¿QUÉ ES LA MATEMÁTICA?

RICHARD COURANT
Jefe del Departamento de Matemáticas
de la Universidad de Nueva York

HERBERT ROBBINS
Profesor de Matemáticas
de la Universidad de Nueva York

¿QUÉ ES LA MATEMÁTICA?

UNA EXPOSICIÓN ELEMENTAL DE SUS IDEAS Y MÉTODOS

Traducción del inglés por
LUIS BRAVO GALA
Licenciado en Ciencias Exactas

AGUILAR

colección ciencia y técnica
sección matemáticas y estadística
obra incorporada con el asesoramiento
de luis bravo gala

edición española

depósito legal m 14137/1979
quinta edición–segunda reimpresión–1979
ISBN 84-03-20032-3
printed in spain impreso en españa por gráficas halar s l
andrés de la cuerda 4 madrid

edición original
what is mathematics?
oxford university press new york and london

PRÓLOGOS

PRÓLOGO A LA PRIMERA EDICIÓN

Desde hace más de dos milenios, una cierta familiaridad con la matemática ha sido considerada como parte indispensable de la formación intelectual de toda persona cultivada. En la actualidad, sin embargo, se halla en grave peligro el puesto ocupado tradicionalmente en la educación por esta disciplina; por desgracia, algunos de los profesionales que la representan comparten la responsabilidad de tal situación. La enseñanza de la matemática ha degenerado con frecuencia en un vacío entrenamiento de resolución de problemas, que si bien puede desarrollar una habilidad formal, no conduce en cambio a una comprensión efectiva ni a una mayor independencia intelectual. La investigación matemática muestra una tendencia hacia la superespecialización y hacia una excesiva insistencia en lo abstracto; las aplicaciones y conexiones con otros campos del saber han sido descuidadas. Sin embargo, tal estado de cosas no debe justificar una política de retraimiento. Por el contrario, la reacción opuesta puede y debe partir de aquellos que se sienten conscientes del valor de la disciplina intelectual. Profesores, estudiantes y público culto piden una reforma constructiva y no una resignación siguiendo la línea de menor resistencia. La meta será una verdadera comprensión de la matemática como un todo orgánico y como base para el pensamiento y la acción científicos.

Algunos libros espléndidos de biografía e historia y otros más populares han estimulado el interés general latente; pero el conocimiento no puede adquirirse utilizando únicamente medios indirectos. La comprensión de la matemática no puede ser transmitida indirectamente o como un juego sin dificultades, como tampoco pueden adquirir una educación musical, a través de reseñas periodísticas brillantes, aquellos que no han escuchado buena música con frecuencia. Un contacto real con el *contenido* de la matemática viva es necesario. Sin embargo, cabe evitar algunos detalles de técnica y muchas digresiones; la presentación de las matemáticas debe estar tan libre de excesos de rutina como del dogmatismo prohibitivo que rehuye revelar el motivo o la meta y que constituye así un obstáculo de

de mala fe para un esfuerzo honesto. Es posible seguir una ruta directa a partir de los elementos fundamentales hasta puntos avanzados, desde los cuales puedan divisarse la sustancia y las fuerzas directrices de la matemática moderna.

El presente libro es un intento en esa dirección. Y en tanto que presupone únicamente los conocimientos que pueden adquirirse en la enseñanza media, puede considerarse como elemental. No constituye, sin embargo, una concesión a la tendencia peligrosa que consiste en soslayar toda dificultad. Su lectura requiere un cierto grado de madurez intelectual y un deseo de tener pensamientos propios. El libro está escrito para principiantes y entendidos, para estudiantes y profesores, para filósofos e ingenieros, como libro de texto y de consulta. Quizá esto encierre una intención demasiado ambiciosa. Bajo la presión de otros trabajos han sido hechas ciertas concesiones al publicar el libro, después de varios años de preparación, antes de estar realmente terminado. Por ello, críticas y sugestiones serán bien recibidas.

De todos modos, se espera que el libro pueda servir un propósito útil y constituir una contribución a la educación superior americana, procedente de alguien que está profundamente agradecido por la oportunidad que le fué ofrecida en este país. Mientras que la responsabilidad por el plan y la filosofía de esta publicación recae enteramente en el que suscribe, todos los méritos que pueda encerrar debe compartirlos con Herbert Robbins. Éste, desde el momento en que fué asociado a esta tarea, la ha considerado generosamente como su propia causa, y su colaboración ha desempeñado un papel decisivo en el trabajo de completar la obra en su forma presente.

Debo reconocimiento también a la ayuda de varios amigos. Discusiones con Niels Bohr, Kurt Friedrichs y Otto Neugebauer han influído en la actitud filosófica e histórica; Edna Kramer nos ha hecho críticas constructivas desde el punto de vista de la enseñanza; David Gilbarg preparó las primeras notas de curso que originaron el libro; Ernest Courant, Norman Davids, Charles de Prima, Alfred Horn, Herbert Mintzer, Wolfgang Wasow y otros prestaron su ayuda en la interminable tarea de escribir varias veces el manuscrito, y contribuyeron a muchas mejoras de detalle; Donald Flanders hizo varias sugestiones valiosas y arregló el manuscrito para la imprenta; John

Knudsen, Hertha von Gumppenberg, Irving Ritter y Otto Neugebauer prepararon las figuras; H. Whitney contribuyó a la colección de ejercicios del Apéndice. La General Education Board de la Fundación Rockefeller ha contribuído generosamente al desarrollo de los cursos y notas que fueron luego la base del libro. Gracias sean dadas también a la Waverly Press, y en particular a Grover C. Orth, por su trabajo, de una competencia extraordinaria; y a la Oxford University Press, en especial a Philip Vaudrin y W. Oman, por su animosa iniciativa y cooperación.

R. COURANT.

NEW ROCHELLE, N. Y.
Agosto de 1941.

PRÓLOGO A LA SEGUNDA, TERCERA Y CUARTA EDICIONES

Durante los últimos años, la fuerza de los acontecimientos condujo a una creciente demanda de información y enseñanza matemáticas. Ahora más que nunca existe el peligro de frustración y desilusión, a no ser que estudiantes y profesores intenten ver más allá del formalismo y manipulación matemáticos, y comprendan la verdadera esencia de la matemática. Este libro fué escrito para esos profesores y estudiantes, y la acogida hecha a la primera edición confirma a los autores en la esperanza de que pueda ser de utilidad.

