbb{R}^n$ .....11
- Marco Blanco Ariza  
Integración numérica al rescate: dos alternativas para integrar  
 $f(x) = e^{-x^2}$  con Python.....19
- Pablo Amster  
Una zambullida en las ecuaciones con retardo.....23
- Dayron Fabián Achury  
MateKuento.....28
- John A. Arredondo  
 $K'$ Problema.....30

*Nada es más importante que ver las fuentes de la invención, que son, en mi opinión, más interesantes que los inventos en sí mismos.*

Gottfried Leibniz

**KONRAD  
LORENZ**  
FUNDACIÓN UNIVERSITARIA