UN TEOREMA DE TRASCENDENCIA EN NÚMEROS

‡‡

En estas notas se demuestra, siguiendo a Serge Lang, una versión débil del siguiente resultado:

Teorema (de las seis exponenciales)**.** *Si* $\{\beta_1, \beta_2\}$ y $\{z_1, z_2, z_3\}$ *son subconjuntos de* $\mathbb{C}$ *que son linealmente independientes sobre* $\mathbb{Q}$*, entonces al menos uno de los números*

$$e^{\beta_1 z_1},\ e^{\beta_1 z_2},\ e^{\beta_1 z_3},\ e^{\beta_2 z_1},\ e^{\beta_2 z_2},\ e^{\beta_2 z_3},$$

es trascendente.

1. Preliminares

Lema 1.1 (C. L. Siegel, 1949)**.** *Sea*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ \qquad\cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = 0, \end{cases} \tag{1}$$

un sistema de r ecuaciones lineales en n incógnitas donde cada uno de los coeficientes a_{ij} *es un número entero. Supóngase que* $A \geq 1$ *es tal que* $|a_{ij}| \leq A$ *para cada* $1 \leq i \leq r$ *y cada* $1 \leq j \leq n$*. Si* $n > r$*, entonces el sistema admite una solución no trivial* $\mathbf{z} := (z_1, \ldots, z_n) \in \mathbb{Z}^n$ *que satisface lo sig.:*

$$\|\mathbf{z}\|_\infty := \max_{1\leq j\leq n} |z_j| \leq 2(nA)^{\frac{r}{n-r}}.$$

Demostración. La matriz $C := (a_{ij})_{1\leq i\leq r,\, 1\leq j\leq n}$ determina una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^r$. Como cada una de las entradas de C es un número entero, se cumple que C manda $\mathbb{Z}^n$ en $\mathbb{Z}^r$. Ahora bien, para cada $H \in \mathbb{N}$, denotemos con $\mathbb{Z}^n(H)$ al subconjunto de $\mathbb{Z}^n$ conformado por aquellos $\mathbf{x} \in \mathbb{Z}^n$ tales que $\|\mathbf{x}\|_\infty \leq H$. En vista de que para todo $\mathbf{x} \in \mathbb{Z}^n(H)$ y $1 \leq \ell \leq r$ se verifica que

$$|a_{\ell 1}x_1 + \cdots + a_{\ell n}x_n| \leq nAH,$$

tenemos que la imagen de $\mathbb{Z}^n(H)$ bajo C es un subconjunto de $\mathbb{Z}^r(nAH)$. Luego, puesto que $|\mathbb{Z}^n(H)| = (2H+1)^n$ y $|\mathbb{Z}^r(nAH)| = (2\lfloor nAH \rfloor + 1)^r$, se tiene que si

$$(2nAH + 1)^r < (2H + 1)^n, \tag{2}$$

entonces $C|_{\mathbb{Z}^n(H)}$ no es inyectiva. Esto permite garantizar la existencia de $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^n(H)$ (distintos) tales que $C(\mathbf{x}) = C(\mathbf{y})$; de esta igualdad y la linealidad de C se sigue que $\mathbf{z} := \mathbf{x} - \mathbf{y}$ es una solución no trivial de (1) cuya norma infinito es

$$\|\mathbf{z}\|_\infty \leq \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty \leq 2H. \tag{3}$$

Para concluir la demostración, basta notar que, eligiendo $2H$ como el único número par tal que

$$(nA)^{\frac{r}{n-r}} - 1 \leq 2H < (nA)^{\frac{r}{n-r}} + 1,$$

se satisface la desigualdad en (2), tal como lo constatan los cálculos siguientes:

$$(2nAH + 1)^r < (nA)^r(2H + 1)^r \leq (2H + 1)^{n-r}(2H + 1)^r = (2H + 1)^n.$$

□

Lema 1.2. *Si* $r_1, \ldots, r_k$ *son números complejos distintos, entonces las funciones* $f_i \colon \mathbb{C} \to \mathbb{C}$ *definidas por* $f_i(t) = e^{r_i t}$ *(para cada* $t \in \mathbb{C}$*) son linealmente independientes sobre* $\mathbb{C}$*.*

‡‡José Hernández Santiago || email: stgo@matmor.unam.mx || Morelia, Mich., 17-18 de noviembre de 2016.

