Revista de Educación Matemática

Registro Nacional de la Propiedad Intelectual N° 168024

ISSN 0326-8780 (versión impresa)

ISSN 1852-2890 (en línea)

Página web: https://revistas.unc.edu.ar/index.php/REM/index
Correo electrónico: revm@famaf.unc.edu.ar

Revista de Educación Matemática

Volumen 38, N° 2 – 2023

ÍNDICE

Artículos

Secciones fijas

Revista de Educación Matemática
Volumen 38, N° 2 (2023), páginas 3 – 15
Unión Matemática Argentina - Famaf (UNC)

Artículo de Matemática

CÁLCULO DE RADIOS GEOCÉNTRICOS POR GRADOS DE LATITUD

Rodrigo Tovar Cabañas, Shany Arely Vázquez Espinosa
e Hipólito Villanueva Hernández

Resumen. El objetivo del presente ejercicio es hacer frente a la pérdida de cognición y por ende pericia para realizar transformaciones geométricas de modelos esféricos a modelos elipsoidales dentro de la enseñanza de la geografía a nivel licenciatura. Se postula que los sistemas satelitales de navegación global basados en elipsoides de referencia aún son concebidos como sistemas esféricos por muchos alumnos debido a la ausencia de ejemplos prácticos que ayuden a su comprensión pedagógica. Para cubrir dicha laguna en este trabajo se calcularon la excentricidad, la latitud geocéntrica y el radio geocéntrico de los 90 principales paralelos del elipsoide de referencia WGS84. Metodológicamente se trabajó sobre un universo de estudio superior a los 324 mil segundos de latitud. De los cuales sólo se muestran los 90 casos principales. Con un criterio analógico se determinó el radio geocéntrico para los puertos de Salina Cruz, Oaxaca, Tampico, Tamaulipas y Ensenada, Baja California. Como resultados se exhibe que la distancia al centro de la Tierra entre Salina Cruz, y Ensenada hay una diferencia de 4 kilómetros.

Abstract. The objective of this exercise is to face the loss of cognition and therefore the ability to perform geometric transformations from spherical models to ellipsoidal models within the teaching of geography at the undergraduate level. It is postulated that global navigation satellite systems based on reference ellipsoids are still conceived as spherical systems by many students due to the absence of practical examples that help their pedagogical understanding. To cover this gap in this work, the eccentricity, geocentric latitude and geocentric radius of the 90 main parallels of the reference ellipsoid WGS84 were calculated. Methodologically, we worked on a study universe greater than 324 thousand seconds of latitude. Of which only the top 90 cases are shown. Using an analog criterion, the geocentric radius was determined for the ports of Salina Cruz, Oaxaca, Tampico, Tamaulipas and Ensenada, Baja California. As results it is exhibited that the distance to the center of the Earth between Salina Cruz and Ensenada there is a difference of 4 kilometers.

Palabras clave: Latitud geocéntrica, radio geocéntrico, elipsoide WGS84.
Keywords: Geocentric latitude, geocentric radius, WGS84 ellipsoid.

§1. Introducción

Los sistemas satelitales de navegación global, tales como: Beidou de la República Popular de China; Galileo de la Unión Europea; GLONASS de Rusia; NavIC de India y Michibiki utilizado en Japón, más el crecimiento exponencial de usuarios de tecnología basada en el Sistema de Posicionamiento Global (GPS, por sus siglas en inglés), dentro de los cuales destacan los investigadores de ciencias como: arqueología, antropología, ecología, botánica, hidrología, y por supuesto la geodesia, quienes necesitan otorgar la máxima precisión posible a sus geolocalizaciones, lecturas de sus receptores GPS que, por ciertas cuestiones y condiciones atmosféricas, en ocasiones otorgan lecturas discordantes.

Por ejemplo, en lugares cercanos a la línea de costa se han dado casos de lecturas inferiores al nivel del mar, por lo tanto, para usuarios no familiarizados con el campo de la geometría satelital, es útil tener a la mano una serie de tablas que contribuyan a calcular y calibrar de forma práctica tales discordancias, o usar algún elipsoide, es decir, una superficie de revolución derivada de la rotación de una elipse, que genere menos errores en la altura elipsóidica, dígase elevación por encima o por debajo de una superficie idealizada que se aproxima a la forma de la Tierra. De modo que el objetivo de este breve ensayo pedagógico es calcular radios geocéntricos para distintas latitudes del elipsoide de referencia del Sistema Geodésico Mundial de 1984, World Geodetic System 84, o simplemente WGS84 y de paso ampliar la escasa literatura, en español, que existe en torno al procesamiento y cálculo matemático de latitudes y radios geocéntricos. La figura geométrica de la Tierra es una categoría trascendental en geodesia, la cual se refiere al tamaño y la forma utilizados para modelar al planeta Tierra. Tanto el tamaño como la forma a los que se refiere dependen del contexto y objetivos buscados, incluidas la precisión tanto en espacio como en tiempo requeridas por el modelado buscado. Es útil señalar que la forma esférica es una aproximación de la figura geométrica de la Tierra bastante satisfactoria para diversos propósitos, con ella se han desarrollado varios modelos de considerable precisión, por ejemplo, para cálculos astronómicos. Sin embargo, la superficie topográfica de la Tierra, con su accidentado relieve y fosas oceánicas, en general, es la preocupación central de topógrafos, hidrógrafos y geofísicos, puesto que su modelado matemáticamente, teniendo en cuenta las irregularidades citadas, es una labor en extremo compleja.

En otras palabras, la concepción pitagórica de una Tierra esférica ofrece una superficie adecuada para un fácil tratamiento geométrico, loable para muchos cálculos astronómicos y de navegación que utilizan ese modelo esférico para estudiar la Tierra. Empero, para otras tareas, a escala regional, se requieren mediciones más precisas de la forma de la Tierra, es allí donde entra el modelado de la superficie representado como un esferoide achatado u oblato.

En efecto, para levantamientos topográficos de áreas pequeñas, como fraccionamientos urbanos o proyectos arquitectónicos, un modelo plano (plano catastral) de la superficie de la Tierra es suficiente, porque la precisión de la topografía local supera los errores derivados de la curvatura de la Tierra, pero para mapear áreas más grandes, como regiones, países o continentes se requiere conocer el tamaño y la forma de la Tierra con mayor precisión, puesto que en esos casos los traspiés derivados de la curvatura de la Tierra superan al método de trigonometría plana.

En esa línea de pensamiento, en términos históricos, milenios después de Eratóstenes, y siglos después de las fórmulas geométricas de Johannes de Sacrobosco (1195-1256), de Francesco Maurolico (1494-1575) y de Francesco Barozzi (1537-1604), para comprobar la distancia de un grado de latitud (De Andrade, 2008), las primeras medidas modernas que contribuyeron a conocer la forma geométrica de la Tierra fueron: las de 1528 de Jean François Fernel, quien a partir del conteo de las revoluciones de las ruedas de su carruaje se conoció la distancia entre lugares separados cientos de kilómetros; y las de Willebrord van Roijen Snell, latinizado como Snellius del año 1615, este último calculó por primera vez, distancias en la superficie terrestre, basándose en el método de la triangulación topográfica (Veloso, 2016). No hay que olvidar la reimpresión que se dio en el siglo XVII de la *Cosmographia* de Maurolico, obra en la que se perfeccionó una metodología para medir la Tierra, la cual fue empleada hacia 1670 por Jean Picard para medir la longitud de un arco meridiano (Gómez, 2007).

Por tal motivo se considera a Picard, como la primera persona moderna en medir el tamaño de la Tierra con un grado razonable de precisión. Picard logró esto al oponer la distancia en línea recta (110 km) que hay de la ciudad de Amiens (ubicada a los $49^\circ\,51'$ de latitud norte) a la ciudad de Paris (ubicada a los $48^\circ\,51'$ de latitud norte) mediante el método de triangulación aplicado a trece polígonos que necesitó para tal hazaña. Sus mediciones arrojaron un resultado de $110{,}4$ km para un grado de latitud, lo que da un radio terrestre correspondiente de $6328{,}9$ km. Años más tarde, cuando se reconoció que el valor del radio polar era más corto que el del radio ecuatorial, comenzaron a surgir las interrogantes que llevaron a la creación de modelos de elipsoide para la representación de la forma geométrica de la Tierra. El elipsoide que representa la Tierra es un elipsoide particular pues es una superficie de revolución: se obtiene por una rotación de una elipse, siendo su ecuación:

$$\frac{(x^2+y^2)}{a^2}+\frac{z^2}{b^2}=1.$$

Donde x, y y z son las coordenadas cartesianas de un punto en la superficie del esferoide, y a y b son los semiejes mayores y menores, respectivamente. En el caso del modelo WGS84, se tiene que:

- Semieje mayor $a = 6.378.137{,}000$ metros,
- Semieje menor $b = 6.356.752{,}314$ metros.

Ahora bien, el primer elipsoide de la Tierra fue propuesto por Maupertuis, Clairaut, Camus, Le Monnier, M. Abbé Outhier, y M. Celsius, quienes determinaron la forma elipsoidal de la Tierra en 1738 (Veloso, 2016). Otros elipsoides que han sido elaborados a lo largo de la historia se muestran en la Tabla 1.

Ahora bien, dado que el planeta Tierra, aunque está ligeramente achatado en los polos y abultado en el Ecuador, tiene una forma relativamente amorfa, sin embargo, la figura geométrica utilizada en geodesia que más se aproxima a su forma es un esferoide oblato. Un esferoide oblato (o elipsoide oblato) es un elipsoide de revolución obtenido por rotación de una elipse alrededor de su eje más corto. Cuando un esferoide representa la forma de la Tierra u otro cuerpo celeste recibe el nombre de elipsoide de referencia (Kopeikin, 2019).

El uso de esferoides resuelve el problema de la forma del planeta Tierra, sin embargo, complican el cálculo de medidas y áreas debido a que como se sabe, las elipses (elipsoides) presentan dos focos y las circunferencias (esferas) uno, De modo que, en términos pedagógicos, para conocer el radio geocéntrico (o la distancia que hay entre el centro de la Tierra y un punto cercano a la superficie de ésta) de una latitud determinada de la Tierra, tomando en consideración un modelo elipsoidal, se requiere conocer la latitud geocéntrica del elipsoide de referencia (la latitud geocéntrica es una latitud elíptica, la cual se mide usando un sistema de coordenadas basado en un elipsoide), para esta última cifra es importante determinar la excentricidad ε del elipsoide de referencia, en otras palabras, para estipular la latitud geocéntrica primero hay que establecer cuál es la excentricidad del elipsoide de referencia.

§2. Metodología

Un elipsoide de referencia puede describirse mediante una serie de parámetros que definen su forma y que incluyen un semieje mayor a, un semieje menor b, de cuya relación se determina su razón de aplanamiento f, dada por

$$f = \frac{a-b}{a}.$$

Además, su distancia focal c, que se define como la distancia entre el centro del elipsoide y uno de sus focos. Su fórmula es:

$$c = \sqrt{a^2 - b^2}.$$

Así como su excentricidad, también denominada primera excentricidad ε que mide la relación entre la distancia focal y la longitud del semieje mayor, en este caso a; y su segunda excentricidad ε' que mide la relación entre la distancia focal y

Elipsoide de referencia	**Radio ecuatorial (m)**	**Radio polar (m)**	**Razón de aplanamiento**	**Lugar donde se calculó**
Maupertuis (1738)	6.397.300	6.363.806,283	191	Francia
Plessis (1817)	6.376.523,0	6.355.862,9333	308,64	Francia
Everest (1830)	6.377.299,365	6.356.098,359	300,80172554	India
Everest 1830 modificado (1967)	6.377.304,063	6.356.103,0390	300,8017	Singapur
Everest 1830 (definición 1967)	6.377.298,556	6.356.097,550	300,8017	Brunéi
Airy (1830)	6.377.563,396	6.356.256,909	299,3249646	Bretaña
Bessel (1841)	6.377.397,155	6.356.078,963	299,1528128	Japón
Clarke (1866)	6.378.206,4	6.356.583,8	294,9786982	Norteamérica
Clarke (1878)	6.378.190	6.356.456	293,4659980	Norteamérica
Clarke (1880)	6.378.249,145	6.356.514,870	293,465	África
Helmert (1906)	6.378.200	6.356.818,17	298,3	EUA
Hayford (1910)	6.378.388	6.356.911,946	297	EUA
Internacional (1924)	6.378.388	6.356.911,946	297	Europa
NAD 27 (1927)	6.378.206,4	6.356.583,800	294,978698208	Norteamérica
Krassovsky (1940)	6.378.245	6.356.863,019	298,3	USSR
WGS66 (1966)	6.378.145	6.356.759,769	298,25	EUA
Nacional de Australia (1966)	6.378.160	6.356.774,719	298,25	Australia
Nuevo Internacional (1967)	6.378.157,5	6.356.772,2	298,24961539	
GRS-67 (1967)	6.378.160	6.356.774,516	298,247167427	
Sudamericano (1969)	6.378.160	6.356.774,719	298,25	Sudamérica
WGS-72 (1972)	6.378.135	6.356.750,52	298,26	EUA
GRS-80 (1979)	6.378.137	6.356.752,3141	298,257222101	Global
WGS-84 (1984)	6.378.137	6.356.752,3142	298,257223563	Global
IERS (1989)	6.378.136	6.356.751,302	298,257	Global
IERS (2003)	6.378.136,6	6.356.751,9	298,25642	Global

Tabla 1. Principales parámetros geométricos de algunos de los principales elipsoides. Fuente: Elaboración personal con base en (Franco, 1999).

la longitud del semieje menor, en este caso b. Se denotan: primera excentricidad

$$\varepsilon = \frac{c}{a},$$

segunda excentricidad

$$\varepsilon' = \frac{c}{b}.$$

Toda excentricidad ε de una elipse, es una medida de la desviación de la elipse respecto a la circularidad. Donde si $\varepsilon = 0$, significa que la elipse es en realidad un círculo; si $\varepsilon = \frac{4}{5}$, la elipse es bastante elíptica; si la relación entre el semieje-menor y el semieje-mayor es $\frac{1}{10}$, la $\varepsilon = 0{,}995$ aproximadamente. Si $\varepsilon = 1$, entonces en realidad la elipse está aplanada, dígase compactada. En efecto, el alargamiento de una elipse se mide por su excentricidad, a través de un rango que va desde 0 en el caso límite de un círculo a 1 en el caso límite de elongación que hace que la elipse aparente tener forma lineal. Analíticamente, la ecuación de una elipse estándar centrada en el origen con ancho a y altura b es:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Asumiendo que $a \geq b$, los focos son $(\pm c, 0)$, mientras que si c se asume como $c = \sqrt{a^2 - b^2}$, se tiene que:

$$\varepsilon = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a}.$$

La primera excentricidad, también puede expresarse geométricamente como

$$\varepsilon = \sqrt{1 - \frac{b^2}{a^2}};$$

en tanto que la segunda excentricidad puede enunciarse como

$$\varepsilon' = \sqrt{\frac{a^2}{b^2} - 1}.$$

De tal modo que para determinar la primera excentricidad ε del planeta Tierra (de utilidad para calcular cualquier latitud geográfica o geodésica), se empleará dicha fórmula, la cual es la forma estandarizada para practicar la conversión de coordenadas geodésicas (Jo, Lee, y Sunwoo, 2015).

