

La Ciencia para Todos

Desde el nacimiento de la colección de divulgación científica del Fondo de Cultura Económica en 1986, ésta ha mantenido un ritmo siempre ascendente que ha superado las aspiraciones de las personas e instituciones que la hicieron posible. Los científicos siempre han aportado material, con lo que han sumado a su trabajo la incursión en un campo nuevo: escribir de modo que los temas más complejos y casi inaccesibles puedan ser entendidos por los estudiantes y los lectores sin formación científica.

A los diez años de este fructífero trabajo se dio un paso adelante, que consistió en abrir la colección a los creadores de la ciencia que se piensa y crea en todos los ámbitos de la lengua española —y ahora también del portugués—, razón por la cual tomó el nombre de La Ciencia para Todos.

Del Río Bravo al Cabo de Hornos y, a través de la mar Océano, a la Península Ibérica, está en marcha un ejército integrado por un vasto número de investigadores, científicos y técnicos, que extienden sus actividades por todos los campos de la ciencia moderna, la cual se encuentra en plena revolución y continuamente va cambiando nuestra forma de pensar y observar cuanto nos rodea.

La internacionalización de La Ciencia para Todos no es sólo en extensión sino en profundidad. Es necesario pensar una ciencia en nuestros idiomas que, de acuerdo con nuestra tradición humanista, crezca sin olvidar al hombre, que es, en última instancia, su fin. Y, en consecuencia, su propósito principal es poner el pensamiento científico en manos de nuestros jóvenes, quienes, al llegar su turno, crearán una ciencia que, sin desdeñar a ninguna otra, lleve la impronta de nuestros pueblos.

ÁLGEBRA EN TODAS PARTES

El álgebra es una rama de las matemáticas que estudia las relaciones entre los números y las letras. Se utiliza para resolver problemas que involucran cantidades desconocidas. El álgebra es una herramienta poderosa que nos ayuda a entender el mundo que nos rodea.

El álgebra se divide en varias ramas, como el álgebra elemental, el álgebra lineal, el álgebra matricial, el álgebra abstracta, etc. Cada rama tiene sus propias reglas y métodos para resolver problemas.

El álgebra es una parte fundamental de las matemáticas y se utiliza en muchas áreas de la ciencia y la tecnología. Sin el álgebra, no podríamos entender muchos de los fenómenos que nos rodean.

El álgebra es una herramienta poderosa que nos ayuda a resolver problemas que involucran cantidades desconocidas. Se utiliza para encontrar soluciones a problemas que no se pueden resolver con los números solos.

El álgebra es una rama de las matemáticas que estudia las relaciones entre los números y las letras. Se utiliza para resolver problemas que involucran cantidades desconocidas.

José Antonio de la Peña

Comité de Selección

Dr. Antonio Alonso
Dr. Francisco Bolívar Zapata
Dr. Javier Bracho
Dr. Juan Luis Cifuentes
Dra. Rosalinda Contreras
Dr. Jorge Flores Valdés
Dr. Juan Ramón de la Fuente
Dr. Leopoldo García-Colín Scherer
Dr. Adolfo Guzmán Arenas
Dr. Gonzalo Halffter
Dr. Jaime Martuscelli
Dra. Isaura Meza
Dr. José Luis Morán López
Dr. Héctor Nava Jaimes
Dr. Manuel Peimbert
Dr. José Antonio de la Peña
Dr. Ruy Pérez Tamayo
Dr. Julio Rubio Oca
Dr. José Sarukhán
Dr. Guillermo Soberón
Dr. Elías Trabulse

Coordinadora

María del Carmen Farías R.

ÁLGEBRA EN TODAS PARTES



la
ciencia/166
para todos

Primera edición, 1999
Sexta reimpresión, 2010

Peña, José Antonio de la
Álgebra en todas partes / José Antonio de la Peña —
México : FCE, SEP, CONACyT, 1999
196 p. : ilus. ; 21 x 14 cm — (Colec. La Ciencia para
Todos ; 166)
Texto para nivel medio superior
ISBN 978-968-16-6052-9

1. Matemáticas 2. Álgebra 3. Divulgación científica I. Ser.
II. t.

LC QA155

Dewey 508.2 C569 V.166

A NELIA

Distribución mundial

Comentarios y sugerencias: laciencia@fondodeculturaeconomica.com
www.fondodeculturaeconomica.com
Tel. (55)5227-4672 Fax (55)5227-4664

 Empresa certificada ISO 9001: 2000

Diseño de portada: Laura Esponda Aguilar / León Muñoz Santini

La Ciencia para Todos es proyecto y propiedad del Fondo de Cultura Económica,
al que pertenecen también sus derechos. Se publica con los auspicios de la
Secretaría de Educación Pública y del Consejo Nacional de Ciencia y Tecnología.

D. R. © 1999, FONDO DE CULTURA ECONÓMICA
Carretera Picacho-Ajusco, 227; 14738 México, D. F.

Se prohíbe la reproducción total o parcial de esta obra
—incluido el diseño tipográfico y de portada—,
sea cual fuere el medio, electrónico o mecánico,
sin el consentimiento por escrito del editor.

ISBN 978-968-16-6052-9

Impreso en México • Printed in Mexico

El tiempo ha llegado, dijo la Morsa,
de hablar de muchas cosas,
de zapatos, de barcos y sobres lacrados,
de coles y reyes.

Alicia a través del espejo, LEWIS CARROLL

Ver el mundo en un grano de arena
y el cielo en una flor silvestre,
Asir el infinito en la palma de tu mano
y la eternidad en una hora.

Augurios de inocencia, WILLIAM BLAKE

INTRODUCCIÓN

El hombre es mortal por sus temores
e inmortal por sus deseos.

PITÁGORAS

Comenzaré refiriendo mis temores. Escribir un libro de matemáticas para un público amplio es un reto muy atractivo del que no parece fácil salir bien librado. La sensación es parecida a la de tocar sonatas de Bach ante un público que creía que iba a escuchar un concierto de rock. O a la de ir a dar una conferencia de divulgación de matemáticas a una preparatoria para descubrir, cuando ya está uno ante un auditorio de 100 personas, que por equivocación anunciaron que la conferencia tendría como tema la sexología. Esto último me pasó hace años en una preparatoria en Puebla. Cuando aclaré al auditorio el error muchos se salieron, pero los que se quedaron, y fueron bastantes, no la pasaron mal.

Sé que la mayor parte del auditorio que tengo enfrente piensa que las matemáticas son feas, frías, aburridas y difíciles. Yo y muchos otros matemáticos sabemos que esto no es cierto: las matemáticas son bellas, cálidas, apasionantes y no siempre difíciles (de hecho, a veces, cuando entiende uno bien las cosas pueden ser claras y sencillas). Este libro constituye un intento de convencer al lector de que las matemáticas pueden ser así. La única manera que tengo de hacerlo es mostrándole algunas de las cosas que me gustan.

En este libro vamos a hablar de *álgebra*. Bueno, de algunos temas de álgebra. Algunos serán familiares para el lector, como el teorema de Pitágoras o las ecuaciones cuadráticas. Otros serán nuevos. Entre otras cosas, el lector encontrará aplicaciones de la teoría de matrices para la predicción de resultados de partidos de baloncesto; encontrará un estudio acerca de la teoría de los autómatas y sus lenguajes y su relación con la reciente derrota de Kasparov "a manos" de una computadora; encontrará una explicación de la demostración del último teorema de Fermat, sobre la que escribieron todos los periódicos del mundo en 1994.

Hay dos cualidades en particular de las matemáticas que queremos mostrar al lector:

Las matemáticas son útiles. Encontramos a las matemáticas en la solución de problemas muy variados. Desde problemas de conteo, hasta problemas físicos, químicos, ornamentales, deportivos y otros. En todos estos campos, el común denominador es la eficiencia con que funciona la maquinaria matemática.

Las matemáticas son una ciencia viva. Por alguna extraña razón se tiene la falsa idea de que las matemáticas forman un cuerpo de conocimiento completo, escrito en libros sólo comprensibles para algunos iniciados: los matemáticos. A este respecto puedo referir una anécdota. Una vez iba a entrar en un país en el que se requería visa. Me preguntaron en la Oficina de Migración:

—¿Cuál es su ocupación?

—Investigación en matemáticas.

—¿En qué área?

—Álgebra.

—¡Imposible! Si me dijera que en cálculo, tal vez lo creería. Pero en álgebra ya se sabe todo.

Por supuesto, esto no es así. El ejemplo reciente más famoso es la solución que dio en 1994 Andrew Wiles a un problema que fue planteado a principios del siglo XVII. Pero menos famosos que este problema, hay miles de problemas en los que los matemáticos trabajaron el día de hoy.

Hablemos ahora un poco sobre el contenido y la organización de este libro. Nuestra presentación de los temas será más o menos cronológica, de manera que a lo largo de los capítulos toca-

remos temas de: *aritmética, teoría de números, teoría de matrices, álgebra abstracta y finalmente álgebra y computación*. Los capítulos son independientes entre sí, aunque en ocasiones se supone que alguna notación o algunas ideas ya son conocidas. Los primeros cuatro capítulos tratan problemas matemáticos cuyo planteamiento data de siglos atrás; éstos son generalmente numéricos (por ejemplo, problemas de conteo y de solución de ecuaciones). Los siguientes cuatro capítulos tratan lo que generalmente se llama *álgebra moderna*, es decir, la que se ha desarrollado de finales del siglo pasado hasta nuestros días. Los problemas que se estudian en estos capítulos comprenden estructuras algebraicas abstractas (por ejemplo, grupos y matrices). En el último capítulo damos algunos detalles biográficos de los matemáticos más importantes que tratamos a lo largo del libro. Al final de algunos capítulos el lector encontrará problemas. Algunos aparecen con soluciones completas, otros no. El libro cierra con una sección de referencias comentadas que debe servir al lector interesado para obtener otras lecturas que le permitan profundizar en temas que hayan despertado su curiosidad.

Algunas partes del libro se pueden leer como un recuento histórico y anecdótico salpicado con unas pocas ideas matemáticas. Otras son más difíciles. El lector no debe desanimarse si no entiende algunas cosas. Si durante una primera lectura encuentra partes demasiado difíciles, puede saltárselas y pasar al siguiente párrafo, sección o capítulo. Tal vez en una segunda lectura las cosas que fueron difíciles la primera vez comiencen a aclararse. Como ayuda para que el lector sepa cuáles son las secciones más difíciles, las hemos marcado con un asterisco, así: (*).

Las verdades matemáticas se llaman *teoremas*. Para llegar a establecer la validez de un teorema se requiere una demostración rigurosa que sea aceptada por cualquier matemático en cualquier lugar del mundo. Por ello, la única manera de tener una impresión del trabajo matemático es ver algunos teoremas y los pasos del razonamiento que llevan a establecer sus demostraciones. A lo largo del libro el lector encontrará muestras de teoremas, algunos con demostraciones, otros sin ellas. El lector puede evitar la lectura de las demostraciones de algunos teoremas sin que esto afecte su comprensión del texto. Sin embargo,

el verdadero espíritu de las matemáticas se halla en los razonamientos que llevan a la prueba de los teoremas; la claridad de estos razonamientos es también la fuente de la belleza matemática. El lector que sienta curiosidad por la demostración de algún resultado que sólo se menciona en este libro, encontrará recomendaciones de otros libros de divulgación y textos matemáticos en las referencias comentadas al final del libro.

Antes de terminar esta introducción, quiero insistir en algunas de las cosas que este libro no es. *Este libro no es un libro de texto*. Es decir, no se trata de que el lector aprenda muchas cosas, que entienda todo y que al final se le haga un examen. Tómese este libro en forma relajada y sin formalidades. Pero, por otra parte, *este libro no es una novela*. Es decir, tampoco se puede leer a ratos, sin concentración y sin prestar mucha atención. Si así se hace, poco se entenderá. Leer este libro exigirá algo del lector. Pretende ser un libro de matemáticas. Si fuera un libro *sobre* matemáticas, hubiera sido más fácil hacerlo y el lector lo podría tomar como una novela. Por lo tanto, el lector debe estar listo para *hacer matemáticas*. Para ello, debe tener lápiz y papel a mano, repetir en el papel lo que se haga en el libro, rehacer y completar sus temas y *pensar*.

Terminaré hablando de mis descos: que después de leer este libro el lector haya entendido algunas cosas y que le hayan parecido interesantes, bellas o divertidas, al menos algunas de ellas. Por supuesto, esto puede ser demasiado ambicioso. Veremos.

Finalmente, quiero agradecer a mi hermano Ricardo, a Michael Barot y a Carlos Daniel Amero por la cuidadosa lectura del texto y sus comentarios. También a Ruth y a Walter que leyeron algunos capítulos; a Elías Vigueira y Omar Guerrero por su trabajo fotográfico y a Gabriela Sanginés que me ayudó con el formato final. Las ilustraciones de los personajes que mencionamos las podrán ver en los dibujos que yo realicé, pues no pudieron hacerse con fotografías por problemas técnicos. Por supuesto, agradezco al Instituto de Matemáticas de la UNAM por el rico ambiente matemático donde trabajo cotidianamente y en particular a mis colegas y estudiantes del grupo de Representaciones de Álgebras.

Al llegar al final del proceso de elaboración de este libro descubro que aún hay muchas ideas que me gustaría desarrollar,

modificaciones que me gustaría hacer. Pero en algún momento tiene uno que concluir.

México, noviembre de 1997.

Un hombre se propone la tarea de dibujar el mundo. A lo largo de los años puebla un espacio con imágenes de provincias, de reinos, de montañas, de bahías, de naves, de islas, de peces, de habitaciones, de instrumentos, de astros, de caballos y de personas. Poco antes de morir, descubre que ese paciente laberinto de líneas traza la imagen de su cara.

El hacedor, JORGE LUIS BORGES

I. De los dedos de las manos a las computadoras

Dios hizo los números naturales,
lo demás es creación de los hombres.

GIUSEPPE PEANO

LAS dos primeras preguntas que hace un adulto a un niño que se encuentra por primera vez son: ¿Cómo te llamas? y ¿cuántos años tienes?

Generalmente, si el niño puede contestar la primera pregunta también podrá contestar la segunda, aunque sea indicando la respuesta con los dedos.

En efecto, el primer contacto de un niño con las matemáticas se da muy pronto en su vida. El pequeño aprende su edad y a contar algunos de los objetos que le rodean. Al menos hasta el diez, el número de dedos de las manos.

Los primeros hombres, como todavía lo hacen algunos pueblos primitivos, sólo necesitaban números pequeños y los formaban con los dedos de la mano. A medida que la sociedad fue evolucionando, hubo que hacer cálculos más complicados. Lo primero que se tuvo que hacer es encontrar la forma de indicar números mayores que diez. Por ejemplo, usando los dedos y otras partes del cuerpo, los miembros de la tribu sibiller de Nueva Guinea, cuentan hasta el 27. En la figura I.1 se ve un



Figura I.1. Niño imitando la forma de contar en la tribu sibiller.

niño que imita la manera de contar de esta tribu y utiliza su índice derecho para señalar los dedos de la mano izquierda para contar del 1 al 5. Después usa su muñeca izquierda, antebrazo, codo, bíceps, clavícula, hombro, oreja y ojo para contar del 6 al 13. La nariz es el 14, luego señalando con el índice izquierdo baja del ojo hasta el menique para los números del 15 al 27.

Un pueblo tan avanzado como el de los romanos tenía un sistema de numeración bastante primitivo y poco práctico. Todos conocemos los números romanos I, II, III, IV, V... Para convencerse de lo poco práctico de este sistema de numeración basta tratar de efectuar una suma como $XLI + XCIX$. En la figura I.2 se puede comparar cómo se escriben los primeros numerales en diferentes culturas.

V	vv	vvv	vvv	vvv	vvv	vvv	vvv	vvv	<
Numerales babilonios									
1	2	3	4	5	6	7	8	9	10
Numerales egipcios									
1	2	3	4	5	6	7	8	9	10
Numeros mayas									
1	2	3	4	5	6	7	8	9	10
Numerales chinos									
1	2	3	4	5	6	7	8	9	10
Numerales griegos									
A	B	Γ	Δ	E	F	Z	H	Θ	I
Numerales romanos									
I	II	III	IV	V	VI	VII	VIII	IX	X

Figura I.2. Los numerales en diferentes sistemas de numeración.

Se cree que la notación que usamos para los numerales —1, 2, 3, 4, 5, 6, 7, 8, 9— tiene origen hindú. Alrededor del siglo x los árabes tomaron estos conocimientos de los hindúes e introdujeron su uso en España, de donde posteriormente pasaron a toda Europa. La forma de nuestros números nos es tan familiar que no estamos conscientes de la lenta evolución por la que pasaron a lo largo de siglos. En la figura I.3 podemos ver algunos pasos de esta evolución.

Hindú, siglo XI १ २ ३ ४ ५ ६ ७ ८ ०

Arábigo (Ghobarí), siglo XI ١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ٠

Arábigo oriental, 1575 ١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ٠

Europeo, siglo XV 1 2 3 4 5 6 7 8 9 0

1 2 3 4 5 6 7 8 9 0

Contemporáneo 1 2 3 4 5 6 7 8 9 0

Figura 1.3. La evolución de los números arábigos.

Parece que el sistema posicional en base 10 que usamos comenzó a usarse en la India alrededor del año 500 de nuestra era. Una vez conocido este sistema, los únicos dígitos importantes son los que denotan del 1 al 9 y el 0, que son los que se conservaron y evolucionaron hasta llegar a los números actuales. Pero la notación posicional no se popularizó sino hasta el siglo IX, después de que el matemático Al-Juarizmi de Bagdad (Mohamed ibn Musa) escribió un tratado de aritmética dirigido a los comerciantes donde recomendaba el uso de este sistema. Luego, poco a poco, el sistema decimal fue siendo aceptado en Europa. Es interesante saber que en el siglo XIII el gobierno de Florencia dictó leyes contra este sistema, pues se decía propiciaba la falsificación de billetes de banco, que podían ser fácilmente alterados para tener otra denominación.

SISTEMAS POSICIONALES

Considere un número positivo en nuestro sistema de numeración, como por ejemplo el 23 107. Sabemos desde la primaria que el primer dígito a la derecha (en este caso, el 7) corresponde a las unidades, el siguiente hacia la izquierda (el 0) a las decenas, luego (el 1) a las centenas y así sucesivamente. De esta

manera tenemos que 23 107 es una abreviatura de la expresión:

$$\begin{aligned} 23\,107 &= 2 \times 10\,000 + 3 \times 1\,000 + 1 \times 100 + 0 \times 10 + 7 \times 1 \\ &= 2 \times 10^4 + 3 \times 10^3 + 1 \times 10^2 + 7 \times 10^0 \end{aligned}$$

donde usamos la convención de que $10^m = 10 \times 10 \times \dots \times 10$ (m veces) representa un 1 seguido de m ceros. Mas en general, cualquier entero positivo n puede representarse en *notación decimal* como:

$$n = a_r a_{r-1} \dots a_0$$

donde cada letra a_i es un dígito entre 0 y 9, de forma que la expresión de n en notación decimal es una abreviatura de:

$$n = a_r \cdot 10^r + a_{r-1} \cdot 10^{r-1} + \dots + a_0.$$

En el ejemplo de arriba, tenemos que $a_0 = 7$, $a_1 = 0$, $a_2 = 1$, $a_3 = 3$ y $a_4 = 2$ para formar el número 23 107. Todo esto puede generalizarse tomando un entero positivo cualquiera $b > 1$ como *base*. Cualquier número n puede escribirse como:

$$n = b_r \cdot b^r + b_{r-1} \cdot b^{r-1} + \dots + b_0$$

con b_r, b_{r-1}, \dots, b_0 números enteros entre 0, 1, ..., $b - 1$. La expresión obtenida de esta manera:

$$n = b_r b_{r-1} \dots b_0$$

se llama la *representación posicional de n en base b* .

Por ejemplo, el número 23 107 se escribiría en base 8 de la siguiente manera: 55 103, ya que:

$$\begin{aligned} 5 \times 8^4 + 5 \times 8^3 + 1 \times 8^2 + 0 \times 8^1 + 3 \times 8^0 \\ = 20\,480 + 2\,560 + 64 + 3 = 23\,107. \end{aligned}$$

Donde, por supuesto, $8^4 = 8 \times 8 \times 8 \times 8$, $8^3 = 8 \times 8 \times 8$, etcétera. En base 16, este mismo número 23 107 se escribiría 5 a43, donde la letra "a" denota el número 10 correspondiente a la base 16, esto es:

$$5 \times 16^3 + 10 \times 16^2 + 4 \times 16 + 3 = 23\,107.$$

Varios sistemas posicionales con diferentes bases han sido usados a lo largo de la historia, aunque la base 10 ha sido la dominante. Por ejemplo, los mayas usaban base 20, los babilonios usaban base 60. Para indicar la hora, nuestros relojes hoy en día usan todavía una combinación de base 12 y base 60 (si decimos que son las 2 horas 23 minutos y 11 segundos de la mañana, queremos decir que han transcurrido $2 \times 60^2 + 23 \times 60 + 11 = 8591$ segundos del día).

El sistema posicional de base 2 se llama sistema binario y es el sistema que utilizan las computadoras electrónicas. Nuestro número 23107 se escribe como 101101001000011 en sistema binario. (¿Por qué?)

La razón por la que el sistema binario se utiliza en las computadoras es la siguiente: podemos pensar en una fila de focos que pueden estar apagados o prendidos. Si un foco está apagado indica que en ese lugar el dígito correspondiente es 0, si está prendido el dígito es el 1. Así nuestro número 23107 se puede ver como la fila de focos siguiente:



Figura 1.4. El número 23107 en sistema binario.

Una computadora funciona por medio del flujo de la corriente eléctrica. De esta manera, un 1 indica que la corriente pasa por una puerta magnética, mientras que un 0 indica que la corriente no puede pasar por la puerta correspondiente.

ÁBACOS Y COMPUTADORAS

“¿Puedes sumar?” Pregunto la Reina Blanca.
“¿Cuánto es uno y uno y uno y uno y uno y uno
y uno y uno y uno?”

“No sé”, dijo Alicia, “perdí la cuenta”.
“No sabe sumar”, interrumpió la Reina Roja...

Alicia a través del espejo, LEWIS CARROLL

Contar es el uso más elemental que se da a los números. También se hacen operaciones con ellos: sumar, restar, multiplicar, dividir, y tal vez otras operaciones más complejas.

La mayoría de los sistemas de numeración que se usaron en la Antigüedad no eran muy adecuados para realizar operaciones. Solamente la introducción de los sistemas posicionales (en particular, el de base 10) facilitó las operaciones aritméticas.

Desde tiempos remotos se ha tratado de diseñar aparatos que simplifiquen y hagan más rápidas las “cuentas” aritméticas. Tal vez uno de los más antiguos es el *ábaco*, que es un invento simple y eficiente que aún se usa en muchos países. Aparentemente fue inventado en Babilonia hace más de 5000 años, pero son los chinos los que lo llevaron a la forma en que se usa actualmente.

El *ábaco* tiene fichas móviles colocadas en filas en un tablero. Cada fila tiene 5 fichas divididas en 2 grupos: un grupo tiene 4 fichas, el otro sólo una. La primera fila indica las unidades, la segunda las decenas, la tercera las centenas y así sucesivamente. La ficha aislada de la primera fila vale 5, todas las otras valen 1; la ficha aislada de la segunda fila vale 50, las otras 4 valen 10 cada una, etcétera. Para escribir un número se hace en sistema decimal pegando las fichas necesarias al travesaño intermedio del *ábaco*. En la figura 1.5 indicamos como se escriben algunos números en el *ábaco*.

Sumar con el *ábaco* es sencillo. Por ejemplo, consideremos la suma de 347 y 282. Escribimos el primer número en el *ábaco*. En seguida tratamos de agregar el segundo número con las fichas. Comenzamos por las centenas: agregamos 2 fichas. Seguimos con las decenas: debemos agregar 8 fichas, pero no las hay disponibles. Pero $80 = 100 - 20$, entonces si agregamos una ficha en la fila de las centenas y quitamos 2 en la fila de las decenas, habremos sumado 80. Sumar 2 unidades es sencillo. El resul-

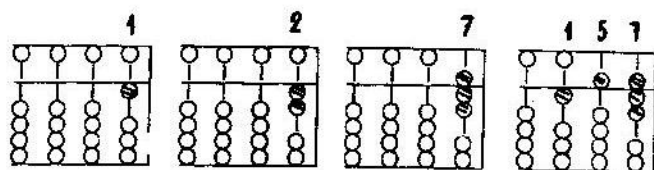


Figura 1.5. Contando con el ábaco.

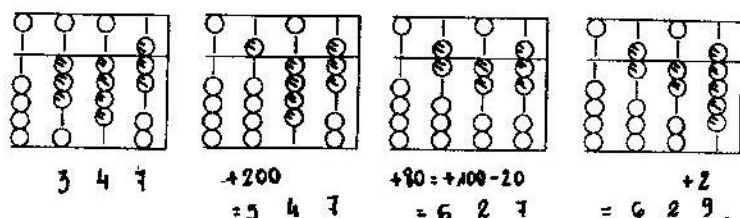


Figura 1.6. Sumando $347 + 282$.

tado de la suma queda escrito en el ábaco. En la figura 1.6 ilustramos los pasos anteriores.

Hasta hace unos 20 años eran frecuentes los torneos aritméticos en que se enfrentaban personas que usaban el ábaco hábilmente (generalmente orientales) en contra de personas con calculadoras electrónicas. El resultado era que el ábaco se imponía siempre. No sé cuán verídicas sean estas historias, pero es cierto que en algunos países, como China y Japón, hay una gran tradición del uso del ábaco y algunas personas lo saben usar con habilidad sorprendente.

Muchas han sido las máquinas que los hombres han inventado para facilitar las operaciones. Es sorprendente que, en 1900, unos pescadores encontraron en el mar Egeo parte de un mecanismo con engranajes que parece datar de la Grecia clásica. Aparentemente este mecanismo formaba parte de una calculadora que permitía hacer operaciones aritméticas. En tiempos más cercanos, en el siglo XVII, John Napier construyó una calculadora de bolsillo para multiplicar. Blaise Pascal, en Francia, construyó una máquina que permitía sumar y restar mecánicamente. Ciertamente el tamaño de esta máquina era mucho mayor que el de una moderna calculadora de bolsillo.

Alrededor de 1830, el matemático e inventor inglés Charles

Babbage diseñó una máquina programable, el "ingenio analítico", que es el precursor de las modernas computadoras digitales. Babbage quería que su máquina tuviera la capacidad de realizar cualquier operación aritmética con base en instrucciones de tarjetas perforadas, una unidad de memoria en donde se almacenaran números, una unidad de control secuencial y casi todos los elementos que contiene una moderna computadora. Su máquina nunca pudo funcionar debido a problemas técnicos con la fabricación de piezas delicadas. Sus proyectos fueron olvidados y se volvió a saber de ellos sólo con el descubrimiento de sus diarios en 1937. Sin embargo, las ideas de Babbage eran correctas y él, junto con Ada Lovelace (hija de Lord Byron), fueron los primeros en idear lenguajes de computadora. A Lovelace se atribuye la frase "las computadoras sólo saben hacer lo que se les indica que hagan", sin embargo creía que el "ingenio analítico" podría componer refinadas piezas de música de cualquier complejidad y extensión.

A partir de los años cincuenta, el acelerado crecimiento y desarrollo de la tecnología de las computadoras han tenido gran

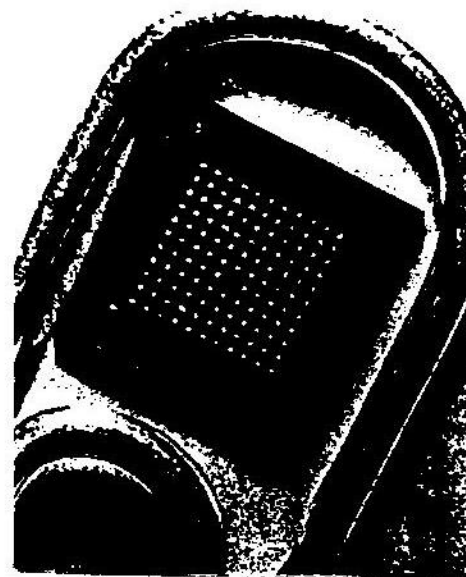


Figura 1.7. Componente de una computadora moderna.

repercusión en el mundo y lo han modificado de manera permanente. En 1946, la primera computadora electrónica, ENIAC, comenzó a funcionar en la Universidad de Princeton; era capaz de realizar 5 000 sumas por segundo. En la actualidad hay supercomputadoras que pueden efectuar millones de operaciones por segundo. Hoy, prácticamente todas las actividades humanas están relacionadas, controladas o apoyadas por alguna computadora. Pero su uso se masificó hace menos de 15 años, con la llegada de las computadoras personales que han permitido a mucha gente el acceso directo a la computación.

ADIVINA EL NÚMERO QUE ESTOY PENSANDO

One and one and one is three.
Try to be good looking cause you're so hard to see.

"Come together", LOS BEATLES

Hay un juego llamado las "Veinte preguntas". Alguien piensa el nombre de un personaje histórico y los demás tienen que adivinar de quién se trata haciendo sólo 20 preguntas que pueden contestarse con "sí" o "no". Proponemos la siguiente variante del juego. Se le pide a alguien que piense un número entre 1 y 1 000 000. Hay que adivinar el número haciendo cuando más 20 preguntas que puedan contestarse con "sí" o "no". ¿Cuáles preguntas hay que hacer para ganar?

Solución. La solución es usar el sistema posicional base 2. Comencemos por números pequeños. ¿Cuántas cifras se requieren para escribir 13 en base 2? En base 2, el número 13 es: 1101. Se requieren cuatro cifras. Podríamos por tanto saber que la otra persona ha pensado el número 13 haciendo cuatro preguntas. ¿Cuáles? El "1" más a la izquierda de 1101 quiere decir que nuestro número es mayor o igual que $2^3 = 8$; el siguiente "1" quiere decir que de los 8 números que quedan entre $2^3 = 8$ y $2^4 = 16$, nuestro número está en la mitad de más arriba (es decir, es mayor o igual que $2^3 + 2^2 = 12$); el siguiente "0" indica que de los números que quedan (12, 13, 14, 15 y 16), el nuestro está en la mitad de abajo (es decir, es menor

que $2^3 + 2^2 + 2^1 = 14$). El último "1" ubica precisamente al número 13.

Regresemos a la pregunta original. Usando una calculadora sabemos que $2^{20} = 1\,048\,576$. Por lo tanto, todo número entre 1 y 1 000 000 requiere 20 cifras para ser escrito en sistema binario. ¿Cuáles son las preguntas que hay que hacer entonces en nuestro juego?

1) ¿Es tu número mayor o igual que $2^{19} = 524\,288$? Si la respuesta es afirmativa, nuestro número comienza con "1" en la posición 19 en base 2. Si es negativa, comienza con "0".

Si la respuesta a la pregunta 1 fue afirmativa, entonces la pregunta 2 debe ser: ¿Es tu número mayor o igual que $2^{19} + 2^{18} = 786\,432$? En caso afirmativo, nuestro número comienza con "11" en base 2. Si es negativo comienza con "10".

Si la respuesta a la pregunta 1 fue negativa, preguntamos ahora: ¿Es tu número mayor o igual que $2^{18} = 262\,144$? En caso afirmativo, nuestro número comienza con "01" en base 2. En caso negativo, con "00".

Al final de las 20 preguntas, conoceremos el número que la otra persona pensó en base 2 (y entonces también, fácilmente, en base 10).

Martin Gardner sugiere la fabricación de "cartas para adivinar el pensamiento" basadas en este principio. Para ello basta reproducir en papel las seis cartas que a continuación damos.

1	3	5	7	9	11	13	15	2	3	6	7	10	11	14	15
17	19	21	23	25	27	29	31	18	19	22	23	26	27	30	31
33	35	37	39	41	43	45	47	34	35	38	39	42	43	46	47
49	51	53	55	57	59	61	63	50	51	54	55	58	59	62	63

4	5	6	7	12	13	14	15	8	9	10	11	12	13	14	15
20	21	22	23	28	29	30	31	24	25	26	27	28	29	30	31
36	37	38	39	44	45	46	47	40	41	42	43	44	45	46	47
52	53	54	55	60	61	62	63	56	57	58	59	60	61	62	63

16	17	18	19	20	21	22	23	32	33	34	35	36	37	38	39
24	25	26	27	28	29	30	31	40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55	48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63	56	57	58	59	60	61	62	63

Se pide a una persona que piense un número entre 1 y 63 y que no nos lo diga. Le damos las 6 cartas y le pedimos que nos indique las cartas donde su número no aparece. ¡Inmediatamente nosotros adivinamos el número que pensó!

¿Cómo lo hacemos? Muy sencillo: la carta que comienza con 1 contiene todos los números entre 1 y 63 que tienen "1" en el lugar de las unidades cuando escribimos el número en sistema binario. La carta que comienza con 2 contiene todos los números que tienen "1" en el lugar de las decenas en sistema binario y así sucesivamente. Por ejemplo, si la persona pensó "30" nos indica entonces las cartas que comienzan con 1 y 32. ¿Qué hacemos nosotros? Sumamos los primeros números de las cartas que no nos indicó (las que sí contienen a 30), esto es, $2 + 4 + 8 + 16 = 30$.

II. Un mundo hecho de números

El número es el origen de todas las cosas.

PLATÓN

LA FILOSOFÍA y la ciencia como ocupaciones válidas y trascendentes nacieron en la Grecia antigua. Lo hicieron cuando el hombre se comenzó a preguntar por el orden de los sucesos de su entorno, cuando se dio cuenta de que no todo sucede al azar. En esta época no había divisiones para las áreas del conocimiento humano, todo formaba parte de una disciplina única: la filosofía.

Según Bertrand Russell, el primer filósofo griego en hacer de su interés una forma de vida fue Pitágoras. Poco se sabe de su vida. Nació en Samos y vivía ahí alrededor del año 544 a.C., cuando reinaba el tirano Polícrates. En algún momento Pitágoras no soportó la tiranía y pasó a vivir en Crotona, en el sur de Italia. Allí fundó una escuela filosófica que floreció hasta 510 a.C. Protestas en contra de la escuela hicieron salir a Pitágoras hacia Metaponto, donde permaneció hasta su muerte.



Figura II.1. Pitágoras.

La escuela pitagórica dio inicio a la tradición científica y en particular a las matemáticas. Otro elemento que tenía un papel importante en su filosofía era la música. Pitágoras descubrió las relaciones numéricas simples que llamamos intervalos musicales: si una cuerda afinada se pisa a la mitad sonará una octava más arriba; si su longitud se reduce a tres cuartos, entonces suena un cuarto más alta; si su longitud se reduce a dos tercios, sonará una quinta más arriba. Parece ser que estos descubrimientos permearon muchas de las ideas posteriores de Pitágoras.

La base de su filosofía es que todo en la naturaleza se puede entender por medio de los números. De hecho, por relaciones simples entre números enteros, como en las fracciones de las cuerdas musicales. Para entender el mundo, primero se debe entender los números. Desde entonces es éste uno de los conceptos centrales de la ciencia.

Los pitagóricos desarrollaron representaciones de números por medio de disposiciones de guijarros. Pensaban que toda la materia estaba formada por partículas indivisibles y que la variedad de formas encontradas en la naturaleza no sólo depende de la cantidad de estas partículas, sino también de la disposición espacial de ellas. Por ello clasificaron los números de acuerdo a los arreglos que se podían formar con el correspondiente número de guijarros. Por ejemplo, 1, 3, 6, 10, 15, etc., son *números triangulares* porque se puede formar triángulos con ellos.

También introdujeron los números cuadrados (1, 4, 9, 16, ...),

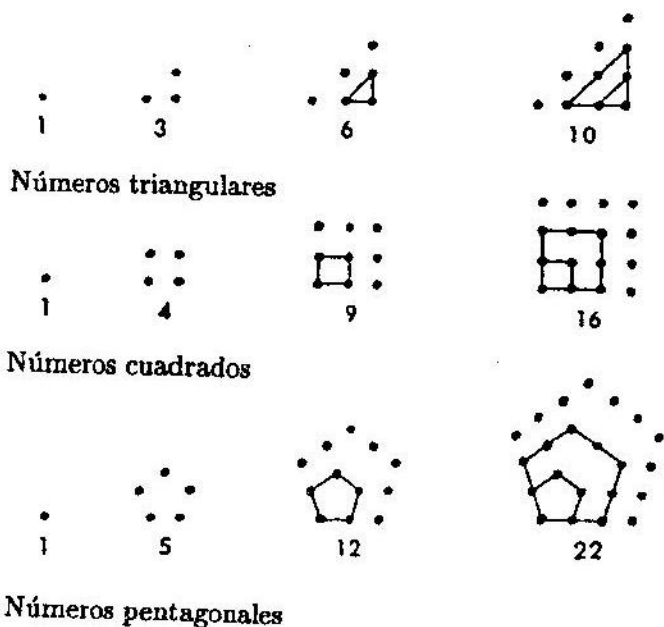


Figura II.2. Números triangulares, cuadrados, pentagonales.

pentagonales (1, 5, 12, 22, ...), hexagonales (1, 6, 15, 28, ...), etcétera. ¿Cuál es el patrón general de estos números? Si llamamos t_n al número triangular n -ésimo, entonces tenemos que:

$$t_n = t_{n-1} + n = 1 + 2 + \cdots + (n-1) + n = \frac{1}{2}n(n+1).$$

El lector interesado puede ver una demostración de la última igualdad en la biografía de Gauss en el capítulo IX.

Si llamamos c_n al n -ésimo número cuadrado, obtenemos:

$$\begin{aligned} c_n &= n^2 = [(n-1) + 1]^2 \\ &= (n-1)^2 + 2(n-1) + 1 \\ &= c_{n-1} + 2(n-1) + 1. \end{aligned}$$

Donde hemos hecho uso de la fórmula del *cuadrado del binomio*: $(a+b)^2 = a^2 + 2ab + b^2$.

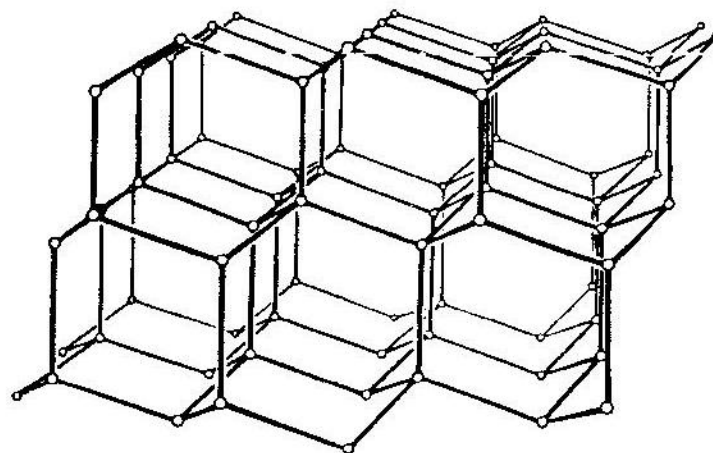


Figura II.3. Arreglos de números hexagonales para formar sólidos.

Esto comienza a mostrar un patrón puesto que $t_n = t_{n-1} + (n-1) + 1$. En efecto, para los números pentagonales p_n se tiene que $p_n = p_{n-1} + 3(n-1) + 1$ y de aquí se calcula:

$$\begin{aligned} p_n &= 1 + 5 + 12 + \cdots + 3(n-2) + 1 + 3(n-1) + 1 \\ &= (1 + 1 + \cdots + 1) + 3(1 + 2 + \cdots + n-2 + n-1) \\ &= (n-1) + \frac{3}{2}(n-1)n = \frac{1}{2}n[2 + 3(n-1)]. \end{aligned}$$

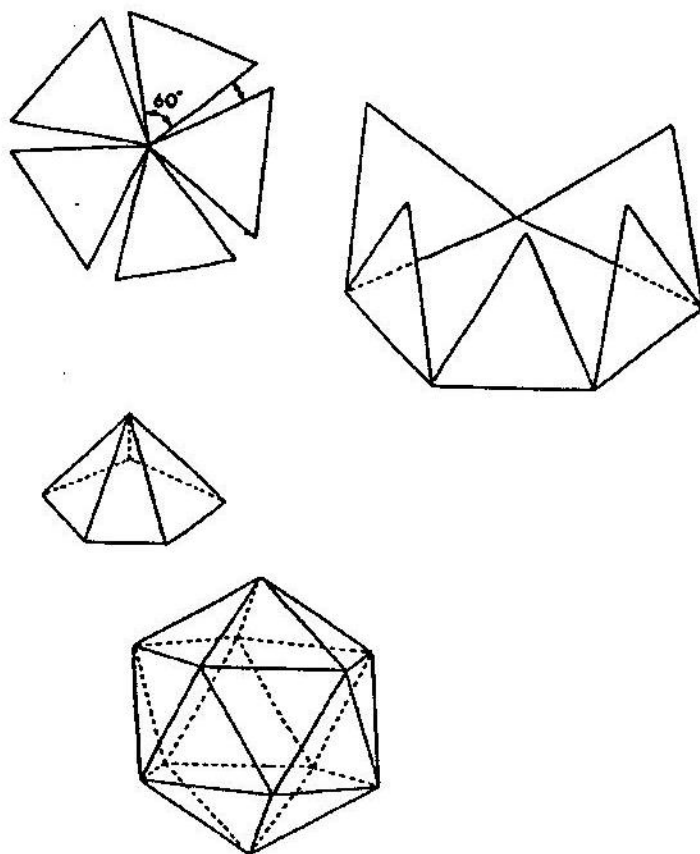


Figura II.4. Armado del icosaedro. Todas las caras son triángulos equiláteros, hay 12 vértices y 20 caras.

Fórmulas similares se encontraron para los números asociados a polígonos regulares de cualquier número de lados. La idea de los pitagóricos es que con estos números como bloques básicos se podían formar todos los cuerpos sólidos. Por ejemplo, sobreponiendo "pisos" de hexágonos se obtiene un prisma hexagonal. También describieron cómo construir los sólidos regulares por principios simples "de armado". Ilustramos esto en la figura II.4 que describe cómo armar el icosaedro.

Se atribuye a Pitágoras el descubrimiento del famoso teorema que lleva su nombre. Aunque no se sabe cuál era la demostración que tenía del teorema, se sabe que conocía una. Su escuela fue la primera en comprender la importancia de una demostración matemática rigurosa. Así, la demostración de una afirmación sobre los triángulos vale para todos los triángulos, no sólo para aquellos que tenemos enfrente. Además, es válida independientemente de los hombres y del tiempo. Es eterna.

Teorema. Dado un triángulo rectángulo con catetos de longitud a y b e hipotenusa de longitud c , se tiene la siguiente relación:

$$a^2 + b^2 = c^2.$$

Primera demostración. Consideremos el triángulo de la figura II.5, donde indicamos los ángulos α , β y γ . Por definición, γ es de 90° .

Como los ángulos internos de un triángulo suman 180° , entonces $\alpha + \beta = 90^\circ$. Entonces podemos disponer cuatro triángulos iguales al dado como en la figura II.6.

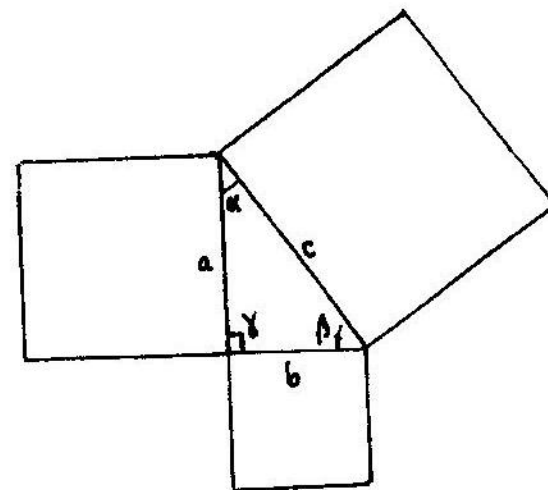


Figura II.5.

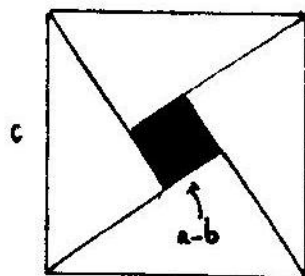


Figura II.6.

Si A es el área del triángulo dado, podemos calcular el área del cuadrado externo como:

$$c^2 = 4A + (a - b)^2 = 4\left[\frac{1}{2}ab\right] + a^2 - 2ab + b^2 = a^2 + b^2. \quad \square$$

Hay muchas otras pruebas del teorema de Pitágoras, la siguiente es corta y elegante.

Segunda demostración del teorema de Pitágoras. Dividimos nuestro triángulo como se muestra en la figura II.7.

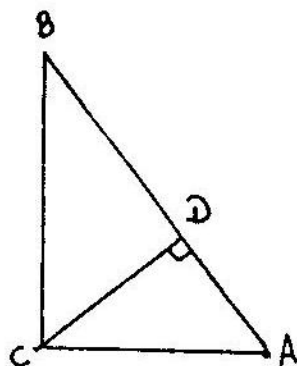


Figura II.7.

