REAL ACADEMIA DE CIENCIAS
EXACTAS, FÍSICAS Y NATURALES DE ESPAÑA

DISCURSO INAUGURAL

DEL AÑO ACADÉMICO 2024-2025

LEÍDO EN LA SESIÓN CELEBRADA EL DÍA 25 DE SEPTIEMBRE DE 2024
POR EL ACADÉMICO NUMERARIO

EXCMO. SR. D. JOSÉ BONET SOLVES

SOBRE EL TEMA

DEMOSTRACIONES MATEMÁTICAS Y EL ORDENADOR

MADRID
DOMICILIO DE LA ACADEMIA
VALVERDE, 22 – TELÉFONO 917 014 230
www.rac.es
2024

ISSN: 1138-4093
ISBN: 978-84-87125-88-1
Depósito legal: M-18452-2024

ÍNDICE

Páginas

1. Introducción.

Excelentísima Señora Presidenta,
Excelentísimas Señoras Académicas, Excelentísimos Señores Académicos,
Autoridades, Señoras y Señores,

Los matemáticos hemos utilizado máquinas para ayudar en nuestro trabajo desde la antigüedad, por ejemplo el ábaco y las tablas trigonométricas o de logaritmos. La aparición y desarrollo de los ordenadores en el siglo XX y, más recientemente los avances en Inteligencia Artificial, han tenido un gran impacto en la investigación matemática, no solo en el análisis numérico, la simulación de modelos y las aplicaciones a la tecnología, sino también en las demostraciones de teoremas matemáticos. Mi propósito en este discurso es presentarles a ustedes algunos ejemplos de la influencia de las innovaciones en computación e Inteligencia Artificial en las demostraciones matemáticas. A la vez, reflexionaremos acerca de lo que los matemáticos consideran una demostración y sobre las pruebas asistidas por ordenador.

2. El juego de mesa SET y la conjetura del cap set.

Me van a permitir ustedes que comience mi exposición con la descripción de un juego de cartas de percepción visual, llamado **SET**. El juego SET está formado por una serie de cartas, cada una de las cuales consta de un diseño sobre fondo blanco. Este diseño está determinado por *cuatro características*:

(1) La forma de los símbolos (óvalos, ondas o rombos).

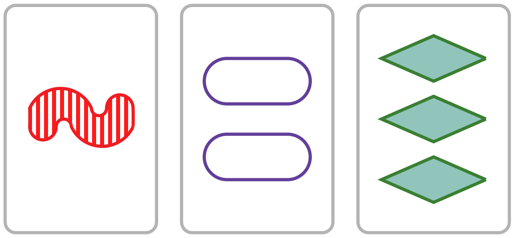
(2) los colores (rojo, verde o morado);

(3) el número de símbolos (uno, dos o tres);

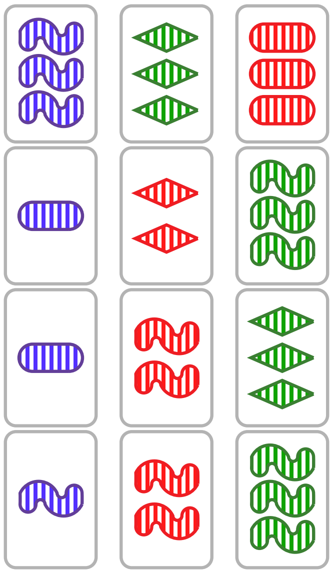
(4) el fondo del símbolo (sólido, rayado o hueco).

Las cartas combinan todas estas características del diseño, por ello existen exactamente $3^4 = 81$ cartas diferentes en el mazo del SET.

El objetivo del juego es conseguir formar un trío especial, llamado SET, de entre las 12 cartas que deben de estar boca arriba sobre la mesa. *Un SET se forma si, respecto a cada una de las cuatro características (forma, color, número y fondo), evaluadas una a una, las tres cartas poseen la característica igual o las tres poseen la característica diferente.* Las tres cartas de la imagen siguiente forman un SET.



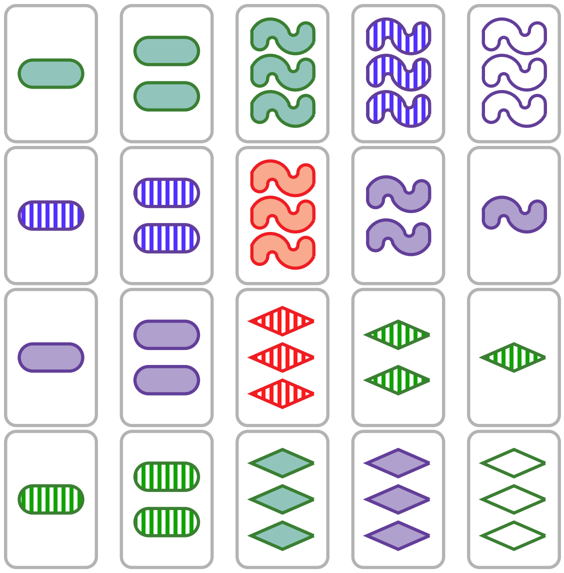
En efecto, las tres tienen diferente forma, diferente color, diferente número y diferente fondo. Si un jugador encuentra un SET entre las 12 cartas, lo coge, lo anuncia y se añaden 3 nuevas cartas. Si ningún jugador puede encontrar un SET entre las 12 cartas, se añaden 3 cartas más y así sucesivamente. Aquí hay otros 4 ejemplos de SET:



En el primero, el número de símbolos es igual en las tres cartas, es 3, el fondo rayado también es común y la forma y los colores son distintas en cada carta. En el segundo, tenemos común el fondo rayado a las tres cartas pero la forma, número y color es distinta en las tres cartas. El tercero, es similar al anterior. En el cuarto, la forma y el fondo son comunes a las tres cartas y son distintos, el número y el color. Es fácil que se puedan encontrar SETs en las 12 primeras cartas. De hecho, la probabilidad de que exista un SET con las 12 cartas primeras es del $96{,}77\,\%$ y con 15 cartas la probabilidad es del 99,96 por ciento. Se pueden consultar más detalles en los artículos de Ibáñez [37] y [38].

La pregunta más interesante para nosotros es *determinar el número máximo de cartas que puede haber en la mesa sin formar un SET.* La respuesta es 20, como fue demostrado por G.

Pellegrino [47] en 1970, antes de que Marsha Jean Falco inventara el juego. Aquí tienen un ejemplo de 20 cartas sin ningún SET:



Falco, que era genetista, diseñó el juego en 1974, cuando investigaba si la epilepsia era hereditaria en los pastores alemanes, para visualizar la combinatoria de los genes relacionados con esa epilepsia. Falco creó unas cartas con símbolos para representar los diferentes genes y cromosomas involucrados. En 1988 fundó Set Enterprises para comercializar el juego, que se publicó en 1991.

¿Por qué Pellegrino demostró el resultado antes de la invención del juego? ¿Qué relación tiene esto con el tema que nos interesa? Para contestar estas dos cuestiones debemos interpretar las cartas como vectores con coordenadas en un cuerpo finito.

En el juego hay 3 valores para cada una de las características, por ello reducimos el conjunto de todos los números enteros $\mathbb{Z}$ y su aritmética al conjunto finito $\{0,1,2\}$, o sea a los enteros módulo 3, que se denotan $\mathbb{Z}_3$. Los restantes enteros se identifican de forma cíclica con los de $\{0,1,2\}$; o sea, $3 = 0, 4 = 1, 5 = 2, 6 = 3, \ldots$ Con las operaciones naturales de suma y multiplicación $\mathbb{Z}_3$ tiene estructura algebraica de cuerpo.

Las cartas del juego SET las identificamos con los vectores $\bar{x} = (a,b,c,d)$ con 4 coordenadas en $\mathbb{Z}_3$, cada una de las cuales corresponde a una de las características (forma, color, número y fondo), a las que se asigna el número $1,2$ o 0 (recordemos que $3 = 0$) en función del valor que tome de los tres posibles. El espacio vectorial de esos vectores se denota por $(\mathbb{Z}_3)^4$.

No es difícil demostrar que un SET de cartas son tres vectores $\bar{x},\bar{y},\bar{z} \in (\mathbb{Z}_3)^4$ tales que

$\overline{x}+\overline{y}+\overline{z}=\overline{0}$. Como $2=-1$, esto es lo mismo que decir que $\overline{y}-\overline{x}=\overline{z}-\overline{y}$, o sea, los tres vectores están en la misma recta. Observe que si las 3 características son iguales, su suma es múltiplo de 3 y por tanto 0 en $\mathbb{Z}_3$, y si las tres son distintas, se tiene $0+1+2=0$.

La reformulación matemática de la pregunta principal es:

¿Qué tamaño máximo puede tener un subconjunto de $(\mathbb{Z}_3)^4$ *que no contenga ninguna recta?*

Si generalizamos la dimensión 4 a n, tenemos formulado el *problema del cap set*, es decir:

¿Qué tamaño máximo puede tener un subconjunto de $(\mathbb{Z}_3)^n$ *que no contenga ninguna recta?*

Es importante observar también que la pregunta se puede formular preguntando por el tamaño máximo de un subconjunto de $(\mathbb{Z}_3)^4$ sin progresiones aritméticas de longitud 3. Lo que nos conduce a la conexión de este problema con la estructura de las progresiones aritméticas en los números primos.