Ahora bien, existen muchos elipsoides de referencia, como el GDA94 que utilizan en Australia; el KGD2002 que emplea Corea del Sur, el SAD69 que usan varios países en América del Sur, o el popularmente conocido *International Terrestrial Reference Frames* o ITRF y sus más de 12 versiones. Sin embargo, en este trabajo vamos a ocupar los valores del elipsoide de referencia Sistema Geodésico Mundial, WGS84, que ha sido por más de 35 años, el de mayor refinamiento y uso en las aulas, su uso es importante, tanto para cálculos orbitales de alta precisión para satélites, como usos geofísicos relacionados con las anomalías gravimétricas del geoide. Al emplear sus parámetros: (a) radio ecuatorial: $6.378.137{,}000$m; (b) radio polar: $6.356.752{,}314$m (Ulziisaikhan y Oyuntsetseg, 2020), para obtener la excentricidad

ε resulta en:

$$\varepsilon = \sqrt{1 - \frac{b^2}{a^2}} = \sqrt{1 - \left(\frac{6.356.752{,}314}{6.378.137{,}000}\right)^2} = 8{,}181919131086976 \times 10^{-2}.$$

Dicho dato ε es perfectamente aproximado al que arroja (Meza, 2011), de $8{,}18191908426 \times 10^{-2}$, por lo que ahora se está en la posibilidad de estimar cualquier latitud geocéntrica mediante el uso de la fórmula propuesta por (Snyder, 1987):

$$\psi = \arctan[(1 - \varepsilon^2) \tan \phi'], \tag{2.1}$$

donde ψ se refiere a la latitud geocéntrica; ϕ' representa a la latitud geográfica geodésica[1], ambas operadas en radianes, y ε, como se ha visto, es la excentricidad.

Entonces, dada la latitud ϕ', por (2.1) se obtiene la latitud geocéntrica ψ. Observemos que si $\varepsilon = 0$, como en el caso de la esfera, ϕ' y ψ coinciden. Ahora bien,

$$1 - \varepsilon^2 = 1 - \left(\frac{c}{a}\right)^2 = 1 - \frac{a^2 - b^2}{a^2} = \left(\frac{b}{a}\right)^2;$$

lo que quiere decir que:

$$\psi = \arctan\left((1 - \varepsilon^2) \tan \phi'\right) = \psi = \arctan\left(\left(\frac{b}{a}\right)^2 \tan \phi'\right).$$

De esa forma, es posible demostrar la relación entre ϕ' y ψ. Aplicando la función tangente a ambos lados de la ecuación obtenemos:

$$\tan \psi = \left(\frac{b}{a}\right)^2 \tan \phi'$$

o, equivalentemente,

$$\tan \phi' = \left(\frac{a}{b}\right)^2 \tan \psi.$$

Veamos como se deduce esta última fórmula. Sea (x_0, y_0, z_0) un punto del elipsoide con $y_0 = 0$, entonces la ecuación normal de la recta en el punto, en el plano $y_0 = 0$, viene dada por

$$z - z_0 = \left(\frac{a}{b}\right)^2 \left(\frac{z_0}{x_0}\right) (x - x_0).$$

Por lo tanto, cuando $z = 0$ obtenemos

$$-z_0 = \left(\frac{a}{b}\right)^2 \left(\frac{z_0}{x_0}\right)(x - x_0) \quad \Rightarrow \quad -z_0 = \left(\frac{b}{a}\right)^2 \left(\frac{x_0}{z_0}\right) = x - x_0,$$

[1]La latitud geodésica de un punto en el elipsoide se define como el ángulo que forma la normal que pasa por ese punto y el plano ecuatorial, mientras que la latitud geocéntrica se define como el ángulo entre el plano ecuatorial y la recta que conecta el centro del elipsoide con un punto de la superficie. Obviamente, en una esfera la latitud y la latitud geocéntrica coinciden, pero ese no es el caso del elipsoide que representa la superficie de la Tierra.

lo cual implica que,

$$x = -z_0\left(\frac{b}{a}\right)^2\left(\frac{x_0}{z_0}\right) + x_0 = x_0 - x_0\left(\frac{b}{a}\right)^2.$$

Con esto se concluye que la normal a la elipse en el punto $(x_0, 0, z_0)$ corta al plano $z = 0$ en el punto $(x_0 - x_0(\frac{b}{a})^2, 0, 0)$ y, por tanto,

$$\tan\phi' = \frac{z_0}{x_0 - \left(x_0 - x_0\left(\frac{b}{a}\right)^2\right)} = \left(\frac{a}{b}\right)^2\frac{z_0}{x_0} = \left(\frac{a}{b}\right)^2\tan\psi,$$

que es la fórmula que se quiere probar. El caso de $y_0 \neq 0$ es similar debido a que el elipsoide se obtiene como superficie de revolución alrededor del eje z a partir de una elipse.

ϕ'	ψ	ϕ'	ψ	ϕ'	ψ	ϕ'	ψ
$1°$	$0°\,59'\,35{,}91''$	$26°$	$25°\,50'\,55{,}26''$	$51°$	$50°\,48'\,41{,}95''$	$76°$	$75°\,54'\,33{,}82''$
$2°$	$1°\,59'\,11{,}84''$	$27°$	$26°\,50'\,40{,}69''$	$52°$	$51°\,48'\,47{,}32''$	$77°$	$76°\,54'\,55{,}42''$
$3°$	$2°\,58'\,47{,}83''$	$28°$	$27°\,50'\,26{,}79''$	$53°$	$52°\,48'\,53{,}51''$	$78°$	$77°\,55'\,17{,}38''$
$4°$	$3°\,58'\,23{,}91''$	$29°$	$28°\,50'\,13{,}59''$	$54°$	$53°\,49'\,0{,}51''$	$79°$	$78°\,55'\,39{,}69''$
$5°$	$4°\,58'\,0{,}11''$	$30°$	$29°\,50'\,1{,}10''$	$55°$	$54°\,49'\,8{,}32''$	$80°$	$79°\,56'\,2{,}33''$
$6°$	$5°\,57'\,36{,}45''$	$31°$	$30°\,49'\,49{,}33''$	$56°$	$55°\,49'\,16{,}92''$	$81°$	$80°\,56'\,25{,}25''$
$7°$	$6°\,57'\,12{,}96''$	$32°$	$31°\,49'\,38{,}31''$	$57°$	$56°\,49'\,26{,}31''$	$82°$	$81°\,56'\,48{,}44''$
$8°$	$7°\,56'\,49{,}68''$	$33°$	$32°\,49'\,28{,}04''$	$58°$	$57°\,49'\,36{,}48''$	$83°$	$82°\,57'\,11{,}87''$
$9°$	$8°\,56'\,26{,}62''$	$34°$	$33°\,49'\,18{,}54''$	$59°$	$58°\,49'\,47{,}41''$	$84°$	$83°\,57'\,35{,}50''$
$10°$	$9°\,56'\,3{,}82''$	$35°$	$34°\,49'\,9{,}81''$	$60°$	$59°\,49'\,59{,}08''$	$85°$	$84°\,57'\,59{,}31''$
$11°$	$10°\,55'\,41{,}31''$	$36°$	$35°\,49'\,1{,}88''$	$61°$	$60°\,50'\,11{,}50''$	$86°$	$85°\,58'\,23{,}27''$
$12°$	$11°\,55'\,19{,}11''$	$37°$	$36°\,48'\,54{,}74''$	$62°$	$61°\,50'\,24{,}64''$	$87°$	$86°\,58'\,47{,}35''$
$13°$	$12°\,54'\,57{,}25''$	$38°$	$37°\,48'\,48{,}41''$	$63°$	$62°\,50'\,38{,}48''$	$88°$	$87°\,59'\,11{,}52''$
$14°$	$13°\,54'\,35{,}75''$	$39°$	$38°\,48'\,42{,}90''$	$64°$	$63°\,50'\,53{,}00''$	$89°$	$88°\,59'\,35{,}74''$
$15°$	$14°\,54'\,14{,}65''$	$40°$	$39°\,48'\,38{,}21''$	$65°$	$64°\,51'\,8{,}20''$	$90°$	$90°$
$16°$	$15°\,53'\,53{,}96''$	$41°$	$40°\,48'\,34{,}35''$	$66°$	$65°\,51'\,24{,}06''$		
$17°$	$16°\,53'\,33{,}72''$	$42°$	$41°\,48'\,31{,}33''$	$67°$	$66°\,51'\,40{,}54''$		
$18°$	$17°\,53'\,13{,}94''$	$43°$	$42°\,48'\,29{,}14''$	$68°$	$67°\,51'\,57{,}64''$		
$19°$	$18°\,52'\,54{,}65''$	$44°$	$43°\,48'\,27{,}79''$	$69°$	$68°\,52'\,15{,}33''$		
$20°$	$19°\,52'\,35{,}88''$	$45°$	$44°\,48'\,27{,}29''$	$70°$	$69°\,52'\,33{,}58''$		
$21°$	$20°\,52'\,17{,}64''$	$46°$	$45°\,48'\,27{,}63''$	$71°$	$70°\,52'\,52{,}39''$		
$22°$	$21°\,51'\,59{,}96''$	$47°$	$46°\,48'\,28{,}81''$	$72°$	$71°\,53'\,11{,}72''$		
$23°$	$22°\,51'\,42{,}86''$	$48°$	$47°\,48'\,30{,}84''$	$73°$	$72°\,53'\,31{,}56''$		
$24°$	$23°\,51'\,26{,}37''$	$49°$	$48°\,48'\,33{,}71''$	$74°$	$73°\,53'\,51{,}87''$		
$25°$	$24°\,51'\,10{,}49''$	$50°$	$49°\,48'\,37{,}41''$	$75°$	$74°\,54'\,12{,}63''$		

TABLA 2. Latitud geocéntrica para latitud geodésica del elipsoide WGS84, grados seleccionados. Fuente: elaboración personal

Ahora bien, como se cuenta con el valor de la primera excentricidad, al sustituir valores para el caso de la latitud geográfica de $30°$ $(0{,}523598776$ radianes$)$, se tiene que:

$$\begin{aligned}\psi &= \arctan\left((1-0{,}08181919131086976^2)\cdot\tan(0{,}523598776)\right)\\ &= \arctan\left((1-0{,}00669438006676470)\cdot 0{,}577350269\right)\\ &= \arctan(0{,}5734852668676651).\end{aligned}$$

Luego,

$$\psi = 0{,}520695172 \text{ radianes},$$

lo cual resulta en $29{,}8336358$ grados decimales de latitud geocéntrica, o sea $29^\circ\, 50'\, 01{,}088''$. Al iterar el proceso se obtienen las cifras de la Tabla 2. Nótese que la mayor diferencia entre latitudes geográficas y latitudes geocéntricas ocurre al centro de la distancia angular, a los 45° de latitud geográfica.

Ahora bien, toda vez que se conoce la latitud geocéntrica de un elipsoide y si son conocidos el semieje mayor a y la razón de aplanamiento f, se pueden obtener los valores del semieje menor b y de la excentricidad ε (o primera excentricidad), además de estar en la posibilidad de calcular el radio geocéntrico de cualquier latitud, puesto que:

$$\frac{x^2}{a^2}+\frac{z^2}{b^2}=1 \quad\Rightarrow\quad r^2b^2+z^2z^2=a^2b^2 \quad\Rightarrow\quad (p\cos\psi)^2\,b^2+(p\operatorname{sen}\psi)^2\,a^2=a^2b^2,$$

luego,

$$p^2=\frac{a^2b^2}{(a\operatorname{sen}\psi)^2+(b\cos\psi)^2} \quad\Rightarrow\quad p=\frac{ab}{\sqrt{(a\operatorname{sen}\psi)^2+(b\cos\psi)^2}}.$$

Por tanto, habida cuenta de que $b^2=a^2(1-\varepsilon^2)$, se llega a la fórmula propuesta por (Corchete, 2009):

$$p=\frac{a\sqrt{1-\varepsilon^2}}{\sqrt{1-\varepsilon^2\cos^2\psi}}.$$

Con la cual se puede deducir la fórmula para calcular el radio geocéntrico de cualquier latitud geocéntrica:

$$p=\frac{a\sqrt{1-\varepsilon^2}\operatorname{sen}^2\psi}{\sqrt{1-\varepsilon^2}\cos^2\psi},$$

donde el radio geocéntrico es p, a el semieje mayor (radio ecuatorial), ε se refiere a la primera excentricidad y ψ, en este caso representa a la latitud geocéntrica. Siendo sus valores para la sustitución: $a = 6.378.137{,}000$ m; $\varepsilon = 0{,}08181919092890633$; y en este caso se va a calcular p radio geocéntrico para la latitud geográfica ϕ' de $23^\circ\, 26'\, 47''$, la cual corresponde a una latitud geocéntrica ψ de $23^\circ\, 18'\, 22{,}42''$, cuyo valor decimal es: $23{,}3062292369669$, y en radianes es $0{,}406770436409638$, de modo que:

$$\begin{aligned}
p &= \frac{6378137{,}000\sqrt{1-0{,}0818191909289063 3^2}\cdot \operatorname{sen}^2(0{,}406770436409638)}{\sqrt{1-0{,}08181919092890633^2}\cdot\cos^2(0{,}406770436409638)}\\
&= \frac{6378137{,}000\sqrt{1-0{,}00669438\cdot 0.156531739}}{\sqrt{1-0{,}00669438\cdot 0.843468261}}\\
&= \frac{6378137{,}000\sqrt{1-0{,}001047883}}{\sqrt{(1-0{,}005646497)}}\\
&= \frac{6378137{,}000\sqrt{0{,}998952117}}{\sqrt{0{,}994353503}}\\
&= \frac{6378137{,}000\cdot 0{,}99947592}{0{,}996895802} = \frac{6354993{,}954}{0{,}996895802} = 6374782{,}54\text{ m}.
\end{aligned}$$

Al iterar el proceso se obtienen las cifras de la Tabla 3, a partir de las cuales se puede conocer y comparar las magnitudes de cada radio geocéntrico, según la latitud del mismo, y observar desde otro punto las concepciones en torno a estos conceptos medulares de la geografía.

Para hacer inteligible dichos cálculos, en lo particular para usuarios no familiarizados con el campo de la geometría satelital ni con la geodesia, pero que usan de forma tangencial información de esta última ciencia, se cita el siguiente ejemplo, sean las latitudes de los puertos de Salina Cruz, Oaxaca, Tampico, Tamaulipas y Ensenada, Baja California: $16^\circ\,09'\,37''$ N; $22^\circ\,15'\,52''$ N y $31^\circ\,51'\,36''$ N, respectivamente, sus radios geocéntricos son: 6.376.497,90; 6.375.096,63 y 6.372.224,61 m, respectivamente. Lo cual, teóricamente, significa que, aunque los tres se encuentran a 0 msnm, entre el nivel del mar de Salina Cruz, Oaxaca (aproximadamente paralelo 16° de la tabla 3) y el de Ensenada, Baja California (aproximadamente paralelo 32° de la tabla 3) hay 4.273,29 m de diferencia, respecto al centro del planeta Tierra, en otras palabras, así como el nivel del mar del Océano Ártico se encuentra 21 mil metros más cerca del centro de la Tierra respecto al nivel medio de los mares ecuatoriales, el nivel del mar del puerto de Ensenada, Baja California se encuentra 4.273,29 m, más cerca del centro de la Tierra que el puerto de Salina Cruz, Oaxaca.