Críticas de varios lectores han dado lugar a numerosas correcciones y mejoras. Gracias cordiales debemos a Natascha Artin por su generosa ayuda en la preparación de la cuarta edición.

R. COURANT.

NEW ROCHELLE, N. Y.
18 de marzo de 1943.
10 de octubre de 1945.
28 de octubre de 1947.

CÓMO DEBE UTILIZARSE ESTE LIBRO

El libro está escrito en un orden sistemático, pero esto no quiere decir que sea necesario que el lector lo estudie página a página y capítulo tras capítulo. Por ejemplo, la introducción histórica y filosófica puede muy bien dejarse para después de haber leído el resto. Los diferentes capítulos son, en gran parte, independientes unos de otros. Frecuentemente, el comienzo de una sección podrá ser comprendido sin dificultad. Luego, el camino conducirá gradualmente hacia adelante, para llegar en el final de cada capítulo y en los suplementos a cuestiones más difíciles. Así, el lector que desee información general más bien que conocimientos específicos puede conformarse con una selección de materias que eluda los análisis detallados.

El estudiante con poca base matemática también puede hacer una selección. Asteriscos y tipos pequeños de letra indican partes que pueden ser omitidas en una primera lectura sin gran perjuicio para la comprensión de las siguientes. Por otra parte, no habrá inconveniente si el estudio se limita a las secciones o capítulos en los que el lector esté más interesado. La mayoría de los ejercicios son de carácter no rutinario; los de mayor dificultad van marcados con un asterisco. El lector no debe alarmarse si no alcanza a resolver algunos de ellos.

En los capítulos sobre construcciones geométricas y sobre máximos y mínimos se encontrará material apropiado para grupos selectos de estudiantes de cursos superiores.

Esperamos que el libro servirá tanto para el estudiante que comienza como para el que esté más avanzado, así como para el profesional que se halle verdaderamente interesado en la matemática. Además, puede servir como base para cursos de tipo no tradicional sobre los conceptos fundamentales de esta ciencia. Los capítulos III, IV y V pueden utilizarse en un curso de geometría, mientras los capítulos VI y VIII ofrecen en conjunto una exposición autónoma del cálculo infinitesimal, en la que se concede atención especial a los conceptos básicos y se trata de evitar los métodos rutinarios. Pueden usarse como un texto inicial por un profesor que desee hacer contri-

buciones propias, completando el material de acuerdo con sus necesidades específicas y, especialmente, añadiendo otros ejemplos numéricos. Numerosos ejercicios distribuídos a lo largo del texto y la colección adicional del final facilitarán el uso del libro en las clases.

Esperamos también que el lector encuentre a menudo detalles de interés en muchas discusiones elementales que contienen el germen de más amplios desarrollos.

ÍNDICE GENERAL

ÍNDICE GENERAL

¿QUÉ ES LA MATEMÁTICA?

INTRODUCCIÓN

¿QUÉ ES LA MATEMÁTICA?

La matemática, como una expresión de la mente humana, refleja la voluntad activa, la razón contemplativa y el deseo de perfección estética. Sus elementos básicos son: lógica e intuición, análisis y construcción, generalidad y particularidad. Aunque diversas tradiciones han destacado aspectos diferentes, es únicamente el juego de estas fuerzas opuestas y la lucha por su síntesis lo que constituye la vida, la utilidad y el supremo valor de la ciencia matemática.

Sin duda, todo el desarrollo matemático ha tenido sus raíces psicológicas en necesidades más o menos prácticas. Pero una vez en marcha, bajo la presión de las aplicaciones necesarias, dicho desarrollo gana impulso en sí mismo y trasciende los confines de una utilidad inmediata. Esta tendencia de la ciencia aplicada hacia la teórica aparece tanto en la historia antigua como en muchas de las contribuciones a la matemática moderna debidas a ingenieros y físicos.

La historia de las matemáticas comienza en Oriente, donde, hacia el año 2000 a. de J.C., los babilonios poseían ya una gran cantidad de material que podría ser clasificado hoy como perteneciente al álgebra elemental. Pero como ciencia, en el sentido moderno, la matemática aparece más tarde, en Grecia, entre los siglos v y iv antes de J.C. El contacto creciente entre el Oriente y los griegos, que comienza en los tiempos del imperio persa y culmina en el período que sigue a las expediciones de Alejandro, puso a los griegos al corriente de los conocimientos de los babilonios en matemática y astronomía. La matemática fué sometida entonces a las discusiones filosóficas que florecieron en las ciudades griegas. Los pensadores griegos se dieron pronto cuenta de las grandes dificultades inherentes a los conceptos matemáticos de continuidad, movimiento e infinitud, así como al problema de medir magnitudes arbitrarias con unidades prefijadas. Entonces fué llevado a cabo un admirable esfuerzo para vencerlas y el resultado, la teoría de Eudoxio del continuo geométrico, fué de tal perfección, que para encontrar algo que pueda comparársele es necesario qué, dos milenios más tarde, aparezca la teoría moderna de los números irracionales. La tendencia axiomático-deductiva en matemáticas tuvo su origen en tiempos de Eudoxio y cristalizó en los *Elementos* de Euclides.

Sin embargo, aunque la tendencia teórica y axiomática de la matemática griega es una de sus más importantes características y ha ejercido una influencia enorme, nunca se insistirá demasiado en que las aplicaciones y conexiones con la realidad física desempeñaron un papel importante como parte de la matemática de la antigüedad, y que en muchas ocasiones fué preferido un modo de exposición menos rígido que el de Euclides.

Es muy posible que el descubrimiento de las dificultades relacionadas con las cantidades *inconmensurables* desviara a los griegos del desarrollo del cálculo numérico, alcanzado con anterioridad en Oriente. En su lugar, se abrieron camino a través de la geometría axiomática pura. Y así comenzó un extraño rodeo en la historia de la ciencia, y quizá se perdió una gran oportunidad. Durante casi dos mil años el peso de la tradición geométrica griega retrasó la inevitable evolución del concepto de número y el desarrollo del cálculo algebraico, que más tarde habían de ser la base de la ciencia moderna.