Demostración. Es un ejercicio sencillo de álgebra lineal. □

Definición 1.1. *Decimos que las funciones $f, g\colon \mathbb{C} \to \mathbb{C}$ son **algebraicamente independientes** (sobre $\mathbb{C}$) si para cualquier $P \in \mathbb{C}[x, y] \setminus \{0\}$ resulta que $P(f, g)$ no es la función cero.*

Proposición 1.1. *Si $\alpha, \beta \in \mathbb{C}$ son linealmente independientes sobre $\mathbb{Q}$, entonces las funciones $f_1(t) = e^{\alpha t}$ y $f_2(t) = e^{\beta t}$ son algebraicamente independientes.*

Demostración. Supongamos que $P(x, y) = \sum_{i,j} a_{i,j} x^i y^j \in \mathbb{C}[x, y]$ es tal que $P(e^{\alpha t}, e^{\beta t}) = 0$ para todo $t \in \mathbb{C}$. Se sigue de esto que

$$0 = P(e^{\alpha t}, e^{\beta t}) = \sum_{i,j} a_{i,j} e^{(\alpha i + \beta j)t}.$$

Como α y β son linealmente independientes sobre $\mathbb{Q}$, los números $r_{ij} := \alpha i + \beta j \in \mathbb{C}$ son todos distintos. Aplicando el lema anterior, obtenemos que cada uno de los a_{ij} es igual 0, de lo cual se concluye que $P(x, y)$ es el elemento cero de $\mathbb{C}[x, y]$. □

Proposición 1.2 (Principio del módulo máximo)**.** *Sean $D \subseteq \mathbb{C}$ una región y $f\colon D \to \mathbb{C}$ una función holomorfa. Si $|f|$ tiene un máximo local en algún $z_0 \in D$, i.e., si $|f(z_0)| = \|f\|_U := \sup_{\zeta \in U} |f(\zeta)|$ en algún disco U centrado en z_0 y contenido en D, entonces f es constante.*

Demostración. Supongamos que f no es constante y que $|f(z_0)| > 0$. Puesto que, para cada $z \in U$ se cumple que $|f(z)| \leq |f(z_0)|$, se sigue que $f(U) \subseteq \overline{\mathrm{B}_{|f(z_0)|}(0)}$. De esto y del hecho de que $f(z_0) \in f(U) \cap \partial(\overline{\mathrm{B}_{|f(z_0)|}(0)})$ se concluye que $f(U)$ no es un abierto de $\mathbb{C}$, lo cual entra en contradicción con lo que establece el teorema de la aplicación abierta. □

Corolario 1.1 (Desigualdad de Jensen)**.** *Sea f una función holomorfa en $\{z \in \mathbb{C}\colon |z| \leq R\}$. Si $f(0) \neq 0$ y los ceros de f en $\mathrm{B}_R(0)$ son $z_1, z_2, \ldots, z_N$, entonces*[†]

$$|f(0)| \leq \|f\|_R (|z_1 \cdots z_N| / R^N).$$

Demostración. Resulta fácil convencerse de que

$$\frac{R^2 - z\overline{z_n}}{R(z - z_n)}$$

es un número complejo de módulo 1 cuando $|z| = R$. Por tanto,

$$g(z) = f(z) \prod_{n=1}^{N} \frac{R^2 - z\overline{z_n}}{R(z - z_n)}$$

es una función holomorfa en $\{z \in \mathbb{C}\colon |z| \leq R\}$ y $|g(z)| = |f(z)|$ si $|z| = R$. Aplicando el principio del módulo máximo a la función g se obtiene que

$$|g(z)| \leq \|f\|_R,$$

de lo cual se concluye que

$$|f(0)| R^N = |g(0)| |z_1 \cdots z_N| \leq \|f\|_R |z_1 \cdots z_N|.$$

□

Corolario 1.2. *Sea f como en el corolario anterior. Para $r > 0$, denótese con $\nu(r) = \nu(f, r)$ al número de ceros de f, contados con multiplicidad, en $\{z \in \mathbb{C}\colon |z| < r\}$. Se tiene entonces que*

$$\int_0^R \frac{\nu(x)}{x}\, dx \leq \log \|f\|_R - \log |f(0)|.$$

[†]Con $\|f\|_R$ denotamos en lo sucesivo al valor máximo de $|f|$ en la frontera del círculo de radio R centrado en el origen.