Esta perspectiva es importante, debido a que recientemente se está comprendiendo qué las anomalías del relieve, así como las anomalías en los distintos niveles del mar están también asociados al campo gravitatorio del planeta Tierra según su región geofísica, ergo, latitud, por lo que es importante conocer esta perspectiva que arroja el conocimiento del radio geocéntrico de los distintos grados de latitud de un elipsoide y su relación con la geodesia relativista (Kopeikin, 2019).

ϕ'	ψ	**R Geocéntrico**	ϕ'	ψ	**R Geocéntrico**	ϕ'	ψ	**R Geocéntrico**
1°	0,993	6378130,563	31°	30,830	6372509,079	61°	60,836	6361813,434
2°	1,986	6378111,258	32°	31,827	6372177,878	62°	61,840	6361499,312
3°	2,979	6378079,108	33°	32,824	6371840,852	63°	62,844	6361192,545
4°	3,973	6378034,151	34°	33,821	6371498,408	64°	63,848	6360893,517
5°	4,966	6377976,439	35°	34,819	6371150,957	65°	64,852	6360602,604
6°	5,960	6377906,039	36°	35,817	6370798,921	66°	65,856	6360320,172
7°	6,953	6377823,033	37°	36,815	6370442,724	67°	66,861	6360046,574
8°	7,947	6377727,518	38°	37,813	6370082,799	68°	67,866	6359782,154
9°	8,940	6377619,605	39°	38,811	6369719,584	69°	68,870	6359527,242
10°	9,934	6377499,42	40°	39,810	6369353,52	70°	69,875	6359282,157
11°	10,928	6377367,103	41°	40,809	6368985,055	71°	70,881	6359047,204
12°	11,921	6377222,809	42°	41,808	6368614,638	72°	71,886	6358822,676
13°	12,915	6377066,707	43°	42,808	6368242,723	73°	72,892	6358608,85
14°	13,909	6376898,979	44°	43,807	6367869,766	74°	73,897	6358405,993
15°	14,904	6376719,822	45°	44,807	6367496,226	75°	74,903	6358214,353
16°	15,898	6376529,445	46°	45,807	6367122,562	76°	75,909	6358034,166
17°	16,892	6376328,073	47°	46,808	6366749,236	77°	76,915	6357865,651
18°	17,887	6376115,941	48°	47,808	6366376,709	78°	77,921	6357709,013
19°	18,881	6375893,299	49°	48,809	6366005,442	79°	78,927	6357564,44
20°	19,876	6375660,41	50°	49,810	6365635,895	80°	79,933	6357432,105
21°	20,871	6375417,547	51°	50,811	6365268,527	81°	80,940	6357312,165
22°	21,866	6375164,997	52°	51,813	6364903,795	82°	81,946	6357204,76
23°	22,861	6374903,06	53°	52,814	6364542,153	83°	82,953	6357110,013
24°	23,857	6374632,045	54°	53,816	6364184,053	84°	83,959	6357028,031
25°	24,852	6374352,274	55°	54,818	6363829,941	85°	84,966	6356958,905
26°	25,848	6374064,078	56°	55,821	6363480,261	86°	85,973	6356902,709
27°	26,844	6373767,8	57°	56,823	6363135,449	87°	86,979	6356859,499
28°	27,840	6373463,794	58°	57,826	6362795,939	88°	87,986	6356829,316
29°	28,837	6373152,422	59°	58,829	6362462,155	89°	88,993	6356812,183
30°	29,833	6372834,057	60°	59,833	6362134,516	90°	90	6356808,107

TABLA 3. Radio geocéntrico en metros para latitud geográfica y latitud geocéntrica del elipsoide WGS84, grados seleccionados. Fuente: elaboración personal.

§3. Consideraciones finales

Actualmente el modelado de la superficie terrestre atraviesa por el debate entre la mecánica clásica y la mecánica cuántica, es decir, para comprender los fenómenos físico geográficos, a escala regional y continental, además de la física newtoniana, se requiere la consideración de las mediciones más precisas de la forma de la Tierra que arrojan los nuevos elipsoides de referencia basados en nuevos procesos. Para ello es importante familiarizarse con la precisión de los distintos elipsoides de referencia geodésica mundial, con la intención de apreciar las considerables variaciones en magnitud que presentan, en el radio, paralelos con tan solo un grado angular de diferencia.

Esta medida proporciona una nueva interpretación de las concepciones en torno a conceptos geográficos, por ejemplo, se puede reinterpretar el tamaño de cimas y valles, midiendo no su altura, sino su distancia al centro de la Tierra, por ejemplo, entre el nivel del mar de Salina Cruz, Oaxaca y el de Ensenada, Baja California hay $4{,}2$ km de diferencia, respecto al centro del planeta Tierra, en otras palabras, el nivel del mar del puerto de Ensenada, Baja California se encuentra 4 km más cerca del centro de la Tierra que el puerto de Salina Cruz, Oaxaca.

Esta perspectiva es de utilidad, debido a que últimamente se está comprendiendo qué, las anomalías del relieve y las anomalías de los distintos niveles del mar están asociados a las variaciones que experimenta el campo gravitatorio del planeta Tierra, por lo que es importante divulgar y ampliar ésta perspectiva que arroja el conocimiento del radio geocéntrico de los distintos paralelos o grados de latitud de cualquier elipsoide de referencia.

Bibliografía

Corchete, V. (2009). *Elipsoide de revolución: superficie de referencia y superficie equipotencial*. Almería, España: Universidad de Almería. doi: http://doi/10.13140/RG.2.2.10394.88003

De Andrade, R. (2008). A herança de Sacrobosco e seus comentadores: desenvolvimentos e erros na astronomia geocêntrica do século XVI. En R. De Andrade, C. Silva, J. Hidalgo, y L. Al-Chueyr (Eds.), *Filosofia e Historia da Ciência no Cone Sul: Seleção de trabalhos do 5o. encontro* (p. 373-382). Associação de Filosofia e História da Ciência do Cone Sul.

Franco, R. (1999). *Nociones de topografía geodesia y cartografía*. Cáceres, España: Universidad de Extremadura.

Gómez, C. (2007). La forma de la tierra: expedición para medir un grado del arco de meridiano en el virreinato del Perú (1735-1744). *EGA. Revista de Expresión Gráfica Arquitectónica*, *1*(12), 128-139.

Jo, K., Lee, M., y Sunwoo, M. (2015). Fast GPS-DR sensor fusion framework: removing the geodetic coordinate conversion process. *IEEE Transactions on Intelligent Transportation Systems*, *17*(7), 2008-2013.

Kopeikin, S. (2019). Reference-Ellipsoid and Normal Gravity Field in Post-Newtonian Geodesy. En D. Puetzfeld y C. Lämmerzahl (Eds.), *Relativistic Geodesy: Foundations and Applications", Fundamental Theories of Physics* (p. 159-228). Springer.

Meza, P. (2011). *El datum en navegación*. Valdivia, Chile: Universidad Austral de Chile, Facultad de Ciencias de la Ingeniería.

Snyder, J. (1987). *Map projections-A working manual* (*Vol. 1395*). Washington D. C, Estados Unidos de América: United States Geological Survey.

Ulziisaikhan, G., y Oyuntsetseg, D. (2020). UAV and terrestrial laser scanner data processing for large scale topographic mapping. *Mongolian Geoscientist*, *50*, 63-73.

Veloso, F. (2016). A elaboração de uma nova descrição geral da Terra nos primeiros séculos da época moderna (1522–1780). *Revista Equador*, *5*(2), 159-189.

Rodrigo Tovar Cabañas
El Colegio de Veracruz, México
(✉) *rod_ geo77@hotmail.com*

Hipólito Villanueva Hernández
Universidad Autónoma de Nuevo León, México
(✉) *polo_ arase@hotmail.com*

Shany Arely Vázquez Espinosa
Universidad Veracruzana, México
(✉) *shanyvaz@gmail.com*

Recibido: *8 de abril de 2022*.
Aceptado: *12 de abril de 2023*.
Publicado en línea: *3 de agosto de 2023*.

Revista de Educación Matemática
Volumen 38, N° 2 (2023), páginas 16 – 27
Unión Matemática Argentina - Famaf (UNC)

Artículo de Matemática

EL DISCRETO ENCANTO DE LA RAÍZ CUADRADA

Pablo Amster y José Ángel Cid

Resumen. Presentamos una versión divulgativa de nuestro artículo (Amster y Cid, 2022) en el que mostramos cómo la raíz cuadrada compleja nos permite demostrar de una forma muy sencilla diversos resultados topológicos en el plano, tan profundos como el Teorema de Brouwer y el Teorema de Invariancia del Dominio, así como del Teorema Fundamental del Álgebra.

Abstract. We present a version for the general readership of our article (Amster y Cid, 2022) in which we show how the complex square root allows us to prove in a very simple manner various topological results in the plane, as deep as the Brouwer Fixed Point Theorem and the Invariance of Domain Theorem, as well as the Fundamental Theorem of Algebra.

§1. Introducción

En cualquier curso de variable compleja, uno de los momentos culminantes es aquel en el que se demuestra el *Teorema Fundamental del Álgebra* (*TFA*), según el cual todo polinomio no constante admite al menos una raíz. Claro que, a esa altura del curso, seguramente ya estamos al tanto de muchas propiedades importantes de las funciones llamadas *holomorfas*, vale decir, aquellas que son derivables (en sentido complejo, por supuesto) en una región abierta del plano. Las propiedades de tales funciones son tan mágicas que el teorema se deduce de varias maneras distintas, como consecuencia inmediata de otros resultados. Sin embargo, esto requiere un buen trabajo antes de llegar al hueso: posiblemente antes de enunciar el teorema hayamos lidiado no solo con el concepto de derivada sino también con el de integración a lo largo de curvas. En realidad, si el único objetivo es llegar al teorema en cuestión, se puede evitar hablar de todo esto y estudiar, en cambio, las propiedades de las funciones *analíticas*, que son las que admiten localmente

Palabras clave: Raíz cuadrada compleja, Teorema de Brouwer, Teorema fundamental del álgebra, Índice de una curva.

Keywords: Complex square root, Brouwer Theorem, Fundamental Theorem of Algebra, Winding number.

un desarrollo en serie de potencias. Sin duda, para hacerlo se requiere una buena dosis de simulación. Con aire distraído, tenemos que fingir que desconocemos un hecho que recién se demuestra en una etapa bastante más avanzada del curso: las funciones analíticas son, ni más ni menos, las funciones holomorfas. Como sea, también este camino tiene sus vericuetos: en principio, debemos aprender lo que es una serie de potencias, el radio de convergencia y luego, para llegar a los teoremas picantes, emplear alguna herramienta algo más sofisticada (por ejemplo, el logaritmo complejo).

Por otra parte, el análisis complejo nos abre también la puerta a las maravillas de la topología. En particular, una noción intuitivamente simple como la de *índice*, que cuenta el número de vueltas que da una curva en torno a un punto, permite probar la versión bidimensional de otro gran teorema, el de Brouwer. En este contexto, dicho teorema se puede formular así: *Sea $D_R \subset \mathbb{C}$ el disco cerrado de radio $R > 0$ centrado en el origen. Si $f : D_R \to D_R$ es una función continua, entonces f tiene al menos un punto fijo, es decir, existe $z \in D_R$ tal que $f(z) = z$.* Un resultado algo más sutil es el de *invariancia del dominio*, que dice: *Si $U \subset \mathbb{C}$ es abierto y $f : U \to \mathbb{C}$ es continua e inyectiva, entonces $V := f(U)$ es abierto y $f : U \to V$ es un homeomorfismo (es decir, la inversa $f^{-1} : V \to U$ también es continua).* Cabe agregar que el índice también permite dar una demostración muy elegante y directa (o bien "sin vueltas", si vale el pequeño contrasentido) del mencionado TFA. Pero, nuevamente, llegar a estas instancias requiere por lo general una definición del índice mediante integrales cuando la curva es suave y –si se pretende ahondar en los aspectos topológicos– un esfuerzo adicional para extender dicha definición a curvas cerradas que solamente son continuas.

El objetivo de este artículo consiste en presentar demostraciones elementales de todos los resultados mencionados, que no requieren derivadas, series ni integrales, sino apenas un ingrediente central que se come con todas las comidas: la raíz cuadrada. Para ser absolutamente precisos, ni siquiera hace falta saber cómo se define la raíz cuadrada compleja: alcanza con conocer la raíz real de toda la vida (exceptuando la infancia más tierna, se entiende) y conceptos básicos de continuidad. Tampoco hace falta saber mucho de los complejos en sí: apenas las operaciones básicas y una noción bien conocida, la de módulo. En resumen, para atacar este artículo solo es indispensable prepararse un mate o un café... y animarse a encarar la lectura, sin complejos. Bueno, a decir verdad con unos cuantos, pero únicamente en el sentido matemático del término.

§2. Pero, ¿qué es una raíz cuadrada?

En la introducción hemos prometido grandes hazañas a partir de la más pura e ingenua noción de raíz cuadrada de un número complejo. Y esto es algo que todo el mundo aprendió en el colegio: dado cierto $z \neq 0$, sabemos cómo encontrar sus dos raíces recurriendo a la escritura binomial o incluso a la magnánima forma polar. Sin embargo, puede ocurrir que nos hayamos despertado poco magnánimos y no queramos saber nada de todo eso. ¡No importa! Podemos llevar a cabo nuestra tarea de manera aún más sencilla.

Por empezar, observemos que si w es una raíz cuadrada de z (es decir, $w^2 = z$), entonces también lo es $-w$, ya que $(-w)^2 = w^2 = z$. Y esas son las únicas dos raíces que existen, pues vale

$$0 = z^2 - w^2 = (z - w)(z + w),$$

de donde $z = w$, o bien $z = -w$. Además, asumiendo que ya conocemos la raíz cuadrada real, entonces alcanza con suponer que z tiene módulo 1. En efecto, dado cualquier $\tilde{z} \neq 0$, si encontramos una raíz cuadrada w de $z := \frac{\tilde{z}}{|\tilde{z}|}$, entonces $\tilde{w} := \sqrt{|\tilde{z}|}w$ verifica:

$$\tilde{w}^2 = |\tilde{z}|w^2 = |\tilde{z}|z = \tilde{z}.$$

De paso, podemos también asumir que $z \neq -1$, cuyas raíces cuadradas ya conocemos: se trata –nada menos– que de las famosas i y $-i$.