Un número **primo** p es un número entero estrictamente mayor que 1 que sólo es divisible por 1 y por sí mismo. Estos son los números primos menores que 150:

$$2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,$$

$$53,59,61,67,71,73,79,83,89,97,101,103,$$

$$107,109,113,127,131,137,139,149.$$

El teorema fundamental de la aritmética asegura que todo número entero $a>1$ se puede escribir como un producto de potencias de números primos. O sea,

$$a=p_1^{e_1}...p_k^{e_k},$$

donde $p_1,...p_k$ son primos distintos entre sí y cada exponente e_i es positivo. La descomposición es única salvo en el orden de los números primos. El sistema criptográfico con clave pública RSA de Ron Rivest, Adi Shamir y Len Adleman está basado en la dificultad de la descomposición de un número dado en factores primos y en el pequeño teorema de Fermat. Más información al respecto puede encontrarse en el Capítulo 7 de [54] y en el Capítulo 5 de [57].

Ha habido importantes progresos recientes acerca de la estructura aditiva de los números primos.

Estos son ejemplos de progresiones aritméticas $p,p+r,...,p+rd$ en los primos:

- 2,3
- 3,5,7
- 5,11,17,23

- 11,41,71,101,131
- 7,37,67,97,127,157
- 7,157,307,457,607,757
- $11410337850553 + 4609098694200 \times d;\ d = 0, \ldots, 21$
  (*Moran, Pritchard, Thyssen, 1995*)
- $56211383760397 + 44546738095860 \times d;\ d = 0, \ldots, 22$
  (*Frind, Underwood, Jobling, 2004*)

Green y Tao [26] demostraron en 2004 que los números primos contienen progresiones aritméticas arbitrariamente largas. Tao fue medalla Fields en el Congreso Internacional de Matemáticos ICM celebrado en Madrid en 2006.

La conjetura de los primos gemelos afirma que existen infinitos pares de números primos de la forma $p, p+2$. Por ejemplo:

$$5,7 \quad \text{o} \quad 17,19 \quad \text{o} \quad 59,61 \quad \text{o} \quad 137,139.$$

Denotamos por p_n el primo n-ésimo. En 2013 Zhang demostró que existe una constante $C > 0$ tal que $p_{n+1} - p_n < C$ para infinitos valores de n. Maynard [42] probó en 2015 que la constante C de Zhang se puede mejorar hasta 246.

Los Elementos de Euclides ya tienen una demostración de que existen infinitos números primos. Pero es conocido que, además, la suma $\sum_n 1/p_n$ de los inversos de los números primos es también infinita. Este hecho, (re)descubierto por Erdös cuando era muy joven, le llevó a la siguiente conjetura, formulada por él en 1940 y que permanece abierta.

Conjetura de Erdös (1940): Si A es un conjunto de números naturales tal que $\sum_{n\in A} 1/n = \infty$, entonces A contiene progresiones aritméticas arbitrariamente largas.

El problema cap set está relacionado con estos resultados que acabamos de mencionar. Para formular la *conjetura del cap set* con precisión introducimos la siguiente notación: para un número natural n, $r_3((\mathbb{Z}_3)^n)$ denota el tamaño de un subconjunto de $(\mathbb{Z}_3)^n$ sin progresiones aritméticas de longitud 3. La conjetura fue formulada por varios autores en 1993 y pregunta si existe $0 < c < 3$ tal que $r_3((\mathbb{Z}_3)^n) \leq c^n$ para todo n. Referimos al artículo de Grochow [27] para diversas motivaciones de la importancia de esta conjetura en varias áreas de matemáticas.

La conjetura del cap set resultó ser sorprendentemente difícil y atrajo la atención de muchos expertos en combinatoria. Tao lo consideraba en 2007 uno de sus problemas favoritos.

Después de varios resultados parciales en los que se usaban técnicas de análisis de Fourier, en 2016 Croot, Lev y Pach [10] resolvieron un problema muy relacionado introduciendo un nueva técnica llamada método polinomial, que en líneas generales es una aplicación de la geometría algebraica a la combinatoria. Basándose en esa técnica, Ellenberg y Gijswijt [18] obtuvieron una solución positiva de la conjetura del cap set. La demostración ha sido simplificada por Tao con una versión más simétrica.

Las razones por la que hemos empezado esta conferencia con el juego SET y con el problema del cap set son las siguientes.

En primer lugar, Dahmen, Hörz y Lewis [11] describen en 2019 una formalización de la demostración de Ellenberg y Gijswijt [18] usando el asistente formal LEAN, del que hablaremos más adelante. La prueba formalizada incluye el caso general y valores concretos. Los autores escriben en el resumen de su artículo que su trabajo “muestra que algunas matemáticas recientes profundas están al alcance de los asistentes interactivos de demostraciones”.

Además, en el artículo [52], publicado el 14 de diciembre de 2023 en *Nature*, se anuncia que *Google DeepMind* ha utilizado una nueva herramienta llamada *FunSearch*, para descubrir nuevas construcciones de cap sets muy grandes por encima de todos los ejemplos conocidos, tanto en el caso finito dimensional como en los casos asintóticos. FunSearch es especial porque no es simplemente un programa de computadora convencional; es un modelo de lenguaje grande, una forma avanzada de Inteligencia Artificial (IA) que hasta ahora se había utilizado más para generar texto que para descubrir verdades científicas. Este logro muestra que estos modelos pueden proporcionar información nueva y verificable. Se llama FunSearch porque busca funciones matemáticas, no porque sea divertido.

La novedad y ventaja de FunSearch reside en su capacidad para generar ideas y soluciones como si estuviera jugando un puzzle infinitamente complejo, pero en lugar de piezas de rompecabezas, utiliza fragmentos de código. El enfoque de FunSearch es muy diferente al de otras herramientas de Inteligencia Artificial de DeepMind, como *AlphaTensor*. Mientras que AlphaTensor y similares están diseñados para tareas específicas, como mejorar la velocidad de ciertos cálculos, FunSearch tiene una aplicación mucho más amplia. Alpha Tensor fue utilizado para encontrar métodos más rápidos para multiplicar matrices, tal como se anunció en el artículo de *Nature* [19]. El método proporcionado era más rápido que el propuesto por Strassen en 1969.

La herramienta FunSearch no solo resuelve el problema que se le presenta, sino que crea una especie de “receta” para resolverlo, una secuencia de instrucciones que puede adaptarse a diferentes problemas. Además, los resultados de FunSearch son más fáciles de comprender para los humanos, a diferencia de otras soluciones matemáticas complejas. Bernardino Romera-Paredes, investigador que trabajó en ambas herramientas, explicaba a la prensa que AlphaTensor es genial multiplicando matrices, en nada más, mientras que FunSearch adopta un enfoque diferente, ya que combina un gran modelo lingüístico llamado *Codey* de *Google*.

Los investigadores empezaron por esbozar en Python el problema que querían resolver, pero omitieron las líneas del programa que especificarían cómo resolverlo. Ahí entra en juego FunSearch, pues consigue que Codey rellene los espacios en blanco, es decir, que sugiera el código que resolverá el problema. Un segundo algoritmo comprueba y puntúa las propuestas de Codey. Las mejores sugerencias, aunque aún no sean correctas, se guardan y se devuelven a Codey, que intenta completar el programa de nuevo. Tras un par de millones de sugerencias y unas cuantas docenas de repeticiones del proceso global, que tardó días, consiguió crear un código que encontraba soluciones correctas del problema del cap set muy grandes, que eran desconocidas anteriormente. Alhussein Fawzi, científico investigador de Google DeepMind, confesaba a la prensa que "tenemos hipótesis, pero no sabemos de manera exacta porqué funciona". Tao ha declarado que "FunSearch es un paradigma prometedor. Es una forma interesante de aprovechar la potencia de los grandes modelos lingüísticos".

Una vez presentados estos avances recientes de los verificadores formales de teoremas y de la Inteligencia Artificial para conseguir resultados significativos en problemas de matemáticas profundos y actuales, vamos a presentar varias demostraciones matemáticas obtenidas con ayuda del ordenador desde 1975 gracias a la extraordinaria capacidad de cálculo por "fuerza bruta".

3. Primeros ejemplos de demostraciones matemáticas con ayuda del ordenador.

Los calculadores humanos han jugado un papel muy relevante en la historia de las matemáticas, por ejemplo en las tablas de logaritmos y de funciones trigonométricas. John Napier, al principio del siglo XVII construyó las primeras tablas de logaritmos y tuvo la ayuda de calculadores humanos. Más tarde, dos de estos calculadores, Fenkel y Vega, publicaron tablas de la factorización de los números enteros positivos hasta 408.000 en 1776 y 1777. Gracias a ello fue posible crear tablas de la función $\pi(n)$, que denota el número de primos menores que un número n. Ello llevó a Lagrange y Gauss a la conjetura que $\pi(n)$ se comportaba como

$$L(n) := \frac{n}{\log n - 1} \qquad \text{o} \qquad \mathrm{li}(n) := \int_2^n \frac{1}{\log t} dt.$$

Esta tabla compara los valores de las tres cantidades.

$n = 50000$	$\pi(n) = 5133$	$L(n) = 5092$	$\mathrm{li}(n) = 5166$
$n = 500000$	$\pi(n) = 41538$	$L(n) = 41246$	$\mathrm{li}(n) = 41607$
$n = 10000000$	$\pi(n) = 664579$	$L(n) = 661459$	$\mathrm{li}(n) = 664918$

La conjetura permaneció abierta durante 100 años y fue demostrada en 1896, usando análisis complejo, por Hadamard y de la Vallée Poussin, independientemente, en lo que se conoce como el *Teorema de los Números Primos*:

$$\lim_{n\to\infty} \frac{\pi(n)\log n}{n} = \lim_{n\to\infty} \frac{\pi(n)}{\mathrm{li}(n)} = 1$$

En otras palabras, la probabilidad de que un número elegido al azar entre 1 y n sea primo es aproximadamente $1/\log n$. Este teorema fue uno de los grandes logros de las matemáticas del siglo XIX.