Los triángulos con vértices BCD y ACD son similares (es decir, tienen los mismos ángulos). También son similares al triángulo completo ABC . Aunque no sepamos la longitud de

todos los lados de los triángulos, sabemos que las longitudes de los lados correspondientes de triángulos similares son proporcionalmente iguales, es decir, para los triángulos ABC y ACD se tiene: $\overline{AC}/\overline{AD} = \overline{AB}/\overline{AC}$, o sea:

$$\frac{b}{x} = \frac{c}{b}.$$

Similarmente, usando los triángulos ABC y BCD tenemos:

$$\frac{a}{c - x} = \frac{c}{a}.$$

Estas ecuaciones se pueden reescribir como:

$$b^2 = cx = c^2 - a^2. \quad \square$$

La primera de las demostraciones que hemos presentado parece ser ¡anterior a Pitágoras! En efecto, en el libro chino *Chou pei suang ching*, que data probablemente de 1000 años a.C., aparece una ilustración muy similar a la de la primera prueba del teorema de Pitágoras. De hecho, en China nadie sabe qué es el teorema de Pitágoras puesto que al famoso enunciado lo conocen como el *teorema de Chou*. Esta demostración también es esencialmente la que presenta Euclides en su libro III de los *Elementos*. La segunda de las demostraciones se atribuye al matemático inglés del siglo XVII John Wallis.

UNA TRAGEDIA GRIEGA

Como hemos dicho, para los pitagóricos toda la naturaleza estaba determinada por números enteros o fracciones. En lenguaje moderno, las fracciones de la forma a/b con a y b enteros se llaman *números racionales*. Entonces podemos decir que los pitagóricos pensaban que toda la naturaleza se podía entender por medio de los números racionales.

Pero, por otra parte, tenían el teorema de Pitágoras que, en el caso más simple, nos dice que un triángulo rectángulo con catetos de longitud 1 tiene una hipotenusa de longitud c que satisface $c^2 = 1^2 + 1^2 = 1 + 1 = 2$. Por supuesto, c se denota $\sqrt{2}$. ¿Qué clase de número es $\sqrt{2}$?

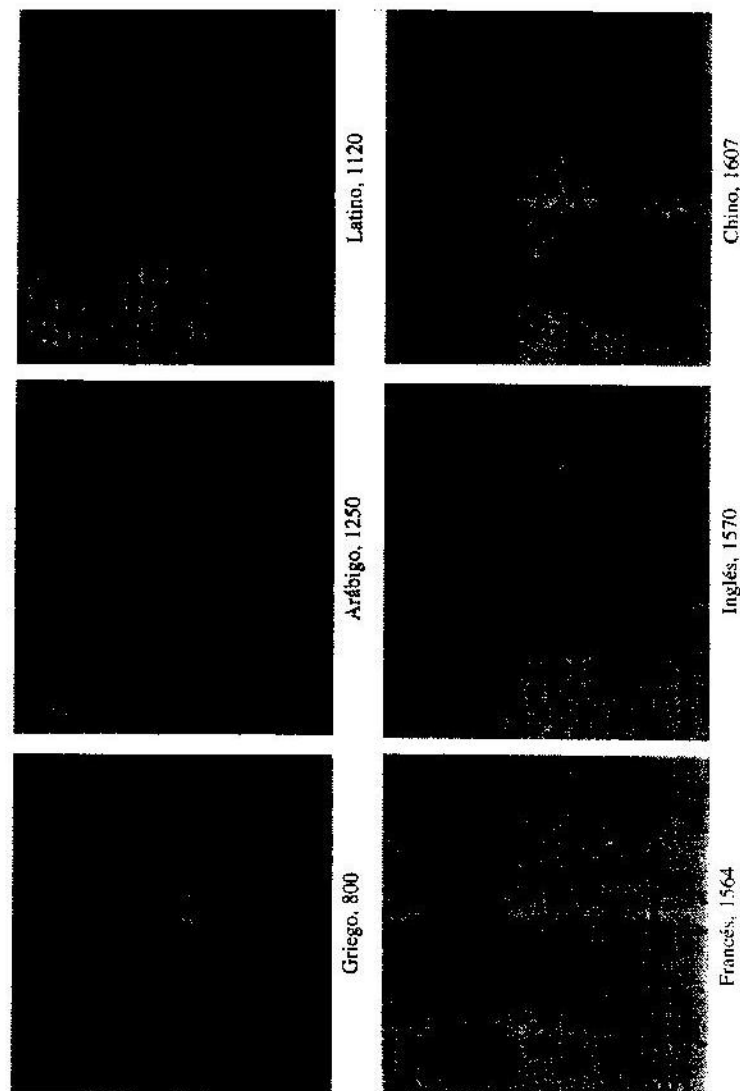


Figura II.8. Ilustraciones que muestran versiones del teorema de Pitágoras en diferentes culturas.

¡Sorpresa! Este número no es racional. Es decir, para Pitágoras y su escuela, el número $\sqrt{2}$ no debería existir porque no se puede construir en forma de fracción a partir de los números enteros. Una demostración de este hecho se encuentra en el libro III de los *Elementos* de Euclides.

Teorema. El número $\sqrt{2}$ no es racional.

La demostración de siempre. Se lleva a cabo por *reducción al absurdo*, esto es, se comienza suponiendo lo contrario de lo que se quiere demostrar y luego por medio de la lógica se deduce una afirmación absurda, lo que muestra que nuestra suposición es insostenible. Procedamos.

Supongamos que $\sqrt{2}$ fuese un número racional, es decir, de la forma $\sqrt{2} = a/b$ con a y b siendo números enteros. Simplificando la fracción podemos siempre suponer que a y b no tienen divisores comunes aparte del 1. Escribamos $a = \sqrt{2}b$ y elevemos al cuadrado. Obtenemos $a^2 = 2b^2$. Esto quiere decir que 2 divide a a^2 . Como un producto de dos números nones es non otra vez, entonces a debe ser par. O sea, a es divisible por 2 y podemos escribirlo como $a = 2d$ para d como otro número entero. Elevando otra vez al cuadrado tenemos: $4d^2 = a^2 = 2b^2$. Cancelamos un 2 de cada lado: $2d^2 = b^2$. Esto es, b^2 es divisible por 2 y como vimos antes, esto implica que b es divisible por 2.

Pero habíamos dicho que el único divisor común de a y b es 1 y ahora encontramos que 2 es un divisor común! Ésta es la contradicción que deseábamos encontrar. \square

Probablemente este resultado es uno de los más fundamentales de las matemáticas. Es interesante que en 1975 el matemático alemán Estermann haya encontrado la siguiente bella demostración.

La prueba nueva. Supongamos que $\sqrt{2}$ es un número racional. Entonces hay un entero positivo mínimo k con la propiedad de que $k\sqrt{2}$ es entero. Por otra parte, sabemos que $1 < \sqrt{2} < 2$, por lo tanto $k < \sqrt{2}k < 2k$ y luego, $k' = (\sqrt{2} - 1)k$ es un entero positivo menor que k . Pero:

$$k'\sqrt{2} = (\sqrt{2} - 1)k\sqrt{2} = 2k - \sqrt{2}k$$

es un entero positivo (por ser diferencia de dos enteros), lo que contradice la minimalidad de k . \square



Figura II.9. Dibujo de Antonio Helguera. Guión de Javier Bracho y José A. de la Peña. Museo Universum. UNAM.

No sabemos en qué momento se dieron cuenta los pitagóricos de la irracionalidad de $\sqrt{2}$, hecho que contradecía sus creencias en las más profundas raíces. Sabemos que el descubrimiento produjo un tremendo escándalo en la escuela. La historia de este escándalo podemos verla caricaturizada por Antonio Helguera en las siguientes páginas.

Haga usted mismo su $\sqrt{2}$

Si tomo una calculadora electrónica y pregunto por el valor de $\sqrt{2}$ la respuesta es (hasta 9 cifras decimales) $\sqrt{2} = 1.414213562$. Pero ésta no es una respuesta exacta, de hecho el número que la calculadora nos da es un número racional ya que:

$$1.414213562 = \frac{707\,106\,781}{500\,000\,000}.$$

Bueno, como todos los resultados que puede dar una calculadora, el valor de $\sqrt{2}$ es sólo una aproximación. Pero podemos describir un método para construir $\sqrt{2}$ con tanta precisión como queramos (más que la precisión de la calculadora con 10 cifras) y esto sólo con operaciones elementales: suma, multiplicación y división.

Tomamos un número cualquiera $x_0 > 0$. Formamos el número x_1 según la siguiente regla:

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

Luego se define $x_2 = \frac{1}{2} \left(x_1 + \frac{2}{x_1} \right)$, esto es, con la misma fórmula con la que antes se había definido x_1 a partir de x_0 . Continuamos así definiendo x_k a partir del número anterior x_{k-1} como $x_k = \frac{1}{2} \left(x_{k-1} + \frac{2}{x_{k-1}} \right)$. Por ejemplo, veamos lo que pasa si comenzamos con $x_0 = 3$. Obtenemos sucesivamente (con aproximación hasta 10 cifras decimales):

k	x_k
0	3
1	1.833333333
2	1.462121212
3	1.4149984299
4	1.4142137800
5	1.4142135624

Donde vemos que ya a la quinta iteración del procedimiento hemos obtenido una aproximación perfecta hasta la novena cifra decimal del número $\sqrt{2}$.

¿Es esto una casualidad? No, de hecho se tiene que para cualquier $x_0 \neq 0$, los números de la sucesión x_k se acercan más y más a $\sqrt{2}$. Esto se escribe como $\lim_{k \rightarrow \infty} x_k = \sqrt{2}$.

Este procedimiento se conoce como el *método de Newton*, pero parece que de manera rudimentaria era ya conocido por los sumerios ¡hace 4 000 años! Veamos por qué funciona el método.

Por la manera en que están definidos, todos los x_k son positivos. Escribamos:

$$x_k = (1 + e_k)\sqrt{2}$$

donde e_k es el "error" del término x_k . Deseamos estimar este error. Por definición:

$$x_{k+1} = \frac{1}{2} \left(x_k + \frac{2}{x_k} \right) = \sqrt{2} \left(1 + \frac{e_k^2}{2 + 2e_k} \right).$$

Entonces el error en el paso $k + 1$ es:

$$e_{k+1} = \frac{e_k^2}{2 + 2e_k}.$$

Como $x_0 > 0$, entonces $e_0 > -1$, lo que hace $e_k > 0$ a partir de $k > 0$. Luego, $x_k > \sqrt{2}$ para $k > 0$. Además vemos que $e_{k+1} < e_k$. Por lo tanto, tenemos una sucesión decreciente:

$$x_1 > x_2 > x_3 > \dots > \sqrt{2}.$$

También $e_{k+1} < e_k^2/2$, luego si e_k es muy pequeño, digamos $e_k < 0.0001$, entonces e_{k+1} es mucho más pequeño, del orden de $e_{k+1} < (0.0001)^2/2 = 0.000000005$. O sea, en cada paso la aproximación obtenida es exacta en (al menos) el doble de cifras decimales que la aproximación anterior.

Como un *ejercicio* para el lector interesado proponemos que demuestre que los números $\sqrt{3}$ y $\sqrt{7}$ son también números irracionales y que obtenga su valor aproximado usando el método de Newton.

EL ÁRBOL DE PITÁGORAS

Hay una interesante manera de colocar triángulos rectángulos uno a continuación del otro. Si consideramos el triángulo con catetos de longitud 1, entonces la hipotenusa tiene longitud $\sqrt{2}$. Ahora podemos considerar un triángulo rectángulo con un cateto de longitud 1 y el otro de longitud $\sqrt{2}$ y colocar este último cateto sobre la hipotenusa del primer triángulo. La hipotenusa de este segundo triángulo medirá $\sqrt{(\sqrt{2})^2 + 1} = \sqrt{3}$. A continuación construimos un triángulo con catetos de longitudes 1 y $\sqrt{3}$ e hipotenusa $\sqrt{4}$ y se coloca junto al segundo triángulo. Continuamos este proceso como se indica en la figura II.10, obteniendo una forma *espiral*.

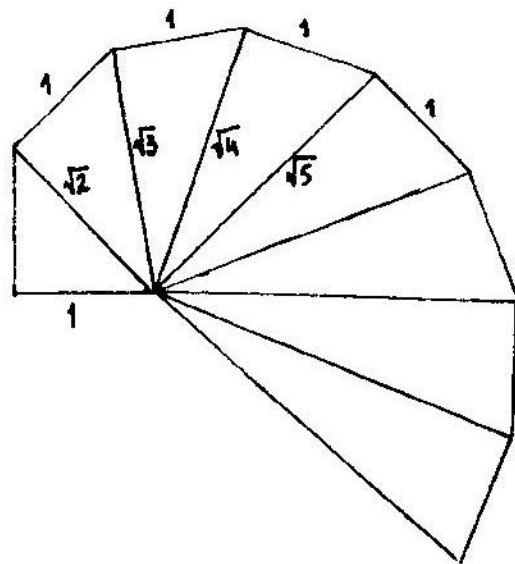


Figura II.10. La espiral pitagórica.

En 1957 apareció en Holanda el singular libro *Geometría en el plano: un milagroso campo de investigación*. Su autor, A. Bosman, trataba de mostrar los sorprendentes arreglos geométricos de la naturaleza. Uno de los más sorprendentes diseños de Bosman fue realizado, según sus palabras, en el mismo pi-

zarrón en que diseñaba submarinos durante la segunda Guerra Mundial. Este diseño que llamaremos el *árbol pitagórico* está relacionado con la construcción de la espiral que hemos hecho antes.

La construcción se hace repitiendo unos cuantos pasos básicos:

Paso 1: dibuje un cuadrado. Paso 2: dibuje un triángulo rectángulo que tenga por hipotenusa uno de los lados del cuadrado. Paso 3: dibuje un cuadrado sobre cada uno de los dos catetos del triángulo dibujado en el paso 2. Paso 4: repita el paso 2 para cada uno de los cuadrados dibujados en el paso 3, usando como hipotenusa el lado opuesto al que ya se usó. Paso 5: repita el paso 3 usando cada uno de los dos triángulos dibujados en el paso 4. Paso 6: ... continúe el proceso tanto como quiera.

Es muy sencillo entender la construcción, en la figura II.11 se ilustra el resultado después de pocos pasos.

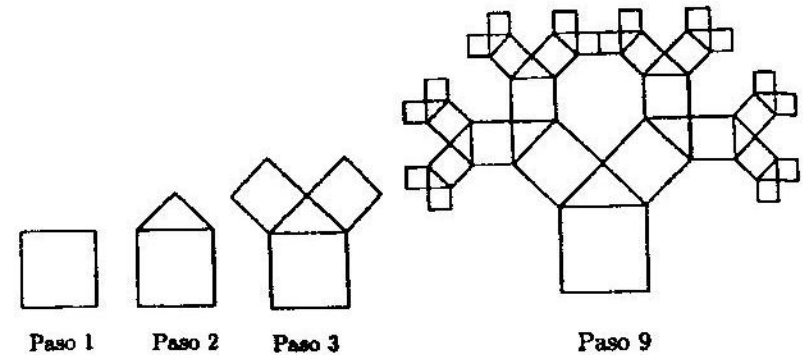


Figura II.11. La idea de la construcción de un árbol pitagórico.

¿Qué sucede si el proceso se repite muchas veces (digamos 50 veces)? Obtenemos figuras como la II.12. Estas figuras se ven sorprendentemente similares a helechos reales.

Que los árboles pitagóricos tengan aspecto de planta no es casual. De hecho, el biólogo Aristid Lindenmayer introdujo el concepto de L-sistemas en botánica más o menos en la forma que hemos definido los árboles pitagóricos. La idea es simple: cuando el tallo de una planta se divide, digamos en dos ramas, el área de la sección transversal de las dos ramas sumadas se debe mantener igual a la sección del área del tallo principal. Si

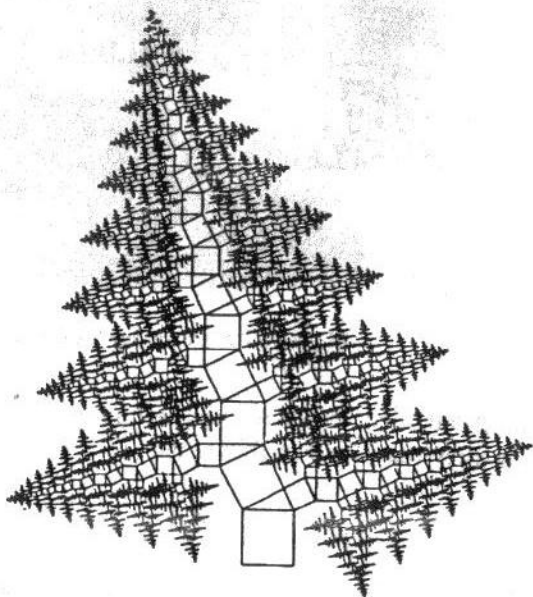


Figura II.12. Árbol pitagórico después de 50 pasos.

pensamos que el tallo es cilíndrico, la sección tendrá un área de $(\pi/4)c^2$, donde c es el diámetro del tallo. Si los diámetros de las dos ramas en que se divide este tallo son a y b , entonces tenemos que $(\pi/4)c^2 = (\pi/4)a^2 + (\pi/4)b^2$, o sea, $c^2 = a^2 + b^2$.
¿Qué sucede en un árbol pitagórico? Un tallo del árbol tiene

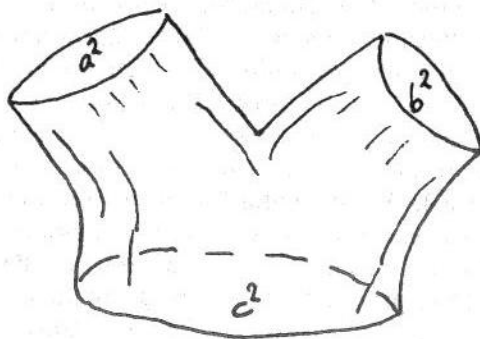


Figura II.13. Esquema de la ramificación de un tallo.

sección de longitud c (el lado de un cuadrado). Cuando se divide el tallo en dos ramas lo hace por medio de un triángulo rectángulo de lados a y b . Luego $a^2 + b^2 = c^2$ por el teorema de Pitágoras y ésta es la misma relación que en el árbol real.

CONSTRUYENDO TRIÁNGULOS RECTÁNGULOS CON LADOS ENTEROS

El interés de los pitagóricos se centraba principalmente en triángulos rectángulos con lados enteros. El caso más sencillo, de longitudes 3, 4 y 5 era ya conocido de los egipcios (observe que $3^2 + 4^2 = 9 + 16 = 25 = 5^2$). Algún tiempo después de Pitágoras, otros matemáticos griegos observaron que había muchas soluciones (a, b, c) para este problema, por ejemplo:

$$(3, 4, 5), (5, 12, 13), (7, 24, 25), (9, 40, 41), (11, 60, 61) \dots$$

Demuestre que hay infinitas soluciones para el problema.

Solución. En el siglo III a.C. Euclides, en sus *Elementos* y poco después Diofanto en su *Aritmética* demostraron que había infinitos triángulos rectángulos de lados enteros (a, b, c) determinados por las fórmulas:

$$a = m^2 - n^2, \quad b = 2mn, \quad c = m^2 + n^2,$$

con $m > n$ números enteros positivos. Esto es fácil de verificar ya que:

$$\begin{aligned} (m^2 - n^2)^2 + (2mn)^2 &= m^4 - 2m^2n^2 + n^4 + 4m^2n^2 \\ &= m^4 + 2m^2n^2 + n^4 = (m^2 + n^2)^2. \end{aligned}$$

Por ejemplo, la solución $(5, 12, 13)$ se obtiene de estas fórmulas haciendo $m = 3$ y $n = 2$. ¿Para cuáles n y m se obtiene la terna $(119, 120, 169)$?

Se puede demostrar (pero no lo haremos aquí) que toda solución (a, b, c) a nuestro problema está dada por las fórmulas de Euclides-Diofanto.

NÚMEROS RACIONALES VS. NÚMEROS IRRACIONALES

Consideremos que los números entre 0 y 1 pueden escribirse como fracciones decimales de la forma $0.a_1a_2a_3\dots$ de forma que todo a_n está definido. Por ejemplo, el 0 es el número $0.0000\dots$ con todo $a_n = 0$. El número $1/3 = 0.3333\dots$ con todo $a_n = 3$. Decimos que una fracción decimal $0.a_1a_2a_3\dots$ es *periódica* si existen dos números positivos M y N de manera que para todo número $n > M + N$ se tiene $a_n = a_{M+r}$ si $n - (M + r)$ es divisible entre N . Por ejemplo, para $1/3 = 0.3333\dots$ se tiene $M = 0$ y $N = 1$; para $21/990 = 0.021212\dots$ se tiene $M = 2$ y $N = 2$.

Demuestre las siguientes afirmaciones:

- Todo número entre 0 y 1 puede escribirse como una fracción decimal de la forma descrita arriba.
- Un número es racional si, y solamente si, la fracción decimal correspondiente es periódica.

III. Calculando lo desconocido

EL ANTICUARIO escocés Alexander Henry Rhind pasaba todos los años largas temporadas en Egipto por recomendación médica, pues padecía tuberculosis y el clima cálido y seco le era favorable. En 1858 compró en la ciudad de Luxor, a orillas del Nilo, un viejo papiro. Poco después se sabía que el escriba Ahmes había escrito este papiro en el año 1650 a.C. y que contenía una muestra importante de los conocimientos matemáticos de los egipcios.

Ahmes explica que él sólo copió el escrito de otros documentos más antiguos. El papiro está escrito en caracteres hieráticos y presenta una colección de 85 ejercicios matemáticos y ejemplos prácticos redactados en lenguaje oscuro. Lo que no se ha podido saber es a quiénes iban dirigidos estos problemas.

Uno de sus problemas dice: "La cantidad, el total y su séptima parte hacen 19." Sin duda, éste es el planteamiento de hacer más

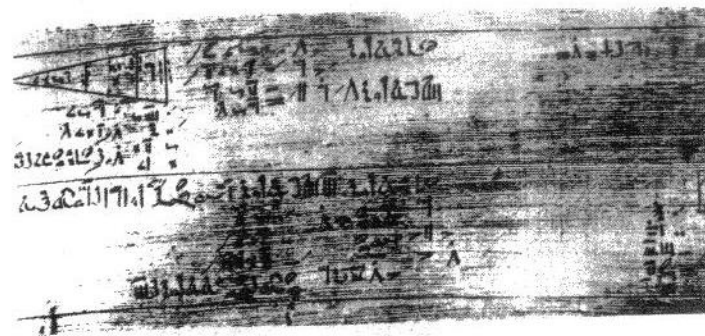


Figura III.1. Parte del Papiro de Rhind.

de 3000 años de un problema que un estudiante de secundaria en nuestros días escribiría así:

$$x + \frac{x}{7} = 19.$$

Por supuesto, cualquier estudiante sabrá rápidamente que la respuesta a este problema es que $x = 16\frac{5}{8}$. Los antiguos egipcios también obtuvieron esta respuesta, aunque la expresaban de una manera mucho más complicada. Otro ejemplo del papiro lo podemos ver en la figura III.2.

Otras culturas antiguas conocieron la solución de los problemas aritméticos que surgían en sus transacciones comerciales. Sin embargo, los siguientes avances conceptuales se dieron en la Grecia clásica. En el siglo III a.C., vivió en Grecia el matemático Diofanto. De su obra principal *Aritmética* sólo nos llegó una parte. Diofanto plantea problemas que deben resolverse encontrando soluciones en valores enteros. Este tipo de problemas se conocen ahora como *ecuaciones diofantinas*. Un problema típico de ecuaciones diofantinas es el problema del granjero que presentamos en el capítulo siguiente.

En la Edad Media fueron los hindúes y los musulmanes los que se encargaron de preservar el estudio del álgebra. En el año 825 Al-Juarizmi, el mismo sabio de Bagdad que había publicado el primer tratamiento posicional de los números, escribió el primer tratado de álgebra. El título de su obra fue *al-jabr al-muqabalah*, que quiere decir "el arte de unir las incógnitas para

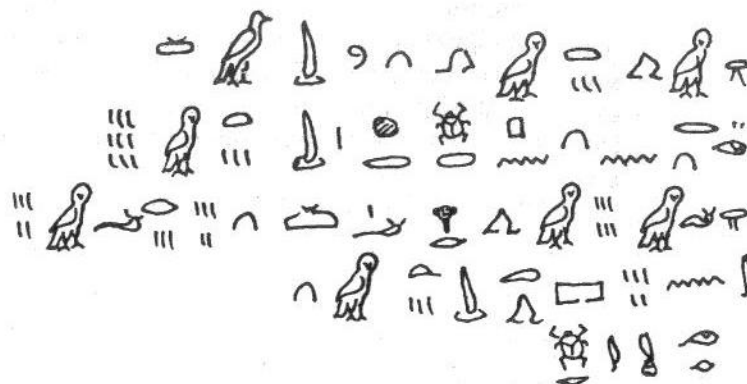


Figura III.2. Dice: "2/3 sumados y 1/3 restado hacen 10. Hallar 1/10 de este 10. El resultado es 1; el resto 9. Se añaden 2/3 de 9, es decir, 6; total 15. Una tercera parte es 5. Era 5 lo que se había restado: resto, 10." El problema "traducido" sería: $x + (2/3)x - (1/3)[x + (2/3)x] = 10$. ¿Puede el lector explicar la solución propuesta en el papiro? En el simbolismo egipcio, las piernas que andaban hacia la izquierda significaban "restar", hacia la derecha "sumar".

encontrar una cantidad". La palabra árabe *al-jabr*, que quiere decir "unir", dio origen a la palabra *álgebra*. En algunas lenguas romances (por ejemplo, en portugués) un *algebrista* puede ser un matemático o una persona que une huesos.

La solución de la ecuación cuadrática $ax^2 + bx + c = 0$ como todo estudiante de secundaria sabe es:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Este resultado se conoció desde tiempo de los griegos. En el siglo XVI, el álgebra recibe un gran impulso con el descubrimiento, por los matemáticos de la escuela italiana, de la resolución *por radicales* de la ecuación cúbica $ax^3 + bx^2 + cx + d = 0$. En este capítulo estudiaremos la solución de esta ecuación y contaremos la interesante historia de su descubrimiento.

El uso de letras para expresar cantidades desconocidas en una ecuación nos parece una idea muy natural hoy en día. Sin embargo, a pesar de la antigüedad de los problemas algebraicos, no fue sino hasta el siglo XVII cuando Descartes introdujo el uso

de las letras a, b, c para denotar números conocidos y el resto de las letras para los números desconocidos. Fue también Descartes el primero en utilizar x^2 en lugar de xx o x^5 en lugar de $xxxxx$.

LA JERARQUÍA DE LOS NÚMEROS

Las matemáticas se inician cuando el hombre comienza a contar. Contamos con los números 1, 2, 3, ... que por eso se llaman *números naturales*. Para el matemático italiano Giuseppe Peano, estos números representaban la única parte de las matemáticas que estaba dada de manera evidente en la naturaleza, todo lo demás tenía que pasar por el proceso de pensamiento y comprensión del hombre (esto es lo que resumió con su frase "Dios hizo los números naturales, lo demás es invención del hombre").

Ya en tiempos de los sumerios, de cuando datan los primeros registros de operaciones aritméticas, aparece la necesidad de *restar*. Con ello se toma conciencia de la necesidad de introducir los números negativos: -1, -2, -3, ... y usarlos con todos los derechos que los números naturales tenían (esto es, se resta y multiplica también con números negativos). Con los números naturales y los negativos se forman los *números enteros*.

Como hemos visto, para los pitagóricos el mundo se podía entender por medio de *números racionales*, esto es, por fracciones de la forma n/m con n y m números enteros. Este concepto recibió un duro golpe cuando descubrieron que el número $\sqrt{2}$ es un número irracional. Pero $\sqrt{2}$ no es excepcional, más bien es la regla. Estamos plagados de números irracionales: $\sqrt{3}$, $\sqrt[3]{2}$, $1 + \sqrt{2}$ y otros muchos números son irracionales (para una demostración de esto véase la última sección de este capítulo). Pero ¿cómo debemos entender estos números irracionales? Una respuesta la da el método de Newton que hemos usado para obtener el valor aproximado de $\sqrt{2}$. Para el número $\sqrt{2}$ existen números racionales r_n (con n número natural) de forma que el "error" de la aproximación $e_n = \sqrt{2} - r_n$ es cada vez más cercano a 0. Llegamos así a la siguiente definición.

Un número a se llama *número real* si existen números racionales r_n (con n número natural) de forma que $a - r_n$ se acerca

a 0 conforme n crece. Con esta definición, es claro que todos los números racionales son reales (para a número racional, elegimos simplemente $r_n = a$ para todo n). Los números irracionales son aquellos números reales que no son racionales.

Uno de los problemas a los que se enfrentaban los algebristas italianos del siglo XVI era que ecuaciones de la forma $x^2 + 1 = 0$ parecían no tener solución, puesto que todo número real elevado al cuadrado es positivo. Pero, por otra parte, era muy conveniente trabajar haciendo de cuenta que una solución para esta ecuación existiese. Fue el matemático italiano Cardano el primero en trabajar "como si la ecuación $x^2 + 1 = 0$ tuviera solución"; sin embargo pensaba que esta solución era un "número imposible" y por ello lo llamó *número imaginario*.

Los matemáticos estaban familiarizados con la idea de que no de toda operación entre números resultaba otro número. Por ejemplo, el resultado de dividir 1 entre 0 no da otro número, ya que si $a = 1/0$ fuera un número, entonces $0 = a \times 0 = 1$, lo que es absurdo. Pero trabajar con números imaginarios no provocaba problemas lógicos.

Llamemos i a una solución de la ecuación $x^2 + 1 = 0$. Éste no es un número como los que conocemos (los números reales) pues satisface que $i^2 = -1$. Sin embargo, adoptémoslo formalmente como un nuevo número. ¿Qué podemos hacer con él?

Dado un número real b podemos multiplicar b por i , este producto lo escribiremos como un nuevo número bi . Dado otro número real a podemos también sumar a con bi para formar el nuevo número $a + bi$. Al conjunto de los números de la forma $a + bi$ los llamaremos *números complejos*. En el conjunto de estos números también podemos sumar como sigue: dados dos números complejos $a + bi$ y $c + di$ tenemos:

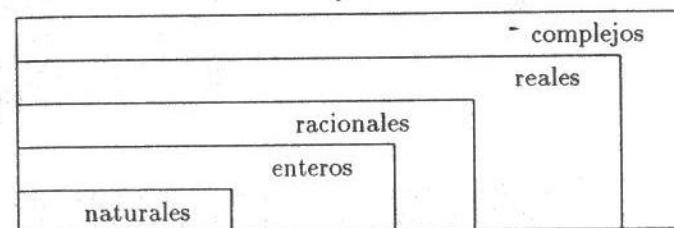
$$(a + bi) + (c + di) = (a + b) + (c + d)i$$

que es otro número complejo. También podemos multiplicar:

$$\begin{aligned} (a + bi)(c + di) &= ac + adi + bci + bdi^2 \\ &= ac + (ad + bc)i - bd \\ &= (ac - bd) + (ad + bc)i, \end{aligned}$$

que es otro número complejo. Obsérvese que hemos usado en la penúltima igualdad que $i^2 = -1$. Por ejemplo, $1 + 3i$ y $2 + (2/3)i$ son números complejos. Su suma es $3 + (11/3)i$ y su producto es $-1 + (20/3)i$. Los números reales son un subconjunto de los números complejos, pues dado un número real a podemos escribirlo como $a = a + 0i$.

Podemos establecer así la siguiente jerarquía de números:



Una manera sencilla de "visualizar" los números complejos es dibujarlos como puntos en un plano, de forma que el número $a + bi$ corresponde a un punto con coordenada a sobre el eje horizontal y coordenada b en el eje vertical. En la figura III.3 mostramos un plano coordenado como indicamos, donde hemos además marcado los dos números complejos $1 + 3i$ y $2 + (2/3)i$. Este plano se llama el *plano complejo*. El eje horizontal consta de todos los números reales. En el eje vertical quedan los números imaginarios de la forma ai , donde a es real.

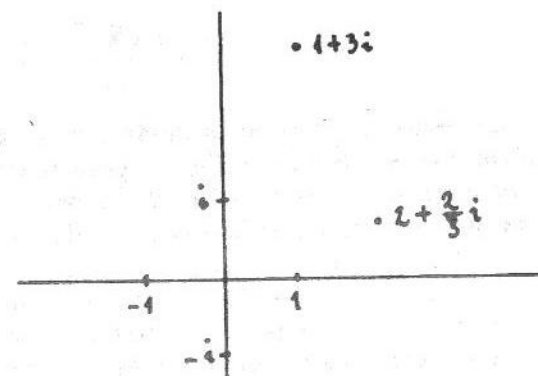


Figura III.3. El plano complejo.

El álgebra de los números complejos tuvo importante papel en el desarrollo de la teoría de ecuaciones y del álgebra toda. La intuición de Cardano respecto a su importancia quedó plenamente justificada cuando el genio matemático del siglo XIX, Carl Friedrich Gauss, demostró que toda ecuación de grado n tiene exactamente n soluciones entre los números complejos. Este profundo resultado se conoce como el *teorema fundamental del álgebra*.

ECUACIONES A LA ITALIANA

Nunca en la historia hubo un grupo más pintoresco de matemáticos que los de la escuela italiana que resolvió las ecuaciones cúbicas y cuárticas.

A finales del siglo XV, la mayoría de los matemáticos eran autodidactas. Se ganaban la vida trabajando en los primeros bancos o jugando a las cartas. Para dar popularidad a sus proezas de agilidad mental, se desafiaban en torneos públicos de resolución de problemas o complicados cálculos mentales.

En 1494, un sacerdote franciscano, Luca Pacioli, publicó un compendio de toda el álgebra conocida hasta sus días. Terminaba observando que los matemáticos aún no podían resolver las ecuaciones cúbicas por métodos algebraicos. Este reto fue parcialmente resuelto por Scipione del Ferro en Bolonia, que encontró la solución general de la ecuación $x^3 + ax + b = 0$, pero no publicó su resultado, seguramente para tener una ventaja sobre otros algebristas en las competencias públicas. Sin embargo, en sus últimos años, confió la solución de la ecuación a Antonio Fior, quien retó a un desafío matemático a otro gran algebrista, Nicolo Fontana, llamado Tartaglia, es decir, el tartamudo.

Tartaglia conocía ya entonces la solución de las ecuaciones de la forma $x^3 + ax^2 + b = 0$. Para su desafío, cada cual pidió al otro resolver problemas de ecuaciones cúbicas. Después de un tiempo en que ambos trabajaron intensamente, Tartaglia pudo resolver los problemas de Fior pero Fior no pudo con los de Tartaglia. Esto estableció a Tartaglia como el más grande calculista de Italia. Hasta que llegó Girolamo Cardano.

Cardano nació en 1501 en Pavia. Hijo natural de un médico y una viuda, tuvo una infancia desventurada. Después de algunos estudios de medicina, vivió en pequeñas poblaciones cam-



Figura III.4. Girolamo Cardano.

pesinas donde se dedicaba a jugar cartas, atender pacientes y escribir tratados sobre quiromancia, medicina y aritmética. Finalmente obtuvo una cátedra de medicina en la universidad y llegó a ser uno de los más renombrados médicos europeos.

En 1539, Cardano se acercó a Tartaglia para pedirle que le diera a conocer su solución de las ecuaciones cúbicas (de la forma $x^3 + ax^2 + b = 0$). Tartaglia se negó, pero tras prolongado intercambio de cartas, cedió. Tartaglia fue a Milán e hizo jurar a Cardano que nunca revelaría su secreto. Éste prometió que escribiría la solución de Tartaglia en lenguaje cifrado para que aún después de su muerte nadie la conociese. Sin embargo, una vez conocido el secreto, Cardano lo publicó en su libro *Ars Magna*. Había circunstancias mitigantes para esta traición: Cardano dio completo crédito a Tartaglia por su descubrimiento; por otra parte, el resultado que publicó no era exactamente el que le había comunicado Tartaglia, ya que Cardano encontró la expresión por radicales de la solución de la ecuación cúbica más general $ax^3 + bx^2 + cx = d$.

Tartaglia consideró imperdonable la traición de Cardano. Después de muchas disputas, se decidió dirimir el asunto en una justa matemática entre Tartaglia y Cardano, quien se hizo representar por su brillante alumno Ferrari. La justa terminó con el triunfo de Ferrari, lo que lo hizo el más connotado algebrista italiano. Posteriormente, Ferrari llegaría a resolver por radicales las ecuaciones de cuarto grado.

Cardano publicó más de 130 libros de medicina, matemáticas, astrología, ajedrez, juegos de azar, filosofía, historia y otros temas. Sin embargo, los últimos años de su vida no fueron afortunados. En 1557, su hijo favorito asesinó a su esposa. El asesino fue ejecutado tres años más tarde. Fue tanta la tristeza de Cardano que abandonó su cátedra en Milán y fue a vivir a Bolonia. Ahí su segundo hijo contrajo tremendas deudas de juego y terminó en la cárcel en varias ocasiones. En 1570, Cardano fue acusado de herejía y encarcelado. El papa le permitió salir de prisión y le concedió una pequeña pensión. Sin embargo, a Cardano no le fue permitido volver a enseñar en la Universidad hasta su muerte, en 1576.

* LA SOLUCIÓN DE CARDANO A LA ECUACIÓN CÚBICA

Deseamos encontrar las soluciones de la ecuación $ax^3 + bx^2 + cx + d = 0$, cuando a no es 0. Sabemos que tendremos que encontrarlas entre los números complejos. Procederemos en forma parecida a como lo hizo Cardano en el siglo XVI.

Hay tres números complejos que satisfacen la ecuación $x^3 = 1$, uno de ellos es, por supuesto, el 1. Los otros números son $\omega = \frac{1}{2}(-1 + \sqrt{3}i)$ y $\omega^2 = \frac{1}{2}(-1 - \sqrt{3}i)$ y se encuentran en el plano complejo como se indica en la figura III.5 (en general tenemos que si $a+bi$ es un número complejo entonces $(a+bi)^3 = (a^3 - 3ab^2) + (3a^2b - b^3)i$).

Tomemos el número complejo ω . Es sencillo verificar que para cualesquiera números A, B, C se tiene que:

$$\begin{aligned}(A + B + C)(A + B\omega + C\omega^2)(A + B\omega^2 + C\omega) \\ = A^3 + B^3 + C^3 - 3ABC.\end{aligned}$$

En un momento haremos uso de esta identidad.

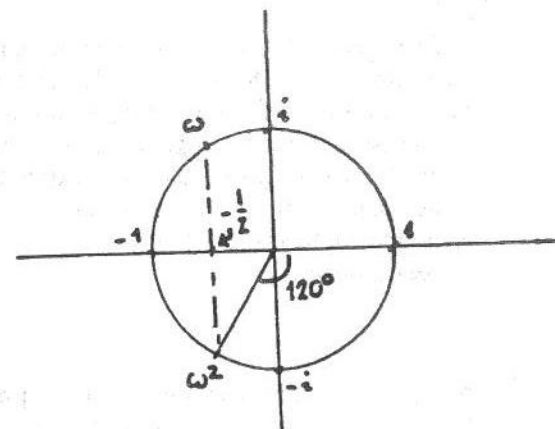


Figura III.5. Las raíces cúbicas de la unidad.

Multipliquemos la ecuación que queremos resolver por $27a^2$, de forma que obtenemos:

$$27a^3x^3 + 27a^2bx^2 + 27a^2cx + 27a^2d = 0.$$

Manipulando un poco esta ecuación, podemos reescribirla:

$$\begin{aligned}(3ax + b)^3 - 9ax(-3ac + b^2) + 27a^2d - b^3 &= 0, \\ (3ax + b)^3 + (27a^2d - 9abc + 2b^3) - 3(3ax + b)(-3ac + b^2) &= 0.\end{aligned}$$

Hagamos $A = 3ax + b$, si podemos además encontrar números B y C con la propiedad de que $B^3 + C^3 = (27a^2d - 9abc + 2b^3)$ y también $BC = -3ac + b^2$, entonces tendríamos que $A^3 + B^3 + C^3 - 3ABC = 0$ y podríamos usar la identidad anterior.

Para calcular B y C , observemos que el sistema:

$$\begin{aligned}B^3 + C^3 &= (27a^2d - 9abc + 2b^3), \\ B^3C^3 &= (-3ac + b^2)^3,\end{aligned}$$

implica que B^3 y C^3 son las raíces de la ecuación cuadrática:

$$t^2 - (27a^2d - 9abc + 2b^3)t + (-3ac + b^2)^3 = 0,$$

que sabemos muy bien resolver.

Una vez calculados B y C , podemos hacer uso de que:

$$(A + B + C)(A + B\omega + C\omega^2)(A + B\omega^2 + C\omega) = A^3 + B^3 + C^3 - 3ABC = 0,$$

para obtener que alguno de los factores $A + B + C$, $A + B\omega + C\omega^2$ o bien $A + B\omega^2 + C\omega$ es cero. Así obtenemos que las tres soluciones de la ecuación cúbica son de la forma $x = (A - b)/3a$, donde $A = -(B + C)$, $A = -(B\omega + C\omega^2)$ o bien $A = -(B\omega^2 + C\omega)$.

Por ejemplo, consideremos la ecuación $x^3 + 24x - 56 = 0$. Primero, debemos hallar los números B^3 y C^3 como raíces de la ecuación cuadrática $t^2 - 27 \times (-56)t + (-3 \times 24)^3 = 0$, esto es:

$$B^3 = -\frac{27 \times 56 + \sqrt{27^2 \times 56^2 + 4 \times 27 \times 24^3}}{2} \\ = -(27 \times 28 + 4 \times 3^5) = -2^6 \times 3^3$$

y:

$$C^3 = -27 \times 28 + 4 \times 3^5 = 2^3 \times 3^3.$$

Luego, tenemos que $B = -12$ y $C = 6$. Las soluciones de la ecuación cúbica son: 2 , $4\omega - 2\omega^2$ y $4\omega^2 - 2\omega$.

EL MATRIMONIO DEL ÁLGEBRA Y LA GEOMETRÍA

Consideremos una ecuación simple como $x - 1 = 0$. Decimos que la ecuación toma el valor 0 en el número 1 (o cuando la evaluamos en 1), el valor 0.5 en el número 1.5, el valor 1 en el número 2, etcétera. Para expresar esto más claramente, podemos escribir la igualdad $y = x - 1$ y decir que, para esta ecuación, $y = 0$ cuando $x = 1$, que $y = 0.5$ cuando $x = 1.5$, etcétera. Esta información podemos expresarla también gráficamente ubicando los puntos $(1, 0)$, $(1.5, 0.5)$ o $(2, 1)$ en el plano coordenado $x - y$ (figura III.6).

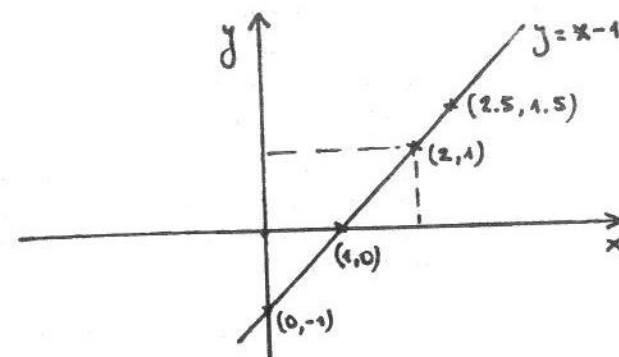


Figura III.6. Puntos sobre la recta $y = x - 1$.

En este plano los puntos de la forma (x, y) tales que $y = x - 1$ determinan la gráfica de la ecuación $x - 1$. Esto nos permite "ver" la ecuación $x - 1$ como una recta en el plano. De igual manera los puntos (x, y) de forma que $y = x^2 - 1$ forman una curva en el plano que se llama *parábola* y que constituye la gráfica de la ecuación $x^2 - 1 = 0$.

Esta simple y bella idea tardó relativamente mucho tiempo en surgir. Fue sólo en el siglo XVII que Pierre Fermat y René Descartes, dos matemáticos franceses, trabajando independientemente, tuvieron la idea de "dar coordenadas" a las ecuaciones. El primero en publicar sus ideas fue René Descartes y a él se le considera generalmente el padre de la rama de las matemáticas que se llama *geometría analítica*.

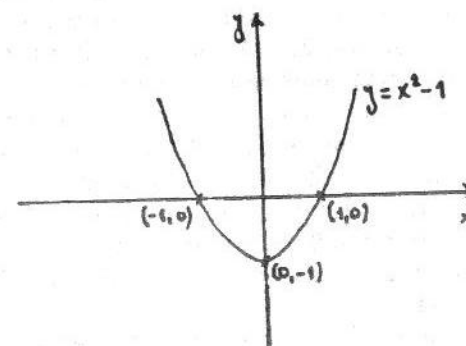


Figura III.7. Gráfica de la parábola $y = x^2 - 1$.

En 1637 Descartes publicó su importante obra filosófica *El discurso del método* que pretendía dar las bases para una revolución en el pensamiento filosófico y científico. Como apéndice a este libro añadió un capítulo titulado *Géométrie*, que comenzaba diciendo:

Cualquier problema de geometría se puede reducir a términos tales que el conocimiento de las longitudes de ciertas líneas rectas baste para resolver el problema [...] y no dudaré en introducir estos términos aritméticos en la geometría.

Los “términos aritméticos” de Descartes no eran otros que los números que indicaban las coordenadas (x, y) de las gráficas de ecuaciones. Desgraciadamente, su trabajo se consideró difícil de leer y no tuvo una acogida inmediata por los matemáticos de la época. Pero eso llegó al poco tiempo.

Descartes había emigrado a Holanda en 1628, donde permaneció hasta un año antes de su muerte en Suecia, en 1650. El año de su salida de Holanda, el matemático Frans van Schooten de Amsterdam preparó una edición comentada de la *Géométrie* y añadió un capítulo donde explicaba con sus propias palabras lo que entendía de la teoría de Descartes. Ésta fue la edición que se hizo conocida en el mundo científico y que tuvo una gran influencia en los desarrollos futuros de las matemáticas. Por ejemplo, tanto Newton como Leibniz reconocieron su influencia en el desarrollo del cálculo diferencial e integral.