Después de un período de preparación lenta, la revolución en la matemática y en la ciencia comenzó su fase vigorosa en el siglo XVII, con la geometría analítica y el cálculo diferencial e integral. Mientras la geometría griega conserva aún un lugar destacado, el ideal griego de cristalización axiomática y de deducción sistemática desaparece durante los siglos XVII y XVIII. Razonamientos lógicos rigurosos, a partir de definiciones claras y no contradictorias, axiomas *evidentes*, fueron cuestiones sin importancia para los nuevos exploradores de la ciencia matemática. En una verdadera orgía de conjeturas intuitivas, de razonamientos convincentes entrelazados con un misticismo sin sentido, con una confianza ciega en el poder sobrehumano de los procesos formales, conquistaron un mundo matemático de inmensas riquezas. Luego, gradualmente, la exaltación del progreso dejó el paso a un espíritu de autocrítica. En el siglo XIX la necesidad inmanente de consolidar, y el deseo de una mayor seguridad en la extensión de la enseñanza superior, que había impulsado la Revolución francesa, condujo inevitablemente a una revisión de los fundamentos de la nueva matemática, en particular del cálculo diferencial e integral, así como del concepto fundamental de límite. Así, el siglo XIX constituyó no sólo un período de nuevos avances, sino que además puede caracterizarse por un afortunado retorno al ideal clásico de precisión y demostraciones rigurosas. Y en este sentido llegó a superar al modelo de la ciencia griega. Una vez más el péndulo se inclinó del lado de la pureza lógica y de la abstracción. Actualmente vivimos aún en este período, aunque es de esperar que la desafortunada se-

paración entre la matemática pura y las aplicaciones a la vida, quizá inevitable en tiempos de revisión crítica, venga seguida de una era de íntima unidad.

La renovada solidez interna, y sobre todo la simplificación enorme alcanzada sobre la base de una comprensión más clara, hacen posible hoy poder dominar la teoría matemática sin perder de vista las aplicaciones. Establecer de nuevo una unión orgánica entre ciencia pura y aplicada y un equilibrio estable entre la generalidad abstracta y la individualidad concreta puede ser muy bien la tarea universal de la matemática en el futuro inmediato.

No es éste el lugar para un análisis filosófico o psicológico detallado de la matemática. Únicamente podemos destacar algunos puntos. Parece existir un grave peligro en el excesivo predominio del carácter axiomático-deductivo de las matemáticas. Ciertamente, el elemento de invención constructiva, de intuición directora, escapa a una simple formulación filosófica; sin embargo, continúa siendo el núcleo de todo resultado matemático, aun en los campos más abstractos. Si la forma deductiva cristalizada es la meta, la intuición y la construcción son, cuando menos, las fuerzas directrices. Una amenaza seria para la verdadera vida de la ciencia aparece contenida en la afirmación de que la matemática no es más que un sistema de conclusiones derivadas de definiciones y postulados que deben ser compatibles, pero que, por lo demás, pueden ser creación de la libre voluntad del matemático. Si esta descripción fuera exacta, las matemáticas no podrían interesar a ninguna persona inteligente. Sería un juego con definiciones, reglas y silogismos, sin meta ni motivo alguno. La noción de que el intelecto puede crear sistemas de postulados plenos de significado de modo arbitrario es una verdad «a medias» decepcionante. Únicamente bajo una disciplina de responsabilidad frente a un todo orgánico, guiada sólo por necesidades intrínsecas, puede la mente libre obtener resultados de valor científico.

Aunque la tendencia pasiva del análisis lógico no puede representar toda la matemática, ha conducido, sin embargo, a una comprensión más profunda de los hechos matemáticos y de su interdependencia, y también a una mayor penetración en la esencia de los conceptos matemáticos. A partir de ella se ha desarrollado un punto de vista moderno en las matemáticas que es característico de una actitud científica universal.

Cualquiera que sea el punto de vista filosófico, para todos los propósitos de observación científica, un objeto agota en sí la totalidad de relaciones posibles respecto del observador o del instrumento. Na-

turalmente, la simple percepción no constituye conocimiento; debe ser coordinada e interpretada con referencia a alguna entidad subyacente, una «cosa en sí», que no es un objeto de la observación física directa, sino que pertenece a la metafísica. Sin embargo, en el proceso científico es importante descartar los elementos de carácter metafísico y considerar los hechos observables como la última fuente de nociones y construcciones. Renunciar a la meta de comprender la «cosa en sí», de conocer la «realidad última», de desentrañar la esencia más íntima del mundo, puede ser psicológicamente penoso para entusiastas ingenuos, pero de hecho es uno de los sacrificios de consecuencias más fecundas en el pensamiento moderno.

Algunos de los mayores avances en la física han sido el premio a una adhesión decidida al principio de eliminar la metafísica. Cuando Einstein consiguió reducir la noción de «sucesos simultáneos que ocurren en lugares distintos» a fenómenos observables; cuando señaló como prejuicio metafísico la creencia de que este concepto debe tener un significado científico en sí mismo, encontró la clave de su teoría de la relatividad. Cuando Niels Bohr y sus discípulos analizaron el hecho de que toda observación física va acompañada de un efecto del instrumento observador en el objeto observado, se hizo claro que el intento de fijar simultáneamente la posición y la velocidad de una partícula no es posible en el sentido de la física. Las consecuencias trascendentes de este descubrimiento, contenidas en la teoría moderna de la mecánica cuántica, son hoy familiares a todo físico. En el siglo pasado prevaleció la idea de que las fuerzas mecánicas y los movimientos de partículas en el espacio eran cosas en sí mismas, mientras que electricidad, luz y magnetismo debían ser reducidos o *explicados* como fenómenos mecánicos, de la misma manera que se hacía con el calor. El *éter* fué inventado como medio hipotético capaz de los movimientos mecánicos no explicados satisfactoriamente y que aparecían bajo las formas de luz y electricidad. Poco a poco se comprendió que el éter era necesariamente inobservable; por consiguiente, no pertenecía a la física, sino a la metafísica. Con pena por algunos y con satisfacción por otros, las explicaciones mecánicas de la luz y la electricidad, y con ellas el éter, debieron ser finalmente abandonadas.