Las ideas de Riemann, publicadas en 1859 en su artículo "Sobre el número de primos menores que una cantidad dada", fueron fundamentales en la demostración. En ese artículo de Riemann apareció formulada por primera vez la *Hipótesis de Riemann*, considerado como uno de los problemas abiertos más importante de las Matemáticas. Se pueden encontrar más detalles acerca de este teorema y de la hipótesis de Riemann en la lección inaugural de Manuel López Pellicer [41].

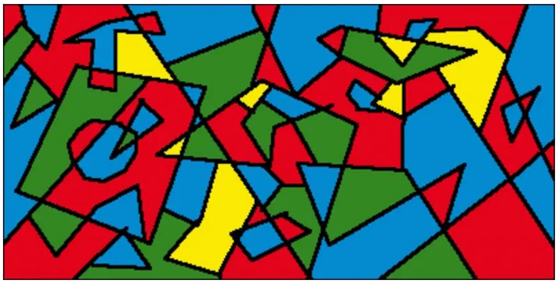
Como escribe Fernando Bombal en su lección inaugural [3] del año académico 2015-2016, "La aparición y desarrollo de los ordenadores digitales durante el siglo XX es probablemente uno de los hechos más revolucionarios de la historia. Su impacto en el desarrollo de la ciencia y la tecnología ha sido y sigue siendo enorme. La posibilidad de disponer de una gran potencia de cálculo, rápido y eficaz, revolucionó los ingeniosos métodos desarrollados por los matemáticos para resolver problemas prácticos de cálculo cuya solución por la mera fuerza bruta sería impracticable." Referimos a ese discurso para una breve y detallada historia de los ordenadores.

Bombal [3] escribe también: "La ventaja de disponer de una elevada potencia de cálculo a gran velocidad ha tenido un gran impacto en la investigación matemática. Y no sólo en áreas la simulación y creación de modelos, el análisis numérico de las soluciones de ecuaciones funcionales y las aplicaciones a la técnica y la ingeniería, sino en los dominios mismos de la matemática pura."

Revisamos algunos ejemplos de utilización de la capacidad de cálculo del ordenador para completar la demostración de varios teoremas.

3.1. El Teorema de los cuatro colores.

La conjetura de los cuatro colores fue propuesta en 1852 por Francis Guthrie, un matemático aficionado inglés como sigue: *"bastan cuatro colores para colorear cualquier mapa plano de regiones conexas, de modo que no haya regiones adyacentes con el mismo color"*.



El hermano de Guthrie, que estudiaba en el University College de Londres, planteó la pregunta a De Morgan. A partir de ese momento muchos matemáticos intentaron resolverlo. En 1879 Kempe publicó un artículo en el *American Journal of Mathematics* en el que sostenía haber resuelto afirmativamente el problema de los cuatro colores. Pero en 1890 Heawood demostró que la prueba de Kempe era errónea y probó que cinco colores eran suficientes. El paso a cuatro colores permaneció abierto. Kenneth Appel y Wolfgang Haken en 1976 lograron reducir el problema a la comprobación de unas 1500 configuraciones básicas. Con la ayuda de un complejo programa de ordenador y el uso de más de 1.200 horas de computación, pudieron anunciar que todas las configuraciones básicas habían sido analizadas y no precisaban más de cuatro colores.

La demostración presentada por Appel y Haken fue seguida de una amplia controversia. La razón para ello es que introducía una nueva concepción de lo que se entiende por demostración matemática, que no depende exclusivamente de un razonamiento humano, ya que contiene etapas en el programa que nunca podrán ser comprobadas por un matemático. Como un ejemplo, Halmos [30] comentaba: "¿Qué hemos aprendido de la prueba de Appel y Haken? ¿Qué sabemos ahora que no supiéramos antes? No me parece una pregunta fácil de contestar." Y Davis y Hersh en la página 384 de [16] escriben "cuando me enteré que la prueba del problema de los cuatro colores estaba basada en reducirla a miles de casos y procesar cada uno de ellos en el ordenador, me sentí descorazonado. Tal vez esto muestra que no era un buen problema después de todo."

En 1996, Neil Robertson, Daniel P. Sanders, Paul Seymour y Robin Thomas [51] crearon un algoritmo de tiempo cuadrático, mejorando un algoritmo de tiempo cuártico basado en la prueba de Appel y Haken. La nueva prueba, basada en las mismas ideas, es similar a la de Appel y Haken pero más eficiente porque reduce la complejidad del problema y requiere comprobar sólo 633 configuraciones reducibles. Tanto la parte de inevitabilidad como la de reducibilidad de esta nueva prueba deben ser ejecutadas por un ordenador y no es práctico comprobarlas a mano.

En 2005, Georges Gonthier y Benjamin Werner presentaron una prueba completamente formalizada del teorema basada en la demostración de Robertson y coautores. Usaron el asistente de pruebas interactivo Coq, que es un sistema de ayuda para la demostración de teoremas que maneja aserciones matemáticas. Fue desarrollado en Francia. La página web es https://coq.inria.fr/

Un asistente de pruebas interactivo es diferente de los programas utilizados por Robertson y coautores y por Appel y Hanken, ya que no se limita a considerar un conjunto de datos con un algoritmo dado, sino que se basa en reglas de inferencia lógica y comprueba la demostración de afirmaciones matemáticas escritas en él. Pueden consultarse más detalles en el artículo de Gonthier [22] publicado en 2008.

3.2. La Conjetura de Kepler.

El matemático y astrónomo Johannes Kepler formuló el siguiente problema matemático, que ya había sido considerado por Walter Raleigh y Thomas Harriot: *"Encontrar cuál es la configuración de esferas de un radio dado que tiene máxima densidad"*. Se puede plantear un problema similar en el plano, considerando disposiciones de círculos con la máxima densidad respecto al área, o, en general, en cualquier dimensión n.

Para disposiciones regulares, es decir, tales que los centros de las esferas forman un retículo simétrico, Gauss probó en 1831 que la mejor configuración en el plano es la hexagonal (es decir, formada por círculos inscritos en hexágonos que teselan el plano), y en el espacio la formada por capas apiladas sobre una configuración cuadrada (o hexagonal, son equivalentes) de modo que las esferas de la capa superior estén colocadas en los huecos que quedan en la capa inferior, configuración de centros no alineados; que es la que suelen utilizar los fruteros para colocar las naranjas y otros frutos esféricos.

El caso general es mucho mas difícil y son muchos los matemáticos que han intentado resolverlo. En 1990, Wu-Yi Hsiang afirmó haber demostrado la conjetura, pero se encontraron errores importantes en la prueba.

En 1998, Thomas C. Hales anunció haber probado la conjetura, mostrando que la configuración regular de centros no alineados es la óptima. Como en el caso del Teorema de los cuatro colores, se trata de un proceso de reducción del número de configuraciones posibles. La demostración de Hales reduce el problema a minimizar una ecuación de 150 variables sobre unas 5.000 configuraciones de esferas. Se trata de un trabajo de más de 250 páginas de razonamientos matemáticos y varios gigabytes de datos y programas de ordenador. La revista *Annals of Mathematics* designó un panel de 12 expertos para revisar la demostración. Tras 4 años de trabajo el presidente del comité de revisión declaró que el comité estaba 99 por cien seguro de la exactitud de la prueba, pero que no iban a continuar. Finalmente, en 2005, apareció publicada la demostración.

En enero de 2003, Hales anunció la creación del Proyecto Flyspeck para transcribir su demostración a una prueba formal que pudiera ser comprobada por alguno de los verificadores formales de teoremas, como Isabelle. En agosto de 2014 se completó el Proyecto, anunciando que el programa creado había verificado la demostración de Hales y que no encontró errores. El artículo con la prueba formal fue publicado en 2017 [29]. En el resumen se explica que el artículo describe una demostración formal de la conjetura de Kepler sobre empaquetamiento denso de esferas combinando los asistentes de pruebas HOL Light e Isabelle.

El problema del empaquetamiento óptimo de hiperesferas en dimensión n tiene considerable importancia en la corrección de errores en la transmisión de mensajes. El conjunto de cadenas binarias de n símbolos forman los vértices de un hipercubo de n dimensiones. Para evitar la transmisión de errores hay que tratar que los vértices que codifican mensajes no sean adyacentes. Una configuración de hiperesferas con máxima densidad maximiza el número de posibles mensajes a enviar, minimizando la posibilidad de error.

3.3. El problema del triplete booleano.

El problema del triplete booleano es una cuestión relacionada con ternas pitagóricas, que se resolvió utilizando una prueba asistida por ordenador en 2016. Se llama *terna pitagórica o triplete pitagórico* a cualquier conjunto de tres números naturales (a, b, c) para los cuales se cumple la relación $a^2+b^2=c^2$. Los primeros ejemplos son $(3,4,5),(5,12,13),(6,8,10),(7,24,25),\dots$.

El planteamiento del problema es el siguiente: *Dado un número natural N, ¿es posible escribir los elementos del conjunto de los números del* 1 *al N utilizando solo dos colores, de modo que cada terna pitagórica de dicho conjunto no esté formada por números del mismo color?* Por ejemplo, en el triplete pitagórico $3,4$ y 5, si 3 y 4 son de color rojo, entonces 5 debe ser de color azul.

Marijn Heule, Oliver Kullmann y Victor Marek investigaron el problema y demostraron que al menos existe un caso en el que tal coloración es imposible. Hasta el número 7824 es posible colorear los números de modo que todos los tripletes pitagóricos sean admisibles, pero la prueba muestra que dicho color no se puede extender para también colorear el número 7825.