Supongamos entonces que $|z| = 1$ y $z \neq -1$. Nuestro método para hallar la raíz cuadrada va a ser puramente geométrico, pero (tal como prometimos) sin recurrir a ángulos. Podemos dejar de lado el caso $z = 1$, que es más bien facilongo. Comencemos calculando el punto medio entre z y 1, es decir, $\frac{z+1}{2}$. Dividiendo este número por su módulo obtenemos

$$w = \frac{z+1}{|z+1|},$$

que también vive en la circunferencia unitaria y equidista de 1 y z, como se puede ver en la Figura 1. Pero w^2 se encuentra exactamente a la misma distancia de w, ya que

$$|w^2 - w| = |w| \cdot |w - 1| = |w - 1|.$$

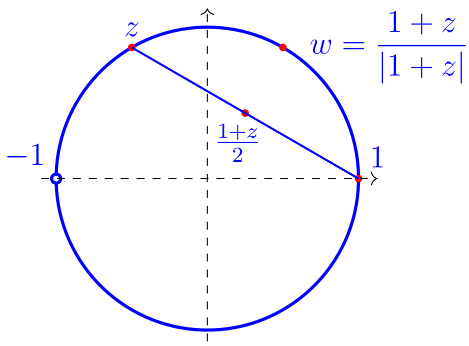

Figura 1

Dicho de otra forma, 1, z y w^2 se encuentran en una única circunferencia centrada en w. Pero, ¡atención! Dos circunferencias no se pueden cortar en más de dos puntos, de donde se deduce que $w^2 = z$. ¡Excelente noticia! Tal como anticipamos, este procedimiento nos permite definir sin mayor esfuerzo una *función* raíz cuadrada sobre el conjunto $A^+ := \mathbb{C} \smallsetminus \mathbb{R}_{\leq 0}$ de manera muy simpática:

$$r_+(z) := \sqrt{|z|}\frac{\frac{z}{|z|}+1}{\left|\frac{z}{|z|}+1\right|} = \sqrt{|z|}\frac{z+|z|}{|z+|z||}.$$

Es claro, además, que r_+ resulta continua. De paso, observemos también que $r_+(A^+) \subset A^+$, así que esta función se puede aplicar sucesivamente para obtener raíces cuartas, octavas, etc. Y, por supuesto, también tenemos una función parecida r_- definida sobre el conjunto $A^- := \mathbb{C} \smallsetminus \mathbb{R}_{\geq 0}$, entretenimiento que quedará para quienes hayan llegado hasta aquí. Pero pasemos a la próxima sección, en la que nos esperan varios asuntos importantes.

§3 Dadme una curva y extraeré sus raíces

En la sección previa vimos que cualquier número complejo $z \neq 0$ tiene exactamente dos raíces cuadradas. Vimos también que es posible definir, por ejemplo en A^+, una función continua que a cada z le asigna una de sus raíces cuadradas. Sin embargo, pronto mostraremos que esto no se puede hacer para un subconjunto arbitrario de $\mathbb{C} \smallsetminus \{0\}$; más en general, una función que no se anula no tiene por qué admitir una raíz cuadrada continua. Para nuestros fines, alcanzará con estudiar las curvas cerradas (lazos) que no pasan por el origen aunque, por comodidad, no vamos a pensar en curvas parametrizadas sino simplemente en funciones continuas $\gamma : S_R \to \mathbb{C} \smallsetminus \{0\}$, donde S_R es la circunferencia de radio R centrada en el origen, vale decir, el borde del disco que antes llamamos D_R. Por conveniencia, vamos a llamar $\mathcal{A}$ al conjunto de dichas funciones,

$$\mathcal{A} := \{\gamma : S_R \to \mathbb{C} \smallsetminus \{0\} : \gamma \text{ es continua}\},$$

que tienen la pinta que se muestra en la Figura 2.

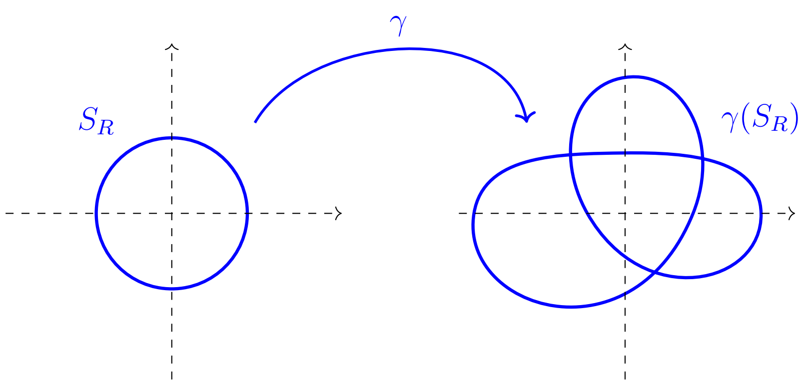

Figura 2

Entonces, dada $\gamma \in \mathcal{A}$, la pregunta es si siempre existe una raíz cuadrada continua, vale decir, cierta $\sigma \in \mathcal{A}$ tal que $\sigma(z)^2 = \gamma(z)$ para todo z. Es fácil ver que no y, en tal caso, no hay por qué enfurruñarse: al fin y al cabo, eso es lo que ocurre con funciones tan renombradas como la identidad o, más en general, con cualquier función *impar*. En efecto, si suponemos que $\gamma \in \mathcal{A}$, $\gamma(-z) = -\gamma(z)$ y γ tiene una raíz cuadrada continua σ, entonces

$$\frac{\sigma(-z)^2}{\sigma(z)^2} = \frac{\gamma(-z)}{\gamma(z)} = -1.$$

Esto dice que $\left(\frac{\sigma(-z)}{\sigma(z)}\right)^2 = -1$ para todo z, de modo que la función $\frac{\sigma(-z)}{\sigma(z)}$ solamente puede tomar los valores i y $-i$. Pero una función continua sobre la circunferencia no puede andar por ahí dando saltitos de un valor a otro, así que hay únicamente dos opciones:

$$\sigma(-z) \equiv i\sigma(z) \qquad \text{o bien} \qquad \sigma(-z) \equiv -i\sigma(z).$$

En cualquiera de los dos casos, esta conclusión nos lleva a un lamentable absurdo, ya que entonces

$$\sigma(1) = \sigma(-(-1)) = (\pm i)^2\sigma(1) = -\sigma(1).$$ [1]

En resumen,

Proposición 3.1. *Si $\gamma \in \mathcal{A}$ es impar, entonces no existe $\sigma \in \mathcal{A}$ tal que $\sigma(z)^2 = \gamma(z)$ para todo z.*

Felizmente, el mundo no acaba con estas funciones tan poco dispuestas a que hurguemos en sus raíces. Vamos a definir ahora, para cada $\gamma \in \mathcal{A}$, un número $v(\gamma)$ del siguiente modo. Si γ no tiene raíz cuadrada continua, ya sabemos que no es la muerte de nadie, lo que nos proporciona el coraje suficiente para anotar $v(\gamma) = 0$. En cambio, si γ nos honra con una raíz σ, entonces nos preguntamos si σ tiene, a su vez, una raíz cuadrada, y así sucesivamente. Si logramos extraer n raíces consecutivas y ninguna más, entonces anotamos, siempre con prolijidad, $v(\gamma) = n$. En cambio, si se pueden seguir extrayendo raíces de manera indefinida, entonces anotamos $v(\gamma) = \infty$. En resumen, $v(\gamma) = n$ significa que existe $\sigma \in \mathcal{A}$ tal que $\gamma = \sigma^{2^n}$ pero no existe $\tau \in \mathcal{A}$ tal que $\gamma = \tau^{2^{n+1}}$; en cambio, $v(\gamma) = \infty$ quiere decir que para todo n existe $\sigma \in \mathcal{A}$ tal que $\gamma = \sigma^{2^n}$. Algunas propiedades de v son muy fáciles de verificar:

[1] Este argumento es el mismo que, con cierta tosquedad, nos enseñaron en el colegio secundario a fin de advertirnos sobre los peligros de una distributividad mal aplicada:

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = -1.$$

Errores similares se describen en el artículo (Benítez y Carena, 2021) donde, si bien se deja de lado la radicación compleja, se menciona por ejemplo que $\sqrt{(-25)\cdot(-4)} = 10 \neq \sqrt{(-25)}\cdot\sqrt{(-4)}$.

1. Si $\mathrm{Im}(\gamma) \subset A^+$ o $\mathrm{Im}(\gamma) \subset A^-$, entonces $v(\gamma) = \infty$. Esto se explica por lo que ya dijimos de la función r_+, que se puede aplicar de manera sucesiva sobre el conjunto A^+, y otro tanto vale para la función r_- definida sobre A^-. La propiedad incluye, claro está, el virtuoso caso de las funciones constantes (no nulas).
2. Si $v\left(\dfrac{\gamma}{\sigma}\right) = \infty$ entonces $v(\gamma) = v(\sigma)$. En efecto, si por ejemplo $v(\gamma) = n < v(\sigma)$, entonces podemos elegir $\tau, v \in \mathcal{A}$ tales que
$$\sigma = \tau^{2^{n+1}}, \qquad \frac{\gamma}{\sigma} = v^{2^{n+1}},$$
de donde se obtiene $\gamma = (\tau v)^{2^{n+1}}$, un nuevo absurdo.
3. $v(\gamma^2) = v(\gamma) + 1$. Esto esconde un pequeño abuso de notación, pero es razonable asumir que "$\infty + 1 = \infty$". Justamente, cuando $v(\gamma) = \infty$ se tiene también que $v(\gamma^2) = \infty$ y no hay nada que probar. Luego, podemos suponer $v(\gamma) = n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Escribiendo $\gamma = \sigma^{2^n}$ resulta $\gamma^2 = \sigma^{2 \cdot 2^n} = \sigma^{2^{n+1}}$, de donde $v(\gamma^2) \geq n + 1$. Por otro lado, si existiese σ tal que $\sigma^{2^{n+2}} = \gamma^2$, entonces $\left(\dfrac{\sigma^{2^{n+1}}}{\gamma}\right)^2 = 1$, que equivale a decir,
$$\frac{\sigma(z)^{2^{n+1}}}{\gamma(z)} = \pm 1 \quad \text{para todo } z \in S_R.$$
En consecuencia –otra vez, por ese asunto de los saltitos– resulta
$$\gamma = \sigma^{2^{n+1}} \qquad \text{o bien} \quad \gamma = -\sigma^{2^{n+1}}.$$
En cualquiera de ambos casos, se deduce que $v(\gamma) \geq n + 1$, lo que es absurdo. Con un poco de espíritu inductivo, fácilmente podemos generalizar la propiedad y concluir que $v(\gamma^{2^m}) = v(\gamma) + m$ para todo $m \in \mathbb{N}_0$.

Llegado este punto, podemos preguntarnos para qué nos tomamos el trabajo de probar esta serie de trivialidades. Una respuesta posible es: porque tenemos muchas ganas de comprobar si, como se prometía en la introducción, realmente todo es tan fácil como aseguran los autores (mejor desconfiar de esa gente). En realidad, la función aquí llamada v tiene relación muy directa con la mencionada noción de índice y permite definirlo de manera muy sencilla, aunque esto requiere un poco más de esfuerzo. En lo que sigue mantendremos un nivel matemático absolutamente elemental y, aún así, lograremos probar varias cuestiones muy bonitas.

Por empezar, veamos que la función $v : \mathcal{A} \to \mathbb{N}_0 \cup \{\infty\}$ antes definida resulta continua. Esto puede parecer un tanto pomposo, pero nos referimos simplemente al hecho de que, para cualquier $\sigma \in \mathcal{A}$, se verifica que si γ está suficientemente cerca de σ, entonces el valor de v es el mismo para ambas. Específicamente, como σ no se anula, podemos fijar $\varepsilon > 0$ como el mínimo valor que alcanza $|\sigma|$, de modo que $|\sigma(z)| \geq \varepsilon$ para todo z. Si ahora $|\gamma(z) - \sigma(z)| < \varepsilon$ para todo z, entonces

$$\left| \frac{\gamma(z)}{\sigma(z)} - 1 \right| < \frac{\varepsilon}{|\sigma(z)|} \leq 1.$$

En consecuencia, la función $\dfrac{\gamma}{\sigma}$ se mantiene siempre en un disco abierto de radio 1 centrado en 1 y, en particular, dentro de A^+. Esto quiere decir que $v\left(\dfrac{\gamma}{\sigma}\right) = \infty$, de donde (por la propiedad 2) se obtiene $v(\gamma) = v(\sigma)$.

Observación 3.2. *El resultado previo se puede expresar, de manera equivalente, diciendo que la función v es constante sobre la bola abierta de radio ε centrada en σ, donde*

$$\varepsilon := \min_{z \in S_R} |\sigma(z)|.$$

Un poco más en general, la anterior demostración muestra en realidad que el valor de v se mantiene constante para cualquier curva γ tal que $|\gamma(z) - \sigma(z)| < |\sigma(z)|$ para todo z. En definitiva, esto prueba una versión elemental de otro teorema célebre, en este caso del análisis complejo. ¿Alguien se anima a decir cuál? ¡Touché!

Una segunda consecuencia de las propiedades que emplearemos se desprende directamente de otro hecho que (¡vaya, qué afortunada casualidad!) vimos al comienzo: si $\gamma \in \mathcal{A}$ es impar, entonces no admite una raíz cuadrada continua y luego $v(\gamma) = 0$.

Finalmente, la clave de nuestras demostraciones futuras se encuentra en otra propiedad crucial que verifica nuestro numerito v. Por comodidad, dada una función $f : D_R \to \mathbb{C}$ continua que no se anula sobre el borde, llamaremos directamente $v(f)$ al valor que toma v sobre la restricción de f sobre S_R, es decir, $v\left(f|_{S_R}\right)$.

Proposición 3.3. *Sea $f : D_R \to \mathbb{C} \setminus \{0\}$ continua. Entonces $v(f) = \infty$.*

Demostración. Consideremos la función continua $h : [0,1] \times S_R \to \mathbb{C}$ dada por $h(s,z) := f(sz)$. Como f no se anula, si llamamos $\gamma_s(z) := h(s,z)$ entonces la función $v(\gamma_s)$ está bien definida y es continua, lo que equivale a decir que es constante. Luego (por la propiedad 1) $v(f) = v(\gamma_1) = v(\gamma_0) = \infty$, ya que $\gamma_0 \equiv f(0)$. $\square$

Aunque parezca una tontería, vamos a formular también un enunciado que es el contrarrecíproco del anterior, pero con un agregado, una pequeña sutileza que va a mostrar la fuerza de este numerito v. Para decirlo más claramente: el hecho de que $v(f) \neq \infty$ no implica solamente que f se anula, sino que su imagen contiene un entorno del origen.

Proposición 3.4. *Sea $f : D_R \to \mathbb{C}$ continua y supongamos que f no se anula en S_R. Sea $\varepsilon := \min_{z \in S_R} |f(z)|$ y supongamos $v(f) < \infty$. Entonces para todo w con $|w| < \varepsilon$ existe $z \in D_R$ tal que $f(z) = w$.*

Demostración. Dado w de módulo menor que ε, para todo z se verifica trivialmente que $|(f(z) - w) - f(z)| = |w| < \varepsilon$ y, por la Observación 3.2, ya sabemos que $v(f - w) = v(f)$. Por hipótesis $v(f) < \infty$ y entonces la Proposición 3.3 implica que $f(z) - w$ tiene que anularse en D_R. □

¡Qué bien, ya tenemos todo lo que necesitamos! Estamos en condiciones de probar todo lo anunciado, aunque debemos tener un poco de paciencia: las reglas más básicas del *suspense* nos exigen pasar antes a la próxima sección.

§4. Lo prometido es deuda

En esta sección demostraremos, por fin, todos los resultados enunciados al comienzo, a partir de las tres propiedades elementales que tiene nuestro mágico "numerito". Hilando más fino, veremos en realidad que los resultados topológicos emplean únicamente las dos primeras, mientras que la tercera, de carácter más algebraico, servirá casualmente para el TFA. Para calentar motores, comencemos por el teorema de punto fijo de Brouwer:

Teorema 4.1. *Sea $f : D_R \to D_R$ continua. Entonces existe $z \in D_R$ tal que $f(z) = z$.*

Demostración. Podemos suponer que $f(z) \neq z$ para $z \in S_R$ pues, en caso contrario, no tendríamos nada que probar. Para $(s, z) \in [0, 1] \times S_R$ consideremos ahora la función $\gamma_s(z) := z - sf(z)$, que no se anula: para $s = 1$, precisamente porque lo estamos suponiendo; para $s \in [0, 1)$, porque

$$|sf(z)| = s|f(z)| \leq sR < R = |z|$$

para todo $z \in S_R$. Como antes, debido a la continuidad de v, esto quiere decir que $v(\gamma_s)$ es constante. El valor para $s = 0$ ya lo conocemos, porque se trata de la identidad; de esta forma,

$$v(z - f(z)) = v(\gamma_1) = v(\gamma_0) = 0$$

y, por la Proposición 3.3, concluimos que $z - f(z)$ tiene que anularse. □

El teorema de Brouwer sigue valiendo si se reemplaza D_R por cualquier conjunto equivalente (en el sentido topológico). Por ejemplo, componiendo la función f con traslaciones, es inmediato verificarlo para un disco cerrado centrado en cualquier punto.