Una situación análoga, quizá más acentuada, existe en la matemática. A través de los tiempos, los matemáticos consideraron sus objetos, tales como números, puntos, etc., como cosas sustanciales en sí. Pero en vista de que estos entes desafiaban siempre los intentos para una descripción adecuada, los matemáticos del siglo pasado llegaron paulatinamente a la convicción de que el problema de la signifi-

cación de dichos objetos como cosas sustanciales no tenía, en modo alguno, sentido dentro de las matemáticas. Las únicas proposiciones relativas a ellos que pueden importar no se refieren a su realidad sustancial; representan únicamente las relaciones mutuas entre «objetos indefinidos» y las reglas que rigen las operaciones con ellos. Lo que «realmente» *son* los puntos, las rectas y los números ni se puede ni es necesario discutirlo en la ciencia matemática. Lo que interesa y lo que corresponde a hechos *comprobables* es su estructura y relación: que dos puntos determinan una recta, que los números se combinan según ciertas reglas para formar otros números, etc. La percepción clara de la necesidad de una *desustanciación* de los conceptos elementales matemáticos ha sido uno de los resultados más importantes y fecundos del desarrollo axiomático moderno.

Por suerte, las mentes creadoras olvidan las creencias filosóficas dogmáticas cuando la persistencia en ellas podría impedir resultados constructivos. Tanto para entendidos como para profanos no es la filosofía, y sí únicamente la experiencia activa en matemáticas, la que puede responder a la pregunta: ¿Qué es la matemática?

CAPÍTULO PRIMERO

LOS NÚMEROS NATURALES

Introducción.—Los números son la base de la matemática moderna. Ahora bien: ¿qué es un número? ¿Qué significado tiene decir $^1/_2 + {}^1/_2 = 1$, $^1/_2 \cdot {}^1/_2 = {}^1/_4$ y $(-1)(-1) = 1$? En la segunda enseñanza se aprende el manejo de las fracciones y de los números negativos, pero una comprensión efectiva de los sistemas de números requiere investigar sus elementos más simples. Mientras los griegos hacen de los conceptos geométricos de punto y recta la base de sus matemáticas, hoy se admite como principio director que todas las proposiciones matemáticas pueden ser reducidas en última instancia a proposiciones sobre los *números naturales*: 1, 2, 3,... «Dios creó los números naturales; el resto es obra de los hombres.» Con estas palabras Leopold Kronecker (1823-1891) señaló la base precisa sobre la cual puede construirse el edificio de la matemática.

Creados por la mente humana para contar objetos agrupados de diversos modos, los números no contienen referencia alguna a las características de los objetos contados. El número 6 es una abstracción obtenida a partir de todas las colecciones que contienen seis cosas; no depende de las cualidades específicas de dichas cosas ni de los símbolos usados para representarlas. Únicamente en etapas avanzadas del desarrollo intelectual llega a percibirse con toda claridad el carácter abstracto de la idea de número. Para los niños, los números están siempre ligados a objetos tangibles, tales como dedos o bolas, y los lenguajes primitivos dan a los números un sentido concreto, designando con palabras distintas los números que corresponden a diferentes tipos de objetos.

Por suerte, los matemáticos no tienen que ocuparse del aspecto filosófico de la transición que da el paso de colecciones de objetos concretos al concepto abstracto de número.

Consideraremos, por tanto, como dados los números naturales, junto con las dos operaciones fundamentales, adición y multiplicación, mediante las cuales pueden ser combinados.

I. CÁLCULO CON NÚMEROS ENTEROS

1. Leyes de la aritmética.—La teoría matemática de los números naturales o *enteros positivos* se conoce con el nombre de *aritmética.*

Se basa en el hecho de que la adición y multiplicación de enteros están regidas por ciertas leyes. Para enunciar estas leyes con toda generalidad no podemos limitarnos a usar símbolos tales como 1, 2, 3, que se refieren a enteros particulares. La proposición

$$1 + 2 = 2 + 1$$

es únicamente un caso particular de la ley general que dice que la suma de dos enteros no depende del orden en que éstos se consideren. En consecuencia, si queremos expresar el hecho de que una cierta relación entre enteros es válida, cualesquiera que sean los valores de los enteros considerados, debemos representarlos simbólicamente mediante letras: a, b, c,... De acuerdo con esto podemos enunciar las conocidas cinco leyes fundamentales de la aritmética:

1) $a + b = b + a$. 3) $a + (b + c) = (a + b) + c$.
2) $ab = ba$. 4) $a(bc) = (ab)c$.
5) $a(b + c) = ab + ac$.

Las dos primeras son las leyes *conmutativas* de la adición y de la multiplicación; indican que el orden de los elementos que intervienen en dichas operaciones puede ser alterado. La tercera, que es la ley *asociativa* de la adición, dice que, en la adición de tres números, se obtiene el mismo resultado si se añade al primero la suma del segundo y el tercero, o al tercero la suma del primero y el segundo. La cuarta es la ley asociativa de la multiplicación. La última, que es la ley *distributiva*, expresa que para multiplicar una suma por un entero se puede multiplicar cada término de la suma por dicho entero y sumar luego los productos obtenidos.

Estas leyes de la aritmética son muy simples e incluso parecen evidentes. Sin embargo, pueden no ser aplicables a otros entes distintos de los números enteros. Así, si a y b son símbolos no de enteros, sino de sustancias químicas, y si *adición* se entiende en su sentido corriente, es evidente que la ley conmutativa no es siempre válida; p. ej., si se añade ácido sulfúrico a agua se obtiene una solución diluída, mientras que la adición de agua a ácido sulfúrico puro puede acarrear consecuencias catastróficas para el experimentador. Análogas consideraciones probarían que, en este tipo de «aritmética» química, las leyes asociativa y distributiva de la adición pueden muy bien no ser ciertas. Asimismo es probable imaginar tipos de aritméticas en las cuales alguna o varias de las leyes 1) a 5) no sean válidas. De hecho, tales sistemas son estudiados efectivamente en la matemática moderna.