Hay 2^{7825} coloraciones posibles para los números hasta 7825. Estas posibles coloraciones se redujeron lógica y algorítmicamente hasta alrededor de un trillón de (aún altamente complejos) casos, y se examinaron utilizando un resolucionador de problemas de satisfacibilidad booleana. El desarrollo de la prueba requirió aproximadamente 4 años de trabajo, así como dos días de cálculo en el superordenador Stampede del *Texas Advanced Computing Center*, usando 800 procesadores en paralelo. Generó una prueba proposicional de 200 terabytes, que se comprimió a 68 gigabytes.

El documento que describe la prueba se publicó en arXiv el 3 de mayo de 2016, y fue aceptado para la conferencia SAT 2016, donde ganó el premio al mejor artículo. En la década de 1980, el matemático Ronald Graham había ofrecido un premio de 100 dólares por la resolución del problema, otorgado a Marijn Heule en 2016.

Heule, Kullmann y Marek son expertos en "propositional satisfiability problems" (SAT, en corto). Este tipo de problemas, del que la coloración de las ternas pitagóricas es un ejemplo, es de tipo NP completo; o sea, (a) no puede ser resuelto por un ordenador ideal (una máquina de Turing) en tiempo polinomial del número de datos que lo definen (problema NP), pero una vez conocida la solución, esta tiene una comprobación polinomial y (b) todo problema NP se puede reducir en tiempo polinomial a uno de tipo SAT. A pesar de lo cual se ha desarrollado toda una artillería computacional para resolver problemas de tipo SAT. Ha sido este tipo de herramientas las que han usado Heule, Kullmann y Marek para dar respuesta negativa a la coloración de las ternas pitagóricas.

Para más información acerca de las clases de complejidad P y NP, referimos al discurso de recepción de David Pérez García [48]. En él escribe: "Las enormes implicaciones filosóficas de suponer que P = NP hacen creer que ambas clases son distintas. Si fueran iguales, cualquier problema en el que se pudiera verificar eficientemente que una solución es correcta, se podría encontrar esa solución también de forma eficiente. Obviamente, esto significaría que la creatividad puede ser automatizada y que, por ejemplo, desde la perspectiva de las matemáticas, sería esencialmente igual de difícil verificar la demostración de un teorema matemático, algo que cualquier persona con la suficiente dedicación y formación matemática puede hacer, que la genialidad de obtener dicha demostración."

Heule, Kullmann y Marek creen, además, que este problema seguirá sin tener solución aumentando el número de colores. De hecho han propuesto la siguiente conjetura: para cada número k mayor que dos hay un número $f(k)$ tal que el conjunto $\{1,2,...,f(k)-1\}$ puede ser coloreado con k colores de manera que no haya tres números del mismo color y que formen una terna pitagórica, mientras que esto es imposible para $\{1,2,...,f(k)\}$. Por ahora sólo sabemos que $f(2) = 7825$.

Mucho se ha debatido desde que, en 1977, Appel y Haken se ayudaran de un ordenador para completar su demostración del teorema de los cuatro colores. La prueba de Heule, Kullmann y Marek debe todavía más al ordenador, ya que incluso no se enseña cuál es la característica que hace que el problema de la coloración de las ternas pitagóricas falle para $\{1,2,3,...,7825\}$.

3.4. El Sudoku.

El Sudoku es un rombecabezas numérico bien conocido que consiste en rellenar un cuadrado de 9×9 casillas con números de modo que cada fila, cada columna y cada sub-cuadrado de tamaño 3×3 contenga todos los números del 1 al 9. Todo problema de Sudoku da ciertos números (pistas) ya colocadas en algunas casillas del cuadrado. Desde que se publicó el Sudoku por primera vez en los años 1980, se consideraba que el menor número de sugerencias necesarias para que hubiera una solución única era 17.

			5		9			
1						7		
	5			7	2			6
7	8				6	9	3	
				4				
	2	9	7				6	5
5			8	1			4	
		4						7
			4		3			

McGuire, Tugemann y Civario [44] mostraron en 2014 que ese era el caso. Escribieron un programa, llamado Checker, para comprobar todas las posibles soluciones de los Sudokus con 16 pistas e implementaron Checker en el Irish Centre for High-End Computing para hacer una búsqueda que tardó un año. El ordenador no encontró una solución única en ninguno de los Sudokus con 16 pistas iniciales y así concluyeron que 17 era efectivamente el número mínimo de pistas.

La táctica de McGuire, Tugemann y Civario muestra uno de los procedimientos usuales de trabajar con asistentes matemáticos: los matemáticos formulan el problema y producen un algoritmo de prueba que incluye un número finito de casos, que deben ser analizados por separado. El asistente o bien encuentra un contraejemplo o bien no encuentra ninguno y los matemáticos concluyen que cierto enunciado ha sido demostrado.

3.5. Una breve reflexión.

En los cuatro ejemplos mencionados en cierto punto de la demostración se necesita computación por ordenador por fuerza bruta que es demasiado amplia para ser realizada por los humanos. Por ello es imposible que estas demostraciones sean verificadas por matemáticos. La característica fundamental de las demostraciones de que sean verificables, que cada paso pueda ser analizado por un agente racional, no se satisface en los ejemplos dados en esta Sección. En el caso del Sudoku la “prueba” no es ni siquiera formalizable, se limita a la comprobación de cada caso con el ordenador.

Las principales objeciones a este tipo de pruebas son la posibilidad de errores en la programación y, sobre todo, la gran dificultad de entender y verificar estas demostraciones obtenidas por el ordenador. El proceso social, que es esencial en la investigación matemática, por el cual los matemáticos interiorizan y luego construyen sobre resultados previos, no aparece en estas demostraciones.

4. Demostraciones matemáticas. Mitos y realidades.

¿Qué es una demostración matemática? Este es el asunto central de esta parte del discurso.

Acerca de este tema sí les puedo hablar con cierta propiedad. He trabajado en análisis matemático, análisis funcional y complejo y teoría de operadores más de 40 años. He demostrado y publicado muchos teoremas con demostraciones más o menos técnicas, más o menos complicadas, largas o cortas, que han tenido en ocasiones influencia en trabajos de otros matemáticos. He resuelto algunos problemas que estaban abiertos desde hacía tiempo. He colaborado con varias decenas de matemáticos de varios países, he actuado como revisor en innumerables ocasiones para muchas revistas de prestigio y he sido editor de revistas con impacto considerable. Les puedo asegurar que conozco bastante bien cómo funciona realmente la investigación matemática. Esta sección refleja opiniones personales basadas en mis propias experiencias.

Vamos a empezar con el mito: Una demostración (prueba) de un teorema matemático es un razonamiento mediante el que se afirma la verdad de una proposición, partiendo de verdades universales y evidentes, aplicando las reglas de la lógica. Es un razonamiento convincente con el que se corrobora la veracidad de una proposición. Teóricamente sería posible verificar todos los pasos a partir de los axiomas de la teoría. Una demostración es una sucesión de símbolos o de frases formales de acuerdo con ciertas reglas de la lógica. Las matemáticas son formales, precisas, ordenadas y abstractas. Se ordenan en definiciones, teoremas, demostraciones y observaciones. Las matemáticas son únicas, universales, proporcionan pruebas rigurosas y ciertas y la verdad matemática es la misma para todo el mundo. Como se menciona en el Prefacio del libro “Proofs from THE BOOK” de Aigner y Ziegler [1], Paul Erdös solía hablar de “El Libro” (The Book) en el que Dios guardaba las pruebas perfectas de los teoremas matemáticos. Y añadía que podías no creer en Dios, pero que, como matemático, tenías que creer en “El Libro”.

Ahora la realidad de la práctica matemática: Una demostración no es más que un argumento que es considerado convincente por revisores competentes. Una demostración es lo que hacemos los matemáticos para convencer a otros expertos en nuestro tema de la verdad de nuestro teorema. No es una verdad eterna a la que se llega solo por puro pensamiento. Las demostraciones son un proceso experimental, de acierto-error. Los matemáticos aprendemos en diferentes etapas movidos por la intuición, a veces equivocada, queremos entender la demostración, incorporarla y compararla con lo que sabemos. Lo que hacemos día a día tiene poco que ver con la lógica formal. La práctica matemática es fragmentaria, informal, intuitiva y tentativa. Hay que conjeturar un teorema antes de probarlo, hay que tener una idea de la prueba antes de completar los detalles, hay que usar observaciones y analogías e intentarlo una y otra vez. Los matemáticos usamos razones variadas para creer que una prueba es correcta, además de la comprobación paso a paso: contraste con ejemplos y casos especiales, analogía con otros resultados, simetría, elegancia e incluso una inexplicable sensación de que debe ser cierta. No hay certeza absoluta en matemáticas, hay certeza virtual, como en todas las áreas de la ciencia y de la vida. La confianza de un matemático en sus teoremas no necesariamente significa que conoce cada paso desde los axiomas de la teoría de conjuntos hasta su resultado, sino que confía en la palabra de otros investigadores expertos, de las revistas y de los revisores. En ese sentido las pruebas matemáticas son una construcción social. Ver por ejemplo [16], [24], [33], [34], [35], [45] y [53], entre otros.

Escribía Ramón y Cajal en su libro [50] (página 158): “En conclusión, caí un poco tarde en la cuenta de que las verdades matemáticas, que rutinarios y secos pedagogos consideran, no sin cierta aristocrática infatuación, cual construcción deductiva (cadena de verdades cuyo primer escalón brota en la esencia del espíritu), surgida a priori, a espaldas y hasta con desdén de la experiencia, representan, por el contrario imposición ineluctable del mundo objetivo, algo así como la quintaesencia de los conceptos derivados de la percepción y escrupulosamente depurados de contingencias, a fin de que la lógica racional pueda depurarlos ágil y cómodamente. Y, sabido esto, no me sorprendió que los axiomas y fórmulas de la geometría y el álgebra se acoplen tan estrechamente a la realidad exterior, puesto que, en último análisis, de la realidad proceden.”