En los primeros años, el uso de los ejes coordenados $x - y$ se hacía de maneras diferentes a las actuales. Por ejemplo, los ejes no se dibujaban siempre perpendiculares, no era frecuente que se usaran números negativos. La versión definitiva de los sistemas coordenados como los usamos hoy día se debe a Isaac Newton, quien en su obra *Enumeratio linearum tertii ordinis* (en esa época los trabajos científicos se escribían en latín), calculó las gráficas de muchas ecuaciones de tercer grado.

El poder de esta unión entre el álgebra y la geometría no ha sido igualado por ningún otro concepto en el campo de las matemáticas. La mejor parte de esas dos ramas actúa como una fuente inagotable de inspiración. Por una parte, las ecuaciones algebraicas abstractas pueden ser visualizadas por medio de figuras geométricas; por otro lado, al estudio de las figu-



Figura III.8. René Descartes.

ras geométricas se le puede aplicar los poderosos métodos del álgebra.

Una muestra del vigor de este método es el siguiente: los griegos estudiaron gran variedad de curvas que eran siempre definidas por medio de ingeniosos métodos, por ejemplo, la elipse es la intersección de un cono con un plano, como se muestra en la figura III.9.

Todas las curvas cónicas (que se obtienen como la intersección de un cono y un plano) corresponden a lugares geométricos de puntos (x, y) que satisfacen ecuaciones de la forma $ax^2 + by^2 = c$ con a, b, c números reales. Pero ¿qué sucede si se desea estudiar ecuaciones de grado más alto, como $y^2 = x^3 + x + 1$ o bien $y^3 = x^7 - x^2$? ¡Difícilmente los métodos griegos nos describirán estos lugares geométricos como intersección de figuras complicadas! De hecho, ecuaciones de esta clase más complicada aparecen en la solución de problemas planteados

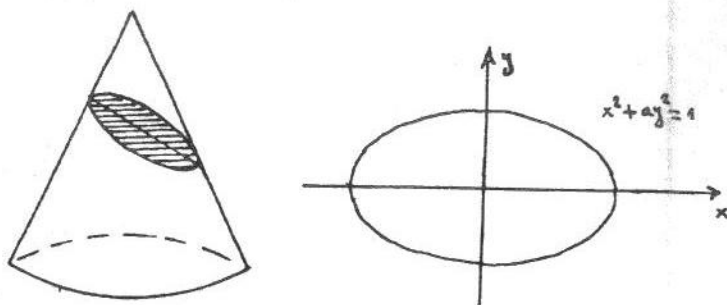


Figura III.9. La elipse como intersección de un cono y un plano. A la derecha, la elipse como el lugar geométrico de los puntos (x, y) que satisfacen $x^2 + ay^2 = 1$, con a un número real positivo.

por los mismos griegos. Consideremos por ejemplo el siguiente problema clásico, que según la leyenda fue planteado por el Oráculo de Apolo en Delfos: *construir un cubo del doble de volumen que el cubo del altar de Apolo.*

Para conocer el volumen de un cubo basta con conocer la longitud de un lado. Si a es la longitud del lado del cubo en el altar de Apolo, el volumen del cubo es $x = a^3$. Se pregunta por la longitud y tal que $y^3 - 2x = 0$, que resulta ser una ecuación de grado 3. Los griegos intentaron resolver este problema por medio de regla y compás, es decir, se trataba de obtener un segmento de recta de longitud y trazando solamente rectas y círculos que se pueden construir si se tiene dado un segmento de longitud a . Por este método el problema no tiene solución, pero los griegos no se dieron cuenta de ello. No fue sino hasta el siglo pasado con el surgimiento de la teoría de Galois que se vio claramente la imposibilidad de encontrar la solución del problema con el uso exclusivo de regla y compás.

Curvas de grado 3 aparecen en otros muchos problemas matemáticos. Por ejemplo, en la solución de otro problema clásico griego que pide encontrar un ángulo de la tercera parte de un ángulo dado, se obtiene la ecuación: $y = 4x^3 - 3x$ (¿por qué?). En la demostración del llamado último teorema de Fermat aparecen ecuaciones de la forma $y^2 = x^3 + Ax + B$, como veremos en el capítulo siguiente. El aspecto general de estas curvas se muestra en la figura III.10.

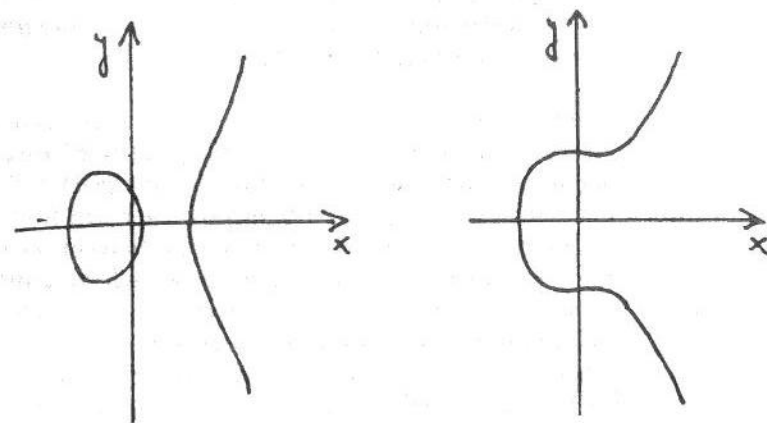


Figura III.10. Las formas posibles de la gráfica de curvas $y^2 = x^3 + Ax + B$ con $4A^3 + 27B^2 \neq 0$.

ENTEROS ALGEBRAICOS

Considere un polinomio cuadrático de la forma $p(x) = x^2 + bx + c$ con b y c números enteros. Las soluciones del problema $p(x) = 0$ se llaman *enteros algebraicos de grado 2*. Demuestre lo siguiente:

- Los números enteros son enteros algebraicos de grado 2.
- Si n, m y D son enteros, entonces el número $n + m\sqrt{D}$ es un entero algebraico de grado 2. Llamemos $\mathbb{Z}[\sqrt{D}]$ al conjunto de los números de esta forma.
- Muestre que $\mathbb{Z}[\sqrt{D}]$ es cerrado bajo las operaciones de suma y multiplicación usuales. Esto es, si $n + m\sqrt{D}$ y $n' + m'\sqrt{D}$ son dos números en $\mathbb{Z}[\sqrt{D}]$ entonces también están $(n + m\sqrt{D}) + (n' + m'\sqrt{D})$ y $(n + m\sqrt{D})(n' + m'\sqrt{D})$.

Observemos que podemos tomar cualquier número D para el caso. En particular podemos tener $D < 0$. Por ejemplo, para el caso $D = -1$ se obtienen números complejos de la forma $n + mi$ con n y m enteros, estos números se llaman *enteros gaussianos* en honor a Gauss.

*d) Un número complejo u se llama un *entero algebraico* si existe un polinomio $x^n + a_{n-1}x^{n-1} + \dots + a_0$ con coeficientes a_0, a_1, \dots, a_{n-1} enteros de forma que $u^n + a_{n-1}u^{n-1} + \dots + a_0 = 0$.

Demuestre que un entero algebraico real que no es un número entero es entonces un número irracional.

Demostración de d). Sea u un entero algebraico real que no es entero. Consideremos el polinomio $p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$ con coeficientes enteros de forma que $p(u) = 0$. Para obtener una contradicción, supongamos que u es un número racional.

Como u no es entero, existe un número entero r tal que $r < u < r+1$. También existe un entero k de forma que ku^i es entero para todo i . Construyamos k : como u, u^2, \dots, u^n son números racionales, entonces existen números enteros k_1, k_2, \dots, k_n tales que $k_1u, k_2u^2, \dots, k_nu^n$ son números enteros. Si elegimos $k = k_1k_2 \dots k_n$, entonces ku, ku^2, \dots, ku^n son números enteros. Pero entonces, $ku^{n+1} = -k(a_{n-1}u^n + \dots + a_0u) = a_{n-1}(k \cdot u^n) + \dots + a_0(ku)$ es suma de productos de números enteros y es por lo tanto un número entero. De la misma manera se muestra que ku^i es entero para toda i . Podemos elegir el entero m más pequeño con la propiedad de que $m \cdot u^i$ es entero para toda i .

Formamos el número $m' = m(u - r)$, que es un entero positivo más pequeño que m . Calculamos para cualquier entero positivo i :

$$m'u^i = m(u - r)u^i = mu^{i+1} - rmu^i,$$

que es la diferencia de dos enteros y por lo tanto un entero. Esto contradice la forma en que se eligió m . \square

NOTA: Se puede demostrar que dados dos enteros algebraicos u y v , entonces la suma $u + v$ y el producto uv son también enteros algebraicos.

IV. La historia en el margen de un libro

"Está bien", dijo Simon y tomó aire.

"Mi pregunta es ésta:

¿es correcto el último teorema de Fermat?"

El demonio tragó saliva.

Era la primera vez que su aire de seguridad se debilitaba.

El demonio y Simon Flagg, ARTHUR PORGES

DIOFANTO DE ALEJANDRÍA fue un prominente matemático griego que probablemente vivió un par de siglos antes de nuestra era. Nada se conoce sobre su vida salvo una rima que apareció en una colección de problemas matemáticos griegos:

La juventud de Diofanto duró una sexta parte de su vida. Se dejó crecer la barba después de un doceavo más. Al pasar un séptimo más de su vida, se casó y cinco años después tuvo un hijo. El hijo vivió exactamente la mitad que Diofanto, que murió cuatro años después que su hijo. Todos estos son los años que vivió Diofanto.

Por supuesto, el problema es encontrar el número de años que vivió Diofanto. El problema es muy sencillo. Se puede expresar así: si x es el número de años que vivió Diofanto, entonces:

$$x = \frac{x}{6} + \frac{x}{12} + \frac{x}{7} + 5 + \frac{x}{2} + 4,$$

que se reduce a $3x/28 = 9$ y finalmente la respuesta es 84 años.

Diofanto estudiaba problemas cuyas soluciones debían ser dadas en números enteros. El problema anterior es un ejemplo muy sencillo, el *problema del granjero* que presentamos más adelante es otro ejemplo un poco más complicado. Las ecuaciones que se plantean en estos problemas se conocen como *ecuaciones diofantinas* y constituyen uno de los temas principales de la rama de las matemáticas conocida como *teoría de los números*.

Diofanto conocía muy bien el teorema de Pitágoras, pero se preguntaba por aquellos triángulos rectángulos cuyos lados

tienen longitud entera. Como hemos visto en el capítulo II, en su libro *Aritmética* demostró que hay infinitas soluciones (a, b, c) para este problema dadas por las fórmulas:

$$a = m^2 - n^2, \quad b = 2mn, \quad c = m^2 + n^2,$$

con $m > n$ números enteros positivos.

A mediados del siglo XVII, el matemático aficionado Pierre Fermat leía una copia de la *Aritmética* de Diofanto y se preguntaba para qué números $n \geq 3$ era posible hallar tripletas de números enteros (a, b, c) que satisficieran la ecuación $a^n + b^n = c^n$. En el margen del libro escribió: "la ecuación $x^n + y^n = z^n$ no tiene soluciones enteras para $n > 2$. Encontré una demostración maravillosa, pero el margen de este libro es muy pequeño para escribirla."

Por supuesto, Fermat se refería a soluciones en enteros positivos y excluía las soluciones triviales del problema como $x = 1, y = 0, z = 1$ o $x = 0, y = 1, z = 1$. La prueba que supuestamente Fermat había encontrado nunca fue publicada y no fue hallada entre los papeles que dejó a su muerte. Lo que sí publicó fue una demostración elemental de que la ecuación $x^4 + y^4 = z^4$ no tiene soluciones enteras. La mayoría de los matemáticos han pensado desde entonces que en realidad Fermat no tenía una prueba del caso general, y que la demostración que creyó tener cuando escribió la nota al margen de su libro estaba equivocada. Sin embargo, la afirmación se conoce como *El último teorema de Fermat*.

Al paso de los años algunos de los matemáticos más importantes trataron vanamente de probar este resultado, pero en su esfuerzo obtuvieron muchas otras cosas. Euler demostró que el caso $n = 3$ tampoco tenía soluciones. Para el principio del siglo XIX, ya se sabía que el teorema de Fermat era también correcto para los casos $n = 5$ y $n = 7$. Para 1964, el teorema había sido demostrado para los números pares y los números menores que 25013. Intentos de resolver este problema dieron inicio a ramas completas de las matemáticas como la teoría analítica de los números y la teoría de ideales en anillos.

Fue sólo en 1983 que fue realizada la siguiente contribución importante. El joven matemático alemán Gerhard Faltings demostró que las ecuaciones de la forma $x^n + y^n = z^n$ con $n > 2$ pueden tener a lo más un número finito de soluciones que no

son múltiplos unas de otras (observe que si (a, b, c) es solución de la ecuación, entonces también (at, bt, ct) es una solución para cualquier entero t). De hecho, Faltings demostró este resultado como parte de un teorema más general que había sido conjeturado tiempo atrás por Mordell. Si bien esto no resolvía el problema de Fermat, mostraba ya que había nuevas herramientas matemáticas que podían decir cosas importantes sobre él.

Hubo tantos intentos fallidos de demostrar el teorema que cuando finalmente la solución llegó, fue una sorpresa para todo mundo. Tal es la fama de este problema, que en los periódicos del mundo entero se informó de la solución al problema. En el verano de 1993, Andrew Wiles, profesor de Cambridge, Inglaterra, anunció que tenía una demostración contenida en un manuscrito de más de 200 páginas. Wiles había pasado los últimos 7 años de su vida trabajando en una línea que creía lo llevaría a la prueba del teorema de Fermat.



Figura IV.1. Andrew Wiles.

Poco después, un panel de matemáticos de varias universidades encontró una falla en la demostración. Uno de los argumentos de la prueba de Wiles estaba incompleto, y por lo tanto

toda la prueba. Wiles pensó que podría reparar fácilmente el problema encontrado, pero no fue así. Pasaron varios meses y sólo en octubre de 1994, con ayuda del matemático Richard Taylor, pudo evitar el uso del resultado equivocado y con ello la demostración quedó completada.

La demostración de Wiles del último teorema de Fermat es una muestra del grado de refinamiento y profundidad que han alcanzado las matemáticas. Mientras que el problema de Fermat requiere una línea para escribirse y cualquier estudiante de secundaria puede entender la afirmación, la demostración requiere técnicas y resultados de álgebra y geometría que se han ido desarrollando a lo largo de siglos. En palabras de Wiles:

Trataba de encontrar patrones generales. Trataba de hacer cálculos que me explicaran un pequeño resultado de matemáticas, luego intentaba hacerlos encajar en mi idea conceptual de ciertas ramas de las matemáticas. A veces esto quería decir que iba a ver cómo se había hecho algo parecido en un libro; a veces había que modificar unas pocas cosas y hacer unos pocos cálculos más. Pero a veces me daba cuenta de que no se había hecho nunca antes algo parecido y tenía que encontrar algo que fuera completamente nuevo.

Barry Mazur, un matemático que ha contribuido con importantes ideas en su campo, describió así sus impresiones de las primeras conferencias de Wiles en Cambridge sobre la demostración del teorema de Fermat:

Nunca había estado antes en una serie de conferencias así. Lo que era único en estas conferencias eran las maravillosas ideas, cuántas nuevas ideas eran presentadas y el dramatismo que se incrementaba. Estábamos en suspenso hasta el final.

Porque Wiles no había anunciado lo que pretendía demostrar: el título de la serie de conferencias era más bien técnico, pero algunos expertos ya sospechaban hacia donde iba. Lo que Wiles en realidad demostró era la validez de la conjetura de Taniyama-Shimura (matemáticos japoneses) que se refiere a las propiedades de ciertas curvas en el espacio y que puede definirse mejor como un problema de *geometría algebraica*. La conjetura fue enunciada a mediados de siglo, pero pasó desapercibida durante cierto tiempo hasta que el matemático francés

André Weil probó una serie de resultados que mostraban que la conjetura era plausible. Demostrarla era suficiente para Wiles, pues Ken Ribet (matemático estadounidense) había demostrado años antes que la validez de la conjetura implicaba el teorema de Fermat. En dos palabras, el razonamiento es como sigue: si una solución entera (a, b, c) existiera para el problema $x^n + y^n = z^n$ con $n > 2$, entonces se podría considerar la curva $y^2 = x(x + a^n)(x - b^n)$ que no cumple con las propiedades que debería satisfacer de acuerdo a la conjetura de Taniyama-Shimura.

Una curva de la forma $y^2 = x(x + a^n)(x - b^n)$ se llama *curva elíptica de Frey* y es un caso especial de las curvas cúbicas $y^2 = x^3 + Ax + B$ que han sido estudiadas en matemáticas desde la Antigüedad. Algunos aspectos de estas curvas los hemos tratado en el capítulo III.

Con más precisión: la conjetura de Taniyama-Shimura afirma que toda curva elíptica es "parametrizable" por funciones modulares. Esto lo podemos entender como una generalización (muy grande) del hecho de que los puntos (x, y) en un círculo de la forma $x^2 + y^2 = 1$ pueden escribirse ("parametrizarse") así: $(x, y) = (\cos t, \sin t)$, donde t corre sobre los números reales y $\sin t, \cos t$ denotan las funciones seno y coseno respectivamente. Ribet demostró que una curva de Frey $y^2 = x(x + a^n)(x - b^n)$ no es "parametrizable" por funciones modulares (es decir, no satisface la conjetura de Taniyama-Shimura). Finalmente, Wiles demostró el siguiente caso especial de la conjetura de Taniyama-Shimura: la curva elíptica de la forma $y^2 = x(x + A)(x + B)$ es "parametrizable" por funciones modulares en caso de que A y B sean enteros tales que $AB(B - A)$ sea divisible entre 16. ¿Por qué es esto suficiente?

Consideremos una supuesta solución entera a, b, c de la ecuación $a^n + b^n = c^n$. Por el resultado de Euler sabemos que $n > 3$; además se sabía que n debería ser non. Haciendo $A = -a^n$ y $B = b^n$, entonces se tiene $AB(B - A) = -a^n b^n (a^n + b^n) = -(abc)^n$. Si los números a y b son nones, entonces a^n y b^n son nones también, lo que implica que c^n es par y por lo tanto también c debe ser par. En consecuencia, alguno de los tres números a, b o c es par y por lo tanto $AB(B - A)$ es divisible entre $2^5 = 32$ (y por lo tanto entre 16). Esto muestra que la curva de Frey es "parametrizable" por funciones modulares,

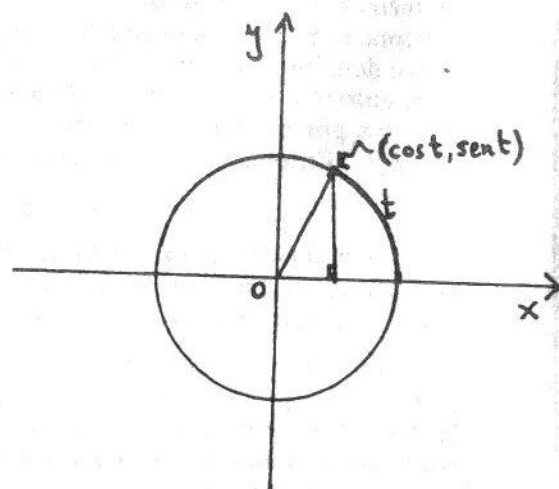


Figura IV.2. "Parametrización" del círculo unitario.

contrariamente a lo que Ribet demostró. Por lo tanto no existen curvas de Frey, o lo que es lo mismo, no existen soluciones (a, b, c) de $a^n + b^n = c^n$.

Los detalles de la demostración del último teorema de Fermat (que debería llamarse ahora teorema de Fermat-Wiles), sólo pueden ser comprendidos por algunos matemáticos especialistas. Sin embargo, podemos decir que su demostración constituye uno de los mayores logros del intelecto humano.

NÚMEROS PRIMOS

Como sabemos, un número entero positivo se llama *primo* si sólo es divisible por 1 y por sí mismo. Los números primos gozan de gran popularidad en las matemáticas desde el tiempo de los griegos clásicos. El estudio de la distribución y propiedades de los números primos forma una de las partes más bellas y profundas de las matemáticas: la *teoría de los números*.

La misma definición de número primo nos da una receta para construirlo: todo número que sea propiamente divisible por 2 no es primo, por ejemplo, 4, 6, 8, 10, ... no son primos; todo número que es propiamente divisible por 3 tampoco es primo,

por ejemplo, 6, 9, 12, 15, ... no son primos; y así podemos seguir eliminando los múltiplos de 4, 5, 6, etcétera. Este procedimiento se conoce como la *criba de Eratóstenes* en honor al matemático griego que la descubrió en el siglo III a.C. Usándolo obtenemos (figura IV.3) los números primos entre 1 y 100.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Figura IV.3. Los números primos entre 1 y 100 son aquellos que no quedan tachados.

Un resultado fundamental nos dice que los números primos pueden verse como los ladrillos constructores de los demás números. En efecto, tenemos el siguiente resultado, conocido como el *teorema fundamental de la aritmética*.

Teorema. *Todo número entero positivo puede ser escrito en forma única como producto de números primos.*

Demostración. Mostraremos primeramente que todo número se puede escribir como producto de números primos. Tomemos un número n . Si n es primo entonces $n = n$ es la factorización buscada (producto con un solo factor). Supongamos entonces que n no es primo. Entonces existe un número m_1 que no es n ni 1 y que divide a n . Esto es, $n = m_1 m_2$, donde m_2 es otro número que no es 1 ni n . Si tanto m_1 como m_2 son primos, ya acabamos. Si no, entonces podemos encontrar números m_{21}, m_{22}, m_{23} que no son 1 y tal que $n = m_1 m_2 = m_{21} m_{22} m_{23}$. Si todos los números m_{21}, m_{22}, m_{23} son primos o 1, entonces hemos terminado. Si no, podemos encontrar cuatro números $m_{31}, m_{32}, m_{33}, m_{34}$ que no son 1 y tales que $n = m_{21} m_{22} m_{23} = m_{31} m_{32} m_{33} m_{34}$. Claramente este proceso no puede seguir indefinidamente puesto que el número n es cuando más el producto de $n/2$ números enteros que no son 1. Luego el proceso se detiene y n es el producto de números primos.

Para demostrar la unicidad de la factorización, supongamos que el número n puede escribirse como producto de los números primos p_1, p_2, \dots, p_s y también de los números primos q_1, q_2, \dots, q_t , de manera que tenemos:

$$p_1 p_2 \dots p_s = n = q_1 q_2 \dots q_t.$$

Puede demostrarse que un número primo p tiene la siguiente propiedad: en caso de que p divida a ab , entonces o bien p divide a a o bien divide a b . Así en la igualdad de arriba tenemos que p_1 debe dividir a alguno de los números q_1, q_2, \dots, q_t , digamos que p_1 divide a q_1 . Como q_1 es primo, entonces $p_1 = q_1$. Dividiendo la igualdad entre p_1 obtenemos $p_2 \dots p_s = q_2 \dots q_t$. Continuamos el proceso hasta demostrar que $s = t$ y que los números primos p_1, p_2, \dots, p_s son exactamente q_1, q_2, \dots, q_t . Lo que termina la demostración del teorema. \square

Lo que el teorema no dice es cómo obtener la factorización de un número en sus factores primos, y éste puede ser un problema difícil. Por ejemplo, la factorización de 83080 en números primos es como sigue:

$$85008 = 2 \times 2 \times 2 \times 2 \times 5 \times 7 \times 11 \times 13.$$

¿Puede el lector calcular la factorización en primos del número 85009? En el capítulo V veremos cómo sacar jugo de las grandes dificultades del problema de factorización en números primos.

Conocemos muchos números primos: 2, 3, 5, 7, 11, 13, ... ¿hasta dónde podemos seguir? Uno de los primeros resultados matemáticos de la Antigüedad de los que se tiene noticia es la demostración por Euclides de la existencia de infinitos números primos. Esta demostración es todavía hoy una muestra de la elegancia que puede tener un argumento matemático.

Teorema. *Existen infinitos números primos.*

Demostración. Supongamos que esto no fuese cierto y que sólo hubiese un número finito de números primos, digamos p_1, p_2, \dots, p_s . Formamos un nuevo número de la siguiente manera:

$$n = p_1 p_2 \dots p_s + 1,$$

es decir, se forma el producto de todos los números primos y se le suma 1. Según el anterior teorema, este número n es el producto de números primos. Puesto que conocemos todos los primos, entonces n debe ser divisible por alguno de p_1, p_2, \dots, p_s , digamos por p_1 . Esto es, existe otro número m de forma que $n = p_1 m$. Entonces podemos reescribir estas igualdades como:

$$p_1(m - p_2 \dots p_s) = 1.$$

Pero obviamente, 1 no puede dividirse por p_1 . Este absurdo muestra que la suposición que hicimos es insostenible. Por lo tanto, hay infinitos números primos. \square

Es interesante notar que el teorema de Euclides nos dice que hay infinitos primos pero no nos dice cuáles son. Sin embargo, la demostración nos da una posible indicación de cómo construir nuevos números primos a partir de otros que ya conocemos: tomamos los primeros números primos que conocemos, los multiplicamos y sumamos 1, ¿el resultado es un nuevo número primo? Veamos: $2 \times 3 + 1 = 7$ que es primo; $2 \times 3 \times 5 + 1 = 31$ es primo; $2 \times 3 \times 5 \times 7 + 1 = 211$ es primo, vamos bien; $2 \times 3 \times 5 \times 7 \times 11 + 1 = 2311$ es primo, esto sigue muy bien; $2 \times 3 \times 5 \times 7 \times 11 \times 13 + 1 = 30031 = 59 \times 509$ no fue un número primo y el procedimiento falló. ¿Por qué falla? Porque la prueba de Euclides indica que el producto de los primeros números primos más 1 es un número que no puede ser divisible entre los números primos que usamos como factores, pero no dice que sea un número primo, por ejemplo, el número $2 \times 3 \times 5 \times 7 \times 11 \times 13 + 1 = 30031$ no puede ser divisible por 2, 3, 5, 7, 11 o 13 y en efecto no lo es, pero es divisible entre 59 y 509, que son números primos mayores que los que utilizamos como factores. Entonces, esta aplicación de la prueba de Euclides sí produce nuevos números primos, pero los deja ocultos.

El uso de las computadoras ha permitido calcular números primos muy grandes. Por ejemplo, en 1968 se demostró que el número $2^{11213} - 1$ es primo. En la figura adjunta vemos la expansión decimal de este número (que tiene 3376 cifras) obtenida por John MacKay de Urbana, EUA y un matasellos de correos conmemorativo de este descubrimiento. En 1986, David Slowinski trabajando con una supercomputadora Cray X-MP/24 demostró que el número $2^{216091} - 1$ es primo. ¡Este número tiene 65050 cifras!

647 38 59

211213-1
IS PRIME

Para cada número n , llamamos $P(n)$ al número de números primos que son menores o iguales que n . Así, $P(10) = 4$, $P(20) = 8$, $P(30) = 10$ y $P(100) = 25$. En la siguiente tabla (calculada con la ayuda de una computadora) vemos cómo varía el cociente $P(n)/n$ conforme n crece:

La observación de Legendre y Gauss era que conforme n crece, el número $P(n)/n$ se parece más y más a $1/\ln n$, donde $\ln n$ es el valor que toma en el número n la *función logaritmo natural* $\ln x$.

- i) $\ln 1 = 0$;
- ii) dados dos números a, b , se tiene que $\ln ab = \ln a + \ln b$;
- iii) hay un único número e tal que $\ln e = 1$. Este número e vale aproximadamente 2.71828182 y es igual a la suma:

donde los puntos suspensivos quieren decir que la suma continúa indefinidamente, en cada paso agregándose el número racional $1/1 \times 2 \dots \times n$, de forma que el valor total de la suma cada vez se acerca más al valor del número e . Esta expresión fue calculada por Euler en 1748.

Por supuesto, no hemos demostrado nada aquí, sólo hemos observado un asombroso parecido entre los valores de dos funciones. Sin embargo, la intuición de Legendre y Gauss era correcta y alrededor de 100 años después, en 1896, el teorema

de los números primos fue demostrado por dos matemáticos al mismo tiempo (trabajando independientemente): en Bélgica, Charles de la Vallée-Poussin, y en Francia, Jacques Hadamard. Enunciemos este resultado de manera formal.

Teorema de los números primos. Sea $P(n)$ el número de números primos que son menores o iguales que n . Entonces se tiene que:

$$\lim_{n \rightarrow \infty} \frac{P(n) \ln n}{n} = 1.$$

EL PROBLEMA DEL GRANJERO

Un granjero gasta \$1 000 en comprar 100 animales de tres tipos diferentes. Cada vaca le cuesta \$20, cada cerdo \$12 y cada oveja \$8. Si suponemos que compró al menos 10 animales de cada tipo, ¿cuántos animales compró?



Figura IV.5.

Solución. Si denotamos por x el número de vacas, por y el número de cerdos y por z el número de ovejas tendremos entonces que:

$$\begin{aligned} 20x + 12y + 8z &= 1000 \\ x + y + z &= 100 \end{aligned}$$

Para eliminar una de las variables, multiplicamos la segunda ecuación por 20. De forma que obtenemos $20x + 20y + 20z =$

2000. A ésta podemos restarle la primera ecuación para obtener $8y + 12z = 1000$. De aquí obtenemos que $y = (1000 - 12z)/8 = 125 - z - z/2$. Como el número de cerdos y y el de ovejas z son enteros positivos, entonces el número $z/2$ es también un número entero, esto quiere decir que z es un múltiplo de 2. Además, como $y \geq 10$, entonces $3z \leq 230$, o sea, $z \leq 76$.

Sustituyendo este resultado en la segunda ecuación, tenemos que $2x = z - 50$. Como x debe ser un número entero mayor o igual a 10, necesariamente $z \geq 70$. De aquí concluimos que $10 \leq x \leq 13$. Las posibles soluciones de este problema son las siguientes:

Vacas	10	11	12	13
Cerdos	20	17	14	11
Ovejas	70	72	74	76

LOS PRIMOS DE FERMAT

Fermat pensaba que todos los números de la forma $2^{2^n} + 1$ con n un número entero eran números primos. En efecto, para $n = 1, 2, 3$ y 4 la fórmula produce los números primos 5, 17, 257 y 65 537. Pero el número $2^{2^5} + 1 = 4\,294\,967\,297$ no es número primo, como fue demostrado por Euler, quien calculó que $4\,294\,967\,297 = 641 \times 6\,700\,417$. No se sabe si hay otros números primos que puedan ser construidos en la forma sugerida por Fermat.

Demuestre que si $2^m + 1$ es un número primo, entonces $m = 2^n$ para algún número n .

Solución. Supongamos que m no es de la forma 2^n , demostraremos que entonces $2^m + 1$ no es un número primo. Dividimos a m entre el número de la forma 2^n con n más grande posible. Esto es, $m = 2^n c$, donde c es un número entero positivo non mayor que 1. Llamemos $a = 2^{2^n}$. Tenemos entonces que:

$$2^m + 1 = 2^{2^n c} + 1 = a^c + 1.$$

Pero para los números nones c se tiene que $a^c + 1$ es divisible entre $a + 1$. Por ejemplo, para $c = 5$ tenemos:

$$a^5 + 1 = (a + 1)(a^4 - a^3 + a^2 - a + 1).$$

En general, $a^c + 1 = (a+1)(a^{c-1} - a^{c-2} + a^{c-3} - \dots - a + 1)$. Como $c > 1$, esto ofrece un factor propio de $2^m + 1$, que no es entonces primo.

V. Enviando mensajes secretos

LA CRIPTOGRAFÍA es la ciencia de las comunicaciones secretas. El problema es transmitir a un destinatario de manera segura un mensaje de forma que sólo él pueda entender el contenido, a pesar de que todo mundo pueda leerlo. Cuando transformamos un mensaje de manera que sólo puede ser entendido por el destinatario, decimos que el mensaje ha sido *codificado* y que el destinatario conoce la *clave de decodificación*.

Todos hemos jugado cuando niños a enviar "mensajes secretos" escribiendo unas letras por otras, de forma que sólo el destinatario sepa cuál es el cambio de letras que utilizamos. Éste es el mismo método que utilizaban los emperadores romanos. Por ejemplo, Julio César usaba un desplazamiento cíclico de las letras del alfabeto de forma que la *A* se escribía como *D*, la tabla completa de las transformaciones sería:

letra: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
codificada: D E F G H I J K L M N O P Q R S T U V W X Y Z A B C

Esta transformación de las letras del alfabeto se llama una *trasposición*. Si decidimos que los espacios en blanco se escriben como *A*, entonces la segunda frase de este capítulo en el código de Julio César se escribiría de la manera siguiente:

HOASUREOHPDAHVAWUDVPLWLUADAXQAGHVWLQDWD
ULRAGHAPDQHJUDAVHJXUDAXQAPHQVDMHAGHAIRUPD
ATXHVRORAHOASXHGDAHQWHQGHUAHOAFRQWHQLGR
ADASHVDUAGHATXHAWRGRAPXQGRASXHGDAOHUOR

¿Qué tan fácil sería para un ejército enemigo descifrar este mensaje? No sabemos qué tan hábiles fueron los enemigos de Julio César, pero este tipo de códigos secretos son fáciles de descifrar.

En efecto, en un idioma la frecuencia con que las diferentes letras aparecen no es la misma. Por ejemplo, la letra *E* es la que más frecuentemente se utiliza y, como bien sabemos, no hay muchos ejemplos de palabras que usen *W*. Pues bien, en el código de Julio César la frecuencia con la que aparece una letra en el mensaje original y la letra correspondiente en el mensaje codificado es la misma. Así, la letra *A* aparece 11 veces en el mensaje original y por lo tanto la *D* aparece 11 veces en el mensaje codificado. Esto nos da la clave para descifrar el código de una transposición: si tomamos un texto suficientemente largo y contamos cuántas veces aparece cada letra tendremos una idea aproximada de la frecuencia con la que esa letra se usa en el lenguaje, luego contamos las frecuencias de cada letra en el mensaje cifrado; frecuencias parecidas indican que probablemente se trate de una letra y su correspondiente codificación.

Tratemos de decodificar el mensaje de Julio César. Para ello usamos el texto de este capítulo (desde el inicio hasta este párrafo) para calcular la frecuencia con que se usan las letras en español. Obtenemos así la siguiente tabla:

Letra	Número de ocurrencias	Frecuencia	Letra	Número de ocurrencias	Frecuencia
espacio	395	.172	U	80	.034
E	316	.137	M	71	.030
A	247	.107	P	39	.017
S	170	.074	F	34	.015
O	162	.070	B	23	.010
I	131	.057	Q	19	.008
N	126	.054	J	18	.008
L	117	.050	G	18	.008
R	112	.048	Y	7	.003
C	105	.045	V	6	.003
T	91	.039	H	6	.003
D	90	.039	Z	4	.002

Contamos enseguida el número de veces que aparece cada letra en el mensaje cifrado. Obtenemos la siguiente tabla:

Letra	Número de ocurrencias	Letra	Número de ocurrencias
A	28	G	10
H	24	X	9
Q, R	12	U	7
D	11	V, P	6

Las demás letras aparecen cuando mucho en 3 ocasiones. Vemos entonces que el signo más usado en el mensaje codificado es A y es lógico esperar que esta letra debe representar el espacio entre palabras que es el signo más utilizado en el lenguaje escrito ordinario. La segunda letra más usada es la H y podemos concluir que probablemente se trate de la letra E. Hasta aquí las cosas van perfectamente. Las cosas comienzan a complicarse después, pues la tercera letra más usada en el mensaje codificado es la Q que no corresponde a la letra A, que es la tercera más usada en el lenguaje ordinario. Sin embargo, las letras más usadas en el mensaje: Q, D, R, G corresponden a las letras N, A, O, D, que están entre las letras más usadas ordinariamente. Después de ensayar entre varias sustituciones diferentes es fácil atinar con la sustitución correcta que nos permite llegar al siguiente mensaje:

eO SUoEOePa eV WUaVPLWLU a Xn deVWLnaWaULo de PaneUa
VeJXUa Xn PenVaMe de loUPa TXe VoOo eO SXeda enWendeU eO
FonWenLdo a SeVaU de TXe Wodo PXndo SXeda OeeUOo

Hemos escrito con minúsculas las letras decodificadas y dejado en mayúsculas las que todavía no hemos tratado de descifrar. Un vistazo a las palabras cortas de este texto nos obliga a pensar que probablemente la X codifique a la u, la O a la l, la V a la s, la T a la q. Hechas esas sustituciones el mensaje puede terminarse de descifrar fácilmente.

Los problemas que encontramos en el desciframiento del mensaje anterior se deben a su longitud. El método que acabamos de ver para descifrar un mensaje sólo es útil si sabemos que el mensaje ha sido cifrado por medio de una transposición de letras y si el mensaje es relativamente largo. Un mensaje relativamente corto puede tener características especiales de forma que la frecuencia de las letras que utiliza no sean parecidas a

las del lenguaje ordinario. Por ejemplo, en un mensaje corto como: ATACA AHORA, la letra que más aparece es la A, mientras que no hay una sola E. Un mensaje así será prácticamente imposible de descifrar si no se conoce la clave de cifrado.

Al paso de los años quedó claro que la codificación de mensajes (largos) en la forma que lo hacían los romanos era fácil de ser descifrada por las personas que no deberían conocer el mensaje. Por ello se comenzaron a usar claves de codificación más y más complejas. Pero la mayor parte de los esfuerzos fueron inútiles pues la historia está llena de ejemplos de éxitos de analistas que logran violar los códigos del enemigo.

Durante la primera Guerra Mundial los británicos interceptaron un mensaje cifrado del ministro de Relaciones Exteriores de Alemania, Arthur Zimmermann, dirigido al embajador en México, Heinrich von Eckardt. Después de muchos esfuerzos los analistas británicos lograron romper el código del mensaje y descubrir un plan alemán de alentar al gobierno de México para que entrara a la guerra como aliado de Alemania asegurándole que, al triunfo, recuperaría los territorios perdidos en la guerra de 1847. El aviso que se envió al presidente de Estados Unidos, Woodrow Wilson, decidió a éste a entrar inmediatamente en la guerra del lado de los aliados, lo que probablemente permitió un fin más rápido del conflicto armado.

CODIFICANDO CON MATRICES

Una *matriz* de tamaño 2×2 es un arreglo de números:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

como las matrices A_0 y B_0 siguientes:

$$A_0 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \quad B_0 = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}.$$

Hay muchas cosas que se pueden hacer con las matrices. Por ejemplo, se pueden sumar como sigue:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}.$$

De esta forma la matriz $A_0 + B_0$ resulta:

$$\begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}.$$

Una matriz también se puede multiplicar por una *columna* $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ de forma que el resultado es otra columna. Esta multiplicación está dada por la siguiente regla:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}.$$

Por ejemplo, si multiplicamos nuestra matriz A_0 por la columna $v = \begin{pmatrix} 2 \\ 11 \end{pmatrix}$, resulta la columna $A_0 \cdot v = \begin{pmatrix} 37 \\ 61 \end{pmatrix}$.

Finalmente, observemos que también podemos multiplicar matrices si pensamos que una está formada por dos columnas. En efecto, si tenemos dadas dos matrices A y B de tamaño 2×2 y b_1 y b_2 son las dos columnas que forman la matriz B , entonces definimos la matriz $A \cdot B$ como la matriz cuyas columnas son $A \cdot b_1$ y $A \cdot b_2$. Podemos así calcular los productos:

$$A_0 \cdot A_0 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 13 & 21 \\ 21 & 34 \end{pmatrix}$$

$$A_0 \cdot B_0 = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Vamos a ver cómo podemos usar estas operaciones de matrices para obtener una clave de cifrado más difícil de descifrar que la de Julio César.

Comencemos con asignar a cada letra del alfabeto un número. Por ejemplo, podemos elegir la siguiente sencilla asignación:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Supongamos que queremos codificar el mensaje: "El problema es comunicar de manera segura un mensaje".

Comenzamos partiendo este mensaje en pares de letras: "El pr ob le ma es co mu ni ca rd em an er as eg ur au nm en sa je", y formamos con estos pares columnas de números según la tabla de cifrado:

$$\begin{pmatrix} 5 \\ 12 \end{pmatrix} \begin{pmatrix} 16 \\ 18 \end{pmatrix} \begin{pmatrix} 15 \\ 2 \end{pmatrix} \begin{pmatrix} 12 \\ 5 \end{pmatrix} \begin{pmatrix} 13 \\ 1 \end{pmatrix} \begin{pmatrix} 5 \\ 19 \end{pmatrix} \begin{pmatrix} 3 \\ 15 \end{pmatrix} \begin{pmatrix} 13 \\ 21 \end{pmatrix} \\ \begin{pmatrix} 14 \\ 9 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} \begin{pmatrix} 18 \\ 4 \end{pmatrix} \begin{pmatrix} 5 \\ 13 \end{pmatrix} \begin{pmatrix} 1 \\ 14 \end{pmatrix} \begin{pmatrix} 5 \\ 18 \end{pmatrix} \begin{pmatrix} 1 \\ 19 \end{pmatrix} \\ \begin{pmatrix} 5 \\ 7 \end{pmatrix} \begin{pmatrix} 21 \\ 18 \end{pmatrix} \begin{pmatrix} 1 \\ 21 \end{pmatrix} \begin{pmatrix} 14 \\ 13 \end{pmatrix} \begin{pmatrix} 5 \\ 14 \end{pmatrix} \begin{pmatrix} 19 \\ 1 \end{pmatrix} \begin{pmatrix} 10 \\ 5 \end{pmatrix}.$$

Escogemos una matriz A de tamaño 2×2 que sea *invertible*, esto es, que exista otra matriz B de forma que $A \cdot B$ sea la *matriz identidad*:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Por ejemplo, nuestra matriz A_0 definida antes es invertible. Para continuar, multiplicamos cada una de las columnas obtenidas por la matriz invertible A_0 , para obtener un sistema de 22 nuevas columnas como siguen:

$$\begin{pmatrix} 46 \\ 75 \end{pmatrix} \begin{pmatrix} 86 \\ 138 \end{pmatrix} \begin{pmatrix} 36 \\ 55 \end{pmatrix} \begin{pmatrix} 39 \\ 61 \end{pmatrix} \begin{pmatrix} 29 \\ 44 \end{pmatrix} \begin{pmatrix} 67 \\ 110 \end{pmatrix} \begin{pmatrix} 51 \\ 84 \end{pmatrix} \begin{pmatrix} 89 \\ 144 \end{pmatrix} \\ \begin{pmatrix} 55 \\ 87 \end{pmatrix} \begin{pmatrix} 60 \\ 97 \end{pmatrix} \begin{pmatrix} 48 \\ 74 \end{pmatrix} \begin{pmatrix} 49 \\ 80 \end{pmatrix} \begin{pmatrix} 44 \\ 73 \end{pmatrix} \begin{pmatrix} 64 \\ 105 \end{pmatrix} \begin{pmatrix} 59 \\ 98 \end{pmatrix} \\ \begin{pmatrix} 31 \\ 50 \end{pmatrix} \begin{pmatrix} 96 \\ 153 \end{pmatrix} \begin{pmatrix} 65 \\ 168 \end{pmatrix} \begin{pmatrix} 67 \\ 107 \end{pmatrix} \begin{pmatrix} 52 \\ 85 \end{pmatrix} \begin{pmatrix} 41 \\ 62 \end{pmatrix} \begin{pmatrix} 35 \\ 55 \end{pmatrix}.$$

Por último reescribimos estas columnas en un arreglo de números consecutivos para borrar toda huella de lo que hemos hecho. El mensaje cifrado que enviaremos es el siguiente:

46 75 86 138 36 55 39 61 29 44 67 110 51 84 89 144 55 87 60 97 48 74
49 80 44 73 64 105 59 98 31 50 96 153 65 168 67 107 52 85 41 62 35 55

La persona que recibe el mensaje y debe descifrarlo procede de la siguiente manera sencilla. Debe conocer la matriz B_0 que tiene la propiedad de que $A_0 \cdot B_0$ es la matriz identidad I , luego divide los números del mensaje en parejas y forma las columnas correspondientes. Luego multiplica las columnas por la matriz B_0 y obtiene columnas que le permiten leer el mensaje. En

nuestro ejemplo, comenzaríamos a descifrar así:

$$\begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 46 \\ 75 \end{pmatrix} = \begin{pmatrix} 5 \\ 12 \end{pmatrix} = \begin{pmatrix} E \\ L \end{pmatrix},$$

$$\begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 86 \\ 138 \end{pmatrix} = \begin{pmatrix} 16 \\ 18 \end{pmatrix} = \begin{pmatrix} P \\ R \end{pmatrix}.$$

Un mensaje cifrado de esta manera es muy difícil de descifrar si no se conoce la matriz A_0 o la matriz B_0 . Pero no es imposible, de hecho el mensaje dirigido al embajador de Alemania en México y descifrado por los servicios de inteligencia británicos durante la primera Guerra Mundial, estaba cifrado por medio de una matriz de tamaño 6×6 en la forma en que hemos trabajado antes.

Finalmente nos preguntamos, ¿de cuántas formas podemos elegir nuestra matriz invertible A_0 ? Contestamos esta pregunta por medio de un sencillo teorema.

Teorema. Una matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ tiene una inversa con entradas enteras si y solamente si $ad - bc$ vale 1 o -1 .

El número $ad - bc$ se llama el *determinante* de A y se denota como $\det A$.