Un modelo concreto para el concepto abstracto de entero indicará bien las bases intuitivas sobre las que reposan las leyes 1) a 5). En

vez de usar los símbolos habituales 1, 2, 3,..., representemos los enteros que dan el número de objetos de una colección determinada (p. ej., colección de manzanas de un árbol particular) mediante puntos situados en casillas rectangulares, de modo que a cada objeto corresponda un punto. Operando con estos rectángulos se pueden estudiar las leyes de la aritmética de los enteros. Para sumar dos enteros a y b, dispongamos los rectángulos correspondientes uno contiguo al otro, suprimiendo la raya de separación.

FIG. 1.—Adición.

Para multiplicar a por b, coloquemos los puntos de los dos rectángulos en filas y formemos un nuevo rectángulo con a filas y b columnas de puntos. Las reglas 1)-5)

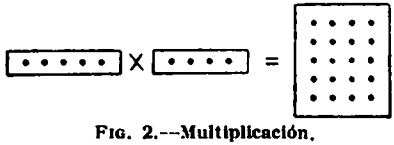

FIG. 2.—Multiplicación.

corresponderán entonces de modo intuitivo a propiedades evidentes de las operaciones indicadas con los rectángulos.

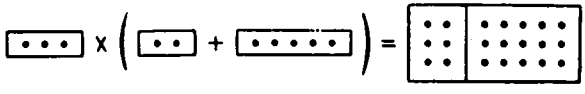

FIG. 3.—La ley distributiva.

A partir de la definición de adición de dos enteros se puede definir la relación de *desigualdad*. Las dos proposiciones equivalentes, $a < b$ (léase «a menor que b») y $b > a$ (léase «b mayor que a»), significan que el rectángulo b puede obtenerse a partir del a mediante la adición de un tercer rectángulo conveniente, c, de modo que se tenga $b = a + c$. Si escribimos esto en la forma

$$c = b - a,$$

definimos la operación llamada *sustracción*.

FIG. 4.—Sustracción.

La adición y la sustracción se llaman *operaciones inversas*, ya que si la adición del entero d al entero a es seguida de la sustracción del entero d de la suma obtenida, el resultado es el entero inicial a:

$$(a + d) - d = a.$$

Debe observarse que el entero $b - a$ ha sido definido únicamente en el caso en que se tenga $b > a$. La interpretación del símbolo $b - a$ como *entero negativo* cuando sea $b < a$ será discutida más adelante (páginas 63 y siguientes).

A menudo es conveniente usar una de las notaciones equivalentes, $b \geqslant a$ (léase «b mayor o igual que a») o $a \leqslant b$ (léase «a menor o igual que b»), para indicar la proposición contraria de la $a > b$; p. ej., escribiremos: $2 \geqslant 2$, y $3 \geqslant 2$.

Haremos ahora una pequeña ampliación del dominio de los enteros positivos, representados por rectángulos con puntos, introduciendo el entero *cero*, que representaremos por un rectángulo vacío. Denotando éste con el habitual símbolo 0, se tiene, de acuerdo con nuestra definición de adición y multiplicación:

$$a + 0 = a,$$
$$a \cdot 0 = 0,$$

cualquiera que sea el entero a. En efecto, $a + 0$ indica la adición de un rectángulo vacío al rectángulo a, mientras que $a \cdot 0$ indica un rectángulo sin columnas; es decir, vacío. Parece entonces natural extender también la definición de sustracción, poniendo

$$a - a = 0$$

para todo entero a. Se tienen así las propiedades aritméticas características del cero.

Modelos geométricos análogos a los rectángulos de puntos, tales como los antiguos ábacos, fueron usados frecuentemente hasta fines de la Edad Media; pero quedaron desplazados poco a poco por otros métodos simbólicos más cómodos, basados en el sistema decimal.

2. Representación de los números enteros.—Ha de distinguirse netamente entre un número entero y los símbolos 5, V, ... usados para representarlo. En el sistema decimal los diez símbolos: 0, 1, 2, 3, ..., 9, se utilizan para representar el cero y los nueve primeros enteros positivos. Un entero mayor, p. ej., el «trescientos setenta y dos», puede expresarse en la forma

$$300 + 70 + 2 = 3 \cdot 10^2 + 7 \cdot 10 + 2,$$

y se representa en el sistema decimal por el símbolo 372. Es de fundamental importancia observar que aquí la significación de los símbolos 3, 7, 2 es *relativa* y depende de su *posición* en el lugar de las unidades, decenas o centenas. Con esta «notación relativa» se puede representar cualquier entero utilizando únicamente los símbolos de los diez primeros números en varias combinaciones. La regla general consiste en expresar todo entero en la forma ilustrada por el ejemplo

$$z = a \cdot 10^3 + b \cdot 10^2 + c \cdot 10 + d,$$

donde los números a, b, c, d son enteros de cero a nueve. El entero z se representa entonces en la forma abreviada

$$abcd.$$

Notemos de paso que los coeficientes d, c, b, a son los restos obtenidos mediante divisiones sucesivas de z por 10; así

$$\begin{array}{r|l} 372 & 10 \\ \cline{2-2} 2 & 37 \quad 10 \\ \cline{2-3} 7 & 3 \quad 10 \\ \cline{2-3} 3 & 0 \end{array}$$

La expresión particular dada antes para z puede servir únicamente para números menores que 10 000, ya que para enteros mayores serán necesarias más de cuatro cifras. Si z es un entero comprendido entre 10 000 y 100 000, lo podremos escribir en la forma

$$z = a \cdot 10^4 + b \cdot 10^3 + c \cdot 10^2 + d \cdot 10 + e,$$

y representarlo por el símbolo *abcde*. Análogas consideraciones valen para enteros comprendidos entre 100 000 y 1 000 000, etc. Será conveniente disponer de un medio de indicar el resultado con toda generalidad mediante una sola fórmula. Para ello designemos los distintos coeficientes: e, d, c,..., con una sola letra afectada de subíndices: a_0, a_1, a_2, a_3, ..., e indiquemos el hecho de que las potencias de 10 pueden ser tan grandes como sea necesario designando la mayor potencia que interviene, no con 10^3 ó 10^4, como en los ejemplos anteriores, sino con 10^n, donde n ha de interpretarse como un entero arbitrario. Entonces el método general para representar un entero z en el sistema decimal consistirá en expresar z en la forma

$$z = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \ldots + a_1 \cdot 10 + a_0, \qquad [1]$$

y representarlo por el símbolo

$$a_n a_{n-1} a_{n-2} \dots a_1 a_0.$$

Como en los casos anteriores, los números a_0, a_1, a_2, ..., a_n, comprendidos entre 0 y 9, serán precisamente los restos sucesivos de dividir z repetidamente por 10.