Los resultados de matemáticas se verifican y se presentan con el método axiomático. El método axiomático en la presentación de las matemáticas ha llegado a ser en algunos círculos una forma de fanatismo, hasta el punto de sacrificar la claridad en aras de la consistencia de la notación, la brevedad del argumento y la linealidad de las inferencias lógicas. Sin embargo, no debe confundirse la presentación de las matemáticas con su contenido. De hecho, no es cierto que el método axiomático sea el instrumento de creación en matemáticas.

Los matemáticos no usamos el método axiomático para resolver problemas y demostrar teoremas. Mazur [43] en 2014 escribía: "Razonar por analogía es la pieza clave, está presente en la mayor parte del pensamiento matemático y es a menudo la inspiración de algunos de los proyectos más ambiciosos en matemáticas." Tao en su publicación acerca del rigor comenta que "la finalidad del rigor no es destruir la intuición; al contrario, debe ser usado para destruir la mala intuición mientras clarifica y eleva la buena intuición. Sólo con una combinación de formalismo riguroso y buena intuición se pueden atacar problemas matemáticos complejos." Está claro que no hay alternativa viable a la presentación axiomática, si queremos que un teorema y su prueba puedan ser aceptados por la comunidad matemática más allá de una duda razonable.

Todos los matemáticos cometen errores. Preferimos una prueba elegante y bonita, aunque pueda contener errores, a una que sea larga, aburrida y correcta. Me atrevo a repetir esa frase oída tantas veces: "los errores de un gran matemático valen mucho más que los resultados correctos de un mediocre." Los matemáticos somos muy tolerantes con esos errores, siempre que sean reparables y no afecten a las ideas centrales de la prueba. Los errores que pueden corregirse son muy útiles para comprobar la estabilidad de una teoría. Los errores en demostraciones son bastante comunes. El hecho de que sean encontrados y corregidos enfatiza que las pruebas sean verificables.

La actividad principal de un matemático es encontrar demostraciones de teoremas, no comprobar que son correctas. La detección de un error en una prueba o una omisión, es bienvenido como un paso para mejorarla. El libro de Lakatos "Pruebas y Refutaciones" [40] es una presentación fascinante de este aspecto de las demostraciones matemáticas. La historia de la práctica matemática real sugiere que cuanto menos se cuestiona una demostración, más susceptible es que aparezca un error en ella. Así pues, es importante encontrar varias maneras de explicar y verificar una determinada prueba y, si se puede, encontrar pruebas alternativas de una teorema. Las demostraciones con ayuda del ordenador pueden dar nuevas perspectivas.

El criterio principal para el valor de una demostración no es ser irrefutable. Depende mucho más de si la prueba proporciona una técnica, un método que podría ser utilizado para obtener otros teoremas. El valor de una demostración reside también en las posibilidades que abre de conexión con otras áreas de matemáticas. Los matemáticos, cuando hacen demostraciones o las comprueban, están interesados en nuevas técnicas, nuevas ideas, nuevas relaciones con otras áreas y nuevos métodos de demostración, no solo en el resultado. Incluso una prueba simple o más corta puede ser una aportación matemática esencial. Como decía Paul Erdös, nadie culpa a un matemático porque su primera demostración de un teorema nuevo sea burda. De hecho, es posible que una primera demostración, no muy elegante, presente las ideas y técnicas nuevas con mayor claridad que una prueba elaborada y precisa. En ocasiones el exceso de rigor y perfección en la exposición oculta las verdaderas novedades. Por otra parte, el proceso de simplificación y mejora de la prueba de teoremas importantes requiere el trabajo serio de generaciones de matemáticos.

La práctica de los matemáticos no se corresponde con la noción formal de una demostración matemática, ni con la forma como las pruebas son expuestas en los artículos y libros. Las demostraciones matemáticas deben ser convincentes, deben poder ser juzgadas y analizadas por la comunidad de expertos. El punto principal de una demostración matemática es persuadir al lector de la verdad de un enunciado. La verificación es clave. La mejor verificación que conocemos en matemáticas es que muchos otros matemáticos estudien la prueba desde distintas perspectivas y les convenza a la luz de lo que conocen previamente. Las demostraciones se aceptan por esas comunidades de expertos. Una demostración es un argumento que fuerza el consenso de la comunidad. De ese modo lo que es cierto y lo que es interesante depende de la comunidad de matemáticos. El valor que los matemáticos asignan a un trabajo de matemáticas depende de la percepción del trabajo, no de su calidad objetiva.

Es un hecho conocido que los matemáticos hacemos juicios probabilísticos acerca de la validez de resultados matemáticos. No les asociamos probabilidades, pero usamos frases para calificarlos del tipo “probablemente”, “casi seguro” o “muy poco probable”. El artículo de Gowers [23] considera en profundidad la cuestión de hacer juicios probabilísticos acerca de teoremas no demostrados y discute cómo se pueden interpretar formalmente estas opiniones.

El contenido central de las matemáticas son las ideas y la comprensión, no solo las pruebas. Cuando un matemático prueba un nuevo teorema lo hace consiguiendo una nueva idea. Por lo general la idea es correcta aunque la demostración no lo sea al principio. Cuando obtiene un resultado realmente nuevo, original y profundo, manejando nuevas ideas, el matemático intenta generalmente incluir a otros expertos tratando de explicar sus ideas de la forma más clara y accesible posible. Compartir argumentos matemáticos es la condición necesaria para la justificación matemática. La formalización de una demostración no garantiza la corrección del argumento. De hecho, hasta que no compruebas tus resultados con otros matemáticos, no puedes estar seguro de que no hayas pasado algo por alto.

Granville en [25] cita a Eugenia Cheng en ese respecto: “la prueba supuestamente establece la verdad de una parte de matemáticas, pero la demostración no es lo que convence a los matemáticos de la verdad, otra cosa lo hace.” La comunicación matemática convierte la creencia del matemático en su prueba en una verdad compartida con el lector de la misma. La comunidad matemática acepta demostraciones si los expertos pueden apreciar completamente la estrategia general y si cada parte de la prueba puede ser comprobada por un experto en las técnicas. Además, el proceso de verificación inspira a menudo nuevas perspectivas en la comunidad matemática. Los revisores intentan incorporar las nuevas ideas y técnicas en su propio repertorio; no solo leen el texto del autor, sino que lo reinterpretan en sus propios términos incorporando las nuevas ideas a lo que ya saben.

Los matemáticos siempre intentamos conseguir demostraciones cortas, simples y elegantes. Estas pruebas son deseables porque las matemáticas son un proceso social. Los matemáticos necesitamos entender las demostraciones no solo para asegurarnos de que son correctas, sino para interiorizar las ideas centrales contenidas en ellas. Las demostraciones son un vehículo para llegar a una comprensión más profunda de la realidad matemática. Nos gustan las demostraciones informativas, pruebas que expliquen, que nos permitan modificar lo que aprendemos y nos ayuden a desarrollar nuevos conceptos.

5. Verificadores formales de teoremas. La Inteligencia Artificial y el aprendizaje automático.

Usar ordenadores para recoger evidencia experimental en matemáticas es usual hoy día. Los sistemas de álgebra computacional, como Maple o Mathematica, se usan habitualmente en la comunidad matemática, se enseñan a los estudiantes de grado y se utilizan en la investigación para experimentación y visualización. El ordenador hace los cálculos y el matemático proporciona el conocimiento y las interpretaciones de los resultados. Las herramientas tecnológicas que intervienen son los verificadores automáticos e interactivos de teoremas, la Inteligencia Artificial, el aprendizaje automático (*machine learning*) y los modelos de lenguaje de gran tamaño (*Large Language Models LLM*). Los cambios recientes se producen a tal velocidad que es casi imposible seguir los progresos.

5.1. Demostradores automáticos de teoremas.

Los *demostradores automáticos de teoremas (automated theorem provers*, en inglés) son sistemas que conocen los axiomas de una teoría, pueden demostrar teoremas usando lógica de primer orden y disponen de un lenguaje lo bastante rico para expresar teoremas de la teoría. Los axiomas y las conjeturas son representados como fórmulas lógicas y las reglas de inferencia se representan como programas para manipular esas fórmulas. Estos sistemas juegan un papel importante en la verificación de software.

Uno de los primeros éxitos de estos sistemas fue la demostración de la conjetura de Robbins, que había presentado en la década de 1930 una colección de axiomas que él conjeturaba debían ser una axiomatización de las álgebras Booleanas. La dificultad era verificar todos los axiomas de éstas a partir de los axiomas de Robbins. William McCune en 1996 probó la conjetura usando el demostrador automático de teoremas EQP (*Equational prover*), que fue desarrollado por la *Mathematics and Computer Science Division of the Argonne National Laboratory*. Dahn [12] obtuvo una prueba simplificada de 7 páginas que se publicó en el *Journal of Algebra*.

Las pruebas generadas por estos sistemas pueden ser muy largas (y difíciles de verificar por un matemático), pero debe señalarse que no son demostraciones por "fuerza bruta", como podríamos calificar las mencionadas en la Sección 3. Hoy día los demostradores automáticos de teoremas se usan como componentes de los demostradores interactivos de teoremas.