Otro de los resultados que teníamos pendientes es el de invariancia del dominio, que se puede enunciar también de la siguiente manera:

Teorema 4.2. *Sean $U \subset \mathbb{C}$ abierto y $f : U \to \mathbb{C}$ continua e inyectiva. Entonces f es abierta, es decir, la imagen de cualquier conjunto abierto es un conjunto abierto.*

Demostración. Alcanza con probar que si $V \subset U$ es un disco centrado en un punto z_0, entonces $f(V)$ contiene un disco centrado en $f(z_0)$. Ya que hablamos antes de traslaciones, se puede suponer, por comodidad, que $z_0 = 0 = f(0)$.[2] Sea entonces $D_R \subset V$ y consideremos ahora para $(s, z) \in [0, 1] \times S_R$ la función

$$\gamma_s(z) := f(z) - f(-sz),$$

que para $s = 0$ coincide con f. Como f es inyectiva, γ_s no se anula y entonces, empleando una vez más la continuidad de la función v, se obtiene

$$v(f) = v(\gamma_0) = v(\gamma_1) = v(f(z) - f(-z)).$$

Pero esta última función es impar, así que $v(f) = 0$, y la Proposición 3.4 nos garantiza que $f(D_R)$ contiene un entorno de 0. $\square$

El teorema anterior nos está diciendo en realidad dos cosas: por un lado, que $f(U)$ es abierto en $\mathbb{C}$, por otro, que f establece un homeomorfismo entre U y $f(U)$.[3] Ninguna de las dos afirmaciones es baladí, pues por ejemplo una función continua y biyectiva puede tener inversa f^{-1} discontinua. Al lector interesado sobre estos temas le recomendamos el artículo (Cao Labora, 2020) de nuestro colega Dani Cao de la Universidad de Santiago de Compostela.

Para terminar, ¿qué tal si intentamos una prueba del Teorema Fundamental del Álgebra? Para tan honrosos fines, podemos antes hacer un cálculo muy sencillo, que nos pondrá en la pista acerca de quién es en verdad ese misterioso numerito v:

Proposición 4.3. *Sea $\gamma(z) = z^n$ para $n \in \mathbb{Z} \setminus \{0\}$ y escribamos $n = 2^m k$ con k impar. Entonces $v(\gamma) = m$.*

Demostración. Basta escribir $z^n = (z^k)^{2^m}$, aplicar la generalización de la propiedad 3

$$v(z^n) = v\left((z^k)^{2^m}\right) = v(z^k) + m$$

y observar que $v(z^k) = 0$, por ser z^k una función impar. $\square$

Teorema 4.4. *Todo polinomio $p(z)$ de grado mayor o igual que 1 tiene al menos una raíz compleja.*

Demostración. Podemos suponer que nuestro polinomio p es mónico, y considerar la función

$$h(s, z) = z^n + s(p(z) - z^n),$$

[2] En términos más precisos, esto se consigue reemplazando $f(z)$ por la función $f(z + z_0) - f(z_0)$.

[3] Esto último se debe al hecho de que, en general, una función entre dos espacios topológicos es continua si y solo si la preimagen de cualquier abierto es abierta. En el caso de f^{-1}, la equivalencia con el teorema anterior es obvia, pues la preimagen de A no es otra cosa que $f(A)$.

que para $R \gg 0$ no se anula si $z \in S_R$, puesto que z^n tiene grado n y $s(p(z) - z^n)$ tiene a lo sumo grado $n-1$. Pero $h(0,z) = z^n$ y $h(1,z) = p(z)$, de modo que

$$v(p) = v(z^n) < \infty.$$

La conclusión se deduce nuevamente de la Proposición 3.4. □

§5. Comentarios finales a modo de *Epílogo*

Por supuesto, quienes tengan alguna experiencia con la variable compleja podrán aspirar a obtener alguna relación más precisa entre la función v y el índice de una curva: sin ánimos de oficiar de *spoilers*, podemos mencionar el hecho de que v no es otra cosa que la valuación 2-ádica del valor absoluto de dicho índice, vale decir, la potencia de 2 que aparece en su factorización. Esto es lo que anticipa, con toda exactitud, la proposición de la sección previa; si hacemos lo mismo para todos los primos $p > 2$, entonces podemos brindar una definición alternativa de lo que significa el valor absoluto del índice de una curva γ, pensada como un elemento de $\mathcal{A}$. Más en general, se puede probar que el valor absoluto de dicho índice se puede definir como el mayor valor n tal que existe una raíz n-ésima de γ, o bien 0, en caso de que haya raíces de orden arbitrariamente alto. A simple vista, puede llamar la atención el hecho de que, en este último caso, existen raíces de todos los órdenes; sin embargo, esto no es una sorpresa: las curvas tales que $v(\gamma) = \infty$ son aquellas que pueden "desenroscarse" para transformarse, de manera continua, en una constante. Todo esto lo explicamos con mucho detalle en nuestro trabajo (Amster y Cid, 2022) cuya escritura nos dio mucha alegría.

Este artículo también hubiera podido titularse "Ser o no ser (un cuadrado), esa es la cuestión", puesto que nuestra curva γ satisface $v(\gamma) > 0$ si y solo si pertenece a la imagen de la aplicación cuadrática

$$\mathcal{A} \to \mathcal{A}, \quad \sigma \mapsto \sigma^2.$$

La importancia de los cuadrados en matemática ha sido glosado por Olga Taussky–Todd en dos bonitos artículos divulgativos, (Taussky, 1970, 1988), y los profundos resultados de H. Hopf sobre las álgebras de división se basan en gran medida en las propiedades de la aplicación cuadrática, (Ebbinghaus y cols., 1991, Capítulo 8).

Por otra parte, nuestra amiga la raíz cuadrada también permite caracterizar las regiones simplemente conexas (es decir, sin agujeritos) como aquellas regiones conexas en la que cualquier función analítica que no se anula admite una raíz cuadrada analítica. Esta propiedad hace una aparición estelar en la demostración de un resultado tan famoso como el Teorema de la Aplicación de Riemann que asegura que cualquier abierto simplemente conexo distinto de $\mathbb{C}$ (y obviamente del vacío) es equivalente al disco unitario abierto mediante una aplicación analítica que preserva ángulos, (Conway, 1978).

Observe también el lector que en nuestra definición de la función r_+ interviene su prima $\sqrt{\cdot}$, la raíz cuadrada real. Si le han entrado sudores fríos intentando recordar el algoritmo escolar para calcular la raíz cuadrada de un número $a > 0$ entonces agradecerá saber que un método geométrico mucho más simple fue ya reportado por Herón de Alejandría en el siglo I, (Gasull, 2022; Hasselblatt y Katok, 2003): partiendo de una estimación inicial x_0 para la raíz construimos un rectángulo de área a, esto es, de lados x_0 y $\dfrac{a}{x_0}$. Luego, no es una locura pensar que el promedio de ambos lados

$$x_1 = \frac{1}{2}\left(x_0 + \frac{a}{x_0}\right)$$

será una mejor aproximación para intentar construir un cuadrado de área a, lo que nos lleva a repetir, sin prisa pero sin pausa,

$$x_n = H(x_{n-1}) := \frac{1}{2}\left(x_{n-1} + \frac{a}{x_{n-1}}\right), \quad n = 1, 2, 3, \ldots$$

cumpliéndose que $\lim_{n\to\infty} x_n = \sqrt{a}$ siempre que $x_0 > 0$ (o más en general, si $x_0 \in \mathbb{C}$ y tiene parte real positiva).

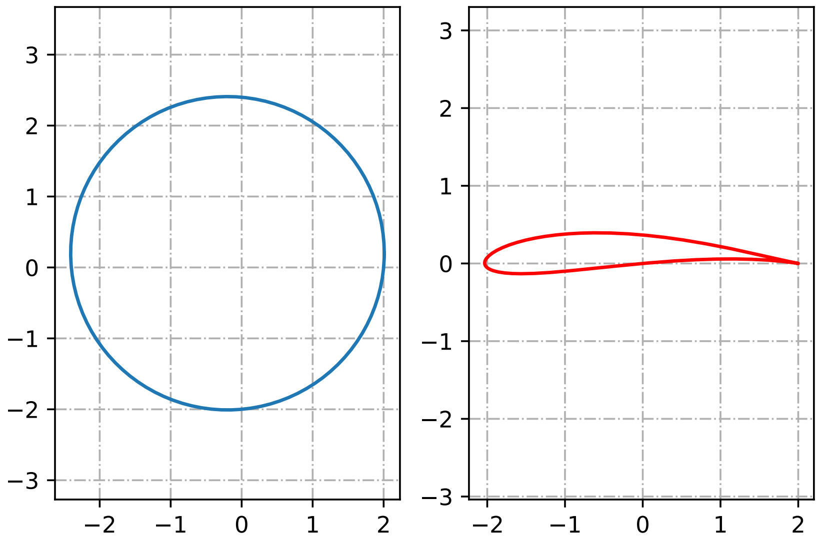

Figura 3. Transformación de Zhukovsky aplicada a la circunferencia de la izquierda.

Sorprendentemente, por aquello de la irrazonable efectividad de las matemáticas, la función H pensada como función de variable compleja

$$H\colon \mathbb{C} \setminus \{0\} \to \mathbb{C}, \quad z \to \frac{1}{2}\left(z + \frac{a}{z}\right),$$

se conoce como transformación de Zhukovsky (ver Figura 3) y encuentra aplicaciones prácticas en aeronáutica, ya que transforma una cierta circunferencia del plano complejo en un perfil similar al ala de un avión, lo que permite estudiar de forma sencilla el flujo de aire a su alrededor, (Penrose, 2005). Esto termina de confirmar un hecho que veníamos sospechando desde el comienzo: ¡la raíz cuadrada es un asunto de altos vuelos!

Bibliografía

Amster, P., y Cid, J. A. (2022). The full power of the half-power. *Expo. Math.*, *40*(4), 994–1013.

Benítez, F., y Carena, M. (2021). Reflexiones acerca de la definición de radicación y su relación con la construcción de nuevos conceptos. *RevEM UMA*, *36*(1), 9–26.

Cao Labora, D. (2020). When is a continuous bijection a homeomorphism? *Amer. Math. Monthly*, *127*(6), 547–553.

Conway, J. B. (1978). *Functions of one complex variable. Second edition*. New York–Berlin: Graduate Texts in Mathematics, 11, Springer–Verlag.

Ebbinghaus, H.-D., Hermes, H., Hirzebruch, F., Koecher, M., Mainzer, K., Neukirch, J., … Remmert, R. (1991). *Numbers*. New York: Graduate Texts in Mathematics, 123, Springer–Verlag.

Gasull, A. (2022). Los babilonios y Newton. *La Gaceta de la RSME*, *25*, 488.

Hasselblatt, B., y Katok, A. (2003). *A first course in dynamics with a panorama of recent developments*. Cambridge University Press.

Penrose, R. (2005). *The road to reality. A complete guide to the laws of the universe*. New York: Alfred A. Knopf, Inc.

Taussky, O. (1970). Sums of squares. *Amer. Math. Monthly*, *77*, 805–830.

Taussky, O. (1988). From Pythagoras' theorem via sums of squares to celestial mechanics. *Math. Intelligencer*, *10*(1), 52–55.

Pablo Amster
Departamento de Matemática, Facultad de Ciencias y Naturales
Universidad de Buenos Aires – IMAS-CONICET
(✉) *pamster@dm.uba.ar*

José Ángel Cid
CITMAga, 15782 Santiago de Compostela, España
Departamento de Matemáticas, Universidade de Vigo
(✉) *angelcid@uvigo.es*

Recibido: *24 de abril de 2023.*
Aceptado: *23 de julio de 2023.*
Publicado en línea: *3 de agosto de 2023.*

por Leandro Cagliero y Ricardo Podestá

si en un cuadrado, por cada vértice trazamos la media diagonal al punto medio del lado opuesto en sentido horario (u antihorario), el cuadrado que queda formado en el centro tiene $\frac{1}{5}$ del área del cuadrado original?

Por media diagonal nos referimos al segmento que une un vértice con el punto medio de un lado opuesto (en un cuadrado hay dos medias diagonales por vértice, que por analogía con el triángulo podemos llamarlas medianas).

¡Qué resultado más sorprendente, que de un cuadrado aparezca de repente el número 5! En la siguiente figura se puede ver la construcción:

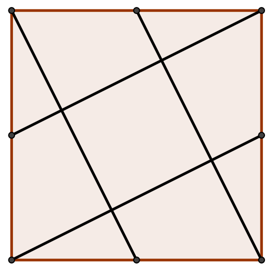

FIGURA 1.

Veamos que se puede deducir el resultado geométricamente de manera muy sencilla. Si rotamos los 4 triángulos rectángulos (mas chicos) por los vértices del cuadrado en sentido horario y apoyamos el punto medio de un lado sobre otro (en la Figura 2 vemos rotado el triángulo $\triangle RBY$), vemos que se forman 4 cuadrados congruentes con el cuadrado central (nos queda una cruz) y por lo tanto el cuadraditoo central tiene $\frac{1}{5}$ del área del cuadrado original (al igual que los otros cuatro).

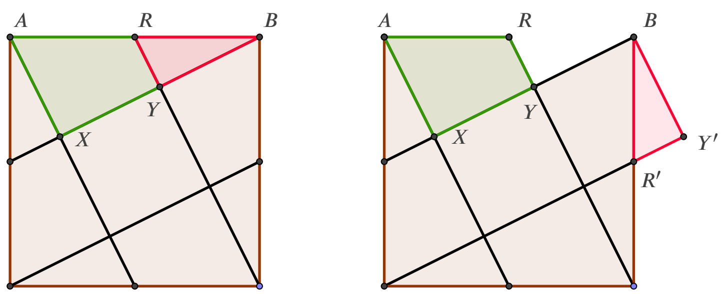

FIGURA 2.

Notar que la misma prueba funciona para paralelogramos (es decir, el paralelogramo interior formado por 2 pares de medianas paralelas que se cortan, tiene $\frac{1}{5}$ del área del paralelogramo inicial). El resultado vale en general para cuadriláteros convexos (trapezoides) y es debido a Rick Mabry (ver (Mabry, 2011), disponible electrónicamente en `https://lsusmath.rickmabry.org/rmabry/corners/CrosscutConvexQuadrilaterals.pdf`).

Si hacemos lo mismo con un hexágono regular, el hexágono interior formado por las medias diagonales tiene área $\frac{1}{13}$ y la estrella hexagonal tiene área $\frac{2}{13}$.