Demostración. Primeramente demostraremos que

$$\sum_{n=1}^{N} \int_{|z_n|}^{R} \frac{1}{x}\, dx = \int_0^R \frac{\nu(x)}{x}\, dx. \tag{4}$$

Para $k \in \{1, \ldots, N\}$, defínase $\chi_k \colon [0, R] \to \{0, 1\}$ como

$$\chi_k(x) = \begin{cases} 1 & \text{si } x > |z_k| \\ 0 & \text{si } x \leq |z_k|. \end{cases}$$

Luego, en vista de que $\sum_{n=1}^{N} \chi_n(x) = \nu(x)$, se desprende que

$$\sum_{n=1}^{N} \int_{|z_n|}^{R} \frac{1}{x}\, dx = \sum_{n=1}^{N} \int_0^R \frac{\chi_n(x)}{x}\, dx = \int_0^R \left(\sum_{n=1}^{N} \chi_n(x) \right) \frac{1}{x}\, dx = \int_0^R \frac{\nu(x)}{x}\, dx.$$

De (4) y la desigualdad de Jensen se concluye que

$$\int_0^R \frac{\nu(x)}{x}\, dx \;\; = \;\; \sum_{n=1}^{N} (\log R - \log |z_n|) = \log \left(\frac{R^N}{|z_1 \cdots z_N|} \right) \leq \log \|f\|_R - \log |f(0)|.$$

□

El corolario 1.2 permitirá establecer una conexión entre el número de ceros de una función en un disco y la *tasa de crecimiento* de la función. Antes de enunciarla es preciso recordar que si f es función entera y ρ, A, B son constantes positivas tales que

$$|f(z)| \leq A e^{B|z|^\rho}$$

para todo $z \in \mathbb{C}$, entonces se dice que f tiene **orden de crecimiento** $\leq \rho$. Al ínfimo de todos los $\rho > 0$ que satisfacen lo anterior se le denomina el **orden** de la función f.

Corolario 1.3. *Sea f una función entera distinta de la función idénticamente 0 que tiene orden de crecimiento $\leq \rho$. Existe $C > 0$ tal que $\nu(r) \leq C r^\rho$ para cada r suficientemente grande.*

Demostración. Puede suponerse que $f(0) \neq 0$ pues, en caso contrario, $f(z) = z^\ell g(z)$ para algún $\ell \in \mathbb{N}$ y una función entera g que no se anula en $z = 0$, que también tiene orden de crecimiento $\leq \rho$ y tal que $\nu(f, r) = \nu(g, r) + \ell$ para todo $r > 0$.

Fijemos $r > 0$. Al aplicar el corolario 1.2 con $R = 2r$ se obtiene que

$$\log \|f\|_{2r} - \log |f(0)| \geq \int_r^{2r} \frac{\nu(x)}{x}\, dx \geq \nu(r) \int_r^{2r} \frac{1}{x}\, dx = \nu(r) \log 2.$$

Por otro lado, de la condición sobre el crecimiento de f se tiene que

$$\log \|f\|_{2r} \leq \log A e^{B(2r)^\rho} = \log A + (2^\rho B) r^\rho.$$

Así las cosas,

$$\nu(r) \leq \frac{\log A + (2^\rho B) r^\rho - \log |f(0)|}{\log 2}$$

y la demostración termina. □

2. Demostración

Etapa 1. Sean $\beta_1, \beta_2, z_1, z_2, z_3$ como en la formulación del teorema. Sean $f, g\colon \mathbb{C} \to \mathbb{C}$ las funciones definidas por las asignaciones $t \overset{f}{\longmapsto} e^{\beta_1 t}$ y $t \overset{g}{\longmapsto} e^{\beta_2 t}$. Probaremos, por *reductio*, que al menos uno de los seis números siguientes

$$f(z_1) = e^{\beta_1 z_1},\ f(z_2) = e^{\beta_1 z_2},\ f(z_3) = e^{\beta_1 z_3},\ g(z_1) = e^{\beta_2 z_1},\ g(z_2) = e^{\beta_2 z_2},\ g(z_3) = e^{\beta_2 z_3}$$

no pertenece a $\mathbb{Z}$. A fines de obtener un absurdo, supongamos lo opuesto. Sean n un cuadrado perfecto *suficientemente grande* y $r := (4n)^{\frac{3}{2}}$. Apelando al lema de Siegel podemos garantizar la existencia de $x_{ij} \in \mathbb{Z}$, no todos cero, tales que la función *auxiliar*

$$F := \sum_{1 \le i,j \le r} x_{ij} f^i g^j \tag{5}$$

se anula en cada uno de los elementos del conjunto

$$\mathcal{K}_n := \{\mathbf{k} \cdot \mathbf{z}\colon 1 \le k_1, k_2, k_3 \le n\}.$$