En el sistema decimal el número 10 ha sido tomado especialmente para servir de base. A primera vista quizá no resalte el hecho de que la elección del 10 no es esencial, y que cualquier entero mayor que uno habría podido servir para el mismo objeto; p. ej., podría usarse un sistema *septimal* (de base 7). En dicho sistema, un entero se expresaría en la forma

$$b_n \cdot 7^n + b_{n-1} \cdot 7^{n-1} + \dots + b_1 \cdot 7 + b_0, \qquad [2]$$

donde las b serían enteros, elegidos entre 0, 1, ... y 6, y se representaría por el símbolo

$$b_n b_{n-1} \dots b_1 b_0 .$$

Así, «ciento nueve» se representaría en el sistema septimal por el símbolo 214, que significaría

$$2 \cdot 7^2 + 1 \cdot 7 + 4.$$

Como ejercicio, el lector puede probar la regla general que permite pasar de la base diez a otra base cualquiera B, y que consiste en efectuar divisiones sucesivas del número z por B; los restos sucesivos serán las cifras del número en el sistema de base B; p. ej.:

$$\begin{array}{r|l} 109 & 7 \\ \cline{2-2} 4 & 15 \mid 7 \\ \cline{2-3} & 1 \mid 2 \mid 7 \\ \cline{3-4} & \quad 2 \mid 0 \end{array}$$

109 (en el sistema decimal) = 214 (en el sistema septimal).

Es natural preguntarse: ¿cuál será la base más conveniente? Puede verse que bases muy pequeñas presentan inconvenientes, mientras que una base grande requiere la utilización de muchas cifras distintas y, en particular, una tabla de multiplicación extensa. La elección de doce como base ha sido defendida por muchos; como ventajas señalan que 12 es divisible por 2, 3, 4 y 6, y, como consecuencia, el cálculo que implique divisiones y fracciones se simplificaría frecuentemente. Para escribir un número en base doce (sistema duodecimal) se necesitan dos nuevas cifras para el 10 y el 11. Indiquemos el 10 con α y el 11 con β.

En el sistema duodecimal, «doce» se escribiría 10; «veintidós» sería 1α; «veintitrés» sería 1β, y «ciento treinta y uno» se escribiría $\alpha\beta$.

La invención de la «numeración relativa», atribuída a los sumerios o babilonios y desarrollada por los indios, fué de enorme trascendencia para la civilización. Los sistemas anteriores de numeración estaban basados en un estricto principio aditivo. En el simbolismo romano, p. ej., se escribía

$$\text{CXVIII} = 100 + 10 + 5 + 1 + 1 + 1.$$

Los sistemas egipcio, hebraico y griego eran de un tipo parecido al de los romanos. Un inconveniente de la notación aditiva es que cuanto mayor es el número mayor es también el conjunto de nuevos símbolos necesarios para representarlo. (Claro está que los científicos antiguos no utilizaban las modernas magnitudes astronómicas o atómicas.) Pero el defecto fundamental de los antiguos sistemas, tales como el romano, residía en el hecho de que el cálculo era tan complicado que únicamente los especialistas podían manejar los problemas de cálculo no triviales. Las cosas pasan de modo distinto con el sistema indio de «valor relativo», hoy en uso. Este sistema fué introducido en Europa, en la Edad Media, por comerciantes italianos, quienes lo habían aprendido de los árabes. Tiene la propiedad cómoda de que todos los números, grandes o pequeños, pueden representarse mediante un pequeño conjunto de cifras diferentes (en el sistema decimal, mediante las «cifras árabes» 0, 1, 2, ..., 9). Esto lleva consigo la importante ventaja de la facilidad de los cálculos. Las reglas de cálculo en los sistemas de notación basados en el valor relativo pueden establecerse en forma de tablas de adición y multiplicación para números de una sola cifra y pueden ser aprendidas de memoria y retenidas para siempre. Existen pocos ejemplos de adelantos científicos que hayan afectado tan profundamente y facilitado tanto la vida diaria como el actual sistema de numeración.

3. El cálculo numérico en sistemas distintos del decimal.—El uso del 10 como base de numeración se remonta a los primeros tiempos de la civilización, y es debido indudablemente al hecho de que son diez los dedos con los que se acostumbra contar. Sin embargo, las palabras que sirven para designar algunos números en distintos idiomas parecen reminiscencias del uso de otras bases, especialmente 12 y 20. En inglés y en alemán las palabras para 11 y 12 no están construídas según el principio decimal de combinar 10 con los números de una cifra, como ocurre para los comprendidos entre 13 y 19, sino que son lingüísticamente independientes de la palabra que corresponde a 10. En francés las palabras *vingt* y *quatre-vingt*, para 20 y 80, sugieren la idea

de que para algunas cuestiones se haya usado un sistema de base 20. En danés, la palabra *halvfirsindstyve*, para 70, significa *a medio camino* (a partir de tres veces) de cuatro veces veinte. Los astrónomos de Babilonia utilizaban un sistema que en parte era sexagesimal (base 60); la creencia en este hecho se apoya en la habitual división de la hora y del grado angular en 60 min.