Demostración. Tomemos dos matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$. Sabemos que el producto de las matrices es:

$$A \cdot B = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

que tiene determinante:

$$\begin{aligned} \det A \cdot B &= (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc') \\ &= (ad - bc)(a'd' - b'c') = \det A \det B. \end{aligned}$$

Ahora estamos listos para la demostración. Si suponemos que B es una matriz con entradas enteras que es inversa de A , entonces $A \cdot B = I$ y $\det A \det B = \det A \cdot B = \det I = 1 \times 1$

$-0 \times 0 = 1$. Como además $\det A$ y $\det B$ son números enteros, debemos tener $\det A = 1 = \det B$, o bien $\det A = -1 = \det B$.

Para el converso, supongamos que $\det A$ vale 1 o -1 . Definimos la matriz B sencillamente de la manera siguiente:

$$B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

El producto de A y B resulta ser:

$$A \cdot B = \begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I,$$

lo que completa la prueba. □

ECHANDO VOLADOS POR TELÉFONO

Juan y María viven en ciudades distintas y desean verse. Juan le habla por teléfono y le pide que lo visite. María piensa que Juan es quien debe viajar. No se ponen de acuerdo.

—Ya sé, echemos un volado y el que pierda hace el viaje —dice Juan.

—Perfecto —dice María sacando de su bolsa una moneda—. Elige: águila o sol.

—Sol —pide Juan, al tiempo que escucha a María decir:— perdiste, tu harás el viaje.

¿Puede Juan confiar en que María no hizo trampa?

La misma situación del volado telefónico se presenta en muchas situaciones de la vida moderna. Por ejemplo, un banco A hace una transferencia al banco B por vía electrónica. ¿Cómo sabe el banco A que la información enviada llegó correctamente? ¿Cómo pueden estar seguros los dos bancos de que alguien no leyó la información enviada y puede hacer mal uso de ella?

Una propuesta para poder echar volados telefónicos sin que nadie sospeche que le están haciendo trampa es de Michael Rabin y se basa en el uso de las computadoras y algunas sencillas ideas matemáticas. Veamos qué deben hacer Juan y María para echar un "volado moderno".

María y Juan encienden sus computadoras. María elige dos números primos p y q y no se los comunica a Juan, pero sí le da el resultado $n = pq$ de multiplicarlos. Digamos que María elige $p = 11$ y $q = 19$, entonces le dirá a Juan el número $n = 209$. La computadora de Juan tiene un programa muy eficiente para saber si un número dado es primo o no lo es. Cuando Juan le da el número 209, su computadora inmediatamente le dice que no es un número primo. Lo que la computadora de Juan no puede hacer es calcular la descomposición de 209 en producto de primos.

En realidad, un número tan pequeño como 209 puede factorizarse rápidamente "a mano". Pero lo que los matemáticos no han podido hallar es una forma eficiente de factorizar números grandes, por ejemplo, un número de 200 cifras puede tardar meses en ser factorizado en primos aun por las más poderosas computadoras del mundo. Para que el ejemplo de Juan y María fuera más realista deberíamos utilizar dos números primos de alrededor de 100 cifras (estos números se conocen, pero no sería cómodo para el lector que aquí los usáramos).

Continuemos. Juan elige un número cualquiera menor que 209 y comprueba si divide a 209. Si esto ocurre, ya ganó el "volado moderno". Pero es muy improbable que esto suceda así. Digamos que escogió el número 17. Juan calcula con su computadora el residuo de dividir entre 209 el número que eligió al cuadrado, esto es, el residuo de dividir $17^2 = 289$ entre 209, que resulta 80. Se dice que 289 es *congruente* con 80 módulo 209 y se escribe $289 \equiv 80 \pmod{209}$. Es el número 80 el que Juan deberá de comunicar a María. Ahora, María utiliza un programa de su computadora para calcular todos los posibles números a tales que $a^2 \equiv 80 \pmod{209}$. El resultado es que los únicos números posibles son 17, 93, 116 y 192 (ya que por ejemplo, $93^2 = 8649 = 41 \times 209 + 80$). Pero observemos que $-17 \equiv 192 \pmod{209}$, al igual que $-93 \equiv 116 \pmod{209}$, de forma que bastará con considerar los números 17 y 93.

Finalmente, María debe comunicar a Juan uno de los dos números que obtuvo de su computadora. Si le da el número 17, Juan no sabrá nada nuevo ya que ése fue el número que él eligió, en ese caso habrá perdido el "volado moderno". ¿Qué pasa si María le da el número 93? En ese caso, Juan calcula la diferencia de los dos números que conoce $93 - 17 = 66$ y usa

su computadora para calcular el máximo común divisor de 209 y 66, que resulta ser 11. Al conocer este factor de 209, el otro factor se calcula inmediatamente para obtener que $209 = 11 \times 19$ y con esto Juan gana el "volado moderno".

Dados números a , b y n , se dice que a y b son *congruentes módulo n* si $a - b$ es divisible entre n . En ese caso escribimos $a \equiv b \pmod{n}$. La aritmética modular es muy útil para esconder secretos. Es fácil codificar módulo un número, pero decodificar no es tan sencillo. Por ejemplo, $200 \equiv 2 \pmod{11}$. Aunque todo mundo sepa que usé el 11 para codificar y que mi resultado fue 2, no sabrán si mi número era 13, 24, 101 o 200, salvo que tengan información adicional. ¿Cómo se hace esto? Supongamos que María quiere comunicarle a Juan el mensaje simple: "ven". Primero lo cambia a un número usando el código: $A = 01, B = 02, \dots, Z = 26$. De forma que su número es 220514. Cada par de cifras en este número es ahora elevado a una potencia fija s , digamos $s = 7$ y escrito módulo nuestro número compuesto $n = 209$. Obtenemos el número 155 168 174 (ya que $22^7 = 2494357888 = 209 \times 11934726 + 155$, etcétera). Los números que María comunica a Juan son: 209, 7 y 155 168 174 sin temor de que su mensaje sea interceptado. Para descifrarlo, Juan debe conocer la factorización $209 = 11 \times 19$ (por cierto, el número s elegido debe no tener divisores comunes con $p - 1 = 11 - 1 = 10$ y con $q - 1 = 19 - 1 = 18$, por esto María eligió $s = 7$). El procedimiento de desciframiento consiste en elevar cada grupo de tres cifras del número 155, 168 y 174 a una potencia t , donde el número t sólo puede ser calculado por alguien que conozca los factores 11 y 19 de 209.

¿Cuál es este número t ? Como s no tiene divisores comunes con $(p - 1) \cdot (q - 1) = 180$, entonces hay números a y b de forma que $as + b(p - 1)(q - 1) = 1$, en nuestro caso $-77 \times 7 + 3 \times 180 = 1$, o sea, $a = -77$. Una aplicación del llamado *pequeño teorema de Fermat* nos dice que si $x^s \equiv y \pmod{n}$, entonces $y^a \equiv x \pmod{n}$. También se tiene que $x^a \equiv x^{(n+a)} \pmod{n}$. En nuestro caso, $n + a = 103$ y éste es el número t que buscábamos. En conclusión tenemos que $155^{103} \equiv 22 \pmod{209}$, $168^{103} \equiv 5 \pmod{209}$ y que $174^{103} \equiv 14 \pmod{209}$, cálculos que Juan lleva a cabo con su computadora para leer el mensaje que María envió.

Hemos oído muchas veces decir que algo "es tan cierto como que $2 + 2 = 4$ ". Qué sorpresa se llevarían muchos si supieran que se puede tener también que $2 + 2 = 1$. Por supuesto, esto no puede pasar en el mundo de la aritmética que hemos aprendido desde niños. Pero puede pasar en el mundo de la aritmética módulo 3.

En la aritmética módulo 3, tenemos que $3 \equiv 0 \pmod{3}$, que $4 \equiv 1 \pmod{3}$, que $5 \equiv 2 \pmod{3}$, que $6 \equiv 0 \pmod{3}$, y así sucesivamente. De hecho, todo número es congruente con algún 0, 1 o 2 módulo 3. Es como clasificar a todos los números en tres clases, dependiendo de con quién son congruentes. Todos los números congruentes con 0 módulo 3 forman una clase a la que llamaremos 0; todos los números congruentes con 1 módulo 3 forman otra clase a la que llamaremos 1; y finalmente los congruentes con 2 módulo 3 forman la clase 2.

¿Qué sucede con las operaciones de sumar y multiplicar en este mundo?

Por ejemplo, $1 + 2 = 3$, luego la suma de 1 y 2 módulo 3 deberá valer 0. Esto lo escribimos $1 +_3 2 = 0$. También tenemos que $2 \times 2 = 4$, luego deberemos tener que el producto de 2 y 2 módulo 3 es 1. Esto lo escribimos como $2 \times_3 2 = 1$. Podemos obtener así tablas de sumar y multiplicar módulo 3 de la siguiente manera:

$+_3$	$\bar{0}$	$\bar{1}$	$\bar{2}$	\times_3	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$	$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Diremos que los números $\bar{0}$, $\bar{1}$, $\bar{2}$ con las operaciones de suma $+_3$ y de multiplicación \times_3 forman el anillo \mathbb{Z}_3 .

Por supuesto, el número 3 no tiene nada de especial. Podemos hacer aritmética módulo cualquier número. Por ejemplo, en la aritmética módulo 8 tenemos el anillo \mathbb{Z}_8 cuyos elementos son los números $\bar{0}$, $\bar{1}$, $\bar{2}$, $\bar{3}$, $\bar{4}$, $\bar{5}$, $\bar{6}$, $\bar{7}$ con la suma $+_8$ y la multiplicación \times_8 . Aquí se tiene, por ejemplo que $\bar{3} +_8 \bar{6} = \bar{1}$ y que $\bar{4} \times_8 \bar{5} = \bar{4}$. ¿Puede el lector calcular las tablas de sumar y multiplicar módulo 8?

Considérese el número 19^{19} . ¿Puede este número escribirse como suma del cubo y la cuarta potencia de dos enteros?

Solución. No es posible. Por supuesto, no deseamos calcular explícitamente el valor de 19^{19} y todos los cubos y cuartas potencias menores.

Trabajaremos en la aritmética módulo 13. Aquí, $19 \equiv 6 \pmod{13}$ y $19^{19} \equiv 6^{19} \equiv 6^{12} \times 6^7 \equiv 1 \times 7 \pmod{13}$.

Calculando los cubos i^3 en \mathbb{Z}_{13} , resulta que un número a^3 , con a entero, puede ser congruente con 0, 1, 5, 8 o 12 módulo 13. Similarmente, un número b^4 , con b entero, puede ser congruente con 0, 1, 3 o 9 módulo 13. Luego la suma $a^3 + b^4$ puede ser congruente con cualquier número módulo 13, excepto el 7. Luego, el problema tiene respuesta negativa.

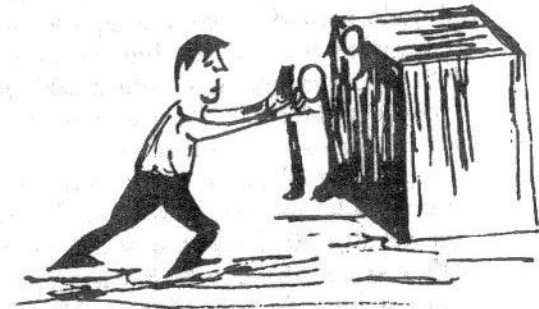


Figura V.1.

¿FUPNWBWBUIRTJKCRDGXLUPP?

Pocas personas pueden creer que es difícil inventar un método de escritura secreta que desafíe la investigación. Sin embargo, puedo asegurar que el ingenio humano no puede crear un cifrado que el ingenio humano no pueda descifrar.

EDGAR ALLAN POE

¿Podrá usted descifrar el mensaje del encabezado de la sección? Para ayudarlo diremos que está codificado de forma que cada

letra tiene un valor numérico de la siguiente manera: $A = 1$, $B = 2$, $C = 3$, ..., $Z = 26$.

El mensaje que se iba a codificar fue dividido en bloques de tres letras y codificado usando la matriz:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Finalmente el resultado numérico volvió a ser "traducido" a letras usando la aritmética módulo 26.

Solución. En primer lugar debemos tener claro que las operaciones que describimos antes para matrices de tamaño 2×2 pueden también efectuarse con matrices de tamaño 3×3 . Por ejemplo, el producto de la matriz A con una matriz columna se lleva a cabo en la siguiente forma:

$$A \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + 2c \\ a + c \\ a + b + c \end{pmatrix}$$

En particular si definimos la matriz:

$$B = \begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 1 \\ -1 & -1 & 2 \end{pmatrix}$$

calculamos que $A \cdot B$ es la matriz identidad:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Es la matriz B la que tenemos que usar para decodificar nuestro mensaje. Hay por supuesto la dificultad adicional de que el mensaje fue "traducido" a letras usando la aritmética módulo 26. Esto quiere decir que la letra F con la que comienza el mensaje puede corresponder a cualquier número n con la propiedad de que $n \equiv 6 \pmod{26}$. Entonces el primer bloque de tres letras FUP por decodificar corresponden a una columna de la forma:

$$v = \begin{pmatrix} 6 + 26a \\ 21 + 26b \\ 16 + 26c \end{pmatrix}$$

para algunos números a, b, c . Al multiplicar esta columna por la matriz B obtenemos:

$$B \cdot v = \begin{pmatrix} 16 + 26a + 52b - 52c \\ -5 - 26b + 26c \\ 5 - 26a - 26b + 52c \end{pmatrix}$$

Sabemos además que estos números deben estar entre 1 y 26. ¿Podemos encontrar los valores de a, b y c ? Del primer renglón de la columna deducimos que $a + 2(b - c) = 0$, del segundo renglón vemos que $-b + c = 1$ y del tercero que $-(a + b) + 2c = 0$. De esto es fácil deducir que $a = 2$, $b = 0$ y $c = 1$. Entonces FUP se decodifica como PUE. El resto del mensaje lo dejamos al lector interesado.

Por supuesto, hay muchas combinaciones de mecanismos para cifrar un mensaje. En 1839 Edgar Allan Poe desafiaba a sus lectores del *Alexander Weekly Messenger* a que le enviaran criptogramas que él descifraría. En febrero de 1840 un lector envió un criptograma que Poe afirmó no tenía ningún sentido. Fue sólo en 1975 que el matemático Winkel descifró el mensaje del lector de Poe. El cifrado era una combinación de métodos, un poco en el estilo del que hemos utilizado para el título de esta sección.

El método descrito para el cifrado del mensaje "ven" de María a Juan es el usado por el ejército de Estados Unidos (y probablemente por el de otros países). Parece ser el método más seguro de cifrado actualmente.

VI. Imágenes de la Alhambra

La urgencia que lleva al decorador a llenar cualquier vacío es generalmente descrita como *horror vacui*, que supuestamente es característico de muchos estilos no clásicos. Tal vez el amor al infinito sería una mejor descripción.

El sentido del orden, E. H. GOMBRICH

COMO en la mayoría de las culturas antiguas, el arte del Islam tiene sus raíces en la religión. Aunque no está escrito en el Corán, la tradición afirma que el profeta Mahoma prohibió las imágenes y los ídolos que representan seres vivos. Sólo Dios puede crear y dar forma a la vida, cualquier imitación hecha por el hombre es considerada idolatría. El maestro religioso El Bourkhari (siglo IX) sentenciaba: "quien imite las criaturas de Dios, tendrá que darles vida: la suya."

Probablemente el resultado más importante de estos mandatos religiosos en la cultura islámica lo encontramos en el arte y en la ciencia. En ambas se desarrolla un gusto marcado por el pensamiento abstracto. A los intelectuales islámicos no les interesa más el aspecto externo de las cosas y los seres, tratan de profundizar en la comprensión de su esencia, en su estructura y funcionamiento. El conocimiento de lo abstracto lo buscan a través de los números, la geometría, las matemáticas, lo buscan en el movimiento de los planetas y otros cuerpos celestes. El Islam continúa de esta manera la rica tradición matemática de los griegos de la Antigüedad. Los eruditos estudian los escritos de Pitágoras, Platón, Euclides, Apolonio y los enriquecen con nuevos descubrimientos y desarrollos. Muchos de estos descubrimientos han llegado hasta nosotros como parte esencial de la cultura, baste recordar que las palabras cero, cifra, álgebra, almanaque y cenit son de origen árabe.

El interés de la cultura islámica por lo abstracto se manifiesta también en la forma artística. En lugar de cuadros naturalistas, como en casi todas las demás culturas, en el Islam florecen las expresiones con motivos geométricos. Para ello se utilizan dos

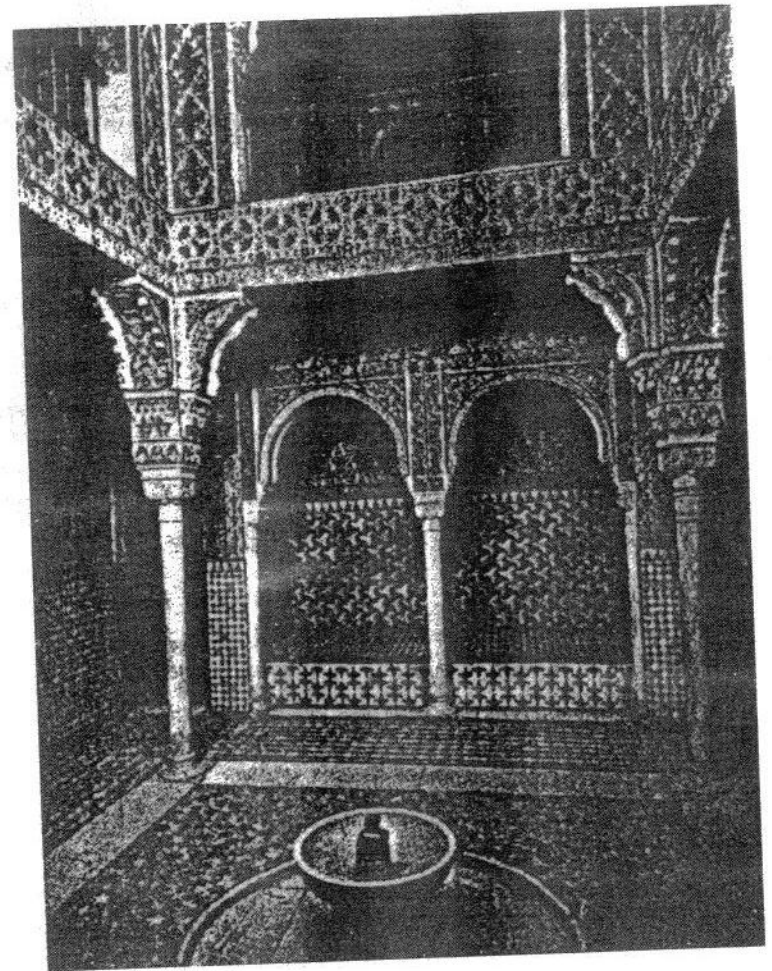


Figura VI.1. Un patio interior de la Alhambra.

formas artísticas: el arabesco floral y el arabesco poligonal (en árabe: *tauriq* y *tastir*). El primero hace uso de figuras y curvas geométricas en forma libre, el segundo usa una geometría de líneas rectas que siguen un patrón bien definido.

La cumbre del arte islámico del arabesco poligonal lo encontramos en la Alhambra, de Granada. La Alhambra es el último de los monumentos de la civilización árabe en Europa. El últi-

mo rey moro reinaba aquí cuando fue vencido por Isabel la Católica en 1492.

La Alhambra está conformada por una serie de patios y cámaras adornados con exquisitas y coloridas figuras geométricas. Estas figuras se presentan en mosaicos o formando la trama de las herrerías de puertas y ventanas. El trabajo de decoración de la Alhambra fue sin duda efectuado por artistas con amplios conocimientos de las matemáticas de su tiempo. Las figuras de sus muros son la muestra que ha quedado de un trabajo de investigación sistemático de estos artistas-matemáticos acerca de las simetrías del espacio plano.

Una de las primeras cosas que llama la atención al observar las fotografías de los mosaicos de las paredes de la Alhambra es que forman patrones repetitivos. Muchos de estos patrones se ven iguales luego de desplazarlos horizontalmente, o de pararlos de cabeza. Estos movimientos que dejan invariante a un dibujo (o patrón) se llaman *simetrías*. Muchas de estas simetrías pueden observarse al primer vistazo, otras son mucho más difíciles de hallar. Los artistas moros se deleitaron en la Alhambra en buscar todas las combinaciones de simetrías posibles, pero su arte fue poco desarrollado por las generaciones que les sucedieron. Sólo en este siglo ha renacido el interés por estudiar seriamente los problemas que ellos, con tanta pasión, consideraron.

Preguntas como: ¿qué formas pueden tener los mosaicos de un embaldosado?, ¿cuáles son sus simetrías?, ¿podemos cubrir una pared usando sólo un tipo de mosaicos?, ¿con cuántos tipos de mosaicos es esto posible?, son las que en los últimos años han vuelto a estudiarse cuidadosamente.

Los embaldosados más sencillos son los de la figura VI.2. Estos embaldosados son sencillos no sólo por estar formados por un solo tipo de mosaico, sino también por el hecho de presentar una estructura repetitiva muy clara. Estudiemos con cuidado esta característica.

Calquemos en una mica transparente la figura VI.2b. Se puede ahora desplazar la mica hasta hacer coincidir otra vez la figura trazada con la figura de la página del libro. Este movimiento es una simetría de la figura. Hay varios tipos de simetrías que nos interesa diferenciar.

En primer lugar queremos considerar las simetrías que dejan

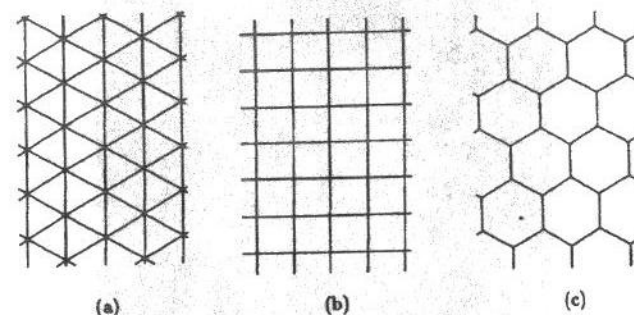


Figura VI.2. Tres embaldosados sencillos.

la mica sin girar. Estos movimientos se consiguen por medio de desplazamientos verticales y horizontales. Llamemos t_v al movimiento que consiste en desplazar la figura verticalmente un cuadro hacia arriba, entonces todas las simetrías verticales hacia arriba se obtienen repitiendo varias veces el movimiento t_v . A un desplazamiento de dos cuadros hacia arriba lo denotaremos por t_v^2 . Un cuadro hacia abajo es el movimiento opuesto a t_v , que llamaremos el *inverso* de t_v y denotaremos por t_v^{-1} . Similarmente, el movimiento t_h , que consiste en desplazar la figura horizontalmente un cuadro hacia la derecha, genera todas las simetrías horizontales. Pero éstos no son los únicos movimientos posibles. Si combinamos un movimiento hacia arriba t_v con uno a la derecha t_h obtenemos otro movimiento (que llamaremos $t_v \cdot t_h$) que también es una simetría de la figura VI.2b. De la misma manera podemos realizar n movimientos hacia arriba o hacia abajo (dependiendo de si $n > 0$ o si $n < 0$) y luego m movimientos hacia la derecha o la izquierda (dependiendo de si $m > 0$ o si $m < 0$) y obtener la simetría $t_v^n \cdot t_h^m$. Observemos finalmente que no realizar movimiento alguno es también una simetría que denotaremos por el símbolo e . Los movimientos descritos se llaman *simetrías traslacionales del embaldosado* y forman un conjunto que denotaremos por la letra T . Los matemáticos dirían que el conjunto T es un *grupo abeliano con dos generadores* t_v y t_h . El lector puede fácilmente observar que las simetrías traslacionales de las figuras VI.2a y VI.2c son también grupos abelianos con dos generadores.

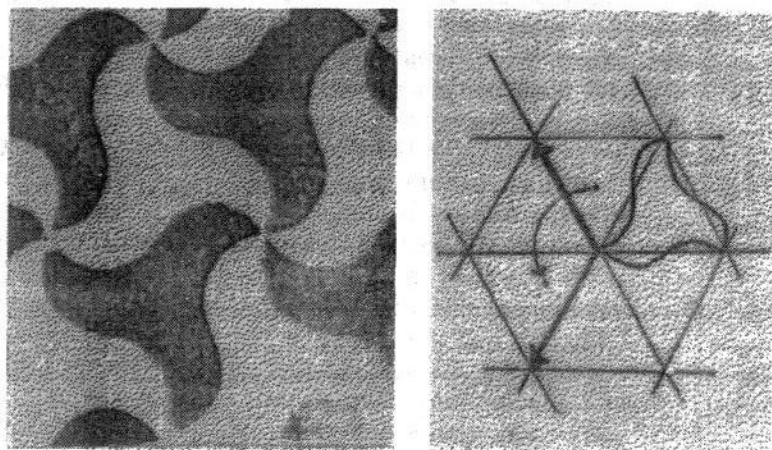
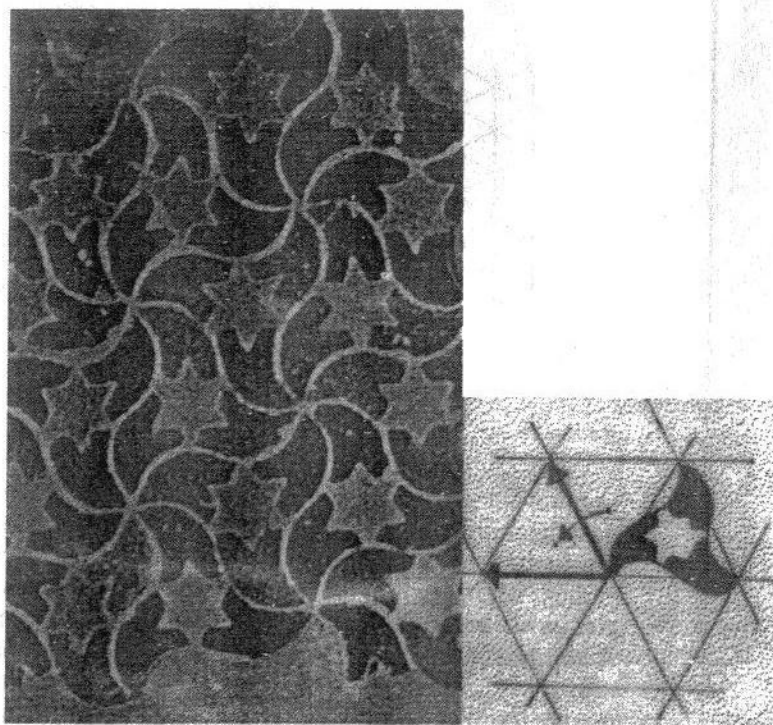


Figura VI.3. Algunos embaldosados de la Alhambra y los generadores de sus grupos de simetría.

¿Qué sucede si dejamos girar la mica? Para simplificar el problema, hagamos girar la mica alrededor de un punto P clavando un alfiler, esto impide cualquier otro tipo de movimiento traslacional. Por ejemplo, elijamos el centro de giro P como alguno de los vértices de los polígonos del embaldosado. En la figura VI.2a, una rotación r de un sexto de vuelta hace coincidir otra vez la figura de la mica con la del libro. Cualquier otra rotación que tenga esta propiedad se obtiene repitiendo varias veces el giro r . Repitiendo seis veces la rotación r regresamos a la posición original, esto es $r^6 = e$. Esto se expresa diciendo que el conjunto de *simetrías rotacionales* R forma un *grupo cíclico de orden 6*. Las simetrías rotacionales de la figura VI.2b forman un grupo cíclico de orden 4 y las de la figura VI.2c uno de orden 3.

Podemos ahora obtener cualquier simetría de una de nuestras figuras por medio de la aplicación de un movimiento traslacional seguido de uno rotacional. Se dice que el conjunto de simetrías S forma un grupo generado por T y R . Más adelante veremos más detalles de estos grupos.

Podemos ahora volver a tomar en cuenta los embaldosados que parecen más complejos, como los de la Alhambra. En las figuras VI.3 y VI.4 vemos algunos ejemplos de grupos diferentes

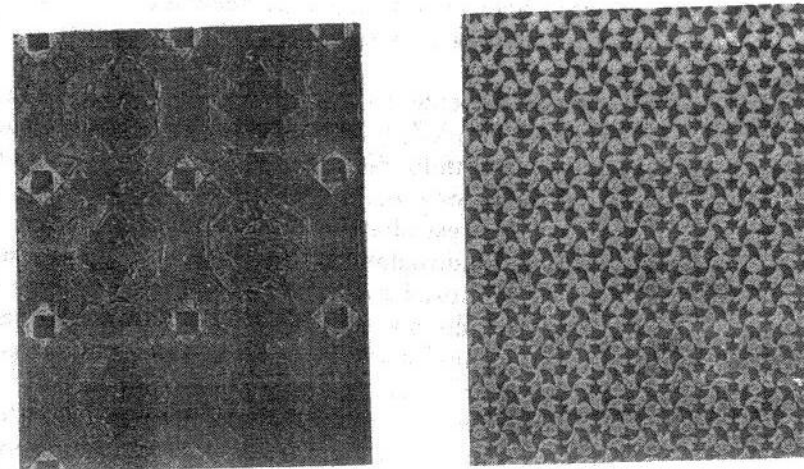


Figura VI.4. Más embaldosados de la Alhambra. ¿Puede determinar sus simetrías?

que aparecen en los embaldosados de la Alhambra. En las figuras indicamos los movimientos que dejan invariantes a los embaldosados, es decir, los generadores de los grupos de simetría. Algunos son tan simples como los que hemos hallado en la figura VI.2, otros son más complicados. Pero, ¿cuántos grupos de simetrías pueden formarse?, ¿cuáles de ellos aparecen en la Alhambra?

El trabajo matemático sobre los embaldosados se inició por el interés de estudiar sus grupos de simetrías. Algunos de los más grandes matemáticos de principios del siglo, como Klein y Hilbert, se interesaban por estos problemas. Sin embargo, los primeros resultados importantes en esta área no vinieron de matemáticos sino de químicos.

LOS CRISTALES: MOSAICOS DE LA NATURALEZA

Los cristales han ejercido una fascinación especial en los hombres. Las excavaciones hechas por investigadores en algunas cuevas de China muestran que el *hombre de Pekín* coleccionaba cristales de cuarzo hace 400 000 años. Los cristales se distinguen por sus colores, sus brillos, pero sobre todo por sus formas.

Cuando se observa un cristal en bruto, su apariencia se distingue claramente de una roca vulgar. Sus caras son prácticamente planas, su cuerpo presenta grandes simetrías. La pirita viene en cubos, la fluorita y los diamantes en forma de octaedro. ¿Por qué?

La forma de un cristal está determinada por los componentes más pequeños, átomos y moléculas, que lo integran. En el espacio, los átomos que forman los cristales se unen como piezas de rompecabezas. Pero hay pocas piezas de rompecabezas que puedan utilizarse: en el diamante, todas las piezas del rompecabezas son átomos de carbono; en la sal, hay átomos de cloro y de sodio.

Los cristales tienen las mismas características que los embaldosados de mosaicos de la Alhambra. Sólo que mientras los mosaicos están diseñados por el hombre para cubrir paredes planas, los cristales han sido diseñados por la naturaleza para llenar el espacio de tres dimensiones.

A mediados del siglo pasado, algunos científicos pensaron que

el comportamiento físico de los diferentes cristales debería estar determinado por su estructura geométrica, por sus simetrías. Los intentos por estudiar los cristales de esta forma fueron iniciados por Weiss en 1804 y Hessel en 1830 (aunque los resultados de este último permanecieron desconocidos por más de 50 años). Bravais redescubrió los resultados de Hessel en 1848 y obtuvo la clasificación de los *grupos cristalográficos puntuales*, es decir, las simetrías cristalinas con respecto a un punto fijo. Finalmente, la clasificación de los grupos cristalográficos fue obtenida por el químico ruso Fedorov en 1885. Pero no fue sino hasta 1913 que Laue, usando el método de difracción por rayos X, pudo describir la estructura de los cristales y demostró que estaban formados por arreglos de átomos como en un rompecabezas.

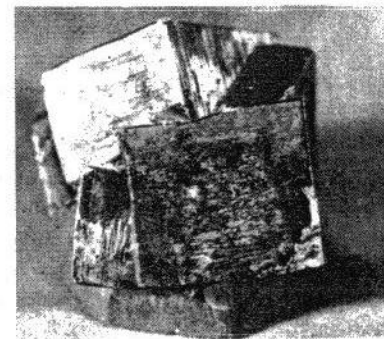


Figura VI.5. Cristales.

Al observar algunos de los diseños de embaldosados planos de la Alhambra vemos que los mosaicos presentan algunas formas regulares, pero no otras. Aparecen triángulos, cuadrados, hexágonos, pero ninguno con figuras regulares de cinco lados, es decir, pentágonos. El lector puede convencerse fácilmente de que no puede cubrirse el plano usando sólo pentágonos (véase la figura VI.6). Este hecho constituye un ejemplo sencillo de una propiedad más general del espacio: no hay simetrías rotacionales de orden cinco, esto es, no podemos construir un embaldosado que al hacerlo girar en un ángulo de 72 grados vuelva a coincidir consigo mismo.

Esta ausencia de simetrías de orden cinco no sólo se da en

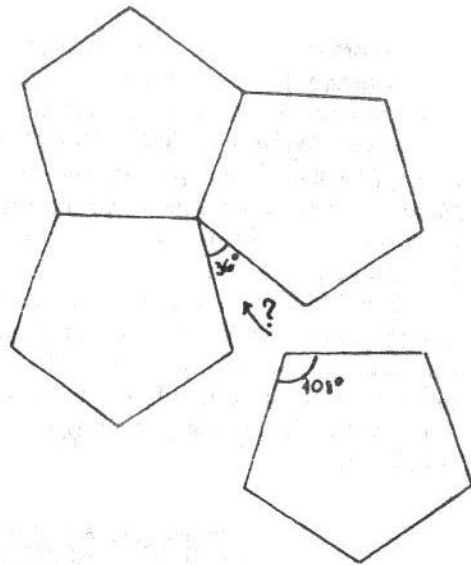


Figura VI.6. No se puede cubrir el plano con pentágonos.

el plano sino también en el espacio. La demostración de esto requiere algunas consideraciones matemáticas que presentamos en el siguiente apartado.

La necesidad de introducir formalismos matemáticos se debe a la amplia riqueza de posibilidades para la construcción de embaldosados y de cristales. Necesitamos tener una forma de estudiarlos todos a la vez y no ir caso por caso, estudiando cada situación que se pueda presentar. Los embaldosados que hemos considerado hasta este momento son sencillos, pero se pueden tener otros más complejos como los de la figura VI.7. Para nuestro estudio nos limitaremos al tipo simple.

Diremos que un embaldosado es *crystalino* si los mosaicos que lo forman satisfacen las siguientes dos propiedades: C1) Hay un número finito de mosaicos M_1, \dots, M_s en el embaldosado de forma que cualquier mosaico M se obtiene como la imagen $g(M_i)$ de algún mosaico M_i con $1 \leq i \leq s$ bajo una simetría g del embaldosado. C2) Cada mosaico del embaldosado posee área finita y no tiene agujeros.

Observamos que los embaldosados de la Alhambra de las figuras anteriores son cristalinos. Los embaldosados de la fi-

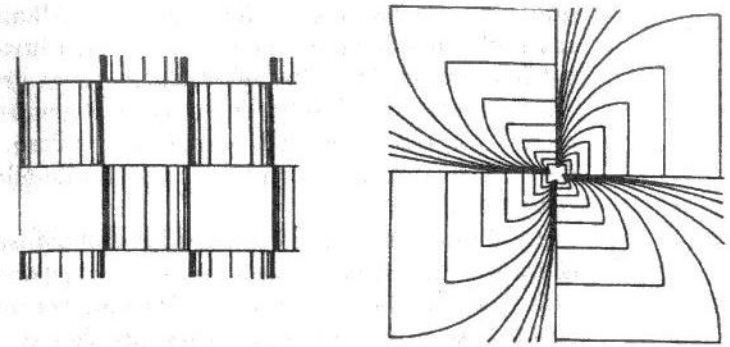


Figura VI.7. Embaldosados no cristalinos.

gura VI.7 no son cristalinos ya que la condición C1) no se satisface. ¿Puede dar ejemplos donde no se satisfaga la condición C2)?

En la figura VI.8 tenemos un embaldosado construido por Kepler (sí, el mismo que estudió las órbitas de los planetas). Este embaldosado es cristalino pues claramente satisface las propiedades C1) y C2), sin embargo podemos notar que no hay simetrías que lleven una estrella en la estrella que está exactamente arriba.

Además de considerar embaldosados cristalinos del plano, también consideraremos embaldosados cristalinos en el espacio. En este caso, nuestros mosaicos serán sólidos que se pegan en el espacio para cubrirlo, como en el caso de los cristales. En la figura VI.9 vemos un embaldosado cristalino en el espacio formado por un solo tipo de mosaicos.

El grupo de simetrías de un embaldosado cristalino se llama *grupo cristalográfico*. Debido a su interés por los cristales, los químicos comenzaron por clasificar los grupos cristalográficos en tres dimensiones y sólo más tarde consideraron el problema de los grupos de embaldosados planos. Las primeras enumeraciones completas de los grupos cristalográficos fueron hechas por el químico ruso Fedorov en 1885 para el caso tridimensional (resultan 230 grupos) y en 1891 para el caso bidimensional (hay 17 grupos).

Si bien la clasificación de los grupos cristalográficos sólo se completó al final del siglo pasado, los moros en España ya

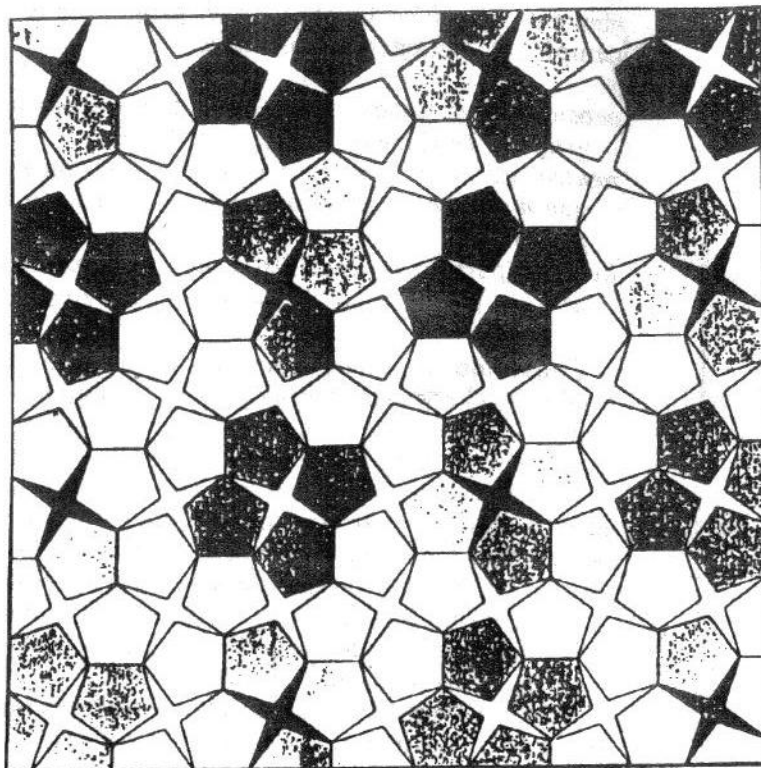


Figura VI.8. Un embaldosado diseñado por Kepler.

sabían bastante de esto: ¡los 17 grupos cristalográficos planos se encuentran en embaldosados de la Alhambra! Seguramente los moros no sabían entonces que ya no podían hallar otros grupos pero sus conocimientos y práctica en el diseño eran tales que encontraron todos.

SIMETRÍAS Y GRUPOS

La noción de simetría es sin duda más antigua que las matemáticas. En el lenguaje cotidiano el adjetivo *simétrico* se usa para calificar un objeto bien proporcionado, bien equilibrado; la simetría denota belleza. Otro significado de este adjetivo entra en la composición *simetría bilateral*, que quiere decir que sus la-

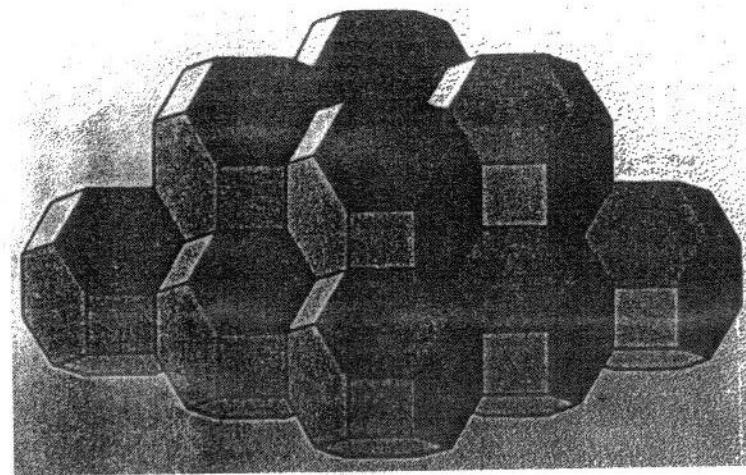


Figura VI.9. Esquema de un cristal formado por un solo tipo de moléculas.

dos derecho e izquierdo son iguales, como en el cuerpo humano. Esta segunda acepción determina un concepto geométrico muy preciso.

Muchas de las culturas antiguas tenían una idea de la simetría acorde con las dos acepciones anteriores. Los griegos pensaban, por ejemplo, que sólo un cuerpo humano simétrico (en el sentido geométrico) podía ser bello. Sin embargo, comenzaron a observar otras formas de simetría. En la naturaleza hay múltiples ejemplos de organismos con formas más complicadas de simetría que la simetría bilateral del cuerpo humano. Por ejemplo, las estrellas de mar (figura VI.10) presentan simetrías alrededor de un centro, si rotamos la estrella en un ángulo de 72° , no notaremos ningún cambio. Las flores presentan también simetrías centrales que pueden ser altamente complicadas. Los copos de nieve, que se presentan en una variedad infinita, presentan siempre simetría hexagonal (es decir, un giro de 60° los hace verse iguales).

Cuando los moros de España desarrollaron sus estudios de simetría ornamental, deben de haber contado con una idea clara de las simetrías y probablemente con métodos sistemáticos de análisis. Sin embargo, la noción matemática de grupo de simetrías se estableció mucho más tarde.

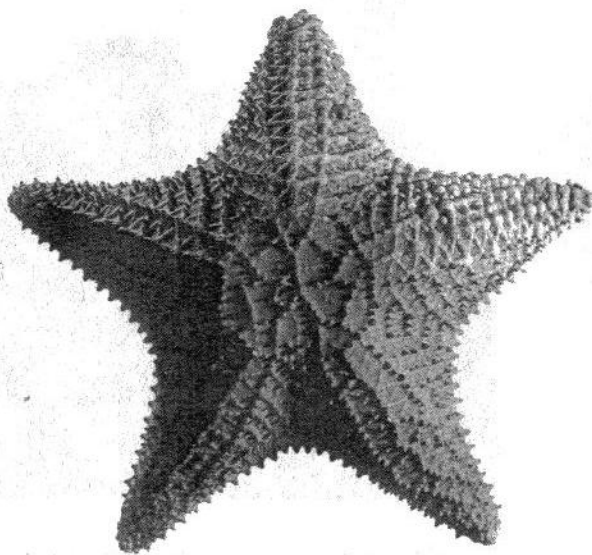


Figura VI.10. Estrella de mar. Ejemplo de simetría rotacional de orden 5.

En 1832, el joven matemático francés Evariste Galois, para formalizar su estudio de las raíces complejas de polinomios, desarrolló por primera vez la noción de *grupo abstracto*. El caso más simple de su estudio consistía en observar que si tomamos las dos raíces $x_1 = \frac{1}{2a}(-b + \sqrt{b^2 - 4ac})$ y $x_2 = \frac{1}{2a}(-b - \sqrt{b^2 - 4ac})$ de la ecuación cuadrática $ax^2 + bx + c = 0$, entonces la transformación del plano complejo g que consiste en enviar cualquier número complejo $z = s + ti$ en su *conjugado complejo* $\bar{z} = s - ti$ tiene la siguiente propiedad: o bien deja fijas ambas raíces x_1 y x_2 (esto se escribe $g(x_1) = x_1$ y $g(x_2) = x_2$, y sucede en caso de que x_1 y x_2 son reales), o bien envía la raíz x_1 en la segunda raíz x_2 (esto se escribe $g(x_1) = x_2$, y sucede en caso de que x_1 y x_2 no son reales). Esta transformación satisface también que $g^2(z) = z$ para todo número complejo z . Esto es, el grupo asociado a la ecuación cuadrática tiene dos elementos e, g que satisfacen $g^2 = e$. Este grupo puede identificarse con el conjunto \mathbb{Z}_2 de enteros módulo 2 con la suma módulo 2.

Un *grupo* es un conjunto G junto con una operación entre

los elementos del grupo. El resultado de la operación entre dos elementos a y b se denota $a \cdot b$. El grupo G con esta operación cumple las siguientes propiedades:

asociatividad: dados cualesquiera tres elementos a, b y c del grupo G se tiene $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

neutro: existe un elemento e en el grupo con la propiedad de que $a \cdot e = a$ y también $e \cdot a = a$, para todo elemento a del grupo.

inverso: para cada elemento a del grupo existe otro elemento a^{-1} con la propiedad de que $a \cdot a^{-1} = e$ y también $a^{-1} \cdot a = e$.

Conocemos muchos grupos (y ya muchos eran conocidos antes de que el concepto fuera definido). Por ejemplo, los números enteros con la suma forman un grupo. En efecto, dados números enteros p, q, r se tiene que $p + (q + r) = (p + q) + r$. El neutro es el número 0 y el inverso de n es $-n$. También los enteros módulo r con la suma módulo r forman un grupo. Si un grupo es finito el número de elementos se llama el *orden* del grupo. Así los enteros módulo n forman un grupo de orden n .