Nótese que el lema de Siegel se está aplicando aquí a un sistema de n^3 ecuaciones lineales en $r^2 = (4n)^3$ incógnitas; luego, puesto que para los coeficientes se tiene la estimación

$$\begin{aligned} |f^i(\mathbf{k}\cdot\mathbf{z})g^j(\mathbf{k}\cdot\mathbf{z})| &\le e^{|i\beta_1(\mathbf{k}\cdot\mathbf{z}) + j\beta_2(\mathbf{k}\cdot\mathbf{z})|} \\ &\le e^{r(|\beta_1|+|\beta_2|)\|\mathbf{k}\|\|\mathbf{z}\|} \\ &\le e^{(|\beta_1|+|\beta_2|)\|\mathbf{z}\| rn\sqrt{3}} \\ &\le C_1^{n^{\frac{5}{2}}} \end{aligned} \tag{6}$$

para alguna constante $C_1 \ge 1$, el lema de Siegel indica que los x_{ij} cumplen que

$$|x_{ij}| \le 2(r^2 C_1^{n^{\frac{5}{2}}})^{\frac{n^3}{(4n)^3 - n^3}} = 2((4n)^3 C_1^{n^{\frac{5}{2}}})^{\frac{1}{63}} \le (2)(4^{\frac{1}{21}}) C_2^{n^{\frac{5}{2}}} \tag{7}$$

para alguna constante $C_2 \ge 1$.

Etapa 2. Lo que se hará a continuación, a grandes rasgos, es utilizar información sobre el *comportamiento* de la función F en $\mathcal{K}_n$ para exhibir un $\omega \in \mathcal{K} := \{\mathbf{k}\cdot\mathbf{z}\colon k_1, k_2, k_3 \in \mathbb{N}\}$ de tal modo que $|F(\omega)|$ sea *pequeño* y $F(\omega) \neq 0$: estos requerimientos sobre ω son clave pues de ellos y del *hecho* de que $F(\omega) \in \mathbb{Z}$ se obtendrá la contradicción con la cual culminará la demostración.

En primer lugar, observemos que de la independencia algebraica de las funciones f y g se desprende que la función F no es idénticamente cero; por otro lado, el supuesto de que $f(z_1), f(z_2), f(z_3), g(z_1), g(z_2), g(z_3)$ son números enteros implica que $F(\mathcal{K}) \subseteq \mathbb{Z}$. Más aún, se tiene que[‡]:

A. Existe $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}$ tal que $F(\mathbf{k}\cdot\mathbf{z}) \neq 0$.

Dem. A fines de obtener una contradicción, supongamos que $F(\mathbf{k}\cdot\mathbf{z}) = 0$ para cada $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}$. Para $r > 0$ la desigualdad $\|\mathbf{k}\| \le \frac{r}{\|\mathbf{z}\|+1}$ implica que $|\mathbf{k}\cdot\mathbf{z}| \le \|\mathbf{k}\|\|\mathbf{z}\| < r$; en consecuencia,

$$\left\lfloor \frac{r}{\sqrt{3}(\|\mathbf{z}\|+1)} \right\rfloor^3 \le \nu(r).$$

Por otro lado, al ser F es una función de orden de crecimiento ≤ 1, el corolario 1.3 implica que $\nu(r) \le Cr$ para algún $C > 0$ y todo número r mayor que una constante absoluta $c > 0$. Puesto que las estimaciones obtenidas para $\nu(r)$ dan lugar a una desigualdad que no tiene sentido cuando r es suficientemente grande, el resultado se sigue.

[‡]Vienen a continuación dos asertos importantes. En los recuadros respectivos bosquejamos las demostraciones de cada uno de ellos.

B. El conjunto $S := \{s \in \mathbb{N}\colon F(\mathbf{k}\cdot\mathbf{z}) = 0 \text{ para todo } \mathbf{k} = (k_1, k_2, k_3) \in \mathbb{N}^3 \text{ t. q. } 1 \leq k_1, k_2, k_3 \leq s\}$ tiene un elemento máximo.