5.2. Demostradores interactivos de teoremas.

Los *demostradores interactivos de teoremas (interactive theorem provers*, en inglés), como Isabelle, HOL, LEAN, Coq, PVS, permiten acceder a demostraciones más profundas. Éstos también conocen las reglas lógicas y los axiomas de una teoría, pero son más ricos y complejos que los demostradores automáticos de teoremas y, por tanto, pueden formalizar matemáticas más profundas. Un matemático que utilice un demostrador interactivo de teoremas deberá teclear las ideas principales del teorema. De alguna forma estos sistemas son como un lenguaje de programación en el que el código corresponde a definiciones, teoremas y pruebas. En este momento, escribir una demostración en uno de estos sistemas consume mucho más tiempo que escribir matemáticas en LaTeX.

Originalmente los demostradores interactivos de teoremas se usaron para verificar resultados matemáticos básicos. Sin embargo, esto ha cambiado. En 2004 un equipo dirigido por J. Avifad verificó el teorema de los números primos en Isabelle. Más tarde G. Gonthier [22] verificó el teorema de los cuatro colores formalmente y Hales y Ferguson [29] verificaron la prueba de la conjetura de Kepler, como ya se ha mencionado en la Sección 3.

En 2019 Dahmen, Hölzl y Lewis [11] formalizaron usando LEAN la solución de Ellenberg y Gijswijt del problema del cap set, tal como fue mencionado en la Sección 2. Los demostradores interactivos de teoremas también han jugado un papel importante en el trabajo [4] en 2019 acerca de la conjetura de Keller, que postula que en cualquier teselado del espacio euclídeo de dimensión n mediante hipercubos idénticos, existen parejas de hipercubos contiguos que comparten entre sí una cara completa de dimensión $(n-1)$.

Más recientemente Commelin y Topaz [9] han dirigido un proyecto llamado el "Experimento Tensorial Líquido" (*Liquid Tensor Experiment*, en inglés) para formalizar un teorema de Clausen y Scholze en LEAN. LEAN es un demostrador interactivo de teoremas que fue escrito por Leonardo de Moura en Microsoft Research, es un software libre y abierto que puede ser ejecutado en cualquier sistema operativo moderno. Referimos al lector a Scholze [55] para más detalles sobre el Experimento Tensorial Líquido, que está basado en la librería matemática de LEAN *mathlib*. Esta librería contiene esencialmente todas las matemáticas preuniversitarias y muchos resultados de nivel de máster de álgebra, teoría de números y geometría algebraica y crece diariamente con aportaciones de muchos matemáticos interesados.

LEAN
THEOREM PROVER

Los asistentes matemáticos pueden decidir si una prueba es consistente con lo que se conoce, pero nunca podrán juzgar si un enunciado matemático es interesante o importante. Por ejemplo LEAN, que ha permitido ya verificar muchas demostraciones, puede además ayudar a los autores a entender mejor su propio trabajo porque pueden transformar algunas de sus ideas en pasos más simples que pueden ser escritos en LEAN.

Para que los asistentes matemáticos sean más utilizados por la comunidad matemática, deberían tener las siguientes características, señaladas por Ganesalingam y Gowers en su artículo [20]: (1) que se puedan introducir los problemas sin necesidad de aprender un lenguaje formal, (2) que el programa proporcione soluciones expresadas en un lenguaje que los matemáticos usan habitualmente, (3) que las soluciones sean argumentos del tipo que los matemáticos utilizan, y (4) que se puedan añadir datos al programa, aumentando su capacidad, sin necesidad de reescribir todo el programa. No es descartable que, en el futuro y con la ayuda de la Inteligencia Artificial generativa, se conviertan en una herramienta indispensable.

Quizá los demostradores interactivos de teoremas conviertan en el futuro las matemáticas en un juego como el ajedrez, el Shogi o el Go. No lo podemos saber ahora. La Inteligencia Artificial ha desarrollado máquinas capaces de derrotar a los mejores maestros en ajedrez, Shogi y Go. Deep Blue derrotó a Gary Kasparov, campeón del mundo de ajedrez, en 1997.

El Go es un juego de mesa tradicional chino con más de 2500 años de antigüedad. Se trata de un juego para 2 personas que, por turnos, van colocando piezas blancas y negras en un tablero estándar de 19×19. El objetivo es capturar las piezas del oponente, eliminándolas así del tablero, o rodear espacios vacíos para hacer puntos de territorio. El Go tiene mucha más complejidad que el ajedrez, porque el número de configuraciones es astronómicamente mayor. Los mejores jugadores de Go fueron derrotados por AlphaGo, que es un algoritmo de

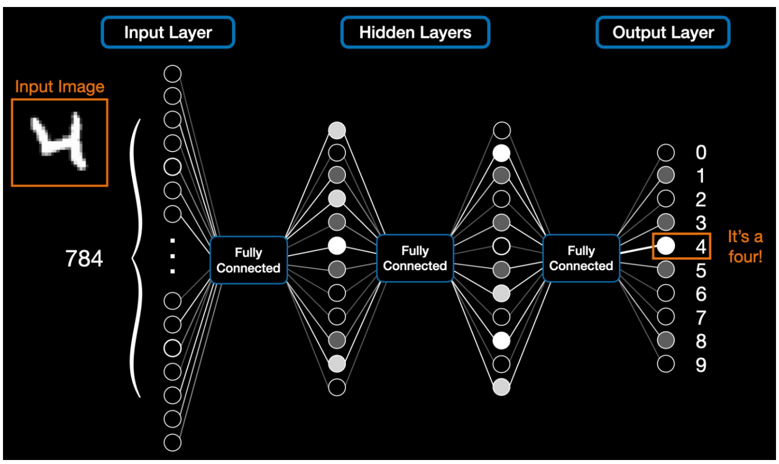
Inteligencia Artificial desarrollado por el equipo de DeepMind, basado en redes neuronales profundas y aprendizaje por refuerzo.

Este algoritmo fue capaz de capturar el aspecto intuitivo del juego que dicen utilizar los grandes maestros del Go mediante el aprendizaje por refuerzo, un algoritmo que se basa en tomar acciones en el Go para maximizar alguna noción de recompensa acumulada. En todos estos juegos la Inteligencia Artificial recibió las reglas, aprendió jugando primero contra aficionados, luego expertos y, finalmente, contra sí misma para terminar siendo imbatible. Una interesante versión novelada de las partidas de AlphaGo puede leerse en la última parte del libro "Maniac" de Benjamin Labatut [39].

5.3. Inteligencia Artificial. Aprendizaje automático. Redes neuronales.

Casi todo el progreso de los últimos 20 años en Inteligencia Artificial ha estado relacionado con el aprendizaje automático (*machine learning*). De modo general, se trata de una arquitectura computacional que es entrenada con una colección de datos y encuentra patrones en los datos que pueden ser utilizados para cierta tarea con un alto grado de precisión.

Una *red neuronal* es un modelo computacional compuesto por capas de nodos interconectados (neuronas artificiales) que trabajan juntas para aprender patrones y relaciones en los datos de entrada. Cada neurona artificial recibe señales de entrada, procesa esas señales y luego produce una señal de salida que se transmite a otras neuronas a través de conexiones llamadas pesos. El *aprendizaje profundo (deep learning)* hace referencia a las muchas capas. Se trata de un nuevo paradigma. En la programación clásica se daban reglas y datos para obtener respuestas. En el aprendizaje profundo se dan datos y respuestas para que se obtengan las reglas.



En los últimos 5 o 6 años la tecnología dominante han sido los modelos de lenguaje de gran tamaño (*Large Language Models LLM*). Son el resultado de la aplicación de las redes neuronales a las tareas orientadas al lenguaje entrenadas con grandes cantidades de texto. El usuario de estos modelos introduce un texto y la tarea del modelo de lenguaje es generar la continuación más probable del texto. La respuesta se genera palabra a palabra. Chatbots como ChatGPT tienen mecanismos para establecer diálogos. En ocasiones el modelo de lenguaje genera texto que resulta incoherente o falso, las llamadas "alucinaciones". El hecho de que estos modelos de lenguaje generen una palabra cada vez hace que no puedan plantear de antemano lo que van a decir como un todo y que no puedan ir hacia atrás y corregir errores. Es peligroso confiar en el texto producido sin comprobar su veracidad. En el artículo de Davis [15] se presentan errores de GPT-4 en pruebas matemáticas.

Vamos a describir brevemente lo que es el aprendizaje profundo (*deep learning*) para un matemático: Tenemos dos espacios vectoriales reales $V = \mathbb{R}^m$ y $W = \mathbb{R}^n$, con n y m grandes, y un subconjunto S de V en el que está definida una función no lineal $f : S \to W$ desconocida. Tenemos una tabla finita de pares $(s, f(s))$ que consiste en muchos elementos $s \in S$ y de sus imágenes por f. Esa tabla puede haber sido muy costosa de conseguir. Lo que se pretende es hacer un programa de ordenador que calcule f con precisión y nos permita evaluar $f(s)$ para muchos más elementos de S y con alta precisión. Una red neuronal después de un periodo de "entrenamiento" produce una función f que permite calcular valores numéricos $f(s) \in W$ de elementos de $s \in V$. En el ejemplo del aprendizaje profundo para la clasificación de imágenes, cada imagen se representa como una matriz 28×28 de valores de grises entre 0 y 255. Si se trata, por ejemplo, de clasificar dígitos escritos a mano, queremos aproximar una función $f : \mathbb{R}^{28\times 28} \to \mathbb{R}^{10}$ que otorgue probabilidades de que cierta imagen con 28×28 píxeles en escala de grises asocie uno de los dígitos $0, 1, \ldots, 9$.

La Inteligencia Artificial ha hecho progresos extraordinarios en los últimos años en muchas áreas, como en los juegos, la generación de imágenes, la transcripción del habla y la traducción entre idiomas. Las aplicaciones del aprendizaje profundo a la ciencia, la tecnología y la industria han sido innumerables y profundas y su impacto en matemática aplicada han sido enormes. Sin embargo, el impacto en matemática pura ha sido de momento modesto. Puede que una de las razones sea que la habilidad de razonar sigue siendo uno de los problemas centrales no resueltos de la Inteligencia Artificial. Se sabe poco de creatividad e intuición.