En un sistema distinto del decimal, las reglas de la aritmética son las mismas que en éste; sin embargo, han de usarse otras tablas para la adición y la multiplicación de los números de una cifra. Acostumbrados al sistema decimal y ligados a él por gran número de palabras de nuestro lenguaje, surge al comienzo alguna confusión. Ensayemos un ejemplo de multiplicación en el sistema septimal; antes será conveniente escribir las tablas de sumar y multiplicar que usaremos en la operación:

Adición

	1	2	3	4	5	6
1	2	3	4	5	6	10
2	3	4	5	6	10	11
3	4	5	6	10	11	12
4	5	6	10	11	12	13
5	6	10	11	12	13	14
6	10	11	12	13	14	15

Multiplicación

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	11	13	15
3	3	6	12	15	21	24
4	4	11	15	22	26	33
5	5	13	21	26	34	42
6	6	15	24	33	42	51

Sea multiplicar 265 por 24, símbolos que representan números en el sistema septimal. (En el sistema decimal sería equivalente a multiplicar 145 por 18.) Las reglas de la multiplicación son las mismas que en el sistema decimal. Comenzamos multiplicando 5 por 4, que da 26, como resulta de la tabla de multiplicar.

$$\begin{array}{r} 265 \\ 24 \\ \hline 1456 \\ 563 \\ \hline 10416 \end{array}$$

Escribimos 6 en el lugar de las unidades y llevamos 2 al lugar de las septenas. Luego buscamos $4 \cdot 6 = 33$, y $33 + 2 = 35$. Escribimos 5, y procedemos de igual forma hasta que hayamos multiplicado todos los números. Sumando $1456 + 5630$, se obtiene $6 + 0 = 6$ en el lugar de las unidades, $5 + 3 = 11$ en el lugar de las septenas. De nuevo escribimos 1 y conservamos 1 para el lugar de los cuarenta y nueves, con lo que tenemos $1 + 6 + 4 = 14$. El resultado final es: $265 \cdot 24 = 10416$.

Para comprobar este resultado, podemos multiplicar los mismos números en el sistema decimal. 10416 (en el sistema septimal) puede escribirse en el sistema decimal calculando las potencias de 7 hasta la cuarta: $7^3 = 49$, $7^3 = 343$, $7^4 = 2401$. De donde $10416 = 2401 + 4 \cdot 49 + 7 + 6$; este cálculo está hecho en el sistema decimal. Sumando estos números, resulta que 10416 en el sistema septimal es igual a 2610 en el sistema decimal. Luego, multipliquemos 145 por 18 en el sistema decimal; el resultado es 2610, lo que confirma el cálculo anterior.

Ejercicios:

1. Constrúyanse las tablas de adición y multiplicación en el sistema duodecimal y háganse algunos ejemplos en este sistema.
2. Exprésense «30» y «136» en los sistemas de bases, 5, 7, 11 y 12.
3. ¿Qué representan los símbolos 11111 y 21212 en dichos sistemas?
4. Fórmense las tablas de adición y multiplicación en los sistemas de bases 5, 11 y 13.

El sistema de valor relativo de base 2 está caracterizado por ser el de base más pequeña. Las únicas cifras en este *sistema diádico* son 0 y 1; cualquier otro número z vendrá representado por una sucesión de estos dos símbolos. Las tablas de adición y multiplicación se reducen a las reglas $1 + 1 = 10$ y $1 \cdot 1 = 1$. El inconveniente de este sistema es evidente; hacen falta expresiones muy largas para representar números pequeños. Así, 79, que puede expresarse en la forma $1 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2 + 1$, se escribe en el sistema diádico 1001111.

Como ilustración de la sencillez de la multiplicación en el sistema diádico haremos la multiplicación de 7 por 5, números que se representan respectivamente por 111 y 101. Recordando que en este sistema se tiene $1 + 1 = 10$, resulta:

$$\begin{array}{r} 111 \\ 101 \\ \hline 111 \\ 111 \\ \hline 100011 \end{array} = 2^5 + 2 + 1,$$

es decir, 35, como debía ser.

Gottfried Wilhelm Leibniz (1646-1716), una de las mayores inteligencias de su tiempo, fué un apasionado del sistema diádico. A propósito de esto, Laplace escribe: «Leibniz veía en el sistema diádico la imagen de la creación. Consideraba que la unidad representaba a Dios, y el cero, la nada; que el Ser Supremo creaba todos los seres de la nada, del mismo modo que la unidad y el cero expresaban todos los números de su sistema de numeración».

Ejercicio: Considérese la cuestión de representar los enteros en la base a. Para nombrar los enteros en este sistema se necesitan palabras para los números 0, 1, ..., $a - 1$ y para las sucesivas potencias de a: a, a^2, a^3, ... ¿Cuántas palabras distintas son necesarias para designar todos los números de 0 a 1000, para $a = 2$, 3, 4, 5, ..., 15? ¿En qué base son necesarias menos palabras? (Ejemplos: Si $a = 10$, se necesitan diez palabras para los números de una cifra, además de las palabras para 10, 100 y 1000, lo que hace 13 en total. Para $a = 20$, se necesitan veinte para los números de una cifra, más las correspondientes a 20 y 400; en total, 22. Si es $a = 100$, se necesitan 101.)

*II. LA INFINITUD DEL SISTEMA DE NÚMEROS ENTEROS. INDUCCIÓN MATEMÁTICA

1. El principio de inducción matemática.—La sucesión de enteros 1, 2, 3, 4, ..., no tiene fin, puesto que después de cada entero n hay uno siguiente: el $n + 1$. Expresaremos esta propiedad diciendo que la sucesión de enteros contiene *infinitos* enteros. La sucesión de enteros constituye el ejemplo más sencillo y natural del infinito matemático, el cual desempeña un papel dominante en la matemática. A lo largo de este libro necesitaremos manejar colecciones o *conjuntos* que contienen una infinidad de objetos matemáticos; p. ej., el conjunto de todos los puntos de una recta o el conjunto de todos los triángulos de un plano. La sucesión de enteros es el ejemplo más simple de conjunto infinito.