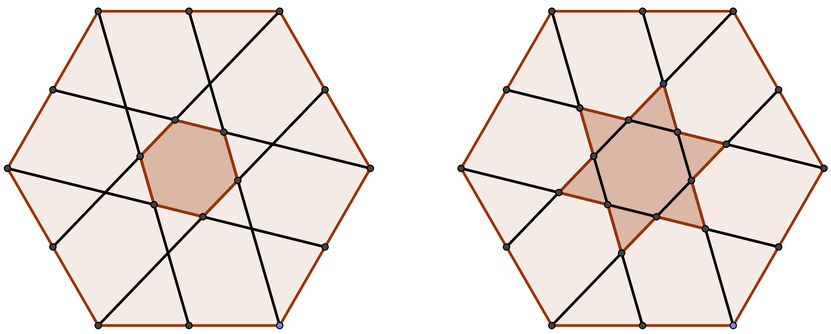

FIGURA 3.

Invitamos al lector a encontrar una demostración de este hecho, similar a la que realizamos con el cuadrado. A modo de ayuda diremos que se puede hacer apoyando 6 hexágonos congruentes al hexágono central H, tanto sobre los lados (ver (Mabry, 2018)) como sobre los vértices de H.

Como somos curiosos, podemos preguntarnos entonces ¿qué sucede con los demás polígonos regulares de un número par de lados? Para ello podemos hacer las cuentas usando coordenadas.

Llamemos P_{2k} al polígono regular de $2k$ lados inscripto en una circunferencia de centro el origen de coordenadas y radio igual a 1. Sea Q_{2k} el polígono determinado por las intersecciones de las medias diagonales de P_{2k}. Es claro que el centro de Q_{2k} será el mismo de P_{2k}, así que Q_{2k} estará inscripto en una circunferencia de centro el origen del sistema de coordenadas y cierto radio ρ. De este modo, la razón de semejanza entre P_{2k} y Q_{2k} será ρ y por lo tanto

$$\text{Área}(Q_{2k}) = \rho^2 \,\text{Área}(P_{2k}).$$

Nuestro objetivo es hallar ρ.

Pensemos que el primer y segundo vértice de P_{2k} son, respectivamente,

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{y} \qquad v_2 = \begin{pmatrix} \cos\left(\frac{\pi}{k}\right) \\ \sin\left(\frac{\pi}{k}\right) \end{pmatrix}.$$

La media diagonal correspondiente a cada uno de ellos resulta de unirlos con el punto medio del lado que está antes (en el sentido antihorario) del respectivo vértice opuesto. Es decir que debemos trazar los segmentos $\overline{v_1w_1}$ y $\overline{v_2w_2}$ donde

$$w_1 = \frac{1}{2}\begin{pmatrix}-1\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}\cos\left(\frac{(k-1)\pi}{k}\right)\\ \sin\left(\frac{(k-1)\pi}{k}\right)\end{pmatrix} \qquad \text{y} \qquad w_2 = \frac{1}{2}\begin{pmatrix}-1\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}\cos\left(\frac{(k+1)\pi}{k}\right)\\ \sin\left(\frac{(k+1)\pi}{k}\right)\end{pmatrix}.$$

Ahora debemos hallar el punto donde se cortan los segmentos $\overline{v_1w_1}$ y $\overline{v_2w_2}$. Ese punto resulta ser

$$u = \frac{1}{3\cos\left(\frac{\pi}{k}\right)+5}\begin{pmatrix}\cos\left(\frac{\pi}{k}\right)-1\\ 2\sin\left(\frac{\pi}{k}\right)\end{pmatrix}.$$

Dado que u es uno de los vértices de Q_{2k}, sabemos que $\rho = |u|$ y por lo tanto

$$\rho^2 = \frac{1-\cos\left(\frac{\pi}{k}\right)}{3\cos\left(\frac{\pi}{k}\right)+5}.$$

Podemos verificar que para $k = 2$ y $k = 3$, el valor de ρ^2 es $\frac{1}{5}$ y $\frac{1}{13}$, como ya habíamos anticipado.

¿Habrá algún otro k para el cual ρ^2 sea un número racional? La respuesta es, un poco sorpresivamente, que no. En efecto, despejando $\cos\left(\frac{\pi}{k}\right)$ en términos de ρ^2 obtenemos que uno es racional si y solo si el otro es racional y el Teorema de Niven (ver por ejemplo (Olmsted, 1945)) nos dice que $\cos\left(\frac{\pi}{k}\right)$ es un número racional solamente cuando $k = 2$ ó $k = 3$.

Dicho de otro modo, el teorema de Niven nos dice que no puede existir una prueba geométrica como las mencionadas para el cuadrado y el hexágono regular para otro polígono regular de un número par de lados.

Bibliografía

Mabry, R. (2011). Crosscut convex quadrilaterals. *Mathematics Magazine*, *84*(1), 16–25.

Mabry, R. (2018). Proof without words: One-thirteenth of a hexagon. *Mathematics Magazine*, *91*(3), 184–185.

Olmsted, J. M. H. (1945). Rational values of trigonometric functions. *Amer. Math. Monthly*, *52*(9), 507–508.

Richad Lee Mabry (16-11-1939 – 22-3-2006), conocido como 'Rick' Mabry, fue un profesor de matemáticas norteamericano, quien recibió su Master en Ciencias por la Universidad de Oregon en 1970, dio clases en escuelas de Portland por mas de 30 años. Estaba interesado en cuestiones de combinatoria geométrica, con muchos aportes a la matemática recreativa o divulgativa en estas áreas. Algunas de sus contribuciones pueden verse en las páginas https://lsusmath.rickmabry.org/rmabry/rdmrsrch.htm y https://lsusmath.rickmabry.org/rmabry/

Revista de Educación Matemática
Volumen 38, N° 2 (2023), páginas 31 – 47
Unión Matemática Argentina - Famaf (UNC)

Artículos de Matemática

SOBRE JUSTIFICACIONES EN MATEMÁTICA: CUANTIFICADORES, IMPLICACIONES Y EQUIVALENCIAS.

Marilina Carena

Resumen. Este trabajo tiene como fin analizar algunas justificaciones acerca de los valores de verdad de proposiciones matemáticas básicas que contienen cuantificadores o implicaciones. En particular, se hace énfasis en las formas incorrectas de hacerlo, puesto que muchas de ellas persisten aún en estudiantes avanzados de carreras de grado. Esto se abordará, principalmente, desde lo coloquial para ser aplicado luego en ejemplos matemáticos básicos.

Abstract. This paper aims to analyze some justifications concerning to the truth value of basic mathematical propositions containing quantifiers or implications. In particular, emphasis is placed on the wrong ways of doing it, since many of them still persist even in advanced undergraduate students. This will be addressed mainly from the colloquial to be applied later in basic mathematical examples.

Introducción

Si bien en la escuela secundaria no suelen abordarse, en general, contenidos de lógica formal, hay ciertas actividades que incluyen razonamientos lógicos básicos. Este es el caso, por ejemplo, de los ejercicios de tipo "Verdadero o Falso" y sus justificaciones.

Teniendo en cuenta la alta probabilidad que existe de acertar la respuesta en tales ejercicios, es natural que la importancia radique en la justificación de la misma. El modo correcto de hacerlo suele presentar dudas aun en estudiantes de carreras de grado, como si las reglas para ello fueran arbitrarias o contraintuitivas: *¿con un ejemplo alcanza?*; *¿puedo enunciar qué sería lo correcto?*; *¿puedo tomar valores particulares?*

Estas son algunas de las preguntas que escuchamos reiteradamente. Por ello, resulta importante poder transmitir que el razonamiento utilizado en la Matemática

Palabras clave: Educación Matemática, implicaciones, cuantificadores.
Keywords: Mathematics education, implications, quantifiers.

no es diferente al aplicado en la vida cotidiana. Esto significa que para poder determinar (y demostrar) si un enunciado matemático es verdadero o falso, debe razonarse de la misma forma en que se lo hace para decidir si algo de la vida cotidiana es cierto o no.

La mayor diferencia es que no toda afirmación de la vida cotidiana es verdadera o falsa, pues eso puede depender del gusto, estado de ánimo o la opinión de cada persona, entre otras cosas. En Matemática, en cambio, se trabaja con oraciones que son enunciados. Un **enunciado** o **proposición** es una oración que admite uno y solo un valor de verdad: verdadero o falso. Por ejemplo, las siguientes afirmaciones *sí* son enunciados:

- Si nació en el año 2005, entonces nació en el siglo XXI.
- Todas las personas tienen cabello color negro.
- Existen personas con cabello color negro.

La primera y la última de las afirmaciones anteriores son verdaderas, mientras que la segunda resulta falsa. Sin embargo, si debemos defender o justificar esta respuesta sobre la veracidad o falsedad de cada una, será necesario usar argumentos correctos, como trabajaremos en las siguientes secciones. Para lograr esto es necesario, a su vez, que las proposiciones se enuncien de forma completa y correcta, conteniendo los cuantificadores e implicaciones que deseamos plantear.

Por ejemplo, consideremos lo siguiente:

$$\sqrt{x+y} = \sqrt{x} + \sqrt{y}.$$

¿Qué se espera que un alumno responda sobre la verdad de esta afirmación? Puesto que no contiene ningún cuantificador, algunas interpretaciones posibles son:

- la igualdad enunciada vale para toda elección de números reales no negativos x e y;
- existen números reales no negativos x e y tales que la igualdad vale;
- la igualdad dada no vale para ninguna elección de números reales x e y.

La interpretación asumida "por defecto" suele ser la primera, pero es preciso enunciar lo que realmente queremos plantear incluyendo los cuantificadores correspondientes ya que, a diferentes cuantificadores, corresponden distintos valores de verdad. Esto puede lograrse de forma coloquial, sin necesidad de recurrir a ninguna simbología extra. Por ejemplo, la siguiente proposición es falsa:

$$\sqrt{x+y} = \sqrt{x}+\sqrt{y} \quad \text{para cualquier elección de reales números no negativos } x \text{ e } y;$$

mientras que resulta verdadero enunciar que:

$$\sqrt{x+y} = \sqrt{x} + \sqrt{y} \quad \text{para cierta elección de números reales no negativos } x \text{ e } y.$$

Similarmente, consideremos la afirmación

Un cuadrado es un cuadrilátero.

Si bien este enunciado no presenta problemas desde el punto de vista lógico puesto que puede determinarse su valor de verdad, quizás no resulte la forma adecuada de expresarlo si estamos describiendo, por ejemplo, las propiedades de un cuadrado. Más precisamente, si un alumno encuentra esta frase por primera vez, podría preguntarse si se refiere a que:

- todo cuadrado es un cuadrilátero;
- todo cuadrilátero es un cuadrado;
- cuadrilátero y cuadrado son conceptos equivalentes.

Es claro que la única opción correcta en este caso es la primera, que también puede reescribirse utilizando los conectores “si” y “entonces”como:

Si es un cuadrado, entonces es un cuadrilátero.

Por la familiaridad que la mayoría de nosotros tenemos con el concepto de cuadrado, quizás el ejemplo previo no ilustre en profundidad la conveniencia de la redacción adecuada. Por ejemplo, vayamos a algo menos básico sobre cuadriláteros, como la afirmación:

Los ángulos interiores opuestos de un cuadrilátero inscripto en una circunferencia son suplementarios.

¿Qué quiere transmitir la oración anterior? Las opciones, nuevamente, pueden ser:

- si los ángulos interiores opuestos de un cuadrilátero son suplementarios, entonces el cuadrilátero puede inscribirse en una circunferencia;
- si un cuadrilátero puede inscribirse en una circunferencia, entonces sus ángulos interiores opuestos son suplementarios;
- un cuadrilátero puede inscribirse en una circunferencia si y solo si sus ángulos interiores opuestos son suplementarios.

Determinar cuál de las interpretaciones anteriores es la correcta puede no ser tan obvio como en el caso del cuadrado, principalmente para aquellos que se encuentren con la propiedad por primera vez.

Una vez dado un enunciado matemático de modo que admita un único valor de verdad (verdadero o falso) es importante poder justificar este valor de forma correcta. En las siguientes secciones trabajaremos ejemplos sobre esto, teniendo en cuenta los cuantificadores, las implicaciones y los conectores de conjunción y disyunción. Más precisamente, este trabajo se organiza de la siguiente forma:

- Sección 1: se trabaja sobre la importancia de los *cuantificadores* involucrados en un enunciado al momento de determinar y justificar su valor de verdad.

- Sección 2: se aborda la diferencia entre *implicaciones* y *equivalencias*, así como los conceptos de recíproco, contrarrecíproco, condiciones necesarias y condiciones suficientes.
- Sección 3: se presentan los conectores de disyunción y de conjunción entre proposiciones, y se analiza el significado de implicaciones verdaderas y falsas que los contienen.

Las tres secciones anteriores se abordan, principalmente, con ejemplos que no involucran matemática. Luego, en la Sección 4, se aplica todo lo trabajado previamente para determinar y justificar el valor de verdad de distintos enunciados matemáticos básicos.

Algunos fragmentos y ejemplos contenidos en este artículo fueron extraídos de (Carena, 2021).

§1. La importancia de los cuantificadores

Para justificar la veracidad o falsedad de un enunciado se debe dar una **demostración**. Las demostraciones tienen diversas formas, dependiendo del enunciado que se quiera probar: la falsedad de algunos enunciados puede demostrarse dando un **contraejemplo**, es decir, un ejemplo para el cual el enunciado no se cumpla, mientras que en ciertos casos bastará con un **ejemplo** para demostrar que el enunciado es verdadero. Nos encontraremos también con enunciados cuyo valor de verdad no podrá demostrarse ni con ejemplos ni con contraejemplos.

Para ilustrar esto, supongamos que queremos demostrar que la afirmación

Todas las personas tienen cabello color negro

es **falsa**. Para ello será suficiente con indicar una sola persona que no tenga el cabello color negro. Notar que ser falsa no significa que *ninguna* persona tiene cabello negro, sino que existe al menos una que no lo tiene. Parece obvia esta aclaración, pero a veces suele olvidarse esta lógica cuando se intenta argumentar sobre un enunciado matemático. Aquí, la persona indicada es nuestro “contraejemplo”.

Sin embargo, para probar que

Existen personas con cabello color negro

es **verdadera**, será suficiente con indicar una persona con cabello negro (será nuestro “ejemplo”). Notar que esta misma persona serviría para demostrar que la siguiente afirmación es **falsa**:

No existen personas con cabello color negro.

Con lo anterior podemos observar que **no** será posible encontrar una receta de cómo demostrar la veracidad o falsedad de un enunciado, sino que dependerá de la forma en la que el mismo esté expresado. Será necesario razonar de manera lógica en cada caso, para determinar si se necesita dar una demostración mediante propiedades, un contraejemplo o un ejemplo, independientemente de que la afirmación sea verdadera o falsa.

Sin embargo, es importante señalar algunas de las formas **incorrectas** que aparecen frecuentemente al momento de intentar probar la validez de algunas afirmaciones. Por ejemplo, si alguien afirma que

Todas las personas nacidas en Argentina tienen cabello color negro,

no alcanzará con exponer ni una, ni dos, ni 500 personas nacidas en Argentina con cabello de color negro, pues esas no son *todas* las personas sobre las cuales se está realizando la afirmación. Exhibir varios ejemplos en un conjunto de objetos satisfaciendo una propiedad, aunque sean muchos, no alcanza para demostrar que dicha propiedad vale para todo el conjunto. De hecho, aunque se exhiban miles de personas argentinas con cabello negro, bastará con encontrar una sola persona que no lo tenga para demostrar que es falsa, tal como se vio en el primer ejemplo de esta sección.