En cualquier caso, las redes neuronales han ayudado realmente a demostrar nuevos teoremas en teoría de nudos, teoría de representación en la conjetura de la invarianza combinatoria y en la búsqueda de contraejemplos en teoría de grafos, tal como se menciona en el artículo de Buzzard [7]. En el primer caso, se usó software libre para generar una gran tabla de invariantes de nudos, y se utilizó la red neuronal para tratar de encontrar relaciones desconocidas entre estos invariantes, ello permitió a los matemáticos Davies, Lackenby, Juhasz y Tomasev formular y probar una conjetura. Véase también el artículo [13] de Davies y coautores. E. Davis opina al respecto en [14] que el papel del aprendizaje profundo en este caso ha sido sobrevalorado y que un análisis estadístico convencional podría haber bastado. Sin embargo, en el caso de la conjetura de la invarianza combinatoria, Davis [14] reconoce que el aprendizaje profundo juega un papel central, pero que debe considerarse como “otra herramienta analítica de matemáticas experimentales y no como una nueva aproximación fundamental a las matemáticas”. Una descripción más detallada del papel del aprendizaje profundo en relación con esta conjetura se encuentra en la Sección 6.2 del artículo de Williamson [60], que también menciona ejemplos de aplicaciones de las redes neuronales para generar contraejemplos en combinatoria y para simplificar pruebas acerca de la clasificación de semigrupos nilpotentes, entre otras.

En 2020 Raayoni y otros [49] usaron un programa llamado Ramanujan Machine para generar fórmulas, algunas conocidas y otras nuevas, acerca de las constantes fundamentales en matemáticas, como e y π. De este modo se generaron conjeturas, algunas de las cuales permanecen abiertas. Aplicaciones de clasificadores de Inteligencia Artificial a curvas elípticas fueron obtenidas por He, Lee, Oliver y Pozdnyakov [32]. También referimos al artículo panorámico de Gómez-Serrano [21] para resultados recientes acerca de demostraciones asistidas por ordenador en ecuaciones en derivadas parciales y, en particular, en problemas de fluidos incompresibles. Ver también [8].

A la vista de todos estos ejemplos, y de los avances en el cap set [52] mencionados en la Sección 2, es claro que la aplicabilidad de estas herramientas es indudable, aunque parece estar limitada de momento a áreas de matemáticas en las que la computación es posible y en las que existen tablas, tal como considera Buzzard [7].

Existen varias obstrucciones para trasladar los éxitos de la Inteligencia Artificial en juegos a las matemáticas: en primer lugar, es difícil dar un valor a una prueba parcial de un teorema matemático. En segundo lugar, en cada paso de una demostración hay, en principio, infinitos caminos posibles, incluidos miles y miles de lemas con muchos parámetros, por ello no es fácil “guiar” al programa hacia la prueba correcta. Sin embargo, la posibilidad de combinar el aprendizaje profundo con los verificadores interactivos de teoremas es muy prometedora.

Granville [25] expresa las siguientes reticencias a la influencia del aprendizaje profundo: “El aprendizaje automático típicamente desarrolla una aproximación a problemas específicos a través de algoritmos simples construidos hábilmente usando datos. Crear una amplia base de datos y analizarla con herramientas específicamente formuladas puede ser muy efectivo (como Google Translate o ChatGPT), pero eso no es lo mismo que desarrollar intuición. Hay una gran cantidad de dinero y propaganda alrededor del aprendizaje profundo y otras formas de Inteligencia Artificial, por ello no es sorprendente que muchos avances sean exagerados o sean fácilmente explicados en términos de algoritmos bien diseñados y potencia computacional extraordinaria.” No obstante, Williamson comenta en la conclusión de su artículo [60]: “El uso del aprendizaje profundo en matemática pura está en su infancia. Las herramientas de aprendizaje automático son flexibles y poderosas, pero necesitan conocimientos previos y experiencia. Las aplicaciones aparecidas hasta ahora apoyan la idea de que son más útiles en las áreas más intuitivas: descubrir patrones, decidir si puede haber contraejemplos o elegir qué partes de un cálculo hacer a continuación. Sin embargo, las posibilidades son infinitas".

Ochigame [45] propone una interesante reflexión de los efectos en autoría, créditos y prioridad del uso de demostradores interactivos de teoremas y de herramientas de Inteligencia Artificial en la investigación matemática.

La Inteligencia Artificial ha realizado progresos impresionantes en algunos problemas matemáticos, tal como hemos comentado. A pesar de ello, es difícil pensar que puedan reemplazar a los matemáticos completamente. Algunas posibles razones son que la Inteligencia Artificial:

(1) Está limitada de momento por los datos y algoritmos que la entrenan, no inventa conceptos o métodos que no existían en los datos del entrenamiento. Por otra parte, los matemáticos usan su intuición, creatividad e imaginación para explorar nuevas áreas de matemáticas.

(2) No es muy buena explicando o entendiendo sus propias soluciones. Puede generar y verificar funciones matemáticas, pero no puede explicar las intuiciones detrás de ellas. Los matemáticos pueden comunicar y justificar sus ideas y construir sobre las aportaciones de otros matemáticos.

(3) Tiene dificultades para adaptarse a diferentes contextos y puede producir errores en situaciones imprevistas, mientras que los matemáticos pueden enfrentarse a la incertidumbre y la complejidad.

La Inteligencia Artificial no reemplaza de momento a los matemáticos, pero los complementa: les ayuda a realizar cálculos tediosos, encontrar nuevas soluciones y sugerir nuevas direcciones de investigación. Por otra parte, los matemáticos pueden ayudar a la Inteligencia Artificial proporcionándole interpretación y generalización. Juntos harán avanzar las fronteras de las matemáticas y la ciencia. La muy interesante entrevista de Drösser a Tao [17] en junio de 2024 presenta muchas aspectos de la fructífera interacción actual y futura de los matemáticos y la Inteligencia Artificial.

6. Conclusión.

En el resumen de su artículo [7] Kevin Buzzard escribía: "Los ordenadores han cambiado la manera como los matemáticos hacemos matemáticas, nos han permitido una gran eficiencia computacional. ¿Nos ayudarán a razonar? ¿Serán capaces de razonar ellos mismos en un futuro próximo?”

¿Cambiarán los ordenadores las matemáticas? Sí. Tanto como lo han hecho hasta ahora, pero aún más, porque amplificarán la habilidad de los matemáticos para asomarse a lo desconocido, permitiéndoles crear y explorar nuevas áreas matemáticas. Las nuevas tecnologías tienen mucho que ofrecer a los matemáticos: ayudan a verificar la corrección de los argumentos, permiten construir librerías amplias y utilizables, exploran conceptos, generan conjeturas, calculan e incluso descubren nuevos teoremas. Para que su impacto en la investigación matemática sea mayor es necesario que las demostraciones asistidas por ordenador puedan ser fácilmente accesibles para los matemáticos, ya que la comprensión, interiorización y comunicación de las pruebas son esenciales. En este sentido los modelos de lenguaje de gran tamaño pueden ayudar.

Los ordenadores no son capaces por el momento de resolver problemas difíciles y profundos de matemáticas, pero pueden ayudar en muchas tareas creativas, como la generación de conjeturas o el descubrimiento de fenómenos inesperados. Son una herramienta muy útil en la verificación de ideas y resultados que no son fáciles de analizar. Pero es difícil de creer que puedan demostrar nuevos teoremas correctos.

Es muy complicado predecir el futuro de las demostraciones matemáticas. Como hemos intentado describir en este discurso, estamos en el inicio de un cambio profundo en la forma como los matemáticos trabajan en problemas en la frontera del conocimiento. Tenemos que estar preparados para una creciente importancia de los ordenadores en la creación matemática y para convivir con distintos tipos de demostraciones con diferentes grados de certeza.

En el futuro que nos espera como investigadores matemáticos y en nuestra relación con el ordenador, no debemos olvidar lo que hemos aprendido acerca de la experiencia matemática en el pasado: la importancia de trabajar juntos desde diferentes perspectivas, la necesidad de que diferentes comunidades matemáticas trabajen conjuntamente para entender y mejorar las demostraciones y contribuir de esa forma al progreso de la ciencia, la tecnología y las matemáticas.

Muchas gracias por su atención.

Bibliografía.

[1] M. Aigner y G. M. Ziegler, *Proofs from The Book, 6th ed.*, Springer, Berlin, 2018. DOI: `10.1007/978-3-662-57265-8`

[2] J. Avigad, *Mathematics and the formal turn*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 225–240. DOI: `10.1090/bull/1832`

[3] F. Bombal, *Una mirada a las matemáticas del siglo XX*, Publicaciones de la Real Academia de Ciencias Exactas, Físicas y Naturales de España, Madrid, 2015, Lección inaugural del año académico 2015-2016. URL: `https://rac.es/ficheros/doc/01155.pdf`

[4] J. Brakensiek, M. Heule, J. Mackey, D. Narváez, *The resolution of Keller's conjecture*, J. Automat. Reason. **66** (2022), no. 3, 277–300. DOI: `10.1007/s10817-022-09623-5`

[5] A. Bundy, *A very mathematical dilemma*, The Computer Journal **49** (2006), no. 4, 480–486. DOI: `10.1093/comjnl/bxl021`

[6] K. Buzzard, *Proving theorems with computers*, Notices Amer. Math. Soc. **67** (2020), no. 11, 1791–1799. DOI: `10.1090/noti`

[7] K. Buzzard, *Mathematical reasoning and the computer*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 211–224. DOI: `10.1090/bull/1833`

[8] J. Cepelewicz, *Deep learning poised to 'blow up' famed fluid equations*, Quanta (2022). URL: `https://www.quantamagazine.org/deep-learning-poised-to-blow-up-famed-fluid-equations-20220412/`

[9] J. Commelin y A. Topaz, *Abstraction boundaries and spec driven development in pure mathematics*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 241–255. DOI: `10.1090/bull/1831`

[10] E. Croot, V. F. Lev y P. P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337. DOI: `10.4007/annals.2017.185.1.7`

[11] S. R. Dahmen, J. Hölzl y R. Y. Lewis, *Formalizing the solution to the cap set problem*, 10th International Conference on Interactive Theorem Proving, LIPIcs. Leibniz Int. Proc. Inform., vol. 141, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019, pp. Art. No. 15, 19.