Dem. En caso contrario existe una sucesión $s_1 < s_2 < s_3 < \cdots$ de números naturales tales que

$$F(\mathbf{k}\cdot\mathbf{z}) = 0$$

para cada $\ell \in \mathbb{N}$ y cada $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}_{s_\ell}$. De esto se colige que $F(\mathbf{k}\cdot\mathbf{z}) = 0$ para todo $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}$ (lo cual es absurdo en vista de lo establecido en **A**): en efecto, sea $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}$; si $m = \text{máx}\{k_1, k_2, k_3\}$ y $\ell \in \mathbb{N}$ es tal que $m < s_\ell$, entonces $\mathbf{k}\cdot\mathbf{z} \in \mathcal{K}_{s_\ell}$ y, en consecuencia, $F(\mathbf{k}\cdot\mathbf{z}) = 0$.

Luego, si s es el elemento máximo de S entonces $s \geq n$; además, si

$$\omega := \mathbf{k}\cdot\mathbf{z} \in \mathcal{K}_{s+1}$$

se elige de tal modo que $k_\nu = s + 1$ para algún $\nu \in \{1, 2, 3\}$, se tiene que $F(\omega) \neq 0$.

Etapa 3. Estimaremos ahora el módulo de $F(\omega)$ utilizando la siguiente identidad

$$F(\omega) = \lim_{t\to\omega}\left(\frac{F(t)}{\prod_{\mathbf{k}\cdot\mathbf{z}\in\mathcal{K}_s}(t - \mathbf{k}\cdot\mathbf{z})}\right)\prod_{\mathbf{k}\cdot\mathbf{z}\in\mathcal{K}_s}(\omega - \mathbf{k}\cdot\mathbf{z}). \tag{8}$$

El número de factores de cada uno de los productos que aparecen en la identidad es s^3. La función que aparece dentro de los paréntesis admite una extensión holomorfa a todo el plano complejo: consiguientemente, para estimar $|F(\omega)|$ podemos aplicar el principio del módulo máximo en el círculo de radio $R = s^{\frac{3}{2}}$ centrado en el origen. Si t es un número complejo de módulo R (y, digamos*, $n \geq 4(|z_1| + |z_2| + |z_3|)^2$), entonces

$$|t - \mathbf{k}\cdot\mathbf{z}| \geq R - |\mathbf{k}\cdot\mathbf{z}| = s^{\frac{3}{2}} - s(|z_1| + |z_2| + |z_3|) \geq \frac{s^{\frac{3}{2}}}{2} = \frac{R}{2}$$

y por tanto

$$\frac{|\omega - \mathbf{k}\cdot\mathbf{z}|}{|t - \mathbf{k}\cdot\mathbf{z}|} \leq \frac{2(|z_1| + |z_2| + |z_3|)s}{R} \leq \frac{C_3}{s^{\frac{1}{2}}}$$

para alguna constante $C_3 > 1$. Así pues, de (8) y la estimación**

$$\|F\|_R \leq r^2(2)(4^{\frac{1}{21}})(C_2^{n^{\frac{5}{2}}})(C_4^{rR}) \leq (2)(4^{\frac{1}{21}})(4n)^3(C_2^{s^{\frac{5}{2}}})(C_4^{(4n)^{\frac{3}{2}}\cdot s^{\frac{3}{2}}}) \leq (2)(4^{3+\frac{1}{21}})(C_5^{s^3}) \leq C_6^{s^3},$$

se llega a que

$$\log|F(\omega)| \leq \log\|F\|_R + \log\left(\frac{C_3}{s^{\frac{1}{2}}}\right)^{s^3} = (\log(C_3C_6))s^3 - \frac{1}{2}s^3\log s.$$

Puesto que la desigualdad en la línea anterior es absurda para s arbitrariamente grande, la demostración termina.

Referencias

[1] S. Lang, "Transcendental numbers and diophantine approximations." Bull. Amer. Math. Soc., vol. **77** (September 1971), number 5, pp. 635–677.

[2] R. Murty & P. Rath, *Transcendental numbers.* Springer-Verlag, 2014.

[3] C. L. Siegel, *Transcendental numbers.* Annals of Mathematics Studies (number 16), Princeton University Press, Princeton, NJ, USA, 1949.

*Aquí es donde cobra sentido el calificativo *suficientemente grande* con el que se adjetivó a n al inicio de esta sección.

**La constante C_2 proviene de (7) y C_4, C_5, C_6 son constantes mayores que 1.