En otras palabras, para demostrar que cierta propiedad vale para **todo** elemento en un cierto conjunto, no basta con chequear que valga para algunos, sino que habrá que verificar que vale para **todos los casos posibles**. Esto puede ser tedioso si estamos hablando de una cantidad elevada de elementos, y requerirá de métodos específicos si se trata de una cantidad infinita de ellos. Esta situación es muy frecuente en las afirmaciones matemáticas, en las que las propiedades se enuncian, por ejemplo, para todos los números naturales o reales. En tal caso, cualquier cantidad de ejemplos que se presenten para demostrar la validez de la propiedad no será suficiente, pues hay infinitos posibles. Sin embargo, un solo ejemplo resultará suficiente para demostrar su falsedad. Para ilustrar esto consideremos los dos enunciados matemáticos siguientes, cuyos valores de verdad se indican:

(i) $(a+b)^2 = a^2 + b^2$, para todo par de números reales a y b. Falso

(ii) $(a+b)^2 = a^2 + 2ab + b^2$, para todo par de números reales a y b. Verdadero

¿Cómo podemos demostrar que (i) es falsa? En este caso, usando la misma lógica que en un debate sobre enunciados no matemáticos: bastará con exponer valores para los cuales la afirmación no se cumple, es decir, un par de números reales a y b tales que $(a+b)^2 \neq a^2 + b^2$. Por ejemplo, podemos tomar $a = 1$ y $b = 2$ para obtener, por un lado, $(a+b)^2 = 3^2 = 9$ y, por el otro, $a^2 + b^2 = 1^2 + 2^2 = 5$. Es claro que $9 \neq 5$, por lo que el par elegido es nuestro contraejemplo.

Cabe señalar que un error frecuente para justificar que un enunciado es falso consiste en exhibir cuál debería ser el resultado correcto. Más precisamente, es frecuente leer que la afirmación (i) es falsa puesto que $(a+b)^2 = a^2 + 2ab + b^2$. ¿Por qué es incorrecto justificar así? Porque seguimos sin demostrar que lo dado es falso salvo que se demuestre, además, que $a^2 + 2ab + b^2 \neq a^2 + b^2$, tomando valores particulares para a y b como arriba, por ejemplo. Es importante comprender esto ya que no siempre expresiones en apariencia distintas representan cantidades diferentes. Para ilustrar esto, podemos considerar la identidad

$$a + \frac{b-a}{2} = \frac{a+b}{2},$$

que propone una manera alternativa de calcular el promedio entre cualesquiera dos números reales a, b. Otro ejemplo es la llamada "identidad pitagórica", que establece que

$$\cos^2(\alpha) + \operatorname{sen}^2(\alpha) = 1,$$

para cualquier número real α.

Así, enunciar y demostrar una verdad relacionada a un enunciado dado *no* es suficiente para probar la falsedad del mismo, ya que podrían ser ambas cosas verdaderas. Si se opta por este camino se debe, además, mostrar que esta verdad contradice a la proposición enunciada. Sin ello, ambas afirmaciones podrían ser verdaderas. Se volverá sobre esto más adelante en casos concretos como, por ejemplo, en el Enunciado 7.

Consideremos nuevamente los valores $a = 1$ y $b = 2$ elegidos para probar que (i) es falsa, y comprobemos que la identidad (ii) sí se cumple para ellos. Ya vimos que $(a+b)^2 = 9$. Por otra parte, para el lado derecho de la igualdad tenemos que

$$a^2 + 2ab + b^2 = 1^2 + 2 \cdot 1 \cdot 2 + 2^2 = 1 + 4 + 4 = 9.$$

Así, vemos que para este par de valores a y b se obtiene el mismo resultado en ambos lados de la igualdad. Podemos probar con otros y ver que también se cumple. Sin embargo, como mencionamos antes, esto no alcanza para probar que (ii) es verdadera ya que, por más que analicemos muchos ejemplos, los posibles valores para a y b son infinitos. Esto solo serviría, en caso de tener que determinar el valor de verdad de la afirmación, como una evidencia empírica de que podría ser verdadera. Esto nos conduce a intentar demostrar que la identidad (ii) vale para valores de a y b generales mediante herramientas adecuadas. Para ello usaremos que z^2 es un símbolo que indica el producto del número z por sí mismo, es decir, $z^2 = z \cdot z$ para todo número real z. Luego, por la propiedad distributiva de la multiplicación respecto de la suma y la propiedad conmutativa del producto, obtenemos

$$(a+b)^2 = (a+b) \cdot (a+b) = a \cdot a + a \cdot b + b \cdot a + b \cdot b = a^2 + 2 \cdot a \cdot b + b^2,$$

para cualesquiera a y b reales, lo que prueba que el enunciado (ii) es verdadero.

Como conclusión, debemos prestar especial atención a los cuantificadores involucrados en un enunciado, ya que los mismos determinan tanto su valor de verdad como las posibles formas de justificarlo. En especial, resulta importante comprender que no alcanza con probar con algunos valores particulares para demostrar la verdad de un enunciado que contiene un cuantificador universal como (ii).

§2. Implicaciones versus equivalencias

Supongamos que se desean encontrar los valores reales de x que satisfacen la siguiente igualdad

$$x + 4 = \sqrt{x + 10}. \tag{1}$$

¿Qué procedimiento y escritura resultan ser los más frecuentes? Lo más probable es que encontremos lo siguiente:

$$\begin{aligned} x + 4 &= \sqrt{x + 10} \\ (x + 4)^2 &= x + 10 \\ x^2 + 8x + 16 &= x + 10 \\ x^2 + 7x + 6 &= 0, \end{aligned}$$

y, luego de aplicar la fórmula de la resolvente, se obtiene $x = -6$ y $x = -1$. Se ve directamente que ninguno de estos valores genera un radicando negativo en la ecuación original. Sin embargo, es fácil chequear que $x = -6$ genera un valor negativo en el miembro izquierdo de (1), y un valor positivo en el miembro derecho (porque $\sqrt{4} = 2$ y $\sqrt{4} \neq -2$, ver (Benitez y Carena, 2021) o la Figura 1 para convencerse), por lo que no es solución de dicha ecuación. Así, solamente $x = -1$ es solución de la ecuación dada.

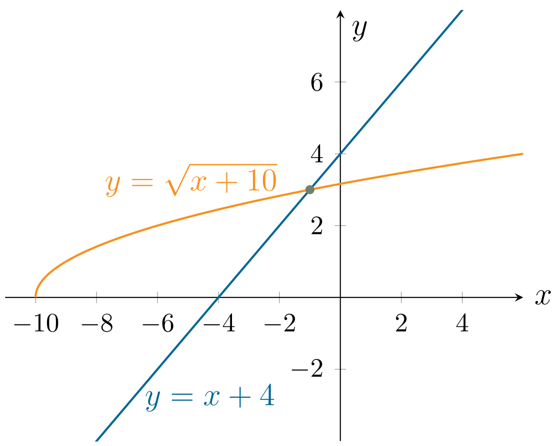

FIGURA 1. Representación gráfica de la ecuación $x + 4 = \sqrt{x + 10}$.

Entonces aquí es donde surgen dos preguntas fundamentales que generan confusión en los alumnos:

- ¿Por qué aparecen soluciones ficticias, si la resolución no posee errores?
- ¿Cuándo una verificación deja de ser opcional para convertirse en parte de la resolución?

La respuesta a ambas preguntas está contenida en el uso de **equivalencias** o **implicaciones**. Remarcar la diferencia entre estos dos conceptos resulta fundamental para un aprendizaje matemático sólido y consistente. Abordaremos esto a continuación.

Aunque las expresiones "si", "entonces" y "solo si" aparecen frecuentemente en frases cotidianas, su significado no siempre es interpretado correctamente. Por ejemplo, supongamos que una persona afirma lo siguiente:

Si *gano la lotería,* ***entonces*** *me compro un auto.*

Aquí, la palabra "si" funciona como introductor de la premisa "ganar la lotería", mientras que "entonces" es un introductor para la conclusión "comprar un auto". Suponiendo que la persona que lo afirma cumple con lo que dice, esto significa simplemente que la validez del primer enunciado implica la validez del segundo. Solo eso. Sin embargo, las siguientes interpretaciones erróneas sobre la afirmación anterior suelen escucharse en el lenguaje cotidiano:

- Si *no* gana la lotería, entonces *no* se compra el auto. ✗
- Si se compró el auto, entonces ganó la lotería. ✗

La primera interpretación es incorrecta, ya que la persona no dijo qué sucederá si no gana la lotería. Solamente afirmó algo que haría si la ganara. Con respecto a la segunda, la persona nunca dijo que ganar la lotería era la única forma de comprarse el auto. Quizás consiguió prestado el dinero para comprarlo, vendió algo para conseguirlo, o cualquier otra posibilidad.

Si A, entonces B. Cuando un enunciado de este tipo es verdadero, significa que si el enunciado A es verdad, entonces también lo es B. No establece nada para B cuando A es falso. Tampoco significa que si B es verdadero, entonces A también. Se denota como $A \Rightarrow B$, y se lee también como "A implica B".

En las dos interpretaciones incorrectas de arriba, la segunda corresponde al error de considerar que vale el **recíproco** de la afirmación (B implica A), mientras que en la primera la falacia surge de pensar que *negar el antecedente implica la negación de la consecuencia* (no ocurre A, entonces no ocurrirá B).

Lo que sí se puede deducir, suponiendo que la persona cumple con lo que afirma como si se tratara de un contrato, es que:

- Si *no* se compró el auto, entonces *no* se ganó la lotería, ✓

ya que afirmó que si ganaba, lo compraba. Esto se conoce como el **contrarrecíproco** de la afirmación hecha, y es posible establecer que $A \Rightarrow B$ es equivalente a

$$\text{negación de } B \Rightarrow \text{negación de } A.$$

Volviendo a la lista anterior de interpretaciones incorrectas, la segunda habría sido correcta si la persona hubiera afirmado que:

Me compro un auto ***solo si*** *gano la lotería.*

De nuevo, la expresión anterior *no* afirma que si gana la lotería entonces se compra un auto. Puede ocurrir que utilice todo el dinero para comprar otra cosa. Solo dice que ganar la lotería sería la única forma de comprarse un auto.

Esto nos conduce a pensar en un conectivo entre dos expresiones, que nos permita concluir que si cualquiera de ellas es cierta, la otra también. Para esto se combinan las frases anteriores para formar lo que se conoce como "si y solo si":

Me compro un auto ***si y solo si*** *gano la lotería.*

De la veracidad de esta frase se deduce que:

- Si se compró el auto, entonces ganó la lotería. ✔
- Si ganó la lotería, entonces se comprará el auto. ✔
- Si *no* gana la lotería, entonces *no* se comprará el auto. ✔
- Si *no* se compró el auto, entonces *no* ganó la lotería. ✔

***A* si y solo si *B*.** Este conectivo se utiliza para relacionar dos enunciados, y significa que la validez de cualquiera de ellos implica la validez del otro. Esto se denota simbólicamente como $A \Leftrightarrow B$, ya que significa ambas cosas a la vez: $A \Rightarrow B$ y $B \Rightarrow A$.

Para considerar otro ejemplo, cercano y real, recordemos un enunciado en relación al COVID-19, que todos conocimos durante la pandemia. Sabíamos que:

Si una persona tiene más de 60 años, entonces es considerada de riesgo.

En la afirmación anterior tenemos:

- A: Tener más de 60 años.
- B: Ser persona de riesgo.

Sabemos que A implica B. ¿Se puede deducir de esto alguna de las siguientes afirmaciones?

(a) Si es persona de riesgo, entonces tiene más de 60 años.
(b) Si no tiene más de 60 años, entonces no es persona de riesgo.
(c) Si no es persona de riesgo, entonces no tiene más de 60 años.

Claramente (a) y (b) son conclusiones incorrectas: una persona puede ser de riesgo por tener alguna enfermedad aunque sea muy joven. En cambio, la afirmación (c) sí es correcta (ya que, si una persona tuviera más de 60 años, entonces sería de riesgo y como no lo es, no tiene más de 60 años), y corresponde al contrarrecíproco del enunciado dado.

Esto, que se ve tan claro con este ejemplo concreto, suele traer problemas cuando se trabaja con proposiciones matemáticas, en donde la situación es más abstracta. Por eso es conveniente tener siempre al alcance ejemplos cotidianos que permitan al alumno trasladar el mismo razonamiento a enunciados matemáticos. No solo en Matemática es importante utilizar las implicaciones de forma correcta para no llegar a conclusiones erróneas, también lo es en lo cotidiano. Es muy frecuente encontrar falacias lógicas en discursos: se utilizan argumentos que parecen válidos para atacar o invalidar otra opinión, pero no lo son. Esto puede ser intencional o, simplemente, por descuido o ignorancia.

En el enunciado sobre el COVID-19, tener más de 60 años era una condición *suficiente* para ser considerada personada de riesgo, pero no *necesaria*. Trabajaremos sobre esto a continuación.

Si el enunciado $A \Rightarrow B$ es verdadero, decimos que la condición A es **suficiente** para que se cumpla la condición B. Por ejemplo, podemos reescribir la frase referida a un polígono:

Si es un cuadrado, entonces es un cuadrilátero,

como

Ser un cuadrado $\Rightarrow$ *ser un cuadrilátero.*

Así, el hecho de que un polígono sea un cuadrado es condición suficiente para que sea un cuadrilátero. Claramente el recíproco no es cierto: existen cuadriláteros que no son cuadrados. Por eso, tampoco podemos concluir nada de cuando no vale A: que un polígono no sea un cuadrado no implica que no sea un cuadrilátero.

Cuando el enunciado $A \Rightarrow B$ es verdadero también se dice que B es condición **necesaria** para que se cumpla la condición A. Así, ser un cuadrilátero es una condición necesaria para ser un cuadrado (y no al revés). Luego, si un polígono *no* es un cuadrilátero, entonces *no* es un cuadrado.

Cuando $A \Leftrightarrow B$ vale, se dice que A y B son **equivalentes**. Así, la validez de la condición A es necesaria y suficiente para la validez de B.

Por ejemplo, puede probarse que las siguientes afirmaciones sobre un número entero son verdaderas:

- que su última cifra sea 4 es condición suficiente para ser un número par (pero no necesaria);

- que sea par es condición necesaria para ser divisible por 6 (pero no suficiente);
- que su última cifra sea 0 es condición necesaria y suficiente para ser divisible por 10.