Como un problema menos sencillo, nos proponemos describir todas las *simetrías de un cuadrado*, permitiendo tanto rotaciones como reflexiones por ejes de simetría.

Primero consideramos todas las rotaciones que envía el cuadrado en sí mismo. Estas rotaciones deberán tener por eje de rotación el centro del cuadrado. Claramente hay tres rotaciones diferentes: las rotaciones por un ángulo de 90° , 180° y 270° , que llamaremos r_1, r_2 y r_3 respectivamente. En la figura VI.11 mostramos estas simetrías y marcamos la forma en que los vértices del cuadrado son permutados.

Después consideramos las reflexiones con diferentes ejes de simetría. Hay cuatro diferentes ejes de simetría: el eje vertical por el centro del cuadrado, el eje horizontal por el centro y las dos diagonales. Denotamos estas simetrías por s_1, s_2, s_3 y s_4 . En la figura VI.12 mostramos estas simetrías y vemos el efecto que tienen en los vértices del cuadrado.

Finalmente, llamamos e a la simetría trivial (no movimientos). Resulta que tenemos la siguiente *tabla de multiplicación* del grupo de simetrías del cuadrado:

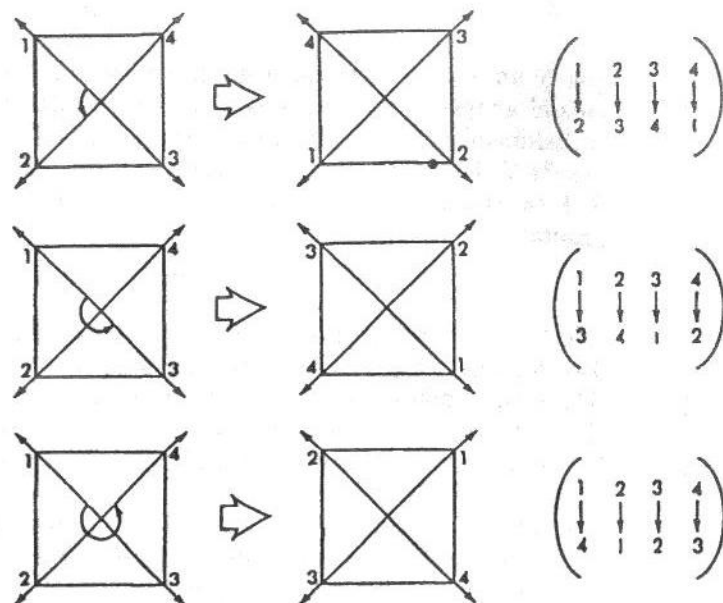


Figura VI.11. Rotaciones del cuadrado.

	e	r_1	r_2	r_3	s_1	s_2	s_3	s_4
e	e	r_1	r_2	r_3	s_1	s_2	s_3	s_4
r_1	r_1	r_2	r_3	e	s_3	s_4	s_2	s_1
r_2	r_2	r_3	e	r_1	s_2	s_1	s_4	s_3
r_3	r_3	e	r_1	r_2	s_4	s_3	s_1	s_2
s_1	s_1	s_4	s_2	s_3	e	r_2	r_3	r_1
s_2	s_2	s_3	s_1	s_4	r_2	e	r_1	r_3
s_3	s_3	s_1	s_4	s_2	r_1	r_3	e	r_2
s_4	s_4	s_2	s_3	s_1	r_3	r_1	r_2	e

Esta tabla se debe entender de la siguiente manera: si aplicamos la simetría a y en seguida la b , entonces leemos a en la primera columna y b en el primer renglón de la tabla. La entrada en la intersección nos dará el resultado $a \cdot b$ de aplicar primero a y luego b . Es interesante notar que por ejemplo, $r_1 \cdot s_1 = s_3$, mientras que $s_1 \cdot r_1 = s_4$.

El problema de calcular las simetrías del pentágono lo dejamos al lector interesado.

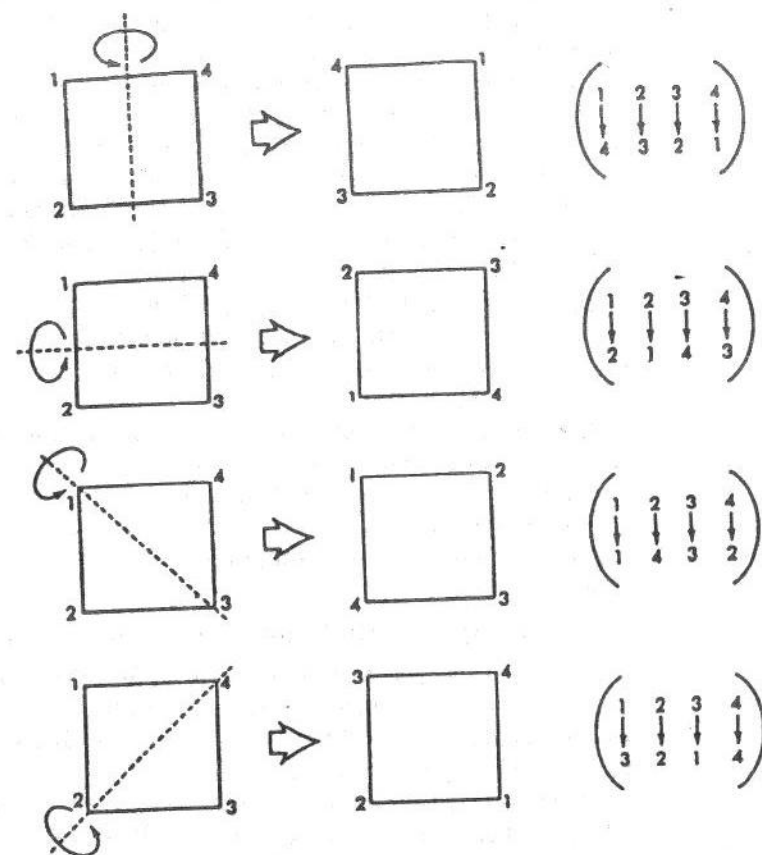


Figura VI.12. Reflexiones del cuadrado.

Otros grupos importantes son muy sencillos de construir. Tomemos los números del 1 a n y consideremos todas las listas de la forma (a_1, \dots, a_n) donde las letras a_1, \dots, a_n son los números $1, \dots, n$ en algún orden (pero es importante el orden, es decir, a_1 va primero, a_2 en segundo lugar, etcétera). El conjunto de estas listas se llama el *grupo simétrico de orden n* y se denota por S_n . Podemos imaginar un juego de sillas: cada una de n sillas tiene asignado un número entre 1 y n y cada persona sentada en la silla i tiene en sus manos el número i . Todo mundo se para de las sillas y se vuelve a sentar en alguna silla (no

cesariamente en la misma silla que antes tenía). El resultado es que a la silla i le queda asignada la persona a_i , o sea, tenemos la *permutación* (a_1, \dots, a_n) .

El conjunto de permutaciones S_n es un grupo: si $a = (a_1, a_2, \dots, a_n)$ y $b = (b_1, b_2, \dots, b_n)$ son dos permutaciones y podemos multiplicarlas de la siguiente manera: $a \cdot b = (b_{a_1}, b_{a_2}, \dots, b_{a_n})$.

Además, el neutro en S_n es $e = (1, 2, \dots, n)$ y el inverso de a es b , cuya entrada a_i es $b_{a_i} = i$. Por ejemplo, para $n = 3$ el grupo S_3 tiene seis elementos $e = (1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(3, 2, 1)$, $(2, 3, 1)$, $(3, 1, 2)$. El producto de $(1, 3, 2)$ y $(3, 2, 1)$ es $(3, 1, 2)$, el inverso de $(2, 3, 1)$ es $(3, 1, 2)$. El lector puede construir de esta manera la tabla completa de multiplicar en S_3 . Observe que el orden de S_3 es $3 \times 2 = 6$ (en general S_n tiene orden $n! = n(n-1) \dots 2$).

Elementos interesantes del grupo S_n son las *trasposiciones* t_{ij} que se obtienen intercambiando i por j y dejando todos los demás números en su lugar. Por ejemplo, si $n = 3$, las trasposiciones son $t_{12} = (2, 1, 3)$, $t_{13} = (3, 2, 1)$ y $t_{23} = (1, 3, 2)$. ¿Qué se puede obtener componiendo estos elementos en S_3 ? Veamos,

$$(1, 2, 3) = t_{12}^2, (2, 3, 1) = t_{12} \cdot t_{13}, (3, 1, 2) = t_{23} \cdot t_{12}.$$

En consecuencia, todo elemento en S_3 se obtiene como producto de trasposiciones: $(1, 2, 3)$ es producto de dos trasposiciones, $(2, 1, 3)$, $(3, 2, 1)$ y $(3, 2, 1)$ son producto de una sola trasposición (por ser trasposiciones ellos mismos) y $(2, 3, 1)$, $(3, 1, 2)$ son producto de dos trasposiciones. Observamos que $(2, 3, 1) \cdot (3, 1, 2) = (1, 2, 3)$, de donde se sigue que los tres elementos de S_3 que son producto de dos trasposiciones forman un subgrupo de S_3 . Esta observación se puede generalizar fácilmente como sigue:

Proposición. Consideremos el grupo simétrico S_n . Entonces:

- a) Todo elemento de S_n se obtiene como producto de trasposiciones.

Si σ es un elemento de S_n que se escribe como $\sigma = t^{(1)} \cdot t^{(2)} \dots t^{(s)}$ con $t^{(i)}$, ($i = 1, \dots, s$) trasposiciones y s el menor número posible, decimos que s es la paridad de σ .

- b) El conjunto de elementos de S_n con paridad par forman un subgrupo de S_n , que se llama el grupo alternante A_n .
c) El número de elementos de A_n es $n!/2$.

Sea S un grupo asociado a un embaldosado plano. Consideremos el grupo de traslaciones T y el grupo de rotaciones R del embaldosado que deja un punto fijo. Ambos T y R son subgrupos de S . De hecho T es un subgrupo normal de S y el cociente S/T es isomorfo a R . Esto se denota por medio de la *sucesión exacta*:

$$0 \rightarrow T \rightarrow S \rightarrow R \rightarrow 1.$$

(NOTA: se dice que un subgrupo T de S es normal si para cada dos elementos $t \in T$ y $s \in S$ se tiene que $sts^{-1} \in T$. El cociente S/T está formado por clases de elementos de S de forma que dos elementos s_1 y s_2 de S están en una misma clase si $s_1 \cdot s_2^{-1} \in T$. La estructura de grupo de S induce una estructura de grupo en S/T .)

Debido a las condiciones C1) y C2) hay dos traslaciones t, t' en T que son linealmente independientes. En realidad T es un grupo abeliano libre generado por t y t' . Probaremos ahora que R es un grupo cíclico generado por un elemento r tal que $r^n = e$ para alguna $n = 1, 2, 3, 4, 6$.

En efecto, sea R el subgrupo de S formado por las rotaciones alrededor del punto P . Sea r un elemento de R que no sea la identidad. Tomemos otro punto Q y fijémonos en los puntos $Q, r(Q), r^2(Q), r^3(Q), \dots$, todos ellos caen dentro del círculo con centro P y radio $|PQ|$. Como este círculo sólo toca un número finito de mosaicos del embaldosado, tenemos que $r^n = e$ para alguna n . Si elegimos n la más pequeña correspondiente a los elementos de R , entonces R está generado por la rotación por un ángulo de $2\pi/n$ alrededor de P .

¿Qué valores puede tomar n ? Elijamos t una traslación en S por un vector v de magnitud mínima posible. Sea $Q = t(P)$ y Q' la imagen de Q por la rotación $2\pi/n$ alrededor de P . El movimiento $t' = rtr^{-1}$ es una traslación en S que lleva P en Q' . El movimiento $t't^{-1}$ en S transforma Q en Q' , véase la figura VI.13.

Como también $t't^{-1}$ es una traslación se tiene que la distancia entre Q y Q' es mayor o igual que la distancia entre P y Q . Esto sólo es posible si $2\pi/n \geq \pi/3$, o sea $n \leq 6$.

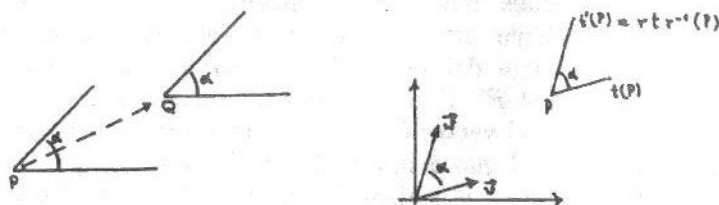


Figura VI.13.

Afirmamos que el caso $n = 5$ no puede suceder. Supongamos, para llegar a una contradicción, que la rotación por $2\pi/5$ alrededor de P está en el grupo S . Entonces la rotación por el ángulo $4\pi/5$ está también en S . Sea Q'' la imagen de Q por esta rotación. Entonces hay una traslación t'' en S que lleva P en Q'' (véase la figura VI.14). La traslación tt'' lleva P en P' y es fácil comprobar que la distancia entre P y P' es estrictamente menor que la distancia entre P y Q , lo que contradice la elección que hicimos de v . En conclusión, n puede sólo tomar alguno de los valores 1, 2, 3, 4 o bien 6.

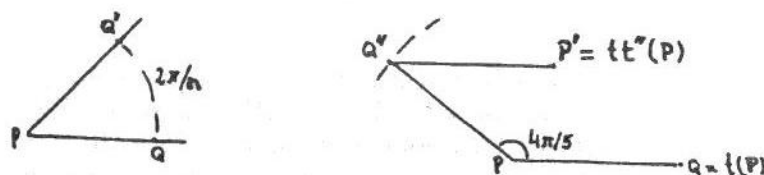


Figura VI.14.

Resulta así que S está generado por las traslaciones t, t' y una rotación r . Hay cinco grupos cristalográficos definidos de esta manera. Damos a continuación la información más importante de ellos.

Grupo	Generadores	Ecuaciones que satisfacen los generadores
T	t, t'	$tt' = t't$
S_{2222}	$t, t', r = \text{rotación de } 180^\circ$	$tt' = t't, \quad r^2 = e$
S_{333}	$t, r = \text{rotación de } 120^\circ$	$r^3 = e$
S_{442}	$t, r = \text{rotación de } 90^\circ$	$r^4 = e$
S_{632}	$t, r = \text{rotación de } 60^\circ$	$r^6 = e$

Hasta aquí hemos tratado solamente los grupos cristalográficos formados por simetrías que *preservan la orientación*. Cuando aceptamos también movimientos que incluyen reflexiones alrededor de ejes aparecen nuevos grupos hasta completar 17. Esbozamos cómo se hace esto.

Consideramos el plano E^2 sobre el que se construye el embaldosado. Dado un punto P en E^2 lo identificamos con $g(P)$ para cualquier g en S . De esta forma obtenemos un *espacio topológico* E^2/S . Consideremos el subgrupo normal S^+ de S formado por las simetrías que respetan la orientación, luego S^+ es uno de los cinco grupos clasificados antes y el cociente es un grupo cíclico de orden 2. El espacio E^2/S^+ puede obtenerse a partir de la clasificación. Resulta que E^2/S^+ es un toro si $S^+ = T$ y E^2/S^+ es una esfera en los demás casos. El grupo S/S^+ induce una involución $\sigma: E^2/S^+ \rightarrow E^2/S^+$ de forma que E^2/S es el espacio cociente $(E^2/S^+)/\langle\sigma\rangle$. Las involuciones del toro y la esfera son conocidas: en el caso del toro, el cociente E^2/S puede resultar una botella de Klein, un anillo o bien una banda de Möbius (figura VI.15).

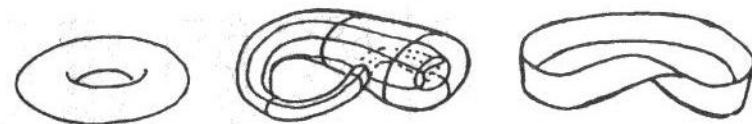


Figura VI.15. Toro, botella de Klein y banda de Möbius.

En el caso de la esfera, la involución σ debe ser una reflexión por un meridiano o bien un cartografiado antipodal; el cociente E^2/S resulta un disco o un plano proyectivo, respectivamente. A partir de esta información se puede reconstruir la forma en que el grupo S opera sobre E^2 y obtener todos los grupos. Ilustraremos esto en un ejemplo.

Consideremos el caso en que $S^+ = T$ y $\sigma: E^2/S^+ \rightarrow E^2/S^+$ es la involución tal que E^2/S es la *botella de Klein*. La botella de Klein resulta de un cuadrado en el que el par de lados verticales se unen identificando los puntos que están frente a frente y el par de lados horizontales se unen punto a punto siguiendo direcciones opuestas para cada lado (véase la figura VI.16).

Consideremos en el plano la reflexión con eje la línea L . Po-

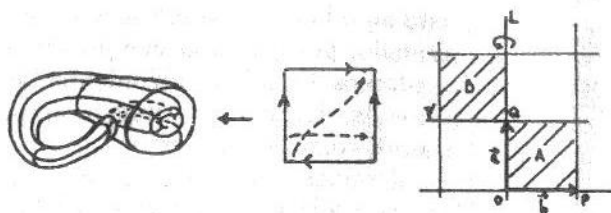


Figura VI.16. La construcción de la botella de Klein.

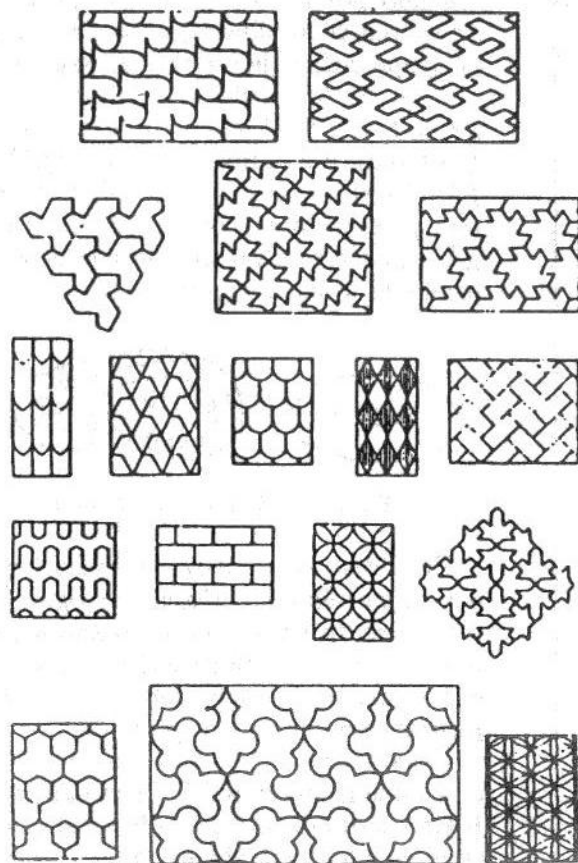


Figura VI.17. Copia de los dibujos originales de Polyá (1924) con ejemplos de los 17 grupos cristalográficos planos.

demostramos definir la transformación $\bar{\sigma}: E^2 \rightarrow E^2$ que consiste en reflejar alrededor de L y trasladar por el vector a . Esta isometría del plano lleva el cuadrado A en el B y el punto P en el P' . Consideremos la traslación $t_b: E^2 \rightarrow E^2$ determinada por el vector b . El grupo abeliano libre T generado por t_b y t_{2a} es tal que $\bar{\sigma}$ induce la involución $\sigma: E^2/T \rightarrow E^2/T$ (en efecto, observe por ejemplo que $t \cdot \bar{\sigma}(P) = Q$). Tenemos entonces que S es el grupo generado por t_{2a}, t_b y $\bar{\sigma}$, de hecho bastan t_b y $\bar{\sigma}$ para generar a S . En la figura VI.17 vemos ejemplos de los 17 grupos cristalográficos planos.

MOLÉCULAS COMO PELOTAS DE FUTBOL

El matemático juega con unas reglas que ha inventado él mismo, mientras que el físico juega con las reglas que determina la naturaleza; sin embargo, a medida que transcurre el tiempo se hace cada vez más evidente que las reglas que el matemático encuentra interesantes son las mismas que la naturaleza ha elegido.

PAUL DIRAC

El uso de la teoría de grupos por los químicos para determinar las propiedades de las moléculas es un procedimiento bien establecido. Desde el punto de vista matemático la mayor parte de los grupos que pueden aparecer como grupos de simetrías de moléculas aisladas son muy sencillos, con excepción hecha del grupo alternante A_5 (en el lenguaje matemático se diría que todos estos grupos —salvo A_5 — son solubles). El grupo A_5 es un grupo de orden 60 que aparece como subgrupo de S_5 (que a su vez tiene orden $5! = 120$). El grupo A_5 es muy especial, entre otras cosas es un *grupo simple*, es decir, no contiene subgrupos normales propios.

En 1985 una nueva familia de moléculas fue descubierta por los investigadores H. Kroto de Inglaterra y R. Smalley y R. Curl de EUA mientras realizaban trabajos de astrofísica tratando de encontrar nuevas moléculas de carbón. Las moléculas que encontraron son arreglos tridimensionales de átomos de carbono (desde 24 átomos hasta miles de ellos) y les dieron el nombre de *fulerenos* en honor al arquitecto Buckminster Fuller, quien

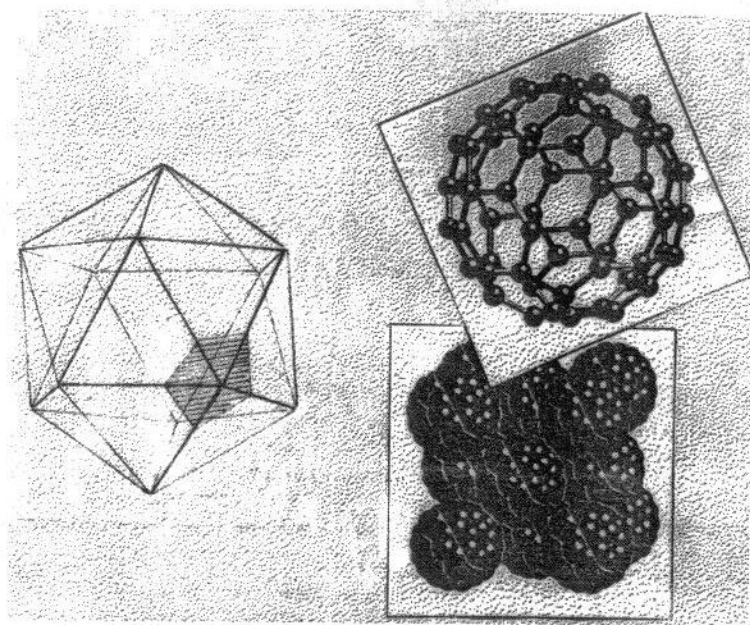


Figura VI.18. Icosaedro, icosaedro truncado y fulerenos.

construyó domos geodésicos con el mismo tipo de estructura. También reciben el nombre de *futboleros* por su gran parecido a una pelota de fútbol. Por su descubrimiento recibieron el premio Nobel de química de 1996.

Los fulerenos exhiben propiedades físicas y químicas sorprendentes (superconductividad, ferromagnetismo y gran estabilidad). Su forma, además, les permite encapsular otras moléculas y formar compuestos de variados usos prácticos. Entre estas moléculas la más estable es el carbono 60 (C_{60}), que tiene una estructura casi esférica con 12 pentágonos y 20 hexágonos unidos en su superficie y los átomos de carbón en los vértices. La estructura de C_{60} se puede obtener a partir del icosaedro por medio de ciertos cortes (o truncamientos). Recordamos que el icosaedro es un poliedro regular con 12 vértices y 20 caras que son triángulos equiláteros, además, en cada vértice concurren 5 aristas. Si dividimos cada arista del icosaedro en tres partes iguales, es fácil demostrar que los 5 puntos marcados sobre las aristas que son más cercanos a un vértice forman un pentágono.

Si cortamos al icosaedro por medio de los 12 planos que contienen estos pentágonos obtenemos el *icosaedro truncado*. Esta figura tiene precisamente $5 \times 12 = 60$ vértices, y sus caras son 12 pentágonos (los de los cortes) y 20 hexágonos (que se forman de cada una de las 20 caras triangulares del icosaedro al cortar).

El grupo de simetrías del icosaedro es A_5 . Como los cortes que hemos hecho para obtener el icosaedro truncado son estables con respecto a este grupo (es decir, un corte va a parar en otro corte cuando se aplica una simetría en el icosaedro), entonces A_5 es el grupo de isometría del fullereno C_{60} (véanse las figuras VI.18 y VI.19).

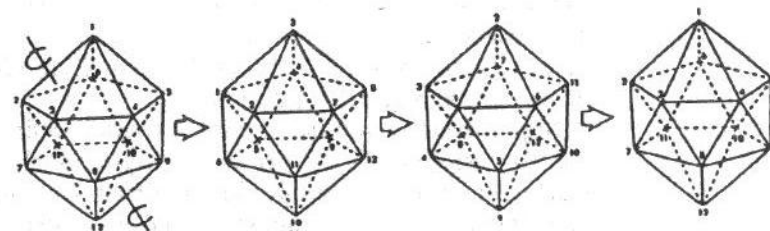


Figura VI.19. Algunos ejes de simetría del icosaedro.

Uno de los usos más recientes de los fulerenos se da en la medicina. Un equipo de investigadores de la universidad de San Luis, Missouri, consiguió proteger con ellos células nerviosas de atacantes moleculares, los *radicales libres*.

Después de una herida grave en la cabeza, el cerebro frecuentemente se ve inundado por neurotransmisores, de modo que las células nerviosas son destruidas como si se tratara de circuitos eléctricos sobrecargados. Algunas enfermedades como el ALS (síndrome de Lou Gehrig) y el Alzheimer se caracterizan por la degradación gradual de las células nerviosas hasta que los pacientes quedan paráliticos o pierden completamente sus facultades mentales. En todos estos procesos destructivos, los radicales libres parecen desempeñar un papel decisivo.

Al encontrarse con un fullereno, un radical libre (que tiene una carga eléctrica negativa) se pega a él fuertemente, pero no puede romperlo por su alta estabilidad. Si se lograra dispersar fulerenos en el sistema nervioso, éstos absorberían y harían

inocuos a los radicales libres. Para poder lograr llevar los fulerenos a lo más profundo del organismo, se requería hacerlos solubles en agua, cosa que no sucede de manera natural. El doctor Dugan y su equipo lo lograron agregando a los fulerenos "agarraderas moleculares" formadas por ácido malónico. Como las moléculas de agua se agarran del ácido malónico, el resultado es soluble en agua.

Las primeras pruebas clínicas de esta idea, realizadas en ratas, parecen muy prometedoras. Por lo pronto, los investigadores están aprendiendo mucho del papel de los radicales libres en los procesos degenerativos del cerebro.

GRECAS

Los más sencillos patrones ornamentales son las bandas unidimensionales donde una misma figura se repite con cierta simetría. Una banda así se llama *greca*. Ejemplos de grecas los vemos en las figuras VI.20 y VI.21.

Supondremos que las grecas se repiten infinitamente en ambas direcciones. En el primer ejemplo es claro que una sola traslación genera el grupo de simetrías: la traslación que lleva una "cresta" en la siguiente. ¿Cuál es el grupo en el segundo ejemplo de la figura VI.20?



Figura VI.20. Dos ejemplos de grecas: la primera es griega, la segunda medieval.

Problema. Encuentre la mayor cantidad de grupos de simetrías de grecas que le sea posible.

Solución. Podemos ilustrar varias situaciones por medio de grecas muy simples, formadas sólo por letras. Así el grupo de la primera greca de la figura VI.20 está generado por una sola traslación y es el grupo de la greca: ...bbbbbb...

La segunda greca de la figura VI.20 es más interesante, se parece a la greca: ...LΓLΓLΓ..., cuyo grupo está generado por una traslación (una L en la siguiente L) y una reflexión a través de un eje horizontal seguida de un desplazamiento (que

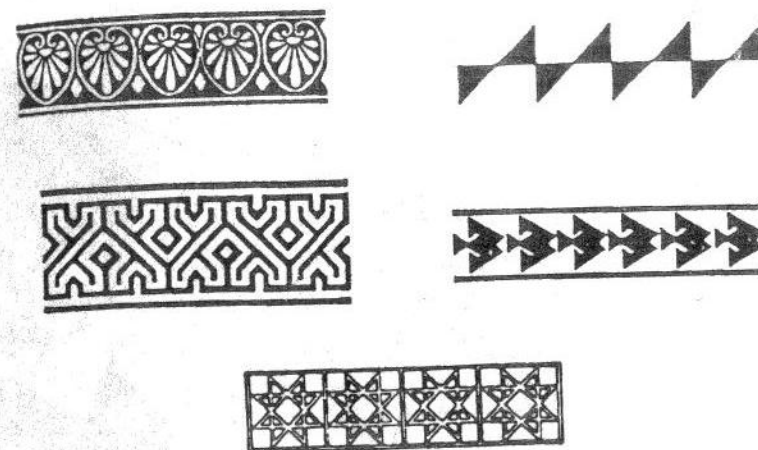


Figura VI.21. Grecas ornamentales correspondientes a los grupos (3) a (7).

lleva una L en la siguiente Γ, a este movimiento le llamaremos *torcimiento*). A los grupos de estas grecas los llamaremos grupo 1 y 2, respectivamente.

Los grupos que podemos encontrar con esta sencilla receta de considerar letras para los patrones son los siguientes:

- 1) ...b b b b b b b b... generado por una sola traslación.
- 2) ...L Γ L Γ L Γ... generado por una traslación y un torcimiento.
- 3) ...b d b d b d b d... generado por una traslación y una reflexión vertical.
- 4) ...N N N N N N N N... generado por una traslación y una rotación de 180° .
- 5) ...b d p q b d p q... generado por una reflexión vertical y una rotación de 180° .
- 6) ...E E E E E E E E... generado por una traslación y una reflexión horizontal.
- 7) ...H H H H H H H H... generado por una reflexión horizontal y dos verticales.

El hecho es que éstos son los únicos grupos que pueden presentarse como grupos de simetrías de grecas. Ejemplos de estos grupos en grecas ornamentales están dados en la figura VI.21.

VII. Pronósticos deportivos

TODO aficionado a un deporte sabe que es muy difícil predecir el resultado de un partido (si no, todo mundo ganaría jugando a los pronósticos deportivos). Hay numerosas historias de personas que han ganado grandes sumas de dinero en ellos. En España, una señora de 80 y tantos años ganó la mayor bolsa de la historia jugando una sola tarjeta de pronósticos. Cuando le preguntaron cómo había hecho para ganar, contestó que ella no sabe nada de fútbol, por lo que tomó una tarjeta ya llena que encontró en la habitación de su hijo (que sí sabe de fútbol) y la copió. Sin embargo, la señora no se percató de que la tarjeta era de los juegos de la semana anterior. Grave error que le valió el premio.

Cuando tratamos de evaluar seriamente las posibilidades de un equipo, es difícil decidir si es mejor tener un buen ataque o una buena defensa. Además, en un deporte como el fútbol las oportunidades de anotar son en general escasas y los resultados entre equipos con fuerzas similares tienden a ser circunstanciales. Sin embargo, hay ciertos principios generales que podemos entender.

Para entender y analizar de qué factores depende el resultado en un deporte (más adelante hemos de analizar con cuidado el baloncesto) necesitamos entender algunas características básicas de los juegos. Comenzaremos por analizar un juego muy simple que depende únicamente de la suerte y no de la habilidad del jugador: el juego de dados.

Juan y Pedro juegan a los dados. Cada uno elige un número entre 1 y 12 y luego arrojan los dados; ganará aquel cuyo número salga primero como suma de los dados. Juan elige 8 y Pedro 10, ¿quién ganará?

La respuesta es que cualquiera de los dos puede ganar, puesto que cualquier combinación de los dados puede salir en la primera tirada. La pregunta correcta es: ¿quién es el ganador más probable? Para contestar esta pregunta necesitamos saber cuáles son los resultados posibles de tirar los dados. Supondremos que un dado es rojo y el otro verde, si el rojo sale 1 y el verde 5 diremos que el resultado de tirar los dados fue (1,5). Con esta convención podemos enumerar todos los resultados posibles:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Esto es, hay 36 posibles resultados de tirar los dados. De éstos vemos que hay cinco maneras posibles de obtener un total de 8: con (2,6), (3,5), (4,4), (5,3) y con (6,2), siendo cada caso igualmente probable. Mientras que sólo hay 3 formas de sumar 10: con (4,6), (5,5) y con (6,4). Claramente, es más probable que Juan gane en el juego. Pero nos gustaría poder decir más: ¿cuál es la probabilidad que tiene Juan de ganar? Tiene cinco posibilidades de 36 posibles resultados, decimos que tiene entonces $P(8) = 5/36$ probabilidades de ganar, mientras que Pedro tiene $P(10) = 3/36 = 1/12$ probabilidades de ganar. En general, decimos que $P(n)$ son las probabilidades que se tienen de que el total de los dados sea n . Entonces calculamos que, $P(1) = 0$, $P(2) = 1/36$, $P(3) = 1/18$, $P(4) = 1/12$, $P(5) = 1/9$, $P(6) = 5/36$, $P(7) = 1/6$, $P(8) = 5/36$, $P(9) = 1/9$, $P(10) = 1/12$, $P(11) = 1/18$, $P(12) = 1/36$. Observemos que la suma de todas las probabilidades $P(1), \dots, P(12)$ es:

$$\sum_{n=1}^{12} P(n) = 1.$$

Queremos ahora saber cuáles son las probabilidades de Juan de que la suma de dos tiradas consecutivas sea 22. Este número se puede obtener de las siguientes formas: $10+12 = 22$, $11+11 = 22$. Si en la primera tirada obtiene 10, entonces en la segunda debe obtener 12, esto puede suceder de 3×1 maneras posibles de un total de 36×36 casos, esto sucede entonces con una probabilidad $P(10,12)$ que es igual a $P(10,12) = P(10) \cdot P(12) = 1/12 \times 1/36 = 1/432$; si en la primera tirada obtiene 11, en la segunda debe obtener también 11, lo que puede suceder con una probabilidad de $P(11,11) = P(11) \cdot P(11) = 1/324$ y finalmente, si en la primera tirada obtiene 12, en la segunda debe obtener 10, lo que puede suceder con probabilidad de $P(12,10) = P(12) \cdot P(10) = 1/432$. La probabilidad de que

una de estas tres cosas pase es la suma de las probabilidades individuales de los tres casos, esto es, Juan obtendrá 22 con la probabilidad de:

$$P(10,12) + P(11,11) + P(12,10) = \frac{5}{648},$$

por supuesto, ésta es la misma probabilidad que tiene de obtener 22 si hace una sola tirada con cuatro dados.

Con estas ideas vamos a tratar de describir matemáticamente el baloncesto. Se juega entre dos equipos de cinco jugadores. Los equipos profesionales se mueven a altas velocidades y con gran eficiencia. Un enceste en una jugada normal vale 2 puntos, salvo que se realice a cierta distancia, marcada en la cancha, en cuyo caso vale 3 puntos. También hay tiros de castigo que de encestar se valen solamente 1 punto. Por supuesto, para describir un modelo matemático del juego necesitamos considerarlo como un juego mecánico donde cada equipo tiene cierta probabilidad de pasar de la defensa al ataque y luego otra probabilidad de anotar. Nuestra aspiración es que este modelo baste para poder hacer un programa de computadora que permita predecir los resultados de partidos reales. Desgraciadamente, un modelo así desconoce todos los pases y jugadas que un jugador puede hacer, de hecho, se está muy lejos de lograr un modelo de Michael Jordan en computadora.

En nuestro partido se enfrentan los equipos A y B. En cualquier momento del juego las siguientes situaciones pueden ocurrir:

- La bola está en posesión del equipo A y va al ataque, diremos entonces que el juego se encuentra en el estado S_1 .
- El equipo A realiza un ataque a fondo, diremos entonces que el juego se encuentra en el estado S_2 .
- El equipo A anota una canasta de 2 puntos. El juego está en el estado S_3 .
- El equipo A anota una canasta de 3 puntos. El juego se encuentra en el estado S_4 .
- La bola se encuentra en posesión del equipo B en la defensa. El juego está entonces en el estado S_5 .
- El equipo B pasa de la defensa al ataque, diremos entonces que el juego se encuentra en el estado S_6 .

- El equipo B realiza un ataque a fondo, diremos entonces que el juego se encuentra en el estado S_7 .
- El equipo B anota una canasta de 2 puntos. El juego está en el estado S_8 .
- El equipo B anota una canasta de 3 puntos. El juego se encuentra en el estado S_9 .
- La bola se encuentra en posesión del equipo A en la defensa. El juego está entonces en el estado S_{10} .

Notamos que en nuestro modelo no hay faltas ni tiros de castigo (ni encestes de 1 punto). Claramente, el juego puede pasar de ciertos estados a otros, pero no a cualquier otro. Cada vez que el juego pasa de un estado S_i a otro S_j , decimos que hay una *transición de estados* (figura VII.1) y que la probabilidad de que esto pase es $P(i, j)$. En el juego de baloncesto las posibles transiciones de estados están dadas como sigue (por supuesto marcamos con una flecha de S_i a S_j cuando es posible la transición del estado S_i al estado S_j , esto es, cuando $P(i, j) \neq 0$).

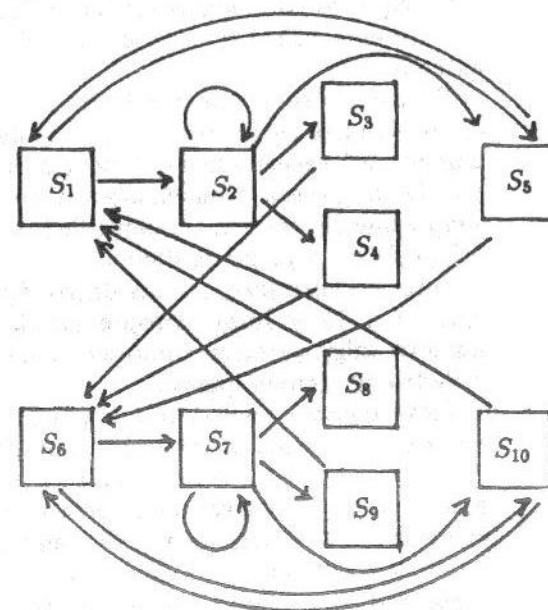


Figura VII.1. Las transiciones del modelo de baloncesto.

Las probabilidades de transición entre los diferentes estados no son todas independientes. Notamos las siguientes relaciones:

a) Supongamos que el juego se halla en el estado S_1 . Entonces o bien el equipo A liga un buen ataque y se pasa al estado S_2 , o bien la bola cae al equipo B en la defensa, o sea, pasamos al estado S_5 . Esto se expresa como:

$$P(1,2) + P(1,5) = 1.$$

De la misma manera se tiene que $P(5,6) + P(5,1) = 1$.

b) Si el equipo A está al ataque, debe realizar un tiro. Pueden suceder varias cosas: hay un enceste de 2 puntos y el juego va en el estado S_3 o hay un enceste de 3 puntos y el juego está en S_4 o la bola rebota en el tablero, la recupera el equipo A y regresamos al estado S_2 o finalmente, la bola pasa al equipo B a la defensiva. O sea:

$$P(2,3) + P(2,4) + P(2,2) + P(2,5) = 1.$$

De la misma manera se tiene que $P(6,7) + P(6,8) + P(6,6) + P(6,10) = 1$.

c) Si el equipo A anota, entonces el B pasa a la ofensiva. Esto es $P(3,6) = 1$, $P(4,6) = 1$ y también $P(7,1) = 1$, $P(8,1) = 1$.

d) Finalmente, si el equipo A recibe la bola a la defensiva, puede pasar a atacar o bien perder la bola y entonces el equipo B estará inmediatamente al ataque. Esto es:

$$P(10,1) + P(10,6) = 1.$$

De la misma manera se tiene que $P(5,6) + P(5,2) = 1$.

Aclaremos un poco más. $P(10,1)$ mide la capacidad del equipo A de pasar de su línea defensiva a media cancha, mientras que $P(1,2)$ mide su capacidad de pasar de media cancha a la línea de ataque y tirar. Por razones prácticas supondremos que $P(10,1) = P(1,2)$ y también que $P(5,6) = P(6,7)$.

Definimos la *matriz de transición* P de tamaño 10×10 correspondiente al partido entre los equipos A y B de la siguiente manera:

$$P = (P(i,j))_{ij},$$

que también denotaremos simplemente como $P = (P(i,j))$.

Supongamos que el juego inicia en el estado S_1 . Esto se puede expresar diciendo que en el momento $t = 0$ la probabilidad $S_1(0)$ de que el sistema (el partido de baloncesto) esté en el estado S_1 es 1. Pero entonces, $S_i(0) = 0$ para $i = 2, \dots, 10$. El siguiente momento $t = 1$ lo medimos una vez que el sistema haya cambiado de estado, entonces estará en el estado S_i con una probabilidad $S_i(1) = P(1,i)$.

En general, después de k transiciones de estado el sistema estará en el estado S_i con una probabilidad $S_i(k)$. ¿Cuánto vale este número? Por ejemplo, si $k = 2$, entonces pudimos pasar del estado S_1 al estado S_2 después de dos transiciones de estado de 10 maneras diferentes (de S_1 a S_i y luego de S_i a S_2 para toda $i = 1, \dots, 10$), con probabilidad $S_2(2) = \sum_{i=1}^{10} P(1,i) \cdot P(i,2)$. ¡Pero ésta es la entrada $(2,2)$ de la matriz $P^2 = P \cdot P$!

Consideraremos las potencias P^k de la matriz P y escribiremos $P^k = (P^k(i,j))$, donde $P^k(i,j)$ es la entrada del renglón i y la columna j de P^k . Lo que observábamos antes para el caso $k = 2$ se puede demostrar en forma general. Tenemos:

Proposición. La probabilidad de que el sistema se encuentre en el estado $S_i(k)$ después de k transiciones de estado es:

$$S_i(k) = P^k(1,i).$$

Demostración. Avancemos paso por paso. Para $k = 1$ sabemos que $S_i(1) = P(1,i)$. Supongamos que después de $k - 1$ transiciones de estado el resultado que deseamos probar todavía es cierto, es decir que $S_i(k - 1) = P^{k-1}(1,i)$ para toda $i = 1, \dots, 10$. ¿De cuáles maneras podemos pasar del estado S_1 al estado S_i con k transiciones de estado? Después de $k - 1$ transiciones estaremos en el estado S_j con probabilidad $S_j(k - 1)$ y podremos pasar del estado S_j al estado S_i con probabilidad $P(j,i)$. Entonces obtenemos:

$$\begin{aligned} S_i(k) &= \sum_{j=1}^{10} S_j(k-1) \cdot P(j,i) \\ &= \sum_{j=1}^{10} P^{k-1}(1,j) \cdot P(j,i) = P^k(1,i), \end{aligned}$$

donde en la última igualdad hemos usado que $P^k = P^{k-1} \cdot P$. \square

La intuición nos dice que siendo un juego de baloncesto una larga sucesión de jugadas (esto es, transiciones de estado), las probabilidades $S_i(k)$ con k grande deben de indicar lo que sucede al final del juego. ¿Cómo podemos usar este modelo para predecir el resultado final de un partido?

Antes de considerar la respuesta a esta pregunta, sería bueno saber si en la realidad podemos conocer las probabilidades de transición $P(i, j)$ de los equipos de baloncesto. La respuesta es afirmativa. Al menos en los partidos de baloncesto profesional se lleva el registro de las jugadas efectuadas y con ellas se hacen estadísticas de los equipos. Los parámetros que necesitamos conocer son los siguientes:

- $P(1, 2)$: que es la probabilidad que tiene el equipo A de pasar de la defensa al ataque. Este número es también $P(1, 2) = 1 - P(1, 5)$, donde $P(1, 5)$ es la probabilidad de que el equipo B detenga el avance del equipo A . Lo que se conoce en las estadísticas de baloncesto profesional es la *eficiencia defensiva* de un equipo, es decir, el número de veces que detiene avances del equipo contrario sobre el total de jugadas. Supondremos que si d_B es la eficiencia defensiva del equipo B , entonces $P(1, 2) = 1 - d_B$.
- $P(2, 2)$: ésta es la probabilidad de que el equipo A pase de estar atacando a volver a atacar, es decir, que después de hacer un tiro que no es anotación, el equipo A recupere la pelota. Esto se mide como el *tablero* promedio del equipo, que denotaremos como t_A . Supondremos entonces $P(2, 2) = t_A$.
- $P(2, 3)$: ésta es la probabilidad de que al atacar se logre una anotación de 2 puntos. Durante un campeonato se mide la *eficiencia promedio en tiros de 2 puntos* $e_A^{(2)}$, calculada como el cociente $E_A^{(2)}/I_A^{(2)}$ del total de encestes de 2 puntos $E_A^{(2)}$ entre el total de intentos $I_A^{(2)}$. Tenemos que $P(2, 3) = E_A^{(2)}/(I_A^{(2)} + I_A^{(3)})$, donde $I_A^{(3)}$ es el total de intentos de anotación de 3 puntos por el equipo A .
- $P(2, 4)$: ésta es la probabilidad de que al atacar se logre una anotación de 3 puntos. Durante un campeonato se mide la *eficiencia promedio en tiros de 3 puntos* $e_A^{(3)}$, calculada como el cociente $E_A^{(3)}/I_A^{(3)}$ del total de encestes de 3 puntos

$E_A^{(3)}$ entre el total de intentos $I_A^{(3)}$. Tenemos que $P(2, 4) = E_A^{(3)}/(I_A^{(2)} + I_A^{(3)})$.