[12] B. I. Dahn, *Robbins algebras are Boolean: a revision of McCune's computer-generated solution of Robbins problem*, J. Algebra **208** (1998), no. 2, 526–532. DOI: `10.1006/jabr.1998.7467`

[13] A. Davies et al., *Advancing mathematics by guiding human intuition with AI*, Nature **600** (2021), 70–74. DOI: 10.1038/s41586-021-04086-x

[14] E. Davis, *Deep learning and mathematical intuition: A review of (Davies et al. 2021)*, arXiv (2021). DOI: 10.48550/arXiv.2112.04324

[15] E. Davis, *Mathematics, word problems, common sense, and artificial intelligence*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 287–303. DOI: 10.1090/bull/1828

[16] P. J. Davis y R. Hersh, *The mathematical experience*, Birkhäuser, Boston, MA, 1980.

[17] C. Drösser, *AI will become mathematicians' co-pilot. Interview of T. Tao* (2024) URL: https://www.scientificamerican.com/article/ai-will-become-mathematicians-co-pilot/

[18] J. S. Ellenberg y D. Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343. DOI: 10.4007/annals.2017.185.1.8

[19] A. Fawzi, M. Balog, A. Huang et al., *Discovering faster matrix multiplication algorithms with reinforcement learning*, Nature **610** (2022), 47–53. DOI: 10.1038/s41586-022-05172-4

[20] M. Ganesalingam y W. T. Gowers, *A fully automatic problem solver with human-style output*, arXiv (2013). DOI: 10.48550/arXiv.1309.4501

[21] J. Gómez-Serrano, *Computer-assisted proofs in PDE: a survey*, arXiv (2018). DOI: 10.48550/arXiv.1810.00745

[22] G. Gonthier, *Formal proof—the four-color theorem*, Notices Amer. Math. Soc. **55** (2008), no. 11, 1382–1393. URL: https://www.ams.org/notices/200811/tx081101382p.pdf

[23] T. Gowers, *What makes mathematicians believe unproved mathematical statements?*, Annals Math. Phil. **1** (2023), 57–110. URL: https://mxphi.com/wp-content/uploads/2023/04/MxPhi-Gowers2023.pdf

[24] A. Granville, *Why mathematical proof is a social compact?*, Quanta Magazine (2023). URL: https://www.quantamagazine.org/why-mathematical-proof-is-a-social-compact-20230831

[25] A. Granville, *Proof in the time of machines*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 317–329. DOI: 10.1090/bull/1826

[26] B. Green y T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547. DOI: 10.4007/annals.2008.167.481

[27] J. A. Grochow, *New applications of the polynomial method: the cap set conjecture and beyond*, Bull. Amer. Math. Soc. (N.S.) **56** (2019), no. 1, 29–64. DOI: 10.1090/bull/1648

[28] J. Hadamard, *Psicología de la invención en el campo matemático*, Espasa-Calpe, 1947. URL: https://archive.org/details/jacques-hadamard-psicologia-de-la-invencion-en-el-campo-matematico

[29] T. Hales et al., *A formal proof of the Kepler conjecture*, Forum Math. Pi **5** (2017), e2, 29. DOI: 10.1017/fmp.2017.1

[30] P. R. Halmos, *Has progress in mathematics slowed down?*, Amer. Math. Monthly **97** (1990), no. 7, 561–588. DOI: 10.2307/2324635

[31] Y.-H. He, *Machine learning in pure mathematics and theoretical physics*, Worl Scientific, 2023. DOI: 10.1142/q0404

[32] Y.-H. He, K.-H. Lee, T. Oliver y A. Pozdnyakov, *Murmurations of elliptic curves*, arXiv (2022). DOI: 10.48550/arXiv.2204.10140

[33] R. Hersh, *What is mathematics, really?*, Oxford University Press, New York, 1997.

[34] R. Hersh, *Experiencing mathematics*, American Mathematical Society, Providence, RI, 2014. DOI: 10.1090/mbk/083

[35] R. Hersh y V. John-Steiner, *Loving and hating mathematics*, Princeton Univ. Press, 2011.

[36] M. J. H. Heule, O. Kullmann y V. W. Marek, *Solving and verifying the Boolean Pythagorean triples problem via cube-and-conquer*, Theory and applications of satisfiability testing—SAT 2016, Lecture Notes in Comput. Sci., vol. 9710, Springer, 2016, pp. 228–245. DOI: 10.1007/978-3-319-40970-2_15

[37] R. Ibáñez, *Las matemáticas del juego de cartas SET (1)*. URL: https://culturacientifica.com/2016/06/15/matematicas-juego-cartas-set-1

[38] R. Ibáñez, *Las matemáticas del juego de cartas SET (2)*. URL: https://culturacientifica.com/2016/06/15/matematicas-juego-cartas-set-2

[39] B. Labatut, *Maniac*, Anagrama, 2023.

[40] I. Lakatos, *Pruebas y refutaciones*, Alianza Universidad, 1978.

[41] M. López Pellicer, *Alrededor de la hipótesis de Riemann*, Publicaciones de la Real Academia de Ciencias Exactas, Físicas y Naturales de España, Madrid, 2012, Lección inaugural del año académico 2012-2013. URL: `https://rac.es/ficheros/doc/01220.pdf`

[42] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413. DOI: `10.4007/annals.2015.181.1.7`

[43] B. Mazur, *Is it plausible?*, Math. Intelligencer **36** (2014), 24–33. DOI: `10.1007/s00283-013-9398-0`

[44] G. McGuire, B. Tugemann y G. Civario, *There is no 16-clue Sudoku: solving the Sudoku minimum number of clues problem via hitting set enumeration*, Exp. Math. **23** (2014), no. 2, 190–217. DOI: `10.1080/10586458.2013.870056`

[45] R. Ochigame, *Automated mathematics and the reconfiguration of proof and labor*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 3, 423–437. DOI: `10.1090/bull/1821`

[46] K. Parshina, *Philosophical assumptions behind the rejection of computer-based proofs*, Kriterion - Journal of Philosophy **37** (2023), no. 2-4, 105–122. DOI: `10.1515/krt-2022-0015`

[47] G. Pellegrino, *Sul massimo ordine delle calotte in $S_{4,3}$*, Matematiche (Catania) **25** (1970), 149–157. DOI: `10.1016/S0304-0208(08)73322-X`

[48] D. Pérez-García, *Los juegos no locales: un nexo entre las matemáticas, la física y las ciencias de la computación*, Publicaciones de la Real Academia de Ciencias Exactas, Físicas y Naturales de España, Madrid, 2023, Discurso leído en el acto de su recepción como académico de número. URL: `https://rac.es/ficheros/doc/c01f35fffcf627ed.pdf`

[49] G. Raayoni et al., *Generating conjectures on fundamental constants with the Ramanujan Machine*, Nature **590** (2021), 67–73. DOI: `10.1038/s41586-021-03229-4`

[50] S. Ramón y Cajal, *Mi infancia y juventud*, Espasa Calpe Colección Austral, 1980.

[51] N. Robertson, D. P. Sanders, P. Seymour y R. Thomas, *A new proof of the four-colour theorem*, Electron. Res. Announc. Amer. Math. Soc. **2** (1996), no. 1, 17–25. DOI: `10.1090/S1079-6762-96-00003-0`

[52] B. Romera-Paredes, M. Barekatain, A. Novikov et al., *Mathematical discoveries from program search with large language models*, Nature **625** (2024), 468–475. DOI: `10.1038/s41586-023-06924-6`

[53] G.-C. Rota, *Indiscrete thoughts*, Birkhäuser Boston, Inc., Boston, MA, 2008, Reprint of the 1997 edition. DOI: `10.1007/978-0-8176-4781-0`

[54] C. Rousseau y Y. Saint-Aubin, *Mathematics and technology*, Springer New York, NY, 2008. DOI: 10.1007/978-0-387-69216-6

[55] P. Scholze, *Liquid tensor experiment*, Exp. Math. **31** (2022), no. 2, 349–354. DOI: 10.1080/10586458.2021.1926016

[56] M. Shulman, *Strange new universes: proof assistants and synthetic foundations*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 257–270. DOI: 10.1090/bull/1830

[57] I. Stewart, *¿Para qué sirven las matemáticas?*, Editorial Crítica, 2022.

[58] T. Tao, *Machine assisted proof*, AMS Colloquium Lectures at the 2024 Joint Mathematics Meetings in San Francisco. URL: https://www.youtube.com/watch?v=AayZuuDDKP0

[59] A. Venkatesh, *Some thoughts on automation and mathematical research*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 203–210. DOI: 10.1090/bull/1834

[60] G. Williamson, *Is deep learning a useful tool for the pure mathematician?*, Bull. Amer. Math. Soc. (N.S.) **61** (2024), no. 2, 271–286. DOI: 10.1090/bull/1829