Retomemos el ejemplo del inicio de la sección para responder, a partir de lo trabajado previamente, las preguntas allí planteadas sobre las soluciones ficticias de la ecuación (1). Cuando se presentan ecuaciones con raíces cuadradas, la forma usual de resolverlas es elevar ambos miembros al cuadrado. Sin embargo, la ecuación que se obtiene al hacer esto *no* es equivalente a la original. Más precisamente, por la propiedad uniforme del producto para números reales se tiene que $x = y$ implica que $x \cdot c = y \cdot c$ para todo número real c. En particular, tomando $c = x$ se obtiene que $x^2 = y \cdot x = y^2$, puesto que partimos del supuesto que $x = y$. Así, probamos que

$$x = y \quad \Rightarrow \quad x^2 = y^2.$$

Es decir, $x = y$ *implica* $x^2 = y^2$, pero el recíproco no es cierto. Esto significa que $x^2 = y^2$ no implica necesariamente que $x = y$. Para probarlo, basta con tomar $x = 1$ e $y = -1$, por ejemplo. Así, las ecuaciones

$$x = y \qquad \text{y} \qquad x^2 = y^2$$

no son equivalentes. Entonces, si en la resolución de la ecuación colocamos el símbolo $\Rightarrow$ para indicar implicación y usamos $\Leftrightarrow$ para equivalencia, tenemos que:

$$\begin{gathered} x + 4 = \sqrt{x + 10} \\ \Downarrow \\ (x + 4)^2 = x + 10 \\ \Updownarrow \\ x^2 + 8x + 16 = x + 10 \\ \Updownarrow \\ x^2 + 7x + 6 = 0. \end{gathered}$$

Así, queda claro que si la ecuación original tiene una solución, esta será solución de la segunda. Sin embargo, no toda solución de la segunda lo será de la primera. En otras palabras, *ser solución de la ecuación cuadrática es condición necesaria para ser solución de la original, pero no es suficiente.*

Esto significa que las soluciones de la ecuación original deberán buscarse entre las de la segunda, siendo esta equivalente a la última ecuación. Es por este motivo que aparecen soluciones ficticias, aun cuando cada paso sea correcto. Por lo tanto, la verificación deja de ser un paso opcional para convertirse en parte de la resolución ya que, de todos los valores hallados, se deben determinar cuáles son soluciones de la ecuación dada y cuáles no.

§3. Disyunción y conjunción

Nos abocaremos ahora a la disyunción "o" y a la conjunción "y". Dejaremos de lado aquí lo que en lógica se conoce como *disyunción exclusiva*, que coloquialmente también se expresa mediante el uso de "o", donde una afirmación es verdadera si vale una y solo una de las opciones. Por ejemplo, si uno gana un sorteo y nos dicen que podemos elegir como premio un automóvil o un viaje, queda claro que no podemos obtener ambas cosas. En un sentido similar se trabaja con proposiciones matemáticas, pero no lo analizaremos aquí.

En este trabajo consideraremos solamente la *disyunción inclusiva* "o", que es un conector entre dos o más proposiciones que hace que el enunciado sea verdadero si al menos una las proposiciones lo es. Para comprenderlo, supongamos que una empresa busca cubrir un puesto de trabajo y el requisito para ello es:

Saber escribir en inglés o en francés.

¿Cuándo una persona cumple con el requisito? Claramente, si sabe escribir en inglés, o en francés, o en ambos idiomas. Si no cumple con ninguna de esas dos cosas, entonces no cumple con el requisito de la empresa. En cambio, supongamos que para el puesto se pide:

Saber escribir en inglés y en francés.

En este caso las condiciones se conectan mediante la *conjunción* "y", por lo que solo cumplirán con el requisito aquellas personan que sepan escribir en ambos idiomas. Quien no sepa hacerlo en al menos uno de ellos, no lo satisface.

Agreguemos ahora una implicación, como en el siguiente enunciado:

Si se obtienen al menos 60 puntos en el examen y se asiste al 80 % de las clases, entonces se aprueba la materia. (†)

En la frase anterior aparece la conjunción "y" entre las dos condiciones suficientes para aprobar la materia. Es decir, si se cumplen *ambas* condiciones, se aprueba la materia. Entonces, ¿qué podemos concluir acerca de un estudiante que *no* aprobó la materia? ¿Que falló necesariamente en ambas cosas? Claramente no, pero sí que falló en al menos una: si un estudiante no aprobó la materia, entonces o bien no obtuvo el mínimo de 60 puntos en el examen, o bien no asistió lo suficiente a clases (o quizás ninguna de las dos cosas). Aparece aquí la disyunción inclusiva "o" entre ambas condiciones.

Consideremos, en cambio, la afirmación

Si se obtienen al menos 60 puntos en el examen o si se presenta una monografía, entonces se aprueba la materia.

Ahora existe una disyunción entre ambas condiciones: con que cualquiera de las dos se cumpla (o ambas, si encontramos a alguien muy aplicado), es suficiente para aprobar la materia. Entonces, si alguien no aprueba la materia, podemos concluir que no hizo ninguna de las dos cosas: no aprobó el examen y no presentó una monografía. En este caso aparece la conjunción "y" entre ambas condiciones.

§4. Enunciados matemáticos

Trabajaremos con algunos enunciados matemáticos variados y básicos, determinando y demostrando su valor de verdad con el objetivo de aplicar y afianzar todo lo trabajado previamente.

Enunciado 1. *No siempre el producto de dos números irracionales es otro número irracional.*

Respuesta: El enunciado puede reescribirse como "existen números irracionales z y w tales que $z \cdot w$ no es un número irracional". Para probar que dichos números existen, podemos considerar $z = \sqrt{2}$ y también $w = \sqrt{2}$. Entonces

$$z \cdot w = \sqrt{2} \cdot \sqrt{2} = (\sqrt{2})^2 = 2,$$

que no es un número irracional. Esto prueba que el enunciado es **verdadero**. Existen muchos otros ejemplos, como $z = \sqrt{3}$ y $w = \sqrt{12}$, o $z = \sqrt{2}$ y $w = \sqrt{8}$. #

Enunciado 2. *La suma de dos números pares es siempre otro número par.*

Respuesta: Notar que la expresión "es siempre" es una de las formas usadas para referirse al cuantificador universal "para todo". Así, lo que afirma el enunciado es que $m + n$ es un número par, *para toda* elección de números pares m y n. Consideremos algunos ejemplos para comprender: $4 + 6 = 10$; $22 + 6 = 28$; $12 - 24 = -12$. Como mencionamos, *no* alcanza con la verificación de algunos casos para concluir que una cierta afirmación sea cierta. Simplemente nos conduce a conjeturar que *puede llegar a ser verdadera*, pero deberemos demostrarla de forma general. Para ello, sean m y n dos números pares. Luego, existen números enteros p y q tales que $m = 2p$ y $n = 2q$. Entonces:

$$m + n = 2p + 2q = 2(p + q).$$

Por la ley de cierre de la suma en los números enteros, $p + q$ es un número entero y, por lo tanto, $m + n$ es un número par. Así, la afirmación es **verdadera**. #

Enunciado 3. $x \leq x^2$ *para todo número real* x.

Respuesta: Como antes, tomemos algunos valores para ver qué ocurre: $2 \leq 2^2$; $3 \leq 3^2$; $-\frac{1}{3} \leq (-\frac{1}{3})^2$. Sin embargo, esto ya no ocurre si tomamos números reales entre 0 y 1. Por ejemplo, si $x = \frac{1}{2}$ entonces $x^2 = \frac{1}{4}$, que es menor que x. Esto demuestra que la afirmación es **falsa**. La primera parte ilustra que a veces una "mala" elección de ejemplos puede conducirnos a un valor de verdad incorrecto, y que no es suficiente con exhibir miles (ni infinitos) ejemplos para probarlo. #

Enunciado 4. $x \leq x^2$ *para todo número real* $x \geq 1$.

Respuesta: A diferencia del enunciado anterior, ahora se agrega la restricción $x \geq 1$. Los ejemplos que tomamos antes muestran que este enunciado *puede ser verdadero*. De hecho lo es, y vamos a demostrarlo de forma general: tomemos un número real $x \geq 1$. En particular, x es positivo, por lo que podemos multiplicar por x esta desigualdad a ambos lados y el orden se mantiene. Así:

$$x \cdot x \geq 1 \cdot x,$$

es decir, $x^2 \geq x$ como afirma el enunciado, que resulta ahora **verdadero**. #

El argumento que usaremos para justificar la respuesta del siguiente enunciado utiliza lo trabajado sobre las implicaciones, y lo que podemos deducir o no de ellas.

Enunciado 5. *Existe un triángulo rectángulo cuyos lados miden* 5, 4 *y* 8 *unidades de longitud.*

Respuesta: Notar que el enunciado no se refiere a la existencia de cualquier triángulo con esas longitudes para sus lados, sino que especifica además que debe ser *rectángulo*. Si dicho triángulo existiera, por ser rectángulo sabemos que cumple con lo establecido en el teorema de Pitágoras. Es decir, recordando que la hipotenusa es el lado de mayor longitud, debería ocurrir que

$$8^2 = 5^2 + 4^2.$$

Pero el lado izquierdo de esta igualdad es 64, mientras que el derecho es 41. Así, puesto que no vale lo que afirma el teorema de Pitágoras, dicho triángulo no puede existir y el enunciado resulta **falso**. #

Al igual que el Enunciado 2, la siguiente afirmación hace referencia a la cerradura de un cierto conjunto numérico bajo una operación determinada.

Enunciado 6. *Si* m *y* n *son números impares, entonces el producto* mn *también lo es.*

Respuesta: Este enunciado puede reescribirse usando cuantificadores como: mn es un número impar *para toda* elección de m y n impares. Como antes, probando con varios valores podemos intuir que es verdadero, así que intentemos demostrarlo. Sean m y n dos números impares cualesquiera. Así, existen números enteros p y q tales que $m = 2p + 1$ y $n = 2q + 1$. Luego

$$mn = (2p + 1)(2q + 1) = 4pq + 2p + 2q + 1 = 2(2pq + p + q) + 1.$$

Introducimos ahora una abreviatura k para la expresión $2pq + p + q$, que resulta ser un número entero por la ley de cierre del producto y de la suma para números enteros. Esto implica que $mn = 2k + 1$, es decir, mn es impar y el enunciado es **verdadero**. #

Consideremos, finalmente, algunos enunciados que contienen disyunción o conjunción.

Enunciado 7. *Si un número entero mayor o igual que* 6 *es divisible por* 2 *o por* 3*, entonces es divisible por* 6*.*

Respuesta: La afirmación es **falsa**. Para demostrar que lo es, basta con exhibir un número entero mayor o igual que 6 que sea divisible por 2 o por 3, pero no por 6. Por ejemplo, el número 15 es divisible por 3 pero no por 6; o el número 14 que es divisible por 2 pero no por 6. Sabemos que el enunciado se convierte en verdadero si cambiamos "o" por "y" pero, como ya mencionamos antes, enunciar el criterio correctamente no alcanza para probar que el dado es falso. #

Enunciado 8. *Si la última cifra de un número entero es* 0 *o* 5*, entonces es divisible por* 5*.*

Respuesta: La afirmación es **verdadera**. Esto es lo que enuncia precisamente el criterio de divisibilidad por 5. Sin embargo, al dar un enunciado así en la escuela secundaria, deberíamos pensar qué justificación se espera por parte de un alumno ya que, en general, no domina una demostración de este hecho en ese nivel , y solo podría limitarse a decir que "así es el criterio". #

Enunciado 9. *Sean* m *y* n *dos números enteros. Si el producto* mn *es un número par, entonces* m *o* n *es un número par.*

Respuesta: En el Enunciado 6 se establece y demuestra que

Si m *y* n *son números impares, entonces el producto* mn *también lo es.*

Imitando el razonamiento realizado debajo de (†), tenemos que si el producto mn *no* es impar (lo que significa que es par), entonces debió haber fallado alguna de las dos condiciones sobre m y n. Es decir, al menos uno de los dos debe ser un número par. Esto prueba que el Enunciado 9 es **verdadero**. #

Finalizamos este trabajo con un enunciado sobre un criterio de divisibilidad que puede ser factible de ser demostrado en la escuela secundaria. Aunque los criterios suelen enunciarse dando condiciones suficientes que garanticen la divisibilidad como en el enunciado anterior, los mismas son, además, necesarias. Esto se verá en el siguiente ejemplo.

Enunciado 10. *Un número es divisible por* 6 *si y solo si es divisible por* 2 *y por* 3*.*

Respuesta: En este caso el enunciado contiene un “si y solo si”, por lo que está diciendo que:

- si un número es divisible por 6, entonces es divisible por 2 y por 3;
- si un número es divisible por 2 y por 3, entonces es divisible por 6.

Para que el enunciado sea verdadero deben cumplirse *ambas* de estas implicaciones. Si al menos una falla, sería falso. Es sencillo ver que la primera se cumple: si un número m es divisible por 6, entonces existe un número entero k tal que $m = 6k$. Luego,

$$m = 6k = (3 \cdot 2)k = 3(2k) = 2(3k),$$

lo que implica que m es divisible tanto por 2 como por 3, ya que $2k$ y $3k$ son números enteros por la ley de cierre para el producto de números enteros.

Para probar la segunda implicación, supongamos que m es divisible tanto por 2 como por 3. Entonces existen números enteros p y q tales que $m = 2p$ y $m = 3q$. De esto se deduce que

$$2p = 3q.$$

Esta igualdad nos dice que $3q$ es un número par que es producto de dos números enteros. El Enunciado 9 establece que, entonces, al menos uno de los dos factores debe ser par. Puesto que 3 no puede serlo, tenemos que q lo es. Es decir, $q = 2t$, para algún número entero t. Así, tenemos que

$$m = 3q = 3(2t) = 6t,$$

lo que prueba que m es divisible por 6. Entonces el enunciado es **verdadero**. #

Conclusiones finales

Como se mencionó en la introducción, la lógica formal y las demostraciones matemáticas no se abordan, en general, en la escuela secundaria. Sin embargo, el desarrollo del sentido de argumentación puede comenzar a desarrollarse de forma coloquial y sin que esto implique pérdida de rigurosidad. Trabajar en enunciados cotidianos que involucren cuantificadores, implicaciones, disyunciones y conjunciones permiten al alumno adquirir y naturalizar la forma de razonar y argumentar lógica y correctamente. Identificar y debatir las formas incorrectas de hacerlo también puede resultar una aporte significativo en esa dirección. Luego, el paso a enunciados matemáticos básicos podría hacerse de manera natural, y constituir así una base sólida para que, quienes lo necesiten posteriormente, puedan construir técnicas para comprender y comunicar demostraciones de mayor nivel y complejidad. El siguiente fragmento, extraído de (Solow, 1993), describe esta idea de una forma muy certera y significativa:

> Para jugar al ajedrez, usted debe aprender primero cómo se mueve cada una de las piezas. Solamente después de que estas reglas han sido asimiladas por su subconsciente, usted podrá concentrar toda su atención en aspectos creativos como las estrategias, tácticas, etc. De igual forma sucede en matemáticas.

Eduardo Sáenz de Cabezón dijo "Las matemáticas nos hacen más libres y menos manipulables". Considero que el desarrollo del sentido lógico que permite analizar, razonar y argumentar de forma correcta sobre la veracidad o falsedad de un enunciado es un claro reflejo de esta afirmación.

Agradecimientos. Se agradece a quienes realizaron la tarea de ser revisores de este artículo por su minuciosa lectura y sus detalladas sugerencias y correcciones que, sin dudas, mejoraron la calidad del trabajo.

Bibliografía

Benitez, F., y Carena, M. (2021). Reflexiones acerca de la definición de radicación y su relación con la construcción de nuevos conceptos. *Rev. Educ. Mat.*, *36*(1), 9–26.

Carena, M. (2021). *Manual de matemática preuniversitaria: edición ampliada y corregida*. 2da ed., Ediciones UNL, Santa Fe.

Solow, D. (1993). *Cómo entender y hacer demostraciones en Matemáticas*. Limusa-Noriega Editores. (Traducción del trabajo original *How to read and do proofs*).

MARILINA CARENA
CONICET - Facultad de Ingeniería Química (UNL)
(✉) *marilcarena@gmail.com*

Recibido: *4 de febrero de 2023*.
Aceptado: *22 de junio de 2023*.
Publicado en línea: *3 de agosto de 2023*.