El número promedio de veces que la bola pasa de un equipo a otro en un partido de baloncesto profesional es aproximadamente 110. Tenemos entonces las siguientes relaciones:

$$2 \cdot e_A^{(2)} \cdot I_A^{(2)} + 3 \cdot e_A^{(3)} \cdot I_A^{(3)} = 2 \cdot E_A^{(2)} + 3 \cdot E_A^{(3)} = f_A$$

y aproximadamente:

$$I_A^{(2)} + I_A^{(3)} + 110 \cdot (-t_A + d_B) = 110,$$

donde f_A es el promedio por partido de los puntos a favor del equipo A . De aquí se pueden despejar fácilmente los valores de $I_A^{(2)}$ e $I_A^{(3)}$ y luego calcular:

$$P(2, 3) = \frac{e_A^{(2)}}{(2 \cdot e_A^{(2)} - 3 \cdot e_A^{(3)})} \left[\frac{f_A}{110 \cdot (1 + t_A - d_B)} - 3 \cdot e_A^{(3)} \right]$$

y:

$$P(2, 4) = \frac{e_A^{(3)}}{(2 \cdot e_A^{(2)} - 3 \cdot e_A^{(3)})} \left[-\frac{f_A}{110 \cdot (1 + t_A - d_B)} + 2 \cdot e_A^{(2)} \right].$$

Según estas observaciones nos basta conocer la información d_A , t_A , $e_A^{(2)}$, $e_A^{(3)}$ y f_A de cada equipo A . La información (en forma algo diferente) de los equipos profesionales de la liga mayor de Estados Unidos (la NBA) se encuentra en periódicos y revistas. En la tabla reproducimos algunos datos importantes de los equipos que participaron en los partidos de *playoffs* de la temporada 1997 de la NBA, estos promedios excluyen los resultados de los partidos de la final entre Chicago y Utah.

Equipo	f	$e^{(2)}$	$e^{(3)}$	t
Seattle	101.4	.428	.354	.124
Houston	100.0	.478	.338	.122
Utah	99.85	.446	.354	.130
Chicago	96.6	.426	.357	.172
L. A. Lakers	96.1	.505	.325	.098
Nueva York	92.1	.464	.344	.075
Atlanta	89.2	.453	.331	.069
Miami	88.8	.442	.387	.090

El partido de la final de la NBA se jugó entre Chicago (equipo A) y Utah (equipo B). Podemos entonces calcular la matriz de transición $P = (P(i, j))$ del partido Chicago-Utah como sigue:

$$P = \begin{pmatrix} 0 & .84 & 0 & 0 & .16 & 0 & 0 & 0 & 0 & 0 \\ 0 & .172 & .325 & .083 & .42 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ .2 & 0 & 0 & 0 & 0 & .8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .8 & 0 & 0 & .2 \\ 0 & 0 & 0 & 0 & 0 & 0 & .130 & .225 & .172 & .573 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .84 & 0 & 0 & 0 & 0 & .16 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Hay un dato adicional que hemos usado en la construcción de la matriz de transición anterior. Esto es que el equipo de Chicago tuvo la segunda mejor defensa de toda la liga y Utah la sexta mejor defensa. Entonces se debe tener que $d_A > d_B$. Obsérvese que en la matriz hemos puesto $d_A = .2$ y $d_B = .16$.

MATRICES ESTOCÁSTICAS Y LOS "TOROS" DE CHICAGO

Un sistema con un número finito de estados y una matriz de transición como hemos considerado en el apartado anterior se llama *cadena de Markov*, en honor del matemático ruso que hizo estas importantes contribuciones a la teoría de probabilidades. Consideraremos ahora con más atención las propiedades de las matrices de transición.

Observemos la matriz de transición $P = (P(i, j))$ definida antes. Ésta cumple las siguientes propiedades: E1) Las entradas $P(i, j)$ de la matriz son números mayores o iguales a 0. E2) La suma de las entradas de cada renglón $\sum_{j=1}^{10} P(i, j) = 1$, ($i = 1, \dots, 10$).

Una matriz con las propiedades E1) y E2) se llama *matriz estocástica*.

En el modelo del juego de baloncesto, en un momento dado k , el sistema debe encontrarse en alguno de los estados $S_i(k)$ con $1 \leq i \leq 10$. Entonces $\sum_{i=1}^{10} S_i(k) = 1$. Esta propiedad es parte del inciso (1) del teorema siguiente.

Teorema. Sea $P = (p_{ij})$ una matriz estocástica de tamaño $n \times n$. Entonces se cumplen las siguientes propiedades:

- 1) Toda potencia P^k es también una matriz estocástica.
- 2) Existe un vector renglón $v^* = (v_1^*, \dots, v_n^*)$ con la propiedad de que $v^* \cdot P = v^*$ y $\sum_{i=1}^n v_i = 1$. En particular, también se tiene $v^* \cdot P^k = v^*$ para toda $k \geq 1$.
- 3) En caso de que exista un número $k_0 > 0$ tal que P^{k_0} tenga todas sus entradas positivas, entonces el vector v^* del apartado 2) es único. En ese caso, para todo vector renglón $w \neq 0$ el límite $\lim_{k \rightarrow \infty} w \cdot P^k$ existe y es igual a λv^* para algún escalar λ .

* *Demostración.* 1) Obviamente para $k = 1$ el resultado es cierto. Supongamos que P^{k-1} es estocástica, esto es:

$$\sum_{j=1}^n P^{k-1}(i, j) = 1,$$

para toda $i = 1, \dots, n$. Entonces para una i fija tenemos:

$$\begin{aligned} \sum_{j=1}^n P^k(i, j) &= \sum_{j=1}^n \left(\sum_{t=1}^n P^{k-1}(i, t) \cdot P(t, j) \right) \\ &= \sum_{t=1}^n P^{k-1}(i, t) \cdot \left(\sum_{j=1}^n P(t, j) \right) \\ &= \sum_{t=1}^n P^{k-1}(i, t) = 1, \end{aligned}$$

donde hemos usado que las sumas las podemos efectuar en cualquier orden sin cambiar su valor, y en los últimos dos pasos usamos que las matrices P y P^{k-1} son estocásticas.

2) Para esta demostración requerimos algunas ideas elementales de álgebra lineal. Si el lector no está familiarizado con ellas, puede seguir adelante sin leer la demostración.

Un vector $v \neq 0$ tal que $v \cdot P = v$ satisface que $v \cdot (P - I) = 0$, donde I es la matriz identidad de tamaño $n \times n$. Esto sucede sólo si las coordenadas del vector $v = (v_1, \dots, v_n)$ son solución no trivial (esto es, no todas las $v_i = 0$) del sistema de ecuaciones:

$$(p_{11} - 1)x_1 + p_{21}x_2 + \cdots + p_{n1}x_n = 0$$

$$p_{12}x_1 + (p_{22} - 1)x_2 + \cdots + p_{n2}x_n = 0$$

$$\dots \dots \dots$$

$$p_{1n}x_1 + p_{2n}x_2 + \cdots + (p_{nn} - 1)x_n = 0.$$

Esto sólo es posible si los renglones del sistema de ecuaciones:

$$w_1 = (p_{11} - 1, p_{21}, \dots, p_{n1})$$

$$w_2 = (p_{12}, p_{22} - 1, \dots, p_{n2}), \dots, w_n = (p_{1n}, p_{2n}, \dots, p_{nn} - 1)$$

son vectores linealmente dependientes. Éste es el caso puesto que la suma de los vectores w_1, \dots, w_n resulta ser:

$$\sum_{i=1}^n w_i = \left(\sum_{i=1}^n p_{1i} - 1, \sum_{i=1}^n p_{2i} - 1, \dots, \sum_{i=1}^n p_{ni} - 1 \right)$$

que es un vector con 0 en todas las entradas, ya que P es estocástica. Esto demuestra que el vector $v \neq 0$ con la propiedad $v \cdot P = v$ existe.

Para construir el vector v^* como en el enunciado, requerimos probar que $\sum_{i=1}^n v_i \neq 0$. Supongamos que esto no es cierto, esto es, $\sum_{i=1}^n v_i = 0$. Podemos reordenar las entradas de v de forma que v_1, \dots, v_t sean todas mayores o iguales a 0 y v_{t+1}, \dots, v_n sean todas negativas. Usamos que $v_i = \sum_{j=1}^n v_j p_{ji}$ para obtener:

$$\sum_{i=1}^t \sum_{j=1}^n v_j p_{ji} = \sum_{i=1}^t v_i = - \sum_{i=t+1}^n v_i,$$

de donde:

$$\sum_{j=1}^t v_j \geq \sum_{j=1}^t v_j \left(\sum_{i=1}^t p_{ji} \right) = - \sum_{j=t+1}^n \left(1 + \sum_{i=1}^t p_{ji} \right) v_j \geq - \sum_{i=t+1}^n v_i,$$

lo cual sólo sería posible si $\sum_{i=1}^t p_{ji} = 0$ para toda j entre 1 y n . Como todas las $p_{ji} \geq 0$, entonces esto implica $p_{ji} = 0$ para $1 \leq i \leq t$ y $1 \leq j \leq n$. Pero entonces también P_0^k tiene entradas 0 (de hecho, la entrada $(1, 1)$). Esto es una contradicción que muestra que $\lambda = \sum_{i=1}^n v_i \neq 0$. Para definir ahora v^* tomamos simplemente $v^* = \lambda^{-1}v$.

3) En lugar de dar la prueba general que requeriría de más preparación, estudiaremos con un poco de detenimiento el caso $n = 2$.

Una matriz estocástica P de tamaño 2×2 se ve:

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

con $0 \leq p, q \leq 1$. Si $p = 0 = q$, entonces la matriz $P = I$ y no hay nada interesante que decir. Supongamos por ello que $p > 0$. Podemos entonces definir la matriz S tal que:

$$S = \begin{pmatrix} q & p \\ 1 & -1 \end{pmatrix}, \quad S^{-1} = \frac{1}{p+q} \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix}$$

de forma que $S^{-1} \cdot S = I$ y además $S \cdot P \cdot S^{-1} = \text{diag}(1, \mu)$, donde $\mu = 1 - p - q < 1$ y $\text{diag}(1, \mu)$ es una matriz diagonal de la forma:

$$\text{diag}(1, \mu) = \begin{pmatrix} 1 & 0 \\ 0 & \mu \end{pmatrix}.$$

Supongamos ahora que P^{k_0} tiene todas sus entradas positivas. Como $P^k = S^{-1} \cdot \text{diag}(1, \mu^k) \cdot S$, entonces $0 < \mu < 1$. Un vector v con la propiedad $v \cdot P = v$ define un vector $w = v \cdot S^{-1}$ de forma que $w \cdot S \cdot P \cdot S^{-1} = v \cdot P \cdot S^{-1} = v \cdot S^{-1} = w$. Pero $w \cdot \text{diag}(1, \mu) = w$ tiene solución $w = (1, 0)$ y ésta es única hasta múltiplos escalares. Por lo tanto, el vector $v = w \cdot S = (q, p)$ es el único hasta escalares que satisface $v \cdot P = v$. Tomando $v^* = (q, p)/p+q$ se obtiene que $v^* \cdot P = v^*$ y además $v_1^* + v_2^* = 1$. Podemos calcular el límite:

$$\begin{aligned} \lim_{k \rightarrow \infty} P^k &= S^{-1} \cdot \lim_{k \rightarrow \infty} \text{diag}(1, \mu^k) \cdot S = \\ &= S^{-1} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot S \\ &= \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix}. \end{aligned}$$

Entonces queda claro que para todo vector $w = (w_1, w_2)$ el siguiente límite existe:

$$\lim_{k \rightarrow \infty} w \cdot P^k = w \cdot \lim_{k \rightarrow \infty} P^k = (w_1 + w_2)v^*,$$

lo que termina la demostración. \square

El vector v^* que satisface que $v^* \cdot P = v^*$ se llama un *vector propio* de la matriz P . ¿Cómo se calcula este vector? Observemos que en caso de que las hipótesis del teorema anterior se satisfagan, podemos elegir este vector cumpliendo las dos siguientes ecuaciones:

$$v^* \cdot (P - I) = 0, \quad \sum_{i=1}^n v_i^* = 1,$$

donde I denota la matriz identidad de tamaño $n \times n$.

Para el caso de nuestro modelo del baloncesto observaremos que se cumplen todas las hipótesis del teorema para la matriz P . Para ello nos basta comprobar que existe un número k tal que P^k tiene todas sus entradas positivas. Esto podemos deducirlo fácilmente de la gráfica asociada a nuestro modelo en la figura VII.1. En efecto, dados dos estados S_i y S_j del sistema siempre hay un camino que va desde S_i hasta S_j . Si r es la longitud de un camino así, entonces $P^r(i, j) > 0$. Una sencilla inspección de la gráfica muestra que hay caminos de longitud 4 desde cualquier estado a cualquier otro, por lo tanto P^4 tiene todas sus entradas positivas.

Es particularmente importante calcular el vector propio v^* ya que tenemos la siguiente interpretación: $v_i^* = \lim_{k \rightarrow \infty} P^k(1, i)$ es la probabilidad de que el partido se encuentre en el estado S_i después de que se han efectuado muchas jugadas. En particular v_3^* y v_4^* son las probabilidades de que el equipo A anote (2 y 3 puntos respectivamente) en una jugada cualquiera una vez que el partido ha alcanzado su situación estable, mientras que v_8^* y v_9^* son las probabilidades de que el equipo B anote (2 y 3 puntos respectivamente) en una jugada cualquiera. De forma que el *marcador más probable* para nuestro partido de baloncesto es $m_A : m_B$ donde el cociente m_A/m_B es:

$$\frac{m_A}{m_B} = \frac{(2 \cdot v_3^* + 3 \cdot v_4^*)}{(2 \cdot v_8^* + 3 \cdot v_9^*)}.$$

Calculemos el vector v^* para el partido Chicago-Utah cuya matriz de transición P hemos definido antes explícitamente. El vector v^* satisface el sistema de ecuaciones:

$$.2v_5^* + v_8^* + v_9^* + .84v_{10}^* = v_1^*$$

$$.84v_1^* + .172v_2^* = v_2^*$$

$$.325v_2^* = v_3^*$$

$$.083v_3^* = v_4^*$$

$$.16v_1^* + .42v_2^* = v_5^*$$

$$v_3^* + v_4^* + .8v_5^* + .16v_{10}^* = v_6^*$$

$$.8v_6^* + .13v_7^* = v_7^*$$

$$.225v_7^* = v_8^*$$

$$.172v_8^* = v_9^*$$

$$.2v_6^* + .473v_7^* = v_{10}^*$$

$$v_1^* + v_2^* + \dots + v_{10}^* = 1$$

La solución de este sistema para las entradas que nos interesan es:

$$v_3^* = .056, \quad v_4^* = .0143, \quad v_8^* = .034, \quad v_9^* = .0264,$$

de donde finalmente el marcador más probable $m_A : m_B$ satisface:

$$\frac{m_A}{m_B} = \frac{.1549}{.1472} = 1.052,$$

que representa con bastante buena aproximación un marcador 82:78 en favor de Chicago. Es interesante notar que la serie de la final de baloncesto de la NBA terminó a mediados de junio (de 1997) en 6 partidos favoreciendo a Chicago 4 juegos a 2. En promedio Chicago ganó por 3 puntos.

¿CUÁNTOS CAMINOS LLEVAN A ROMA?

El señor X es un hombre de negocios que tiene empresas en cuatro ciudades: México, Roma, París y Londres. Cuando viaja entre estas ciudades siempre toma uno de los vuelos indicados en la figura VII.2.

¿Cuántas rutas diferentes desde México hasta Roma puede tomar pasando por cuatro ciudades? ¿Pasando por 10 ciudades?

Cada vez que hace un viaje, el señor X permanece una semana en cada ciudad antes de hacer uno nuevo. Una vez en una

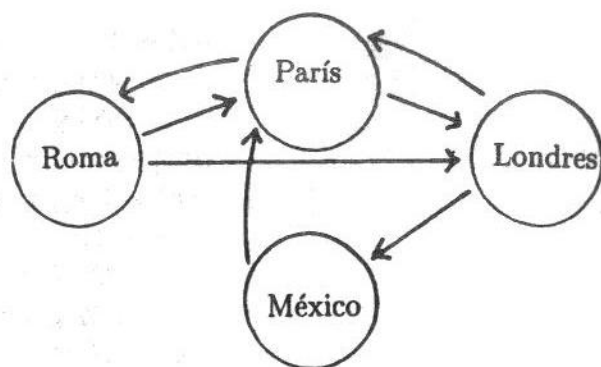


Figura VII.2. La gráfica de viajes del señor X.

ciudad, elige el siguiente destino con igual probabilidad de entre los vuelos permitidos por la gráfica. En una semana ¿cuál es la probabilidad de que el señor X esté en Roma?, ¿en México?

Solución. Construyamos la matriz de vuelos del señor X: si 1 representa a Roma, 2 a París, 3 a Londres y 4 a México, entonces la matriz $V = (v_{ij})$ de vuelos tiene tamaño 4×4 y la entrada v_{ij} es 1 si hay vuelo de la ciudad i a la j y es 0 si no hay vuelo. Entonces:

$$V = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

El número de rutas de longitud 1 que puede tomar el señor X lo dan las entradas de la matriz V . Las de longitud 2 las entradas de la matriz $V^2 = (v_{ij}^{(2)})$, las de longitud 3 el cubo $V^3 = (v_{ij}^{(3)})$ de la matriz V , etcétera. En efecto, si las rutas de longitud $k - 1$ de la ciudad i a la j están dadas por $v_{ij}^{(k-1)}$, que es la entrada (i, j) de la matriz V^{k-1} , entonces las rutas de longitud k de la ciudad i a la j se construyen como sigue: primero tiene que viajar una ruta de longitud $k - 1$ desde la ciudad i hasta alguna otra ciudad t y luego de t a j en un vuelo directo, esto es:

$$v_{ij}^k = \sum_{t=1}^4 v_{it}^{(k-1)} \cdot v_{tj},$$

que es precisamente la entrada (i, j) de la matriz V^k . Podemos entonces calcular V^{10} :

$$V^{10} = \begin{pmatrix} 97 & 177 & 149 & 81 \\ 96 & 178 & 149 & 81 \\ 81 & 149 & 125 & 68 \\ 53 & 96 & 81 & -44 \end{pmatrix}$$

En particular, hay 53 rutas de México a Roma pasando por 10 ciudades.

Para considerar el problema de la probabilidad que el señor X esté en la ciudad de México, debemos considerar la matriz de probabilidades P de los vuelos. Se nos ha indicado que estando en la ciudad i es igualmente probable que tome cualquiera de los vuelos que salen de i , esto es:

$$P = \begin{pmatrix} 0 & .5 & .5 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .5 & 0 & .5 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Estamos considerando una cadena de Markov con 4 estados: donde S_i sucede si el señor X está en la ciudad i . La matriz P es la matriz de transición entre los estados de este sistema. Como hemos visto, la probabilidad de que el sistema se encuentre en el estado S_i después de mucho tiempo está dada por la entrada i del vector renglón v^* que es vector propio de P , esto es $v^* \cdot P = v^*$. Podemos resolver el sistema de ecuaciones que resulta o podemos calcular potencias P^k con k grande (recordemos que $v^* = \lim_{k \rightarrow \infty} v \cdot P^k$, donde v es el primer renglón de P). Obtenemos aproximadamente:

$$v^* = (.19, .38, .285, .142)$$

Así, la probabilidad de que el señor X esté en Roma en una semana cualquiera es de .19. O sea, más o menos 1 de cada 5 semanas el está en Roma y 1 de cada 7 en México.

* OTRAS APLICACIONES

Las cadenas de Markov tienen variadas aplicaciones. En economía, biología, física y otros campos. Consideraremos aquí un par de aplicaciones.

1) La instrucción primaria en México tiene una duración de seis años. Al final del año escolar, cada estudiante se enfrenta a las siguientes opciones: dejar la escuela, repetir el curso en el que estuvo registrado ese año o bien pasar al siguiente curso del ciclo.

Al final del año j -ésimo ($1 \leq j \leq 6$), el escolar enfrenta las diferentes opciones de acuerdo con las siguientes expectativas: a) con una probabilidad de $p_j < 1$ no continuará los estudios primarios; b) con una probabilidad de $q_j < 1$ repetirá curso.

Si $j < 6$, con una probabilidad de r_j , el estudiante será promovido al siguiente curso; o bien si $j = 6$, con una probabilidad de r_6 , el estudiante recibirá su certificado de estudios primarios.

En particular, tenemos que $p_j + q_j + r_j = 1$. Para simplificar el problema supondremos que la probabilidad de deserción y de repetición de curso son constantes a lo largo de la educación, esto es, $p_j = p$, $q_j = q$ y por tanto $r_j = r$ para toda $j = 1, \dots, 6$.

Problemas. a) Calcule la probabilidad que tiene un estudiante que se encuentra en el j -ésimo año del ciclo primario de obtener su certificado después de algún tiempo. Para resolver el problema supondremos que no hay límite de tiempo en que el estudiante puede estar inscrito en el ciclo primario.

b) Dado un índice de deserción p fijo, calcule cuál es el máximo valor de q que permite que al menos tres cuartas partes de la población del país obtengan el certificado de primaria.

Solución. La situación de un escolar en el ciclo de educación primaria puede entenderse como un sistema con ocho estados, que numeraremos en la forma siguiente: 1) El escolar recibe su certificado de educación primaria; 2) El escolar suspendió sus estudios.

$2 + j$ ($1 \leq j \leq 6$): El escolar está en el j -ésimo año del ciclo de instrucción.

La matriz de transición del sistema es por lo tanto:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & p & q & r & 0 & 0 & \dots & 0 & 0 \\ 0 & p & 0 & q & r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & p & 0 & 0 & 0 & 0 & \dots & q & r \\ r & p & 0 & 0 & 0 & 0 & \dots & 0 & q \end{bmatrix}$$

Nuestro problema consiste en demostrar que $\lim_{k \rightarrow \infty} A^k$ existe y calcularlo. En efecto, suponiendo que la matriz $\lim_{k \rightarrow \infty} A^k$ está bien definida, la entrada $(1, j+2)$ de ella es la probabilidad de que un escolar en el año j del ciclo primario reciba alguna vez el certificado de primaria.

Se puede obtener que el límite $\lim A^k$ existe y que es igual a:

$$\lim A^k = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ b_1 & 1 - b_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_6 & 1 - b_6 & 0 & \dots & 0 \end{bmatrix}$$

donde $b_j = qb_j + rb_{j+1}$ para $1 \leq j \leq 5$ y $b_6 = r + qb_6$.

De aquí se sigue que las probabilidades que conciernen a la solución de nuestro problema son las siguientes:

$$b_6 = \frac{r}{1 - q} = \frac{r}{p + r}$$

$$b_j = \frac{rb_j + 1}{1 - q} = \left(\frac{r}{p + r} \right)^{7-j}$$

Es decir, la probabilidad de que un estudiante en el grado j -ésimo obtenga alguna vez el certificado de primaria es $\left(\frac{r}{p+r} \right)^{7-j}$.

En el problema b) se nos pregunta cuál es la máxima probabilidad de reprobación q tal que $\frac{3}{4} \leq b_1 = \left(\frac{r}{p+r} \right)^6$. Entonces si hacemos $c = \sqrt[6]{3/4}$, la solución es:

$$20.27p = \left(\frac{c}{1 - c} \right) p \leq r,$$

$$q = 1 - p - r \leq 1 - (21.27)p.$$

Para que sea posible alcanzar $(3/4) \leq b_1$ se debe tener que $p \leq .047$, esto es, cuando mucho 1 de cada 21 estudiantes pueden desertar en cada ciclo escolar. En caso de que $p = .04$ (o sea, cada ciclo escolar 1 de cada 25 alumnos deserta), entonces q deberá ser menor que .14, es decir, a lo más podrían reprobar en cada ciclo 3 de cada 20 alumnos.

Otro problema. Una fábrica tiene n máquinas del mismo tipo y desea mantener trabajando la mayor cantidad posible. Por ello, la política de la fábrica es que toda máquina que se encuentre descompuesta al principio de una semana, sea reparada para el inicio de la semana siguiente. La probabilidad de que una máquina se descomponga a lo largo de una semana es de p con $0 < p < 1$. Calcule: i) La probabilidad de que en una semana fallen exactamente j de las máquinas que al inicio de la semana estaban en buen estado. ii) Encuentre la matriz de transición de este proceso (definiendo el estado i como la situación en la que al principio de una semana hay i máquinas funcionando correctamente y $n - i$ descompuestas). iii) ¿Cuál es la probabilidad de que una máquina dada se encuentre funcionando correctamente después de mucho tiempo? ¿De que se encuentre descompuesta?

VIII. ¿Sueñan los androides con ovejas eléctricas?

Yo he visto cosas que ustedes no creerían. He visto naves en llamas más allá de Orión. He visto rayos C brillar en la oscuridad cerca de la puerta de Tannhäuser. Todos esos momentos se perderán en el tiempo como lágrimas en la lluvia. Es hora de morir.

Replicante Roy, en el film *Blade Runner*

El 16 de mayo de 1997, el campeón mundial de ajedrez, Gari Kasparov, fue vencido por una computadora. La *Deep Blue* fa-



Figura VIII.1. Gari Kasparov después de su derrota frente a *Deep Blue*.

bricada por IBM ganó un torneo de seis partidas por 3.5 a 2.5 puntos después de derrotar a Kasparov en dos juegos habiendo sólo perdido uno (con varios empates). Era la primera serie en su vida en la que Kasparov salía derrotado. El torneo se dio en medio de una gran publicidad, parcialmente financiada por IBM y en gran medida por el interés de presenciar el enfrentamiento hombre-máquina. Una vez terminada la contienda, los periódicos hablaban de una derrota histórica para el hombre. Por su parte, Kasparov se mostraba indignado ya que decía, "la máquina ha sido especialmente preparada para ganarme" y hablaba de que en una revancha seguramente saldría vencedor.

Lejos quedaron los tiempos de los primeros enfrentamientos de grandes maestros del ajedrez con computadoras en las que el humano siempre (y fácilmente) resultaba vencedor. El maestro David Levy, quien fuera uno de los pioneros en estos duelos, perdió con el programa Fritz 4 para PC luego de que 10 años

antes venciera sin problemas a la primera versión Fritz 1. Un año antes de su duelo fatal, Kasparov había vencido a *Deep Blue* y pronosticaba que volvería a hacerlo. Pronóstico fallido.

Sin duda, las máquinas han evolucionado mucho, llegando a niveles que eran sólo fantasías de ciencia ficción hace pocos años. Pero, a pesar de todo, no ha llegado aún el día en que pensemos que las máquinas han superado al hombre en inteligencia y habilidades creativas. Pero, ¿llegará el día en que las máquinas superen al hombre? ¿Podemos hacer un pronóstico?

Nuestro sentimiento de superioridad sobre las máquinas se debe a que pensamos que sólo hacen lo que les ordenamos hacer y de la manera en que son instruidas para hacerlo, es decir, pensamos que las máquinas carecen de inteligencia, ingenio, creatividad e iniciativa, entre otras cosas. Después de todo, *Deep Blue* sólo juega ajedrez y analiza (a una velocidad tremenda de dos millones de posibilidades por segundo) cuál es la mejor jugada que puede hacer siguiendo las instrucciones de un programa preparado por un ser humano. Pero esto no va a ser así siempre. Uno de los padres de la era de las computadoras, el matemático John von Neumann, concibió desde sus inicios a las computadoras como una clase especial de autómatas. Un autómata en el sentido de Von Neumann es cualquier pieza de maquinaria cuyo comportamiento puede definirse con precisión en términos estrictamente matemáticos. Su idea era fundar las bases de una teoría del diseño y funcionamiento de máquinas aplicable a máquinas más complejas que las que tenemos o que no hemos concebido. Estaba convencido de que a través de esta teoría se entendería cómo construir las y comprenderíamos mejor el funcionamiento de los seres vivos.

Von Neumann no vivió lo suficiente como para ver una teoría de los autómatas completa, pero sí lo suficiente para que algunas de sus ideas de los mecanismos vivos fueran comprobadas por los biólogos. En 1948, Von Neumann explicaba las componentes que deberían caracterizar a un autómata capaz de reproducirse. Cinco años después, los biólogos Crick y Watson descubrieron y explicaron la estructura del ácido desoxirribonucleico (ADN) y con ello, la estructura del sistema reproductor de todo organismo vivo. La estructura tiene los elementos esperados por Von Neumann: un sistema que recoge materias primas del medio y las procesa para obtener un producto especificado



Figura VIII.2. John von Neumann. Uno de los precursores de la era de las computadoras.

por una instrucción precisa que llega de su exterior (los ribosomas); un mecanismo duplicador que recibe una instrucción escrita y la copia (las enzimas de polimerasa); un mecanismo controlador enganchado a los dos mecanismos anteriores que se encarga de pasar las instrucciones que recibe al duplicador y a los ribosomas (moléculas de control); finalmente, una instrucción escrita que contiene todas las especificaciones que hacen que el sistema duplicador fabrique un doble del organismo (éste es el material genético, el ARN y el ADN).

Los resultados obtenidos por el matemático inglés Alan Turing y Von Neumann demuestran la existencia (al menos teórica) de un *autómata universal*, esto es, una máquina que puede hacer todo lo que cualquier otra pueda. Esta máquina no es necesariamente más compleja que las concebibles, lo que necesita es que las instrucciones que se le den sean más elaboradas. Además, este autómata universal puede ser autorreproductor. ¿Podrán estos autómatas ser construidos alguna vez? ¿Cuál será la repercusión científica, económica y social de estos artefactos?

En 1920, el escritor checo Karel Capek escribió la obra de teatro *R.U.R.* que describe una sociedad que se ha hecho dependiente de trabajadores mecánicos a los que él llamó *robots* (del checo *robota*, "trabajo forzado"). Desde entonces, muchos avances tecnológicos han superado la fantasía: hay fábricas de automóviles completamente automatizadas, donde las computadoras controlan y coordinan el trabajo de los robots que construyen el automóvil. Esto parece ser trabajo rutinario y eso es lo que se espera de los robots. Otra cosa es el triunfo de *Deep Blue* sobre Kasparov. Entre otras declaraciones, al final de su primera derrota, Kasparov dijo que alcanzó a vislumbrar destellos de inteligencia en el juego de *Deep Blue*. Ése es el tema central, ¿pueden los robots tener comportamiento inteligente?

En la película *Blade Runner* (basada en el libro de Philip K. Dick cuyo título dimos a este capítulo) aparece una sociedad futura donde los androides (robots con apariencia humana) han alcanzado un alto grado de perfección y pueden confundirse con los seres humanos. Es más, algunos androides se creen seres humanos, pues se les dieron recuerdos y vivencias ajenos. La película refiere la lucha por identificar y destruir a unos androides en rebelión al descubrir que están fabricados para alcanzar sólo cuatro años de "vida".

Si queremos estudiar el problema de la inteligencia de las máquinas, debemos comenzar por tratar el problema de la comunicación entre una máquina y el hombre. En general, los únicos indicios que podemos tener del "pensamiento" de cualquier ser (máquina o persona) es lo que nos es comunicado por alguna forma de *lenguaje*. Comencemos por estudiar el lenguaje de las máquinas.

LOS LENGUAJES DE LAS MÁQUINAS

Trataremos, pues, de entender las máquinas más sencillas. No nos interesa (al menos para nuestros fines) su estructura interna, sino tener un modelo matemático general de su funcionamiento. Por ejemplo, un radio: sin importar la constitución de sus circuitos, éstos sólo tienen la posibilidad de encontrarse en un número finito de "estados" que dependen de si el radio está prendido o apagado, la estación que sintonizamos, el vo-

lumen que le damos y de acuerdo al estado en que el radio se encuentre obtendremos una "salida" (lo que oímos).

En la teoría de Kleene y Moore, una *máquina finita* M consta de un conjunto finito de estados S , un conjunto finito de entradas I y un conjunto finito de salidas F , además, dado un estado de la máquina y una entrada, obtenemos otro estado de la máquina y una salida. Esto se simboliza diciendo que tenemos *funciones* $t: I \times S \rightarrow S$ de *transición* y $f: S \rightarrow F$ de *salida*. Indicaremos además un estado inicial i_0 , el de la máquina al comenzar a funcionar.

Supongamos que nuestro radio tiene sólo dos estaciones, 1 y 2. Sus estados serán tres: 0, apagado, 1 y 2 encendido y sintonizado en la estación correspondiente. Hay tres botones que determinan las entradas posibles: a para encender/apagar, b para sintonizar en 1 y c para sintonizar en 2. Al encender el radio se sintoniza automáticamente en la estación 1. Digamos que tenemos dos posibles salidas: n si no se escucha nada y m si se escucha. El estado inicial es 0. Entonces las tablas de transición y de salida son:

t	a	b	c
0	1	0	0
1	0	1	2
2	0	1	2

f	
0	n
1	m
2	m

La tabla de transición puede ilustrarse por medio de la figura VIII.3.

Consideremos otro ejemplo de máquina finita: la máquina sum_3 que tiene tres estados: 0, 1 y 2. Las entradas posibles

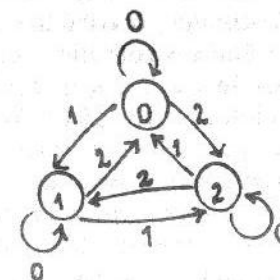


Figura VIII.3.

son también tres que se denotarán con los mismos números, al igual que las salidas. Las tablas de transición y de salida son las siguientes:

t	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

f	
0	1
1	2
2	0

Vemos que la tabla de transición de sum_3 es la tabla de sumar en la aritmética módulo 3. La salida es simplemente el resultado de sumar 1 módulo 3. La gráfica de transición de sum_3 se muestra en la figura VIII.4.

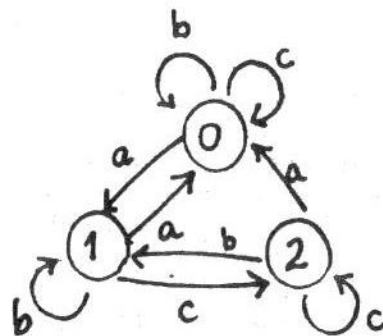


Figura VIII.4.

Por supuesto, a cualquier máquina finita M podemos asociarle una *gráfica de transición* G_M . Estas gráficas desempeñarán un papel importante en lo que sigue.

Dado un subconjunto F_0 del conjunto F salidas, definimos el *lenguaje* $\mathcal{L}(F_0)$ de la máquina M como el conjunto de sucesiones de entradas que podemos efectuar (a partir del estado i_0) para producir una salida en F_0 . Por ejemplo, si elegimos $F_0 = \{m\}$ en el caso de nuestro radio, podemos formar la sucesión de entradas $abaacb$ (es decir, encendemos, sintonizamos la estación 1, apagamos, encendemos, sintonizamos la estación 2 y luego la 1) que está en $\mathcal{L}(\{m\})$ (ya que el estado final es 1 y por tanto la salida es m). ¿Cuáles son todos los elementos de $\mathcal{L}(\{m\})$? Como el único estado que no tiene salida m es el



Figura VIII.5.

estado 0, entonces toda *palabra* en las letras a, b, c que contenga un número impar de estados a es una palabra en $\mathcal{L}(\{m\})$. Observe que estas palabras pueden leerse como *camino dirigidos* en la gráfica G_M .

En general, para cualquier máquina finita M con conjunto de entradas I y un subconjunto F_0 del conjunto de salidas F , diremos que $\mathcal{L}(F_0)$ es un *lenguaje regular* en el *alfabeto* I y sus elementos son *palabras* en $\mathcal{L}(F_0)$. En particular, siempre tenemos la *palabra trivial* que no tiene ninguna letra. Daremos algunos ejemplos.

1) El lenguaje $\mathcal{L}_1 = \{a, b\}^*$ formado por todas las palabras posibles usando las letras a y b , es regular.

En efecto, podemos considerar una máquina con un solo estado 0, dos entradas a y b y una salida 0. La gráfica de transición sería la de la figura VIII.5.

2) El lenguaje \mathcal{L}_2 formado por un número par de a y un número par de b es también regular.

En efecto, podemos definir una máquina de cuatro estados y gráfica de transición como en la figura VIII.6.

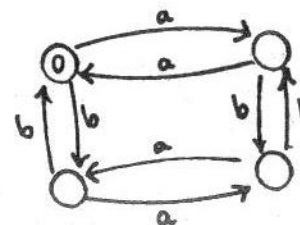


Figura VIII.6.

3) Sea k un número mayor o igual a 2 y consideremos letras a_1, \dots, a_k . Consideremos el *lenguaje palindrómico* \mathcal{L}_3 formado por palabras en las letras a_1, \dots, a_k que se pueden leer igual para atrás que para adelante (por ejemplo, si $k = 2$ entonces

$a_1 a_2 a_1 a_2 a_1$ está en \mathcal{L}_3 ; si $k = 3$, podemos formar palabras como radar, solos). En todos los lenguajes (humanos) se ha jugado a elaborar palíndromas (del griego *palin dromo*: "corriendo hacia atrás otra vez"). El complicado palíndroma (es palíndroma horizontal y verticalmente) en latín se encontró en un muro de Pompeya:

S	A	T	O	R
A	R	E	P	O
T	E	N	E	T
O	P	E	R	A
R	O	T	A	S

Se traduce "Arepo el campesino lleva las ruedas con cuidado". Durante siglos el palíndroma tuvo un sentido sagrado y aún en el siglo pasado las mujeres encintas usaban amuletos con esta inscripción. Por otra parte, los jeroglíficos egipcios eran siempre palindrómicos, por ejemplo, las dos frases de la figura VIII.7 tienen idéntico significado. La razón es simple: no existía la convención de escribir de izquierda a derecha o de derecha a izquierda, la única era que todos los jeroglíficos deberían "ver" hacia el principio del texto.

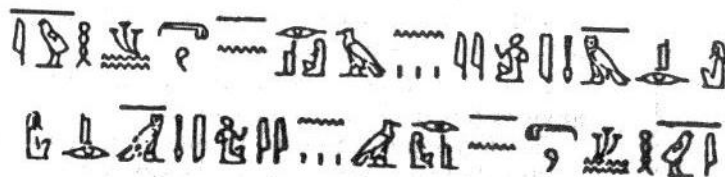


Figura VIII.7.

Nuestro interés en el lenguaje palindrómico \mathcal{L}_3 se debe a que no es un lenguaje regular. Esto es una sencilla consecuencia del siguiente resultado.

Proposición. Si \mathcal{L} es un lenguaje regular, existe un número positivo n tal que toda palabra $w \in \mathcal{L}$ de longitud $\geq n$ puede escribirse como $w = xyz$, donde x, y, z son palabras que cumplen las siguientes propiedades: i) y no es la palabra trivial; ii) xy es una palabra de longitud $\leq n$; iii) toda palabra de la forma $xy^m z$, con $m \geq 1$ está en \mathcal{L} .

Demostración. Supongamos que \mathcal{L} es un lenguaje regular correspondiente a una máquina finita M con m estados. Sea $w \in \mathcal{L}$ una palabra de longitud $\geq m + 1$, entonces w pasa dos veces por el mismo vértice de la gráfica de transición G_M . Escribamos $w = xyz$, de forma que xy es una palabra más corta con la que comienza w que pasa dos veces por el mismo estado y y comienza y termina en el mismo estado. Podemos representar esta división de la palabra w como se ve en la figura VIII.8.



Figura VIII.8.

Dejamos al lector verificar que se cumplen las condiciones i), ii) y iii) para la elección $n = m + 1$. \square

Para ilustrar el principio de esta proposición consideremos el ejemplo 2) anterior. La palabra $w = abaaba$ pertenece al lenguaje \mathcal{L}_2 y tiene longitud 6, mientras que la máquina asociada al lenguaje tiene sólo cuatro estados. La palabra $abaab$ con la que comienza w es de la forma ay con $y = baab$, de forma que toda palabra $ay^m a$ pertenece a \mathcal{L}_3 (por ejemplo, $ay^3 a = abaabbaabbaaba$).

¿Cómo se usa la proposición anterior para ver que el lenguaje palindrómico no es regular? Tomemos un número n cualquiera (tal vez elegido por nuestro peor enemigo), veamos que éste no satisface las condiciones i), ii) y iii) de la proposición, luego el lenguaje palindrómico no es regular. En efecto, como $k \geq 2$, podemos tomar las dos primeras letras a_1 y a_2 y formar la siguiente palabra: $w = a_1^{(n+1)} a_2 a_1^{(n+1)}$, donde claramente, $a_1^{(n+1)}$ significa que repetimos la letra a_1 consecutivamente $n + 1$ veces. Si la proposición fuera cierta para la palabra w , podríamos encontrar una palabra xy con longitud $\leq n$ de forma que $w = xyz$ y $xy^m z$ es palindrómica para toda $m \geq 1$. Pero en nuestro caso, $x = a_1^s$ y también $y = a_1^t$, para algunas $1 \leq s, t \leq n$, lo que implica que $xy^2 z = a_1^{(n+1+t)} a_2 a_1^{(n+1)}$, que obviamente no es una palabra palindrómica.

¿Es posible construir todos los lenguajes regulares? Para responder esta pregunta requerimos hacer un poco de álgebra con conjuntos de palabras.

Denotemos I^* al conjunto de todas las palabras en el alfabeto I . Definimos las *operaciones regulares* $+$, \cdot , $*$ en el conjunto de subconjuntos de I^* . Para E, F subconjuntos de I^* escribimos:

$$E + F = E \cup F;$$

$E \cdot F = \{ef : e \in E, f \in F\}$; en particular, $E^2 = E \cdot E$, $E^3 = E \cdot E \cdot E$, etcétera; obviamente, ef denota la palabra formada poniendo f a continuación de e .

$$E^* = E + E^2 + E^3 + E^4 + \dots$$

Decimos que un subconjunto E de I^* es *constructible* si existen elementos i_1, \dots, i_s de I de forma que E se obtiene a partir de los conjuntos $\{i_1\}, \dots, \{i_s\}$ por medio de las operaciones regulares.

Por ejemplo, si $I = \{a, b\}$, entonces $\{a\}^*$ son todas las palabras que sólo contienen a ; este conjunto puede construirse; el conjunto E de palabras de la forma $aa \dots abb \dots b$ que comienzan con a y siguen con b también, ya que $E = \{a\}^* \cdot \{b\}^*$.

El siguiente teorema de Kleene es uno de los resultados principales de la teoría de los autómatas. Fue demostrado en 1956 y aún es de gran importancia. Aunque la demostración no es difícil nos tomaría demasiado tiempo hacerla.

Teorema. *Un subconjunto de I^* es un lenguaje regular si y solamente si es constructible.*

Los resultados anteriores nos muestran que los lenguajes de las máquinas finitas son bastante limitados. Pero era de esperarse tal como las hemos definido. Las máquinas finitas son las que sólo reaccionan ante la entrada que se les da, no tienen memoria de lo que hicieron, sus respuestas no dependen de lo que ha pasado antes. Por supuesto, las máquinas modernas (cualquier tipo de computadora) son más hábiles. Una computadora tiene cierta capacidad de memoria y, en general, sus respuestas dependen de otras respuestas anteriores. Veremos ahora cómo entender estas máquinas más complicadas y cómo definir sus lenguajes.

Consideraremos primero una generalización muy simple de las máquinas finitas.

Una *máquina secuencial* M tiene un conjunto finito de estados S , un conjunto de entradas I que supondremos coincide con el conjunto de salidas. Tenemos, como antes, una función de transición $t: S \times I \rightarrow S$ y la función de salida es $f: S \times I \rightarrow I$ que depende tanto del estado como de la entrada. Podemos imaginar que esta máquina trabaja de la siguiente manera: le damos una lista de instrucciones, es decir, una palabra w en el "alfabeto" I (el conjunto de todas estas palabras las denotaremos I^*). Esta palabra la podemos dejar escrita en una cinta dividida en casillas de forma que los símbolos de w están escritos uno tras otro, uno en cada casilla; la cinta va siendo tomada casilla por casilla por la máquina M de izquierda a derecha. La máquina lee un símbolo, cambia de estado de acuerdo con la función de transición t , escribe en la casilla que acaba de leer el símbolo correspondiente a la función salida y corre la cinta a la siguiente casilla. Al final el resultado de nuestras instrucciones queda escrito en la cinta. Esto se representa en la figura VIII.9, que muestra dos pasos consecutivos del accionar de la máquina M .

Pensar en las máquinas de esta manera tiene la ventaja de que podemos entender más sencillamente lo que pueden hacer. Por ejemplo, construiremos una máquina *sum* que suma dos

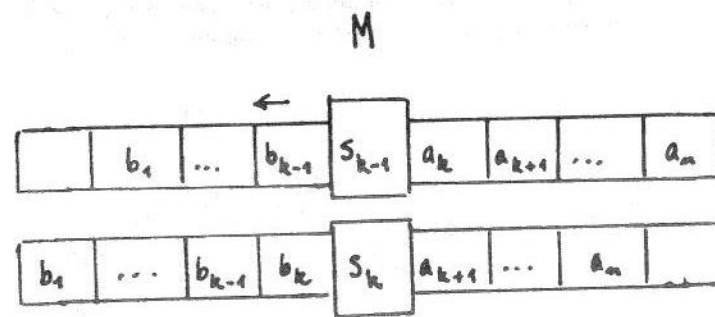


Figura VIII.9. Dos pasos consecutivos en el accionar de una máquina secuencial.

números cualesquiera. Para ello escribiremos los números por sumar como una sucesión de números 1 rodeados de 0 adecuadamente; por ejemplo, 0111101110 denota que queremos sumar el número 4 y el número 3. Estos números van escritos en una cinta que es tomada por nuestra máquina *sum*. La máquina lee el "0" de la izquierda cuando está en estado S_1 . Pasa a la siguiente casilla y lee "1", pasa entonces al estado S_2 , escribe "0" en la casilla y pasa a la siguiente. Mientras la máquina lee "1" escribirá "1" y se quedará en el estado S_2 . Al leer el siguiente "0" la máquina deberá escribir "1" y pasar al estado S_3 . En el estado S_3 la máquina se quedará siempre y respetará el símbolo que lea en la casilla. En nuestro ejemplo, la máquina actúa sobre la cinta como sigue:

	T
	0 1 1 1 1 0 1 1 1 0
0	1 1 1 1 0 1 1 1 0
0 0	1 1 1 0 1 1 1 0
0 0 1	1 1 0 1 1 1 0
0 0 1 1	1 0 1 1 1 0
0 0 1 1 1	0 1 1 1 0
0 0 1 1 1 1	0 1 1 0
0 0 1 1 1 1 1	1 1 0
0 0 1 1 1 1 1 1	1 0
0 0 1 1 1 1 1 1 1	0

de forma que el número que queda escrito en la cinta es el $7 = 4 + 3$. También podemos dibujar la gráfica de transición (con la información de la función de salida) de la manera como se ve en la figura VIII.10.

Las máquinas secuenciales archivan los resultados de sus cálculos, es decir, tienen memoria. Pero aún no interaccionan

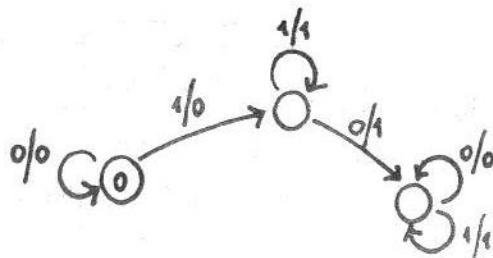


Figura VIII.10. Gráfica de transición de la máquina *sum*.

con el medio. ¿Qué haría falta para que esto sucediera? Muy simple: pensemos qué ocurriría si en la máquina secuencial que hemos considerado antes la cinta pudiera ir tanto a la derecha como a la izquierda. Sucedería que nuestra máquina podría leer casillas que ella misma ha modificado, produciéndose una situación de *retroalimentación*. Esto se escribe formalmente de la siguiente manera.

Una máquina de Turing T tiene un conjunto finito de estados S , un conjunto finito de entradas I que es el mismo conjunto de salidas. Una función de transición $t: S \times I \rightarrow S$ y una función de salida $f: S \times I \rightarrow I \cup \{\leftarrow, \rightarrow\}$.

A una máquina de Turing T podemos imaginarla exactamente como la máquina secuencial leyendo sobre la cinta, si la máquina está en el estado s y lee la casilla i , entonces pasa al estado $s' = t(s, i)$ y en caso de que $f(s, i) = (j, \leftarrow)$ escribe el símbolo j en la casilla donde acaba de leer y la cinta se mueve hacia la izquierda, mientras que si $f(s, i) = (j, \rightarrow)$, entonces la cinta se mueve a la derecha. Si el número n está escrito en la cinta como una sucesión de números 1, entonces el resultado final del proceso de lectura de la cinta (o sea, la cinta como queda escrita) se denota por $T(n)$. Las máquinas de Turing son el dispositivo automático más sencillo que interacciona con el medio.

Como ejemplo de una máquina de Turing que no es secuencial, definiremos la máquina *max* que calcula el máximo entre dos números. Probaremos primero que esta operación no puede efectuarse por medio de una máquina secuencial. En efecto, supongamos que M es una máquina secuencial con estados S y entradas I . Para simplificar nuestra exposición, supondremos que S tiene sólo dos elementos. Damos a nuestra máquina la cinta marcada 01101110, es decir, por los números 2 y 3. Como la máquina es secuencial cuando lee los primeros "1" a la izquierda, no sabe cuántos quedan a la derecha y no puede regresar a corregir, luego debe cambiar todos los "1" por "0" hasta llegar al final de la secuencia. En ese momento no le bastan dos estados para "recordar" que el número máximo era 111.

La máquina de Turing *max* funciona de la siguiente manera: en la cinta están escritos los números a y b como listas de números 1 separados por 0, la máquina *max* va de izquierda a derecha y borra un "1" de cada número, al terminar el segundo

Nos preguntamos ahora, ¿cuánto hemos ganado considerando las máquinas de Turing? Mucho, no es difícil demostrar que el trabajo que realiza cualquier máquina computadora (como las concebimos actualmente) puede ser efectuado por alguna máquina de Turing. De hecho, ¡hay una máquina de Turing que puede realizar lo que haga cualquier computadora!

designemos el símbolo \leftarrow por 01 y el símbolo \rightarrow por 011. Entonces podemos definir un número binario para cada máquina de Turing escribiendo en el orden creciente de los estados s y las entradas i los valores correspondientes a s , i , $t(s, i)$, $f(s, i)$. Así, por ejemplo la máquina *sum* tiene tres estados que corresponden a los números binarios 1, 10 y 11. Para el estado 1 y la entrada 0, se transita al estado 1 y la salida es 0011 (el primer "0" corresponde a la salida en I y el "011" corresponde al movimiento \leftarrow); antes de pasar al estado 1 y la entrada 1, escribimos 0 para distinguir claramente el cambio. Entonces seguimos con 11100011 ("1" por el estado, "1" por la entrada, "10" por $t(S_1, 1)$, "0" por la salida y "011" por el movimiento \leftarrow); antes de pasar al estado 10, escribimos 00. El número binario de la máquina de Turing *sum* es:

Un número horriblemente grande, pero lo que nos importa es el hecho que podamos encerrar el funcionamiento de una máquina de Turing en un solo número.

Teorema. Existe una máquina de Turing T_u de manera que para todos los números n y m se tiene que $T_u(g(n, m)) = T_n(m)$, donde $g(n, m) = \frac{1}{2}(n + m - 1)(n + m - 2) + m$.

Este resultado es importante no sólo en la teoría de autómatas, sino en lógica, inteligencia artificial y otras ramas de las matemáticas modernas. Una máquina T_u como en el teorema se llama *máquina universal de Turing*. Esta máquina puede reproducir el comportamiento de cualquier otra máquina de

Turing y por lo tanto, el de cualquier computadora que exista (o que pueda existir).

Roger Penrose ha calculado el número que le corresponde a una máquina universal en la enumeración de todas las máquinas de Turing: ¡es un número de más de 1400 cifras! Si bien la manera en que Penrose enumera sus máquinas es diferente de la nuestra, el número que obtiene no debería diferir mucho del que se tendría en nuestra enumeración.

Dada una máquina de Turing T con entradas $I = \{0,1\}$ podemos definir el lenguaje \mathcal{L}_T asociado a T como el conjunto de todas las palabras en la cinta que producen por respuesta un solo 1 y todas las demás casillas en 0. Un campo interesante de exploración actualmente es el estudio de los lenguajes \mathcal{L}_T y su comparación con los lenguajes regulares y los lenguajes humanos.

¿PUEDE PENSAR UNA MÁQUINA?

En 1950, Alan Turing escribió un artículo que llevaba por título "Computing machinery and intelligence". Su artículo comienza diciendo "Me propongo tomar en cuenta la pregunta: ¿Pueden las máquinas pensar?" La respuesta que da a esta pregunta no ha sido aún superada. En efecto, debido a las obvias complicaciones y discusiones en que nos meteríamos si queremos definir lo que es "pensar", Turing adopta un punto de vista operacional. Su método lo llama "juego de la imitación" y lo presenta de la siguiente manera.

El juego de la imitación lo juegan tres personas: un hombre A , una mujer B y un interrogador C , de cualquier sexo. El interrogador permanece en una habitación separado de las otras dos personas; su objetivo es determinar cuál de las dos personas en la otra habitación es la mujer. Las respuestas a las preguntas son transmitidas por una cuarta persona, de manera que C no puede ayudarse de la voz u otros medios. Supongamos ahora que el objetivo de A es confundir al interrogador. Por más que B quiera ayudar a C , el hombre A puede insistir en que es la mujer sin que haya forma de corroborarlo para C . Pregunta Turing:

¿Qué sucederá si A es sustituido por una máquina en este juego?
¿Las equivocaciones del interrogador tendrán la misma frecuencia,

jugando de este modo, que cuando los contendientes son un hombre y una mujer?

¿Qué sucede si el interrogador debe decidir si A es una máquina o un hombre? Sigue Turing:

Quizá el juego pueda ser criticado desde el punto de vista de que las disparidades gravitan en exceso en contra de la máquina. Si el hombre se pusiera a tratar de ser la máquina, daría sin duda una mala exhibición; sería descartado de inmediato por su lentitud e inexactitudes aritméticas. ¿Las máquinas no pueden ejecutar determinadas cosas que deben ser descritas como pensamiento, pero que son muy diferentes de lo que hace el hombre? Esta objeción es poderosa, pero no debemos preocuparnos por ella si, pese a todo, puede llegar a construirse una máquina que juegue satisfactoriamente al juego de la imitación.

Pienso que a finales de este siglo el uso de las palabras y la opinión general de la gente educada se habrán modificado tanto que será posible hablar de máquinas pensantes sin esperar que se susciten discusiones.

La computadora *Deep Blue* efectúa una cantidad impresionante de operaciones para evaluar la posición de la partida y decidir la siguiente jugada. ¿Es esto pensar? Tal vez no, pero es una buena imitación. Cómo nos indica Turing, esto es lo único que podemos juzgar. Podemos imaginar una máquina del futuro que cumpla muchas funciones humanas con la destreza que *Deep Blue* juega ajedrez. ¿Diremos que esa máquina piensa? El físico inglés Freeman Dyson piensa que ese día llegará en menos de 50 años y agrega:

No veo un peligro de que la inteligencia humana llegue a ser sustituida por la inteligencia artificial. La inteligencia artificial se mantendrá bajo control humano. El hombre no sólo vive para resolver problemas, la inteligencia artificial nos dará libertad y tiempo para ejercitar aquellas cualidades humanas que las máquinas no pueden tocar.

En la película *Blade Runner*, un policía trata de descubrir por medio de un interrogatorio si una mujer es humana o "replicante". Ella piensa ser humana, ya que tiene recuerdos de su infancia y otras vivencias, además tiene sentimientos. ¿Quién

puede decir si son sentimientos verdaderos o si sólo ejecuta satisfactoriamente el juego de la imitación? Un momento antes de morir, el replicante Roy salva la vida al policía Rick Deckard que lo ha seguido para matarlo. Deckard piensa: "No sé por qué me salvó la vida. Quizá en esos momentos amaba la vida más de lo que la había amado nunca. No solamente su vida, sino la de todos. Todo lo que él quería eran las mismas respuestas que todos buscamos: ¿de dónde vengo?, ¿a dónde voy?, ¿cuánto tiempo me queda?"

El siglo termina y las predicciones de Turing no se han confirmado, pero tal vez no falte mucho tiempo.

DEEP BLUE VS. KASPAROV. DEFENSA KARO-KANN

Ya que hablamos tanto del torneo Kasparov vs. Deep Blue, no resistimos la tentación de reproducir aquí la última y decisiva partida. Después de todo, el juego del ajedrez es un juego de lógica (aunque con ingredientes de psicología).

Blancas: Deep Blue.

Negras: Kasparov.

1. e4	c6	11. Af4	b5
2. d4	d5	12. a4	Ab7
3. Cc3	d x e4	13. Te1	Cd5
4. C x e4	Cd7	14. Ag3	Rc8
5. Cg5	Cgf6	15. a x b5	c x b5
6. Ad3	e6	16. Dd3!	Ac6
7. C1f3	h6	17. Af5	e x f5
8. C x e6!	De7	18. T x e7	A x e7
9. 0-0	f x e6	19. c4	Rinde
10. Ag6+	Rd8		

La posición final es la que se muestra en el tablero (figura VIII.11).

Según los comentarios de los expertos, durante las primeras partidas de la serie Kasparov se mostraba más bien conservador contra Deep Blue. Contrariamente a su estilo explosivo y arriesgado procuraba hacer movimientos seguros y cautelosos. La razón parece ser más bien psicológica: cuando Kasparov

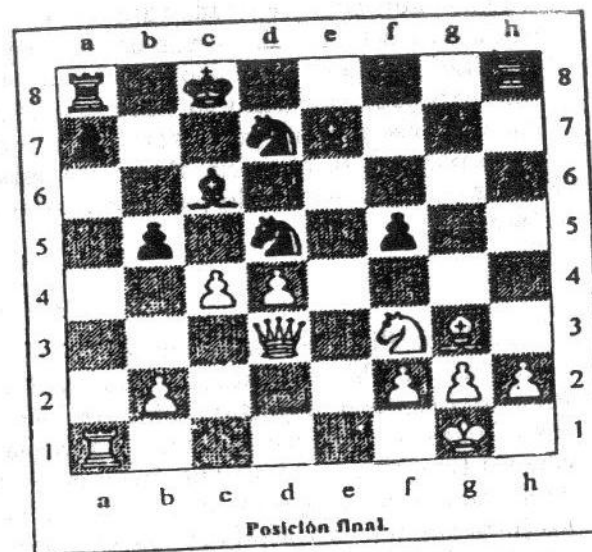


Figura VIII.11. Posición final de la partida.

hace una jugada agresiva y sorprendente contra un contrincante humano, éste se sorprende y no reacciona bien. En cambio, la máquina no se sorprende y sigue trabajando igual que siempre y puede explotar cualquier debilidad del oponente debida al juego arriesgado.

Según el G.M.I. mexicano Marcel Sisniega, la jugada h6 del séptimo movimiento de Kasparov es un error irreparable y todo practicante de la defensa Karo-Kann sabe que lo que debe jugarse es 7... Ad6, 8. De2, h6 (ahora sí, porque el rey negro está protegido) 9. Ce4, Ce4, 10. De4, Cf6... En lo que todos los comentaristas están de acuerdo es que Deep Blue no comete errores y hace algunos movimientos sorprendentemente eficientes como 8. Ce6. También la jugada 16. Dd3 es profunda. Comentaristas en la sala como el G.M. Yaser Seirawan opinaron que la máquina seguramente jugaría 16. De2 porque con esto ganaba un peón.

Después de entregar su dama en la jugada 17, Kasparov colocó el reloj en su muñeca e hizo gestos de que todo estaba perdido. Sin embargo, el G.M. J. Fedorowicz dijo que se sorprendió de que Kasparov se retirara, porque no parecía estar derrotado.

Respecto al comentario de Kasparov en el sentido de que la máquina había sido preparada para vencerlo (lo cual, se supone, resta mérito a la máquina), uno de los programadores de *Deep Blue* dijo que no sabía cómo hacerlo. *Deep Blue* está programada para jugar ajedrez y lo único específico con que contaba para su torneo era una gran cantidad de partidas de Kasparov almacenadas en su memoria. Por cierto, ésta fue otra queja de Kasparov: no haber tenido partidas previas de *Deep Blue* que estudiar.

UNA MÁQUINA PARA CONTAR CONEJOS

Leonardo de Pisa, también conocido como Fibonacci, fue uno de los más notables pensadores de la Edad Media. Su obra *Liber Abaci* contiene importantes innovaciones matemáticas. En sus páginas se encuentra el siguiente problema. Hay un par de conejos (macho y hembra) que nacen en el tiempo 0. Después de un mes estos conejos maduran y tienen una pareja de conejos (otra vez macho y hembra) y así el proceso de reproducción continúa indefinidamente mes con mes. Es decir, al pasar un mes la pareja recién nacida madura y tiene otra pareja de conejos, al igual que tienen otra pareja los conejos del principio. Suponemos que los conejos viven indefinidamente. ¿Cuál es el número de parejas de conejos después de n meses?

Diseñe una máquina de Turing para contar estos conejos y dibuje la gráfica. ¿Cuál es el número binario de esta máquina? Demuestre también que el trabajo de esta máquina no puede ser efectuado por una máquina de estados finitos.

Solución. En la población de conejos distingamos dos tipos: los adultos maduros (que tienen más de un mes de vida) y los jóvenes (que nacieron hace menos de un mes). Si en el mes n la población total de parejas de conejos es a_n y la de parejas de conejos jóvenes es de b_n , entonces tenemos que en el mes $n+1$ las poblaciones serán:

$$a_{n+1} = a_n + b_n,$$

$$b_{n+1} = a_n.$$

La primera fórmula se debe a que los conejos jóvenes son adultos después de un mes, además ningún conejo muere. La segunda fórmula proviene del hecho de que todas las parejas

adultas producen cada mes una pareja joven. Finalmente:

$$a_{n+1} = a_n + a_{n-1},$$

para todo tiempo $n \geq 1$. Con esta fórmula podemos calcular los primeros números de la sucesión de Fibonacci:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ...

donde cada término se obtiene como suma de los dos números previos.

Diseñaremos ahora la máquina de Turing *fib* que alimentada con la cinta 0101 produce toda la sucesión de Fibonacci. Lo haremos en dos partes, primeramente diseñamos la máquina *rep* que reproduce un número anotado sobre la cinta, es decir, queremos construir una máquina que ante la cinta 0110 produzca la cinta 0110110.

La gráfica de *rep* es la de la figura VIII.12, que actúa (paso a paso) sobre 0110 en la siguiente forma:

0 1 1 0	1 1 0 1 0
0 1 1 0 0	1 0 0 1 0 0
0 0 1 0 0	1 0 0 1 0 0 0
0 0 1 0 0 0	1 0 0 1 0 0 0
0 0 1 0 0 0 0	1 0 0 1 1 0
0 0 1 0 1 0	1 0 0 1 1 0
0 0 1 0 1	1 0 0 1 1 0
0 0 1 0 1	1 1 0 1 1 0

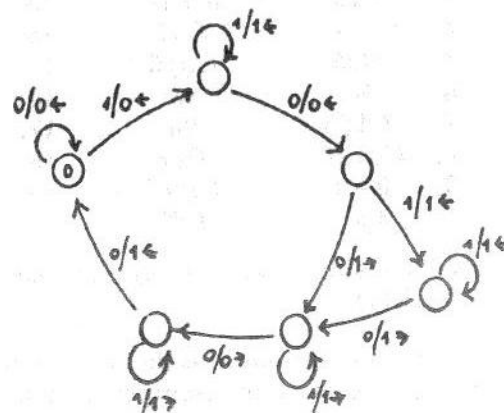


Figura VIII.12.

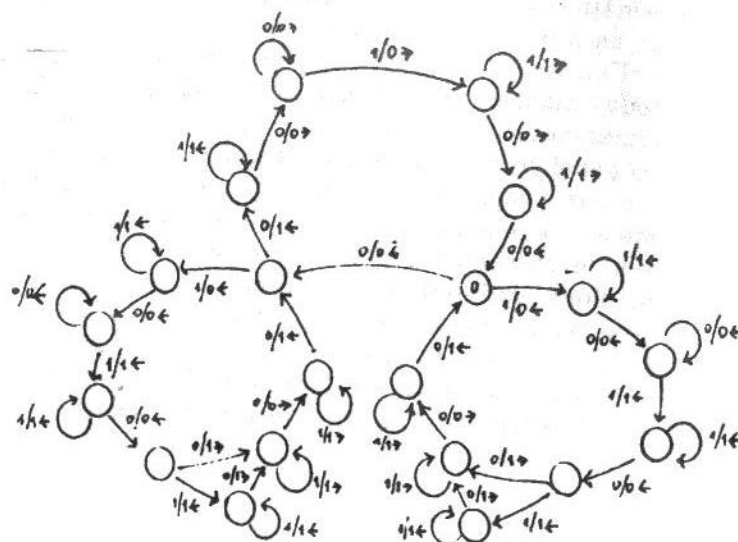


Figura VIII.13.

Ahora, teniendo las máquinas de Turing *sum* y *rep* es fácil construir la máquina *fib*. En efecto, la operación de *fib* sobre 0101 sería:

- *rep* opera sobre el primer número y produce 0101010;
- *rep* opera sobre el segundo número y produce 010101010;
- *sum* opera sobre los últimos dos números y produce 01010110;
- *rep* opera sobre el segundo número y produce 0101011010...

suponemos que es claro cómo continuar.

La gráfica de la máquina *fib* es la de la figura VIII.13.

IX. Algunos algebristas y sus teoremas

Penetran en un país de maravillas.
Soñando mientras pasan los días,
soñando mientras mueren los estíos.

Alicia a través del espejo, LEWIS CARROLL

A LO largo de este libro hemos encontrado los nombres de algunos personajes que han contribuido de manera importante al desarrollo de las ideas y técnicas matemáticas, en particular del álgebra. Pitágoras, Diofanto, Cardano, Descartes, Fermat, Euler, Gauss, Galois y algunos otros que nos son ya en mayor o menor medida familiares, son algunos de los nombres centrales en la historia de las matemáticas, pero no sólo ellos han hecho aportaciones notables al álgebra. Algunos no han sido mencionados por la particular selección de temas que hemos hecho, o la forma como presentamos algunos capítulos. Viète, Jacobi, Bézout, Abel, Lagrange, Sylvester, Hamilton, Cayley, Jordan, Kummer, Frobenius, Dedekind, Kronecker, Hilbert, Lie, Noether, Steinitz, Weyl, Cartan, Pierce, Artin, Weil, son de primera importancia en cualquier historia del álgebra. También las contribuciones de matemáticos como Newton, Leibniz, Lagrange, Weierstrass, Klein, Poincaré son importantes en álgebra (aunque sus contribuciones principales se hayan dado en otros campos de las matemáticas).

Por su naturaleza, los resultados matemáticos no dependen de los contextos históricos o del carácter o ideología del matemático que los demuestra. Sin embargo, en este capítulo queremos presentar algunos datos de la vida de algunos de los algebristas importantes y algunos de los teoremas que los hicieron famosos. Antes queremos enfrentar las siguientes interrogantes: ¿qué hubiera pasado de no haber nacido Fermat?, ¿qué sería de las matemáticas si Gauss hubiera sido médico o carpintero? En una sola pregunta: ¿hasta qué grado depende el desarrollo de las matemáticas de los matemáticos que las hacen?

Por supuesto, no hay respuesta definitiva a estas preguntas

y cualquier cosa que se diga es una simple especulación. Pero seré breve en mi especulación. Creo que si Fermat o Gauss (u otros matemáticos) no hubiesen hecho sus importantes contribuciones a la ciencia, las matemáticas serían esencialmente las mismas que conocemos hoy en día. Esto no quiere decir que su trabajo no fue importante, quiere decir que los descubrimientos que ellos hicieron hubieran sido hechos por otros matemáticos algunos años después. Hay dos razones que considero poderosas para creer esto:

a) *Los resultados en las matemáticas se descubren.* Si ahondamos históricamente en las raíces de los principales resultados matemáticos llegaremos siempre a preguntas relacionadas con el mundo real, con el comportamiento de la naturaleza. Para explicar o generalizar algunas de las propiedades observadas en el mundo real, los matemáticos desarrollan conceptos y definen estructuras matemáticas abstractas (por ejemplo, la noción de grupo, la noción de matriz). Para un matemático las estructuras abstractas tienen una realidad propia, y el matemático las investiga sistemáticamente, casi como un biólogo investigaría un animal. Por ello, las preguntas que el matemático se hace no son arbitrarias, son las mismas (más o menos) que se hacen otros matemáticos.

En pocas palabras, *las matemáticas son una ciencia* como cualquier otra (la biología o la física), con la ventaja de que los resultados que se demuestran en esta ciencia son verdades universales que no dependen del tiempo o de los gustos de los hombres.

b) *Los resultados importantes en las matemáticas se dan como respuesta a problemas que son considerados importantes en un momento histórico y cultural determinado.* Los matemáticos pertenecen a una comunidad científica mundial que indica, más o menos claramente, cuáles son los problemas que se deben tomar en cuenta. Por ello, en general, hay varias personas pensando simultáneamente en los mismos problemas y tantos ejemplos de descubrimientos hechos al mismo tiempo por dos investigadores que trabajan independientemente.

Por supuesto, hay muchos ejemplos que no siguen exactamente el esquema planteado antes; es decir, casos de matemáticos que desarrollan ideas en las que nadie más está pensando y aún más, ideas que el resto de la comunidad no está preparada para recibir. Pero estas explosiones de genialidad son más

bien aisladas y creo que el desarrollo histórico global de las matemáticas cubriría perfectamente estas genialidades por medio de un proceso más pausado.

Una vez contestadas nuestras interrogantes digamos que, de todas maneras, es importante considerar la vida de los matemáticos. ¿Por qué? Porque los grandes individuos desempeñan un papel de guía para otros y los elementos de sus vidas pueden ilustrar y explicar periodos completos del desarrollo del pensamiento. Las matemáticas son hechas por matemáticos que viven en un contexto histórico, social y cultural que en muchas ocasiones es interesante conocer para entender mejor las motivaciones y origen de los problemas considerados. En última instancia, un poco de calor humano nunca sobra.

Los matemáticos que hemos elegido para presentar, así como sus teoremas, están estrechamente relacionados con el material que hemos tratado en los otros capítulos.

EL ABOGADO
Y LOS NÚMEROS

Pierre de Fermat nació en Beaumont de Lomagne en el sur de Francia en 1601 y nunca viajó a más de 200 kilómetros de distancia de su lugar natal. Murió en 1665 en Toulouse sin haber conocido París.

Fermat pertenecía a una familia adinerada. Estudió leyes y obtuvo rápidamente el puesto de magistrado de Toulouse que le dio seguridad económica, cierta nobleza (que le permitía usar el "de" antes de su apellido) y tiempo para otras ocupaciones. En efecto, en el tiempo que su labor de magistrado le dejaba (aparentemente mucho), Fermat escribía poesía en latín y críticas a textos griegos. Pero sobre todo hacía matemáticas.

El interés de Fermat por las matemáticas era muy amplio. Consideraba problemas en teoría de números, probabilidad y cálculo de valores máximos y mínimos de funciones. Desgraciadamente, Fermat no publicaba la mayor parte de sus resultados, se contentaba con escribir cartas a matemáticos refiriendo algunos de sus descubrimientos, en muchas ocasiones sin dar detalles de las demostraciones. Entre sus correspondientes se contaban algunos de los matemáticos más importantes de su



Figura IX.1. Pierre de Fermat.

tiempo: Descartes, Pascal, Huygens, Wallis y Mersenne. En ese sentido, Fermat nunca fue un profesional de las matemáticas, pero probablemente sí el matemático más grande de su tiempo.

Uno de sus logros más importantes fue el desarrollo del concepto de plano coordenado independientemente de Descartes. Esta noción fundamental que fusiona el álgebra y la geometría fue establecida por Fermat en 1636, pero nunca publicada. Luego, Descartes la desarrollaría en su *Géométrie* (1637) y se le acreditaría la gloria de la idea. Por ello al plano coordenado X - Y se le llama *plano cartesiano* (que por otra parte se oye mejor que *plano fermatiano*).

El campo de las matemáticas donde Fermat concentró la mayor parte de las energías de sus tiempos de ocio fue la *teoría de los números*. Para poner en contexto la importancia de sus contribuciones en esta área, baste decir que se le considera el padre de la moderna teoría de los números y a sus contribuciones las más importantes desde la Grecia clásica. Parece que el interés de Fermat por la teoría de números nació de su interés por la literatura de los clásicos griegos y latinos que leía en su

idioma original. Este interés le llevó a los libros de matemáticas griegos, en particular a la *Aritmética* de Diofanto, que le sirvió como fuente de inspiración.

Aunque de muchas de sus afirmaciones no quedan demostraciones completas, las pruebas que hizo y que han llegado hasta nosotros son precisas y correctas. Por ello el matemático contemporáneo André Weil afirma que "cuando Fermat dice que tiene una demostración de una afirmación, hay que tomarlo en serio". Y fue tomado en serio cuando dijo en el margen de un libro que tenía una prueba maravillosa de que la ecuación $x^n + y^n = z^n$ no tiene soluciones en enteros positivos si $n \geq 3$. La historia de este problema (sin duda, uno de los más influyentes en la historia de las matemáticas) puede verse en el capítulo V.

Ya en sus últimos años, en carta a un amigo, Fermat externa su esperanza de que "tal vez la posteridad le estará agradecida por mostrar que los antiguos no sabían todo". La posteridad le está agradecida a Fermat por esto y muchas otras cosas.

Como ejemplo de las contribuciones de Fermat a la teoría de los números hemos elegido el siguiente teorema que empleamos al tratar acerca de los "volados telefónicos".

El pequeño teorema de Fermat. *Si p es un número primo y n es un entero positivo cualquiera, entonces p divide a $n^p - n$.*

Por ejemplo, si $p = 3$ y $n = 3$, entonces $3^3 - 3 = 24 = 8 \times 3$; si $p = 7$ y $n = 8$, entonces $8^7 - 8 = 2097144 = 299592 \times 7$. En términos de congruencias, el enunciado se puede escribir así:

$$\text{si } p \text{ es primo, entonces } n^{p-1} \equiv 1 \pmod{p}.$$

Demostración. Una demostración combinatoria muy sencilla se puede hacer como sigue. Supongamos que tenemos bolitas de colores que se pueden insertar en un hilo. Tenemos n colores diferentes de bolitas. Formamos hilos con p bolitas de forma que no todas las bolitas son del mismo color. ¿De cuántas maneras es esto posible? Como primera bolita podemos elegir una de n colores diferentes, lo mismo para la segunda y así hasta la bolita p , esto es, podemos formar n^p hilos con p bolitas. Pero hay que excluir las que son de un solo color, que son n hilos. Luego la respuesta es $n^p - n$ maneras. Un hilo así lo denotamos

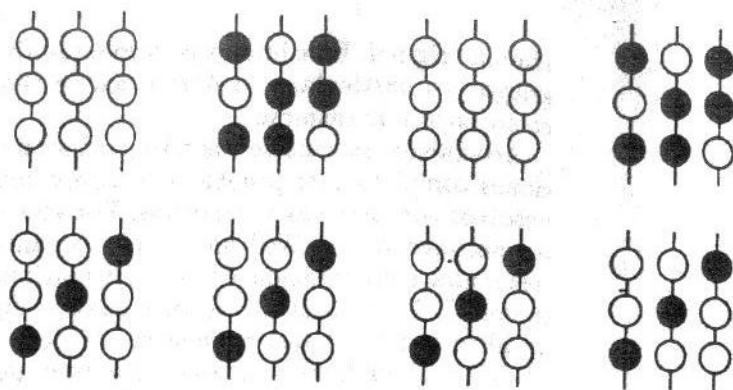


Figura IX.2.

(b_1, b_2, \dots, b_p) , donde b_i denota la bolita i -ésima del hilo. En la figura IX.2 vemos el caso $p = 3 = n$, que da 24 hilos.

Ahora, formamos brazaletes con estos hilos, es decir, unimos los extremos de los hilos. ¿Cuántos brazaletes (no monocromáticos) se pueden formar? Llamemos $B(p, n)$ al número de estos brazaletes. Observemos que en el caso $p = 3 = n$, cada grupo de 3 hilos determina el mismo brazalete, de forma que los brazaletes en este caso son los siguientes (tenemos $B(3, 3) = 8$) (figura IX.3).

Mostraremos que éste es siempre el caso, es decir, para cualquier p y n , cada brazalete proviene de p hilos. En efecto, tome-

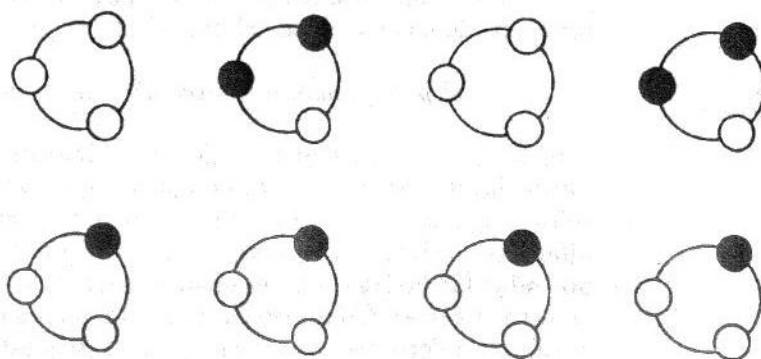


Figura IX.3.

mos un hilo cualquiera $h = (b_1, b_2, \dots, b_p)$. Si formamos el hilo $h^{(2)} = (b_2, b_3, \dots, b_p, b_1)$, pasando la bolita del final al principio del hilo, obtendremos el mismo brazalete. Podemos hacer esto sucesivamente y obtener p hilos $h = h^{(1)}, h^{(2)}, \dots, h^{(p)}$ que determinan el mismo brazalete. ¿Son estos hilos diferentes entre sí? Sí, pues si dos de los p hilos formados fueran iguales, podríamos $h^{(k)} = h^{(m)}$ con $0 < m - k < p$ con $m - k$ de valor mínimo posible y entonces $m - k$ divide a p (pues $h^{(k+1)} = h^{(m+1)}, \dots, h^{(m)} = h^{(2m-k)}$, etc.) Pero es imposible que $m - k$ divida a p que es primo. Luego, cada brazalete se obtiene de p hilos posibles. Esto es:

$$pB(p, n) = n^p - n.$$

Esto es lo que se deseaba demostrar. □

EL CIEGO QUE
VIO MÁS LEJOS

Leonhard Euler (pronúnciese Oiler) nació en Basilea, Suiza, en 1707. Se graduó como matemático y físico. A la edad de 20 años viajó a San Petersburgo, Rusia, a ejercer una cátedra en la recién creada Academia y permaneció allí hasta 1741, cuando pasó a la Academia de Berlín, que en ese momento reunía a algunos de los científicos y humanistas más notables de Europa: D'Alembert, Maupertuis y Voltaire. En 1766 regresó a San Petersburgo, donde permaneció hasta su muerte en 1783. A los 76 años todavía se encontraba activo, a pesar de haber sufrido graves problemas de la vista desde 1735, que lo llevaron a la ceguera total en 1771.

Ya ciego, Euler dictaba a un amanuense sus artículos que no disminuyeron en profundidad o complicación de los cálculos. Se dice que podía efectuar mentalmente operaciones aritméticas complejas que implicaban una precisión de hasta 50 cifras decimales. El día de su muerte calculó la órbita del planeta Urano que había sido descubierto unos días antes.

Euler es uno de los matemáticos más influyentes de la historia y, ciertamente, el más prolífico. Publicó más de 900 artículos, tratados o libros de matemáticas. Sus obras completas (*Opera*



Figura IX.4. Leonhard Euler.

Omnia) comenzaron a publicarse en 1911 a razón de un volumen al año, y la publicación continuará hasta bien entrado el próximo siglo. Esas obras incluyen ¡más de 40 000 páginas! Su contenido es amplísimo: teoría de números, teoría de funciones, análisis, inicio de ramas completas de las matemáticas como el cálculo de variaciones, la teoría de gráficas y la topología. En la aplicación de las matemáticas consideró todo tipo de problemas, desde el movimiento de la Luna hasta la estructura de la música. Sobre su libro *Introductio in analysis infinitorum*, el historiador de las matemáticas, Carl Boyer, dice que “hizo por el análisis lo que los *Elementos* de Euclides hicieron por la geometría y la tendencia resultante de expresar las matemáticas y la física en términos aritméticos ha continuado desde entonces”. Baste mencionar que gran parte de la notación matemática que se usa hoy día en análisis proviene de los trabajos de Euler.

Una de las áreas favoritas de trabajo de Euler fue la teoría de números (1 700 páginas de la *Opera Omnia* se dedican a

este tema). Alcanzó resultados primordiales en la teoría de los números primos, descubrió la importante *ley de la reciprocidad cuadrática*. Como ejemplo de un resultado de teoría de números demostrado por Euler, presentamos la siguiente generalización del pequeño teorema de Fermat. Para ello requerimos la siguiente definición: dado un número entero positivo n , denotamos por $\phi(n)$ la cantidad de números m entre 1 y n que son *primos relativos* con n (esto es, el único divisor común de n y m es 1). Los valores que toma la función ϕ son interesantes, como muestra la siguiente tabla:

n	$\phi(n)$	n	$\phi(n)$
1	1	11	10
2	1	12	6
3	2	13	12
4	2	14	6
5	4	15	8
6	2	16	8
7	6	17	16
8	4	18	6
9	6	19	18
10	4	20	8

Teorema. a) Si a y n no tienen divisores comunes aparte del 1, entonces $a^{\phi(n)} \equiv 1 \pmod{n}$. b) $\sum_{d|n} \phi(d) = n$.

Por ejemplo, si $a = 21$ y $n = 10$, entonces $\phi(10) = 4$ y $21^4 = 194\,481 \equiv 1 \pmod{10}$. Para la segunda afirmación, podemos observar por ejemplo que los divisores de 20 son 1, 2, 4, 5, 10 y 20, de forma que $\sum_{d|20} \phi(d) = 1 + 1 + 2 + 4 + 4 + 8 = 20$. También observemos que el pequeño teorema de Fermat es un caso particular del inciso a), ya que si p es primo, $\phi(p) = p - 1$ y cualquier número entero a es primo relativo con p , luego por a) se tiene $a^{p-1} \equiv 1 \pmod{p}$, o lo que es lo mismo $a^p \equiv a \pmod{p}$.

Demostración. a) Tomemos los números $r_1, r_2, \dots, r_{\phi(n)}$ entre 1 y n que son primos relativos con n . Como a es primo relativo con n , entonces $ar_1, ar_2, \dots, ar_{\phi(n)}$ son también primos relativos con n . Esto es, cada $r_i \equiv 1 \pmod{n}$ y $ar_i \equiv 1 \pmod{n}$.

Entonces $r_1 \dots r_{\phi(n)} \equiv 1 \pmod{n}$ y también $a^{\phi(n)} r_1 \dots r_{\phi(n)} = ar_1 ar_2 \dots ar_{\phi(n)} \equiv 1 \pmod{n}$. Cancelando se obtiene $a^{\phi(n)} \equiv 1 \pmod{n}$, como deseábamos.

b) Para un divisor d de n , definimos el conjunto $T(d, n)$ de los números m entre 1 y n tales que el máximo común divisor de m y n es d . Observemos que todo número entre 1 y n está en uno y sólo uno de los conjuntos $T(d, n)$. Si $t(d, n)$ es la cantidad de elementos de $T(d, n)$, entonces:

$$\sum_{d|n} t(d, n) = n.$$

Calculemos ahora otra expresión para $t(d, n)$. Observemos que si el máximo común divisor de m y n es d , entonces el máximo común divisor de m/d y n/d es 1. O sea, $t(d, n) = \phi(n/d)$. Finalmente, se puede probar que:

$$\sum_{d|n} \phi(n/d) = \sum_{d|n} \phi(d),$$

esta última aserción la dejamos como ejercicio para el lector interesado. La demostración está completa. \square

La siguiente fórmula es también muy interesante:

$$\phi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

donde p corre sobre todos los números primos que dividen a n . Por ejemplo, en el caso $n = 20$, se tiene:

$$\phi(20) = 20 \times \left(1 - \frac{1}{2}\right) \times \left(1 - \frac{1}{5}\right) = 20 \times \frac{1}{2} \times \frac{4}{5} = 8.$$

De esta fórmula se obtiene una forma general de calcular el valor $\phi(n)$ usando la factorización prima de n . En efecto, por el teorema fundamental de la aritmética podemos escribir $n = p_1^{e_1} p_2^{e_2} \dots p_s^{e_s}$, donde p_1, p_2, \dots, p_s son primos diferentes dos a dos. Entonces usando la fórmula anterior tenemos:

$$\begin{aligned} \phi(n) &= p_1^{e_1} p_2^{e_2} \dots p_s^{e_s} \prod_{i=1}^s \left(1 - \frac{1}{p_i}\right) \\ &= \prod_{i=1}^s (p_i^{e_i} - p_i^{e_i-1}) = \prod_{i=1}^s \phi(p_i^{e_i}) \\ &= \phi(p_1^{e_1}) \phi(p_2^{e_2}) \dots \phi(p_s^{e_s}). \end{aligned}$$

Por ejemplo, para $n = 20 = 2^2 \times 5$, se tiene:

$$\phi(20) = \phi(2^2) \times \phi(5) = 2 \times 4 = 8.$$

EL PRÍNCIPE DE LOS MATEMÁTICOS

Carl Friedrich Gauss es considerado, junto con Arquímedes y Newton, uno de los tres matemáticos más importantes de la historia. Aunque contribuyó de manera decisiva en varias ramas de las matemáticas puras, abriendo paso a la modernización iniciada en el siglo XIX, también hizo aportaciones importantes, en su ramo, a la astronomía y la física. El espíritu de su trabajo se puede resumir usando sus palabras: "Las matemáticas son la reina de las ciencias y la aritmética la reina de las matemáticas."



Figura IX.5. Carl Friedrich Gauss.

Gauss nació en Brunswick, Alemania, en 1777, de familia campesina. Impresionados por sus dotes en matemáticas y lenguas, sus profesores lo recomendaron con el duque de Brunswick, que le concedió asistencia económica para que pudiera continuar sus estudios a nivel secundaria. Se cuentan varias anécdotas de su infancia y primera juventud. Se afirma que, teniendo 10 años, su profesor dejó a la clase el siguiente problema: sumar todos los números del 1 al 100. El profesor se disponía a salir por algún tiempo de la clase mientras los estudiantes hacían las sumas; antes de que pudiese salir, Gauss se levantó y le dijo:

—La respuesta es 5 050.

—¿Cómo pudiste hacer 100 sumas tan rápidamente?—, preguntó el maestro.

—No es necesario sumar, sólo se requiere multiplicar 50 por 101—, contestó Gauss.

En efecto, podemos sumar dos veces del 1 al 100 de la siguiente manera:

$$1 + 2 + 3 + \dots + 98 + 99 + 100 = x$$

$$100 + 99 + 98 + \dots + 3 + 2 + 1 = x$$

que es lo mismo que sumar 100 veces el número 101. (Esto es $2x = 100 \times 101$).

Otra anécdota cuenta que a los 17 años se encontraba indeciso entre estudiar matemáticas o filosofía. Pero entonces descubrió cómo construir usando sólo regla y compás el polígono regular de 17 lados. Este problema había sido planteado por los griegos 2 000 años antes y se conocían construcciones para los polígonos de 3, 4, 5 y 15 lados, pero desde tiempos de Euclides no se había hecho ningún progreso. Esto lo decidió por las matemáticas. Como recuerdo de este descubrimiento, la lápida de la tumba de Gauss tiene forma de un polígono regular de 17 lados.

Entre 1795 y 1798 estudió matemáticas en la Universidad de Gotinga y a los 22 años obtuvo su doctorado con un trabajo donde demostraba el llamado *teorema fundamental del álgebra*, esto es, que para todo polinomio $p(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$ existe un número complejo z tal que $p(z) = 0$. De hecho dio tres pruebas diferentes de este resultado. Con ello dejó establecidos a los números complejos como el lugar natural

para efectuar álgebra. A los 24 años, publicó el libro *Disquisitiones Arithmeticae* donde estableció la fundamentación de la moderna teoría de números y que es considerado uno de los mayores logros de la historia de las matemáticas. Sobre este libro el matemático alemán Kronecker dice:

Es absolutamente sorprendente pensar que un solo hombre a tan temprana edad fue capaz de dar a luz esa cantidad de resultados, pero sobre todo de presentar un tratamiento tan profundo y organizado de una disciplina enteramente nueva.

En 1801 ganó reconocimiento público al poder calcular la órbita del asteroide Ceres recién descubierto. Para hacerlo utilizó el *método de los mínimos cuadrados* que había desarrollado en 1794 (a los 17 años). Este método todavía se usa hoy en la misma forma. En 1807 se le nombró director del nuevo Observatorio de Gotinga, puesto en el que permaneció hasta su muerte.

Sería muy largo enumerar los múltiples intereses académicos que Gauss tuvo a lo largo de su vida. Baste mencionar que desarrolló la *geometría no euclidiana* 30 años antes que lo hicieran los matemáticos Lobachevski y Bolyai, pero nunca publicó sus resultados (por eso generalmente no se le acredita con este descubrimiento); desarrolló la teoría de curvas y superficies que después llevarían a su alumno Riemann a sus importantes trabajos; hizo estudios importantes en probabilidad (la curva de Gauss), en astronomía y en electromagnetismo.

Gauss murió en 1855. Poco tiempo después se comenzaron a acuñar en Alemania monedas con su efigie; los billetes de 10 marcos llevan la efigie de Gauss.

Presentar ejemplos de teoremas de Gauss resulta difícil, pues la mayoría son demasiado complicados para el nivel de este libro. Sin embargo, demostraremos parcialmente un resultado relacionado cercanamente con el último teorema de Fermat que hemos tratado en el capítulo V.

Observemos que las raíces cúbicas de -1 son -1 , $\mu = (1 + i\sqrt{3})/2$ y $\nu = (1 - i\sqrt{3})/2$. Consideremos los *números complejos* de la forma $n\mu + m\nu$, donde n y m son enteros. Números de esta forma serán llamados *G-números*.

* **Teorema.** No existen G -números u, v, w satisfaciendo que $u^3 + v^3 = w^3$.

Observemos que este resultado implica el caso del exponente $n = 3$ del último teorema de Fermat. En efecto, si existieran números enteros a, b, c satisfaciendo $a^3 + b^3 = c^3$, entonces $a = a\mu + a\nu$, b y c serían G -números.

Antes de indicar los pasos de la demostración, haremos algunas consideraciones generales acerca de los G -números. La mayoría de las siguientes afirmaciones son sencillas de probar y las dejamos como ejercicios a los lectores interesados.

1) El conjunto de los G -números es cerrado bajo la suma, resta y multiplicación de números complejos.

2) La *norma* de un número complejo $z = x + yi$ se define como $N(z) = x^2 + y^2$. Dados dos números complejos z, z' , se tiene que $N(zz') = N(z)N(z')$, es decir, la norma es multiplicativa.

3) Si $z = n\mu + m\nu$ es un G -número, entonces $N(z) = n^2 + m^2 - nm$, que es un número entero positivo. Los únicos G -números con norma 1 son:

$$\mu, -\mu, \nu, -\nu, 1, -1.$$

No hay G -números con norma 2.

4) Llamamos r al G -número $r = \mu - \nu = i\sqrt{3}$. Entonces el G -número $z = n\mu + m\nu$ es divisible entre r (esto es, existe otro G -número z' tal que $z = rz'$) si y solamente si $n + m$ es divisible entre 3.

5) Si un G -número z no es divisible entre r , entonces $z^3 = 9z' + e$, con z' otro G -número y $e = 1$ o $e = -1$.

Demostración del teorema de Gauss. Supongamos que u, v, w son tres G -números tales que $u^3 + v^3 = w^3$. Entonces podemos considerar el G -número $z = -w$, de forma que $u^3 + v^3 + z^3 = 0$. Además, si u, v, z tienen un divisor común d que es un G -número con $N(d) > 1$, entonces $u = du_1, v = dv_1, z = dz_1$, de forma que $u_1^3 + v_1^3 + z_1^3 = 0$ y por 2), $N(u_1) < N(u)$. Como este proceso no puede continuarse indefinidamente, podemos suponer que en la ecuación $u^3 + v^3 + z^3 = 0$, los números u, v, z no tienen divisores comunes d que son G -números con $N(d) > 1$. A esta ecuación $u^3 + v^3 + z^3 = 0$ le llamamos *ecuación*

de Gauss. La imposibilidad de la existencia de las ecuaciones de Gauss se obtiene de las siguientes dos afirmaciones:

a) En una ecuación de Gauss $u^3 + v^3 + z^3 = 0$, uno y sólo uno de los números u, v o z es divisible entre r (ese número se llamará una *base especial*).

Si y es la base especial de $u^3 + v^3 + z^3 = 0$ y $y = ry'$, a y' le llamamos *divisor especial*.

b) Para cada ecuación de Gauss $u^3 + v^3 + z^3 = 0$ con divisor especial y' , hay una segunda ecuación de Gauss $u_1^3 + v_1^3 + z_1^3 = 0$ con divisor especial y'_1 de forma que $N(y'_1) < N(y')$.

Supongamos que ya hemos demostrado a) y b), entonces por medio de repetidas aplicaciones de b), podemos encontrar ecuaciones de Gauss con divisores especiales y', y'_1, y'_2, \dots con normas decrecientes. Esto es imposible puesto que la norma siempre es un entero positivo. Esto probaría el teorema.

No daremos una prueba de b). Pero sí el argumento de a).

a) Si ninguno de los tres números u, v, z fuera divisible entre r , por 5) tendríamos que $u^3 = 9u' + e_1, v^3 = 9v' + e_2$ y $z^3 = 9z' + e_3$, donde cada uno de los números e_1, e_2, e_3 es 1 o -1 . Pero entonces, $e_1 + e_2 + e_3 \equiv 0 \pmod{9}$, lo cual es imposible dado que la suma $e_1 + e_2 + e_3$ sólo puede tomar los valores $-3, -1, 1$ o 3 . Esto muestra que al menos uno de los tres números u, v, z es divisible entre r .

Supongamos que dos de los tres u, v, z son divisibles entre r , digamos u y v . Entonces $z^3 = -(u^3 + v^3)$ es divisible entre r y es fácil concluir que también z es divisible entre r . Luego, los tres u, v, z son divisibles entre r . Pero $N(r) = 3$, lo que contradice que $u^3 + v^3 + z^3$ es una ecuación de Gauss. \square

EL ELEGIDO
DE LOS DIOS

Evariste Galois nació en Bourg-la-Reine, cerca de París en 1811. Murió a los 20 años de edad en 1832 ("los elegidos de los dioses mueren jóvenes").

Galois nació en el seno de una familia acomodada y políticamente activa durante los agitados tiempos de los gobiernos



Figura IX.6. Evariste Galois.

de Napoleón, Carlos X y Luis Felipe de Orleans. Su madre lo instruyó hasta que ingresó al colegio Louis-le-Grand en 1823. A los 15 años y bajo la guía de uno de sus maestros había leído los trabajos de Euclides y Legendre en geometría y Lagrange en álgebra.

A los 16 años comenzó a considerar un problema fundamental de álgebra: la resolución de ecuaciones por radicales. En el capítulo III hemos visto que las ecuaciones polinomiales de grado 2 y 3 pueden resolverse por *medio de radicales*, esto es, expresiones en los coeficientes de la ecuación que incluyen sumas, multiplicaciones, divisiones y extracción de raíces (cuadradas y cúbicas). Esto también es posible para las ecuaciones de grado 4. Pero no para todas las ecuaciones de grado 5. Esto había sido probado recientemente por otro joven matemático (noruego) Niels Abel, pero Galois no lo sabía. En realidad lo que intentaba encontrar era un resultado más general. Asociando cada polinomio al grupo de todas las posibles permutaciones de las raíces que determinan automorfismos en los números

complejos (el llamado *grupo de Galois de la ecuación*), buscaba condiciones en el grupo que implicaran la solubilidad de la ecuación polinomial por medio de radicales.

A partir de 1827 la vida de Galois estuvo llena de infortunios. Mientras estaba en el Louis-le-Grand, Galois preparó tres publicaciones que corrieron con pésima suerte. En 1829 envió la primera memoria destinada a su publicación a Cauchy, el matemático francés más influyente del momento. Éste perdió el manuscrito. Igual suerte corrió su segundo artículo que esta vez envió a Fourier, quien murió a los pocos días de haberlo recibido. Mientras tanto, su padre se había suicidado por problemas políticos y él fue rechazado para ingresar a la Escuela Politécnica, la universidad de mayor prestigio donde podía estudiar un matemático. Decidió entrar a la Escuela Normal Superior para prepararse como maestro. En esos días su activismo político se incrementó al grado de escribir feroces ataques contra el rey Luis Felipe de Orleans. Galois fue expulsado de la Escuela Normal y puesto en prisión en dos ocasiones. En 1831, su tercer artículo fue rechazado por Poisson, quien le envió una nota diciendo que era incomprensible.

El 29 de mayo de 1832, Galois fue retado a un duelo a pistola. Las razones son poco claras. Pudo haber sido provocado por la pasión hacia una mujer, o montado por sus enemigos políticos. El caso es que Galois sabía que probablemente moriría al día siguiente. Dedicó toda la noche a escribir sus ideas acerca de los grupos de permutaciones de raíces y la solubilidad por radicales de ecuaciones polinomiales, así como vagas menciones a muchas otras ideas. Este escrito, que dirigió como carta a su amigo Auguste Chevalier, contenía unas pocas páginas garabateadas con comentarios en los márgenes que decían "no tengo tiempo". La carta terminaba así:

Pídele a Jacobi o a Gauss que públicamente den su opinión, no sobre la verdad de los resultados, sino sobre su importancia. Después habrá, espero yo, alguien que encontrará provechoso descifrar todos estos garabatos.

Al día siguiente, Galois fue herido y murió un día después de peritonitis. Fue enterrado en la fosa común el 2 de junio.

El apoyo y comprensión de los matemáticos que Galois no

tuvo en vida, lo encontró después. Joseph Liouville estudió la carta a Chevalier y logró entender las ideas principales de Galois. El 4 de julio de 1843, Liouville se dirigió a la Academia de Ciencias de París con estas palabras:

Espero interesar a la Academia anunciando que entre los papeles de Evariste Galois he encontrado una solución tan precisa como profunda de este bello problema: ¿cuándo es una ecuación soluble por radicales?

Si bien no podemos aquí entrar en detalles sobre la *teoría de Galois*, podemos dar un ejemplo de un polinomio que no es resoluble, por radicales e indicar superficialmente las razones de que esto suceda.

* **Teorema.** El polinomio $x^5 - 6x + 3$ no tiene solución por radicales.

Idea de la demostración. Indicamos los principales pasos:

1) La ecuación $x^5 - 6x + 3 = 0$ tiene tres soluciones reales y dos complejas.

Para observarlo, podemos graficar el polinomio $p(x) = x^5 - 6x + 3$ en el plano X - Y . Observando que $p(-2) = -17$, $p(-1) = 8$, $p(1) = -2$ y $p(2) = 23$, y sabiendo que la gráfica tiene un sólo máximo local en $-\sqrt[5]{6/5}$ y un sólo mínimo local en $\sqrt[5]{6/5}$ (para la obtención de máximos y mínimos hace falta un poco de análisis), entonces la gráfica se ve como en la figura IX.7.

2) El polinomio $p(x) = x^5 - 6x + 3$ no puede escribirse como producto de dos polinomios con coeficientes enteros de grado menor que 5.

Supongamos que esto no fuera el caso y tuviésemos $q(x)$ y $r(x)$ de grado menor que 5 y tales que $p(x) = q(x)r(x)$. Como el grado de $q(x)$ más el grado de $r(x)$ es 5, entonces podemos suponer que $q(x)$ tiene grado 1 o 2, mientras que $r(x)$ tiene grado 4 o 3, respectivamente.

Si $q(x)$ tiene grado 1, entonces $q(x) = nx + m$ y $r(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$ con $n, m, a_4, a_3, a_2, a_1, a_0$ números enteros. Luego, $na_4 = 1$, que implica que n es 1 o -1 y además

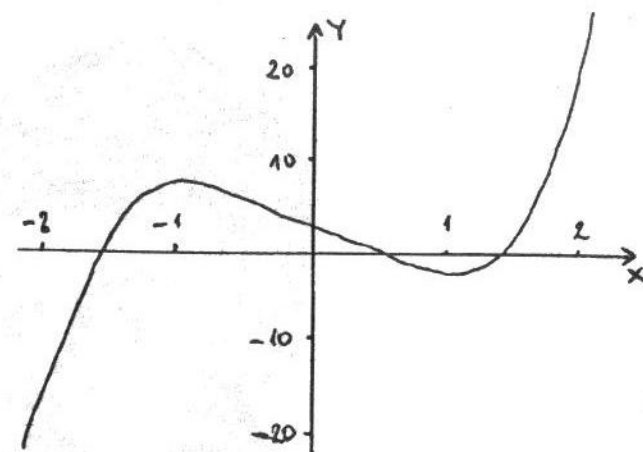


Figura IX.7.

$-m/n = -m$ es una solución de $p(x) = 0$. Pero, por 1, sabemos que $p(x) = 0$ no tiene soluciones enteras.

Supongamos ahora que $q(x) = b_2x^2 + b_1x + b_0$ y $r(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ con coeficientes enteros. Entonces, igualando los coeficientes de los polinomios $p(x) = q(x)r(x)$, tenemos:

$$\begin{aligned} a_0b_0 &= 3 \\ a_0b_1 + a_1b_0 &= -6 \\ a_0b_2 + a_1b_1 + a_2b_0 &= 0 \\ a_1b_2 + a_2b_1 + a_3b_0 &= 0 \\ a_2b_2 + a_3b_1 &= 0 \\ a_3b_2 &= 1 \end{aligned}$$

De la primera ecuación concluimos que uno (y sólo uno) de a_0 y b_0 es divisible entre 3. Supongamos que b_0 es divisible entre 3. De la segunda ecuación obtenemos que b_1 es también divisible entre 3, y finalmente, de la tercera ecuación tenemos que b_2 es divisible entre 3. Esto contradice obviamente la última ecuación. Luego, podemos suponer que $a_0 = 3$ y $b_0 = 1$. Procediendo como antes tenemos que a_1, a_2 y a_3 son divisibles entre 3, lo que otra vez contradice la última ecuación. Esto prueba la afirmación 2).

3) Un polinomio de grado 5 que satisface las condiciones 1) y 2) tiene grupo de Galois S_5 . El grupo de Galois G es un grupo de permutaciones de las cinco raíces complejas del polinomio $p(x)$ (son cinco raíces por el teorema fundamental del álgebra). Luego, G es un subgrupo de S_5 . Un polinomio con la propiedad 2) tiene siempre un elemento de orden 5 en el grupo de Galois, esto es, hay una permutación g en G tal que $g^5 = 1$, pero $g^i \neq 1$ para $1 \leq i \leq 4$. Por otra parte, como hay dos raíces no reales de $p(x)$, éstas tienen la forma z y \bar{z} , lo que produce un elemento de grado 2 en G .

Podemos suponer que a G pertenecen la trasposición $(1, 2)$ y la rotación $(5, 4, 3, 2, 1)$. Como hemos visto en el capítulo 6, estos elementos generan a S_5 . Luego $G = S_5$.

4) El grupo S_5 no es soluble. Esto es, no hay una cadena de subgrupos $S_5 = G_0, G_1, \dots, G_m$ de forma que G_i sea normal en G_{i-1} y el cociente G_{i-1}/G_i tenga orden primo.

En efecto, esto es así porque S_5 tiene como subgrupo normal a A_5 , que no tiene subgrupos normales.

5) El teorema de Galois indica que: si $p(x) = 0$ es soluble por radicales, entonces el grupo de Galois G del polinomio $p(x)$ es soluble.

La idea de la prueba del teorema de Galois es la siguiente: si G es un grupo soluble, tenemos una cadena $G = G_0, G_1, \dots, G_m$ de forma que G_i sea normal en G_{i-1} y el cociente G_{i-1}/G_i tiene orden primo. Asociada a la cadena hay una cadena de subconjuntos de los números complejos $F_0 \subset F_1 \subset \dots \subset F_m$ tales que F_i es el conjunto de números complejos z que quedan fijos bajo la acción de G_i (esto es, $g(z) = z$ para toda g en G_i). El conjunto F_0 contiene a los coeficientes de $p(x)$, mientras que todas las raíces de $p(x) = 0$ están en F_m . El hecho de que G_{i-1}/G_i tenga orden primo implica que los elementos de F_i se pueden construir por radicales a partir de los elementos de F_{i-1} . Luego, las raíces de $p(x) = 0$ se pueden construir por radicales a partir de los coeficientes del polinomio $p(x)$.

En nuestro caso, y debido a que S_5 no es soluble 4), se sigue que el polinomio $p(x) = x^5 - 6x + 3$ no es soluble por radicales. \square

Arthur Cayley y James Joseph Sylvester fueron dos grandes matemáticos ingleses de la época victoriana. Grandes amigos, la obra de cada uno inspiró parte del trabajo del otro. Según E. T. Bell, "la vida de estos dos hombres debe escribirse simultáneamente, pues cada uno hace juego perfecto con la otra y la vida de cada uno suple lo que le hace falta a la otra."

Cayley nació en 1821 en Richmond, Surrey. En contraste con los maestros de Galois, los de Cayley reconocieron tempranamente su genio para las matemáticas y lo apoyaron decididamente. A los 16 años inició sus estudios universitarios en Cambridge y a los 21 obtuvo un puesto especial en la universidad que le permitía un ingreso económico impartiendo algunas clases y seguir sus estudios. Pero no sólo se interesaba por las matemáticas. Leía a los clásicos griegos en su idioma original, novelas inglesas y pintaba acuarelas.

A los 25 años se retiró de su puesto en Cambridge, que le exigía ordenarse clérigo de la iglesia ortodoxa inglesa si quería progresar. Estudió derecho y tres años después despachaba asuntos de testamentos, arriendos y traspasos, lo que hizo durante 14 años. Pero en sus ratos libres continuó con su verdadero interés, las matemáticas. En este tiempo, Cayley publicó más de 200 artículos, algunos de los cuales son clásicos. También tuvo su primer encuentro con Sylvester.

Sylvester nació en 1814 en Londres. Su verdadero apellido era Joseph (de origen judío) pero adoptó el de Sylvester hacia 1820 debido a las persecuciones religiosas (es interesante notar que el Sylvester que eligió la familia es el de unos papas cristianos que se distinguieron por su hostilidad hacia los judíos). A los 15 años pasó a estudiar matemáticas en la Royal Institution de Liverpool y en 1831 ingresó a Cambridge, donde no podía obtener un doctorado por no ser cristiano. Sólo en 1871 Cambridge le otorgó el *Doctorado honoris causa* después de que las disparatadas leyes religiosas fueron abolidas.

En 1838, Sylvester emprendió la aventura de aceptar ser profesor en la Universidad de Virginia, EUA. A los tres meses se vio obligado a renunciar al negarse las autoridades a castigar a un estudiante que lo ofendió. Regresó a Inglaterra y decidió no buscar empleo como matemático, sino como actuuario para una

compañía de seguros de vida. En 1846 comenzó a estudiar leyes e inició el ejercicio de la carrera legal en 1850. Se relacionó con Cayley en 1852.

A partir de entonces Cayley y Sylvester no dejarán nunca de discutir. El primer tema que llamó su atención fue el de la teoría de invariantes. Por ejemplo, si consideramos la ecuación de segundo grado $ax^2 + bx + c = 0$, sabemos que tiene dos raíces iguales solamente si $b^2 - 4ac = 0$. Veamos con cuidado esta expresión $b^2 - 4ac$, que se llama el *discriminante* de la ecuación. Aplicamos un cambio de variable a la ecuación, de forma que la nueva variable y se relaciona con la antigua x por medio de una ecuación lineal $y = px + q$. Entonces, la ecuación de segundo grado en x se convierte en $Ay^2 + By + C = 0$, donde

$$A = c, \quad B = bq - 2cp, \quad C = aq^2 - bpq + cp^2.$$

El discriminante de la nueva ecuación en y es:

$$B^2 - 4AC = (bq - 2cp)^2 - 4c(aq^2 - bpq + cp^2) = q^2(b^2 - 4ac).$$

Es decir, después de aplicar la transformación $y = px + q$, el discriminante de la ecuación sólo cambia por un factor q^2 que depende completamente de los coeficientes de la transformación aplicada. Se dice que el discriminante es un *invariante*.

Los problemas de invariantes demostraron pronto ser de primera importancia en las matemáticas y en la física. Por ejemplo, según la *teoría de la relatividad*, las leyes de la física son las mismas en diferentes *sistemas inerciales* (o sea, sistemas que se mueven con velocidad relativa constante). En otras palabras, las expresiones matemáticas de las leyes de la física deberán ser ecuaciones invariantes bajo cambios de coordenadas entre sistemas inerciales. Con ello el problema de encontrar expresiones matemáticas para las leyes de la física es remplazado por un problema de teoría de invariantes.

Uno de los grandes descubrimientos de Cayley se llevó a cabo en 1858. Trabajando con transformaciones lineales y las correspondientes invariantes, Cayley se dio cuenta de que si ordenaba en ciertos arreglos los coeficientes de las transformaciones podía determinar sencillas operaciones entre los arreglos que correspondían a la composición de transformaciones. Estos arreglos fueron llamados *matrices* con la multiplicación que hemos descrito en los capítulos V y VII. Desde entonces las matrices han

encontrado una variedad amplísima de aplicaciones dentro de las mismas matemáticas y en otras ciencias.

En 1863, la Universidad de Cambridge ofreció un puesto de profesor de matemáticas a Cayley, quien habría de conservarlo hasta su muerte. Por su parte, en 1855, Sylvester consiguió un puesto de profesor de matemáticas en la Royal Military Academy y lo mantuvo hasta 1870, cuando lo obligaron a retirarse. A los 62 años aceptó un puesto en la Universidad Johns Hopkins de Baltimore, EUA.

Las vidas de Cayley y Sylvester cruzaron en otro par de ocasiones. En 1881, Cayley pasó un semestre dando un curso en Johns Hopkins. En 1883, la Universidad de Oxford en Inglaterra invitó al anciano Sylvester a ocupar un puesto de profesor. Entonces reanudaron su amistad y sus trabajos conjuntos hasta la muerte de Cayley en 1895. Sylvester murió en 1897.

Los intereses de Sylvester eran muy diversos, de no haber sido matemático se le conocería quizá como poeta. Algunos poemas suyos se publicaron en memorias que incluían también resultados matemáticos. Según Sylvester, los grandes matemáticos, excepto en caso de accidente, han vivido una larga vida manteniendo su mente vigorosa, dice:

[...] no hay en el mundo ningún estudio que haga actuar de una manera más armónica todas las facultades de la mente como las matemáticas [...] el matemático vive mucho y vive joven; las alas del alma no se desprenden de él tempranamente, ni sus poros se entorpecen con las primeras partículas que vuelan en los caminos polvorientos de la vida vulgar.

Una sencilla combinación de la teoría de invariantes y la teoría de matrices se presenta en el llamado *teorema de Cayley-Hamilton* (el matemático irlandés Hamilton fue otro destacado algebrista de la época victoriana). Para enunciarlo requerimos introducir un poco de notación. Si:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

es una matriz de tamaño 2×2 , entonces $\det A = a_{11}a_{22} - a_{12}a_{21}$. Dejamos al lector interesado el trabajo de demostrar que este determinante es invariante de las matrices bajo las transformaciones lineales en renglones de las matrices. Para

cualquier matriz $A = (a_{ij})$ de tamaño $n \times n$ se puede definir el determinante de A por medio de la siguiente regla debida a Leibniz:

$$\det A = \sum_{\sigma \in S_n} (-1)^{s(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

donde S_n es el grupo de permutaciones de n elementos, y para cada σ en S_n , el signo de σ es $s(\sigma) = 1$ en caso de que σ pertenezca al subgrupo alternante A_n y $s(\sigma) = -1$ en caso contrario. Por ejemplo, S_2 consta de dos elementos, la identidad e y la trasposición $(2, 1)$. Es claro que si A es una matriz 2×2 , se obtiene:

$$\sum_{\sigma \in S_2} (-1)^{s(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} = \det A.$$

El polinomio característico $\chi_A(t)$ de la matriz A de tamaño $n \times n$ es $\chi_A(t) = \det(tI_n - A)$, donde I_n es la matriz identidad de tamaño $n \times n$. Por ejemplo, si A es la matriz:

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix},$$

entonces:

$$\begin{aligned} \chi_A &= \det(tI_2 - A) = \det \begin{pmatrix} t-1 & -2 \\ 2 & t-3 \end{pmatrix} \\ &= (t-1)(t-3) + 4 = t^2 - 4t + 7. \end{aligned}$$

Teorema de Cayley-Hamilton. Sea A una matriz de tamaño $n \times n$ y $\chi_A(t) = t^n + \mu_1 t^{n-1} + \mu_2 t^{n-2} + \cdots + \mu_n$ su polinomio característico. Entonces:

$$A^n + \mu_1 A^{n-1} + \mu_2 A^{n-2} + \cdots + \mu_n I_n$$

es la matriz con todas sus entradas 0.

En el ejemplo anterior, $\chi_A(t) = t^2 - 4t + 7$, entonces:

$$\chi_A(A) = A^2 - 4A + 7I_2 = \begin{pmatrix} -3 & 8 \\ -8 & 5 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ -8 & 12 \end{pmatrix} + \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix},$$

que es claramente la matriz con todas sus entradas 0.

La demostración de este teorema puede verse en cualquier libro de álgebra lineal. Pero invitamos al lector que intente hacerla usando explícitamente todas las definiciones en el caso de matrices de tamaño 2×2 .

EL ESPÍRITU DEL SIGLO XX

Uno de los matemáticos más importantes del siglo XX es, sin duda, *David Hilbert*. Su vida está marcada por esfuerzos de síntesis que sirvieron de fuente de inspiración para muchos de los trabajos emprendidos por otros matemáticos.

Hilbert nació en Königsberg, Prusia, en 1862. Después de una brillante pero dura carrera como *Privatdozent* en Königsberg (que le daba derecho a enseñar sin recibir sueldo), obtuvo una plaza de profesor en la Universidad de Gotinga en 1895 y ahí pasó el resto de su vida.

A finales del siglo XIX, la Universidad de Gotinga tenía gran reputación en el campo de las matemáticas gracias al linaje de



Figura IX.8. David Hilbert.

Gauss, Dirichlet y Riemann. Pero durante los 30 primeros años del siglo XX, alcanzó todavía mayor reputación, debido especialmente a Hilbert y al gran atractivo que ejercía sobre profesores y estudiantes de Alemania y otros países. En esos años estuvieron en Gotinga: Klein, Minkowski, Heisenberg, Born, Landau, Carathéodory, Schmidt, Toeplitz, Haar, Noether, Weyl, Hecke, Courant y, como visitantes, más o menos todos los matemáticos y físicos importantes de la época.

Parte de la atracción de Hilbert se debía a sus logros matemáticos y otra a la sencillez y calidez de su trato personal, algo desacostumbrado en el jerárquico ambiente académico de Alemania. Sobre esto, Richard Courant dice:

En Alemania era considerado imposible que un profesor se rebajara a ser amigo personal de sus estudiantes. Hilbert rompió completamente esta tradición, y esto fue un enorme paso hacia la creación de un ambiente científico; sus estudiantes jóvenes venían a su casa para el té o la cena. La señora Hilbert preparaba grandes cenas para los asistentes y los estudiantes...

Las clases de Hilbert eran inspiradoras. Sus conferencias no eran perfectas en el sentido formal, y sucedía frecuentemente que no había preparado lo suficiente, de modo que al final de la hora tenía que improvisar, lo que hacía de manera torpe [...] teníamos la oportunidad de verlo luchando a veces con problemas simples de matemáticas hasta que encontraba la solución. Esto era mucho más inspirador que una cátedra perfecta.

Los primeros logros de Hilbert se dieron en el campo del álgebra. Antes de llegar a Gotinga había probado un problema en teoría de invariantes que había desafiado a los algebristas por mucho tiempo. Ya en la universidad publicó su tratado sobre números algebraicos que ha sido un libro fundamental desde entonces. En 1899 presentó la primera axiomatización completa de la geometría euclidiana en su libro *Fundamentos de la geometría*.

Gran parte de la fama de Hilbert descansa en la serie de problemas que enunció durante el Congreso Mundial de Matemáticas de París en 1900. En su lista de 23 problemas Hilbert incluyó los que consideraba primordiales en las matemáticas y en los que los matemáticos deberían concentrar sus esfuerzos en los años futuros. Muchos ya han sido resueltos, pero cada vez que un problema fue resuelto significó un acontecimiento ma-

temático y esto se dio a lo largo del siglo XX. Los problemas que quedan por resolver dependen esencialmente de la demostración de la conjetura de Riemann, que se considera en este momento como el problema abierto más importante en matemáticas.

El éxito en fundamentar axiomáticamente la geometría llevó a Hilbert a tratar de lograr una síntesis similar con las matemáticas todas. Sus esfuerzos desde 1905 trataban de lograr la definición de un conjunto de axiomas que fuese congruente (es decir, que no implique contradicciones) y completo (es decir, que todo teorema pueda ser derivado desde los axiomas por medio de la lógica). Finalmente, en 1931, el lógico Kurt Gödel demostró que esto no es posible: en cualquier sistema axiomático se pueden formular afirmaciones indecisibles, esto es, que tanto su validez como su negación son congruentes con los axiomas dados.

Como ejemplo de un resultado de Hilbert presentamos el siguiente teorema, que es uno de los fundamentos de la *geometría algebraica*.

Teorema de los ceros. Consideremos polinomios p, q_1, \dots, q_s en las variables x_1, \dots, x_n . Supongamos que para todo vector complejo (a_1, \dots, a_n) tal que $q_i(a_1, \dots, a_n) = 0$ para toda i , se tiene también que $p(a_1, \dots, a_n) = 0$. Entonces existe un número m y polinomios h_1, \dots, h_s tales que

$$p^m = \sum_{i=1}^s q_i h_i.$$

No intentaremos dar la demostración del teorema. Como un ejemplo explícito de este resultado, tomemos los siguientes polinomios: $p(x_1, x_2) = x_2$, $q_1(x_1, x_2) = x_1 + x_2$, $q_2(x_1, x_2) = -x_1 x_2 + x_2$ y $q_3(x_1, x_2) = -x_2^2 + x_2^3$, en las variables x_1, x_2 .

Si una pareja (a_1, a_2) de números complejos es un cero de los tres polinomios q_1, q_2, q_3 , entonces, $a_1 + a_2 = 0$, $(-a_1 + 1)a_2 = 0$ y $(-1 + a_2)a_2^2 = 0$. De la última ecuación obtenemos que $a_2 = 1$ o $a_2 = 0$. Si se diera el primer caso, entonces de la segunda ecuación tendríamos que $a_1 = 1$ y llegaríamos en la primera ecuación a $1 + 1 = 0$, lo que no es posible en los números complejos. Entonces, $a_2 = 0$, o lo que es lo mismo, (a_1, a_2) es un cero del polinomio $p(x_1, x_2)$. Por otra parte, tenemos que:

$$q_1(x_1, x_2)x_2^2 + q_2(x_1, x_2)x_2 + q_3(x_1, x_2) = x_2^3 = p(x_1, x_2)^3.$$

Para la elaboración de este libro hemos utilizado material de muy diversas fuentes: libros de divulgación, libros de matemática elemental y especializada, libros de filosofía, enciclopedias, artículos en revistas de matemáticas, artículos de periódicos, y otros. Haremos un breve repaso del principal material utilizado en todo el libro y en particular de cada capítulo. Si una referencia está marcada con un asterisco *, significa que el nivel no es elemental. Si lleva dos **, que el material es especializado.

Los libros de divulgación de las matemáticas que hemos usado son los siguientes:

- [1] D. Bergamini. *Matemáticas*. Colección Científica de Life (1981).
- [2] W. Dunham. *The Mathematical Universe*. Wiley (1994).
- [3] E. Maor. *To infinity and beyond*. Princeton (1991).
- [4] J. Newman. *Sigma. El mundo de las matemáticas*. 6 vols. Grijalvo (1968). De la edición de Simon and Schuster (1956).
- [5] I. Peterson. *The Mathematical Tourist*. Freeman & Co. (1988).
- [6] H. Resnikoff y R. Wells. *Mathematics in Civilization*. Dover (1984).
- [7] Scientific American. *Mathematics: An Introduction to its Spirit and Use*. Freeman & Co. (1979).
- [8] A. Wilson. *The infinite in the finite*. Oxford (1995).

Respecto a estos libros tenemos un par de observaciones que hacer. En primer lugar, la mayoría está en inglés. Se da muy poca divulgación de las matemáticas en español. En segundo lugar, y más importante, estos libros no son de matemáticas sino *acerca de las matemáticas*. Con pocas excepciones (sobre todo el libro de Wilson) no se encontrará en estos libros tratamientos matemáticos completos o ejemplos

matemáticos desarrollados. Sin embargo, son libros agradables que pueden leerse como novelas. Los libros de Newman y el de Scientific American son recopilaciones de muchos artículos de divulgación publicados antes en lugares diferentes.

Para algunos *datos históricos* (sobre todo en el capítulo IX) se consultaron también los siguientes libros:

- [9] E. T. Bell. *Men of Mathematics*. Simon and Schuster (1965).
- [10] N. Bourbaki. *Elementos de historia de las matemáticas*. Alianza Universidad (1972).
- [11] C. Boyer y U. Merzbach. *A History of Mathematics*. Wiley (1969).
- [12] *Enciclopedia Biográfica Universal*. "Ciencias exactas". Promexa (1982).
- [13] *Encyclopaedia Britannica*. 30 vols. (1986).
- [14] * P. Gabriel: *Matrizen, Geometrie, Lineare Algebra*. Birkhäuser (1996).
- [15] B. Russell. *Wisdom of the West*. Crescent (1978).

El libro de Bell es muy fácil de leer. Bourbaki trata de ser preciso en sus datos. Gabriel escribió un texto de álgebra lineal en alemán que contiene muchas notas biográficas interesantes. Russell presenta una introducción a las ideas filosóficas en el mundo occidental.

Ahora mencionaremos el material usado en cada capítulo.

CAPÍTULO I

El material referente a la evolución de los sistemas de numeración recurrió a [1] y [6]. Las máquinas de calcular a [7], [13] (el artículo sobre Babbage). El juego de tarjetas con el que se "adivina el pensamiento" se basó en el artículo "The unlisted phone number" en el libro:

- [16] M. Gardner. *Aha! Insight*. Freeman & Co. (1978).

Este librito trae muchos problemas de teoría de números elemental y lógica.

CAPÍTULO II

La vida de Pitágoras y su filosofía en [15]. Los números figurados (triangulares, cuadrados, etc.) en [6] y [8]. La demostración del teorema de Pitágoras que proviene de *Los elementos* de Euclides es

estándar y algo que todo matemático sabe hacer. La demostración de Wallis puede verse en [2]. Comentarios acerca del origen chino del teorema de Pitágoras pueden leerse en [14]. La "nueva" demostración de la irracionalidad de $\sqrt{2}$ puede encontrarse en:

- [17] H. Estermann, en *Math. Gazette* (1975), p. 110.

Los dibujos de Helguera forman parte de la exposición de la Sala de Matemáticas del Museo de las Ciencias *Universum* de la UNAM. El autor del libro tuvo el gusto de trabajar en la dirección del diseño de la sala.

La espiral pitagórica y la construcción de los árboles de Pitágoras pueden verse en los siguientes libros:

- [18] * H. Laurant. *Fractals*. Princeton (1991).
- [19] * H.-O. Peitgen, H. Jürgens y D. Saupe. *Fractals for the classroom*. Springer Verlag (1991).

El método de Newton se trata a fondo en [19]. La exposición acerca de las ternas de Pitágoras es elemental. Puede verse en [6] o en [8].

CAPÍTULO III

Sobre el papiro de Rhind puede consultarse el artículo de J. Newman (1972) en [7]. Consideraciones sobre los métodos algebraicos usados por algunos pueblos de la Antigüedad aparecen en [6] y [8] y la derivación de la solución de la ecuación cúbica en muchos libros. Seguimos aproximadamente la hecha en:

- [20] J. L. Sáenz. "Identidades para la resolución de ecuaciones cúbicas y cuárticas". *Miscelánea Matemática*, núm. 24 (1996).

La colorida historia del descubrimiento de la solución de la ecuación cúbica se narra en el artículo de Ore en [7]. La exposición elemental de las propiedades de las curvas resultantes de ecuaciones cúbicas en:

- [21] S. Lang. *El placer estético de las matemáticas*. Alianza Universidad, núm. 737 (1984).

El resto del material del capítulo es estándar, salvo la demostración del último ejercicio, que es una aplicación de la prueba de Estermann de la irracionalidad de $\sqrt{2}$.

CAPÍTULO IV

Un análisis de las ecuaciones diofantinas puede leerse en el artículo de Gardner en [7] o, por supuesto, en un libro acerca de la teoría de números. Uno relativamente elemental es:

[22] * G. Andrews. *Number Theory*. W. Sanders (1971).

El relato y los detalles matemáticos de la demostración de Wiles al último teorema de Fermat provienen de las siguientes fuentes:

[23] ** G. Faltings. "The proof of Fermat's Last Theorem by R. Taylor and A. Wiles". *Notices of the American Math. Soc.*, vol. 42, núm. 7 (1995).

[24] A. Granville. "Review of the BBC's program, 'Fermat's Last Theorem'". *Notices of the American Math. Soc.*, vol. 44, núm. 1 (1997).

El debate acerca de los números primos es esencialmente elemental (por supuesto, con la excepción del teorema de los números primos). Puede consultarse cualquier libro de teoría de los números.

Resulta interesante ver que la prueba de Euclides de la existencia de números primos infinitos es siempre usada como una muestra de la *belleza* de las matemáticas, por los matemáticos (véase por ejemplo [7]) y los legos en la materia. Arthur Koestler en su autobiografía, *La escritura invisible*, Alianza Editorial, núm. 545 (1974) nos cuenta de su primer día en una prisión en Málaga:

Desde que en la escuela conocí la demostración de Euclides, ésta siempre me llenó de profunda satisfacción, más de orden estético que intelectual. Pues bien, mientras trataba de recordar la demostración y garabateaba los símbolos en la pared, me sentí invadido por el mismo hechizo [...] Debo de haber permanecido ahí algunos minutos, como transportado en un rapto y teniendo conciencia, aunque sin expresarlo con palabras, de que "esto es perfecto".

Los datos sobre los números primos de la forma $2^m - 1$ provienen de artículos en [7] y [2]. Demostraciones del teorema de los números primos se hallan en el artículo de D. Hawkins [7] y [22].

Problemas cuya solución incluye ecuaciones diofantinas en:

[25] A. Dunn. *Mathematical Bafflers*. Dover (1980).

La mayoría de los problemas en este librito son sencillos, pero de agradable presentación.

CAPÍTULO V

Los datos históricos sobre la criptografía y el cifrado por medio de trasposiciones provienen del artículo correspondiente en la *Macropaedia* de [13]. La codificación con matrices es bien conocida. Un tratamiento elemental en:

[26] * S. Grossman. *Aplicaciones del álgebra lineal*. Grupo Ed. Iberoamericano (1988).

Problemas similares al discutido en "Echando volados por teléfono" en [5]. El estudio de la aritmética módulo n aparece en casi todas partes, y la historia del mensaje que Edgar Allan Poe no pudo descifrar en:

[27] M. Gardner. *Penrose tiles to trapdoor ciphers*. Freeman & Co. (1989).

CAPÍTULO VI

Varios artículos escritos por matemáticos y en los que se analiza el embaldosado de la Alhambra pueden encontrarse en:

[28] "La Alhambra". Revista *Epsilon*. Granada (1987).

En particular se recomienda el artículo de J. M. Montesinos que abre la revista. Muchas de las ilustraciones fueron tomadas de él. Una presentación formal de los grupos cristalográficos y de los teoremas de clasificación que mencionamos en el texto puede leerse en:

[29] ** B. Grünbaum y G. Shepard. *Tilings. An introduction*. Freeman (1989).

Un libro clásico sobre las simetrías es:

[30] H. Weyl. *Symmetry*. Princeton (1972).

Los lectores más avanzados que quieran leer sobre la teoría de grupos pueden consultar muchos libros. Dos clásicos son:

- [31] * N. Jacobson. *Lectures in Abstract Algebra*. Princeton (1974).
 [32] * S. Lang. *Álgebra*. Aguilar (1973).

La teoría de Galois amerita consultar el bello libro:

- [33] * I. Stewart. *Galois Theory*. Chapman and Hall (1973).

El grupo de simetrías del icosaedro en [8]. El descubrimiento y propiedades de los fulerenos en:

- [34] ** B. Kostant. "The graph of the truncated icosahedron and the last letter of Galois". *Notices of the American Math. Soc.*, vol. 42, núm. 9 (1995).

Sobre las grecas puede verse en nivel elemental [3], más avanzado [29].

CAPÍTULO VII

Muchos ejemplos de modelos matemáticos empleados en el deporte en:

- [35] * L. E. Sadovskii y A. L. Sadovskii. "Mathematics and Sport". *American Math. Soc.* (1993).

El modelo del juego de baloncesto es una modificación del presentado en [35]. La teoría de matrices que se requiere (cadenas de Markov) puede verse en forma más general en:

- [36] * J. A. de la Peña. *Álgebra lineal avanzada*. Fondo de Cultura Económica (1996).

que es un libro de texto para los estudiantes de una carrera científica (no necesariamente matemáticas).

Los datos de los equipos de baloncesto de la NBA que aplicamos en nuestro modelo fueron obtenidos en:

- [37] Diario *Reforma*. 21 de mayo de 1997.
 [38] *MVP. Pro basketball 96-97*. Revista. Junio (1997).

Problemas de aplicaciones de cadenas de Markov similares a los tratados en las últimas secciones del capítulo en [36].

CAPÍTULO VIII

El debate acerca de si las máquinas pueden pensar es ponderado en varios libros. Mencionamos aquí los siguientes:

- [39] D. Hofstadter. *Gödel, Escher y Bach*. Tusquets (1995).
 [40] A. Ross. *Controversia sobre mentes y máquinas*. Tusquets (1964).

El primero es ya un clásico (fue publicado en inglés en 1975). Por supuesto no considera únicamente el problema mente-máquina, pero lo trata con alguna extensión. El segundo constituye una recopilación de artículos sobre el tema. Inicia con el importante artículo de Turing que aparece en el texto.

- [41] A. Turing. "Computing machines and intelligence". *Mind*, vol. LIX (1950).

El análisis de la muy recomendable película *Blade Runner* en:

- [42] *Blade Runner*. Tusquets (1988).

La teoría de las máquinas de estados finitos en:

- [43] * J. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall (1971).
 [44] * S. Eilenberg. *Automata, Languages and Machines*. Vol. A. Academic Press (1974).

En estos libros viene la prueba completa del teorema de Kleene. Las máquinas de Turing son tratadas en [39], [44] y, dentro del contexto de su comparación con la mente humana en:

- [45] R. Penrose. *The Emperor's New Mind*. Oxford (1989).

Es éste un libro muy interesante al que se le ha criticado su carácter especulativo. En él se da el número correspondiente a una máquina universal de Turing. Una prueba más abstracta de la existencia de la máquina universal en [43]. La forma de tratar las máquinas de Turing por medio de gráficas es similar al tratamiento que se hace en [44]. La construcción, en forma detallada, de las máquinas *max* y *fib* aparece por primera vez en este libro.

Algunas ideas de Dyson pueden leerse en su muy disfrutable autobiografía:

- [46] F. Dyson. *Trastornando el Universo*. Fondo de Cultura Económica.

La partida *Deep Blue*-Kasparov y los comentarios técnicos se obtuvieron de:

- [47] M. Sisniega. "El ogro azul". *El Universal* del 17 de mayo de 1997.

CAPÍTULO IX

Los datos históricos de los algebristas presentados provienen de las fuentes ya mencionadas. La demostración del pequeño teorema de Fermat esencialmente en [22], al igual que el análisis de la función ϕ de Euler. La prueba del teorema de Gauss en:

- [48] * H. Dörrie. *100 Great Problems of Elementary Mathematics*. Dover (1965).

En este libro puede encontrarse también una prueba analítica sencilla del teorema fundamental del álgebra. Probablemente la mejor conocida usa el pequeño teorema de Picard de la variable compleja.

Una biografía extensa y emotiva de la vida de Galois es *El elegido de los dioses* de L. Infeld, físico y antiguo colaborador de Einstein. La aplicación de la teoría de Galois a la irresolubilidad por radicales de la ecuación de quinto grado puede verse con detalle en [33]. La demostración de la imposibilidad de los tres problemas griegos (duplicación del cubo, trisección del ángulo y cuadratura del círculo) usando la teoría de Galois en [33] y [48]. En [33] también se da una prueba del teorema fundamental del álgebra como una aplicación sencilla de la teoría de Galois.

Más detalles de la vida de Sylvester y Cayley en [11] y [14]. También es interesante la vida de Hamilton. Algunos datos sobre la teoría de invariantes aparecen en [31] y [32]. La historia de la noción de determinante y sus aplicaciones en [14]. La prueba más sencilla del teorema de Cayley-Hamilton requiere el uso de las formas canónicas de Jordan (como se efectúa en [36]). Otra prueba puede verse en [14].

Los comentarios de Courant sobre Hilbert provienen de:

- [49] R. Courant. "Reminiscences from Hilbert's Göttingen". *Mathematical Intelligencer*, vol. 12 (1990).

Un recuento más o menos detallado de los 23 problemas de Hilbert y la solución (de los que han sido resueltos) en [11]. Una demostración algebraica del teorema de los ceros en [32].

La más bella experiencia es lo misterioso.
Es la verdadera fuente de todo arte y ciencia.

ALBERT EINSTEIN

INTRODUCCIÓN	11
I. DE LOS DEDOS DE LAS MANOS A LAS COMPUTADORAS	17
Sistemas posicionales	20
Ábacos y computadoras	23
Adivina el número que estoy pensando	26
II. UN MUNDO HECHO DE NÚMEROS	28
El teorema de Pitágoras	33
Una tragedia griega	35
El árbol de Pitágoras	42
Construyendo triángulos rectángulos con lados enteros	45
Números racionales <i>vs.</i> números irracionales	46
III. CALCULANDO LO DESCONOCIDO	46
La jerarquía de los números	49
Ecuaciones a la italiana	52
La solución de Cardano a la ecuación cúbica	54
El matrimonio del álgebra y la geometría	56
Enteros algebraicos	61
IV. LA HISTORIA EN EL MARGEN DE UN LIBRO	63
Números primos	68
La distribución de los números primos	72
El problema del granjero	74
Los primos de Fermat	75
V. ENVIANDO MENSAJES SECRETOS	76

Codificando con matrices	79
Echando volados por teléfono	83
Un mundo donde $2 + 2 = 1$	86
Cómo partir un número en cubos	87
¿FUPNWBWBUJRTJKCRDGXLUP?	87
VI. IMÁGENES DE ALHAMBRA	90
Los cristales: mosaicos de la naturaleza	96
Simetrías y grupos	100
Los grupos cristalográficos planos	107
Moléculas como pelotas de fútbol	111
Grecas	114
VII. PRONÓSTICOS DEPORTIVOS	116
Matrices estocásticas y los "Toros" de Chicago	124
¿Cuántos caminos llevan a Roma?	129
Otras aplicaciones	132
VIII. ¿SUEÑAN LOS ANDROIDES CON OVEJAS ELÉCTRICAS?	134
Los lenguajes de las máquinas	138
Las máquinas de Turing	145
¿Puede pensar una máquina?	150
<i>Deep Blue vs. Kasparov</i> . Defensa Karo-Kann	152
Una máquina para contar conejos	154
IX. ALGUNOS ALGEBRISTAS Y SUS TEOREMAS	157
El abogado y los números	159
El ciego que vio más lejos	163
El príncipe de los matemáticos	167
El elegido de los dioses	171
Dos amigos ingleses	177
El espíritu del siglo XX	181
NOTAS Y REFERENCIAS COMENTADAS	185

Álgebra en todas partes, de José Antonio de la Peña,
 núm. 166 de la colección La Ciencia para Todos,
 se terminó de imprimir y encuadernar en junio de 2010
 en Impresora y Encuadernadora Progreso, S. A. de C. V. (IEPSA),
 Calzada San Lorenzo 244; 09830 México, D. F.
 Se tiraron 1 600 ejemplares.