El proceso de ir paso a paso, de n a $n + 1$, que engendra la sucesión infinita de los enteros, forma también la base de uno de los tipos fundamentales de razonamiento matemático: el principio de inducción matemática. La *inducción empírica* de las ciencias naturales procede de una serie particular de observaciones de un cierto fenómeno para establecer una proporción o ley general que debe regir todas las posibilidades del fenómeno. El grado de certeza con que se establece dicha ley depende del número de observaciones particulares y de confirmaciones del fenómeno. Este tipo de razonamiento inductivo es con frecuencia plenamente convincente; la predicción de que mañana el Sol hará su salida por Oriente tiene toda la certeza posible; pero el carácter de esta proposición no es el mismo que el de un teorema probado con razonamientos estrictamente lógicos o matemáticos.

De modo completamente distinto se utiliza la *inducción matemática* para establecer la certeza de un teorema matemático en una sucesión infinita de casos: el primero, el segundo, el tercero, y así sucesivamente, sin excepción. Designemos con A una proposición que se refiera a un entero arbitrario n; p. ej., A puede ser la proposición: «La suma de los ángulos de un polígono convexo de $n + 2$ lados es n veces 180°.» A' puede consistir en la afirmación: «Trazando n rectas en

un plano no se puede dividir éste en más de 2^n partes.» Para probar uno de estos teoremas para *cualquier* entero n no es suficiente probarlo para los 10 ó 100, ni aun para los 1000 primeros valores de n. Este modo de proceder correspondería precisamente a la inducción empírica. En su lugar, debemos usar un método estrictamente matemático y no un razonamiento empírico, cuyo carácter indicaremos en lo que sigue, al probar los ejemplos especiales A y A'. En el caso A sabemos que para $n = 1$ el polígono es un triángulo, y por geometría elemental se sabe que la suma de sus ángulos es $1 \cdot 180^o$. Para un cuadrilátero, $n = 2$, se traza una diagonal que dividirá al cuadrilátero en dos triángulos. Entonces se ve, de modo inmediato, que la suma de los ángulos del cuadrilátero es igual a la suma de los ángulos de los dos triángulos, lo que da $180^o + 180^o = 2 \cdot 180^o$. Procediendo análogamente para el pentágono, $n = 3$, se descompone en un triángulo y un cuadrilátero. Puesto que la suma de los ángulos del último es $2 \cdot 180^o$, como acabamos de probar, y siendo la suma de los ángulos del triángulo 180^o, obtenemos para el pentágono $3 \cdot 180^o$. Ahora bien: resulta claro que podemos proceder indefinidamente en la misma forma, probando el teorema para $n = 4$; luego, para $n = 5$, y así sucesivamente. Cada proposición se deduce de la precedente en la misma forma, de modo que el teorema general puede ser establecido para todo n.

Análogamente podemos probar A'; para $n = 1$ es evidente, ya que una recta divide al plano en dos partes. Añadamos una segunda recta. Cada una de las partes anteriores quedará dividida en dos nuevas partes, salvo que la nueva recta sea paralela a la primera. En ambos casos, para $n = 2$ no resultan más de $4 = 2^2$ partes. Añadamos una tercera recta; cualquiera de las regiones anteriores quedará, o bien dividida en dos partes o sin dividir. Así, la suma de partes no podrá ser mayor que $2 \cdot 2^2 = 2^3$. Sabiendo que esto es cierto, podemos probar el caso siguiente en la misma forma, y así indefinidamente.

La idea esencial de los argumentos precedentes consiste en establecer un teorema general A para todos los valores de n, probando sucesivamente una sucesión de casos especiales A_1, A_2, ... La posibilidad de hacerlo depende de dos hechos: *a)* Existe un método general para probar que si cualquier proposición A_r es cierta, la proposición siguiente, A_{r+1}, será *también* cierta. *b)* Se *sabe* que la primera proposición A es cierta. Que estas dos condiciones son suficientes para establecer la verdad de *todas* las proposiciones A_1, A_2, A_3, ..., es un principio lógico que resulta tan fundamental para las matemáticas como lo son las reglas clásicas de la lógica aristotélica. Por ello vamos a enunciarlo explícitamente como sigue:

Supongamos que queremos establecer una sucesión infinita de proposiciones matemáticas

$$A_1, A_2, A_3, \ldots$$

que juntas constituyen una proposición general A. *Supongamos*: a) *que por un razonamiento matemático se prueba que, si* r *es un entero cualquiera, de la verdad de la proposición* A_r *se sigue la verdad de la* A_{r+1}, *y* b), *se sabe que la proposición* A_1 *es cierta. Entonces todas las proposiciones de la sucesión son ciertas y queda probada* A.

No debe haber duda en aceptar esto, del mismo modo que no la tenemos para aceptar las reglas elementales de la lógica ordinaria, como un principio del razonamiento matemático, ya que se puede establecer la verdad de cualquiera de las proposiciones A_n partiendo de la aserción *b)* de que A_1 es cierta, y procediendo por uso repetido de la aserción *a*), para establecer sucesivamente la verdad de A_2, A_3, A_4, y así hasta llegar a la proposición A_n. El principio de inducción matemática se basa en el hecho de que después de cada entero r hay un siguiente $r + 1$, y que todo entero n puede ser alcanzado mediante un número finito de pasos, a partir del 1.

Con frecuencia, el principio de inducción matemática se aplica sin mencionarlo explícitamente, o viene indicado simplemente por un «etc.» o un «y así sucesivamente». Así sucede, en particular, en la enseñanza elemental. Pero el uso explícito del razonamiento inductivo es indispensable en demostraciones más sutiles. Damos a continuación algunos ejemplos de carácter sencillo, pero no trivial.

2. Progresiones aritméticas.—*Para todo valor de* n, *la suma* $1 + 2 + 3 \ldots + n$ *de los* n *primeros enteros es igual a* $\dfrac{n(n+1)}{2}$. Para probar este teorema por inducción matemática debemos demostrar que para cualquier n la proposición A_n:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2} \qquad [1]$$

es cierta. *a)* Observemos que si r es un entero y si se sabe que la proposición A_r es cierta; es decir, si sabemos que

$$1 + 2 + 3 + \ldots + r = \frac{r(r+1)}{2},$$

sumándole el número $(r + 1)$ a los dos miembros de esta igualdad obtenemos la ecuación

$$1 + 2 + 3 + \ldots + r + (r+1) = \frac{r(r+1)}{2} + (r+1) = \frac{(r+1)(r+2)}{2}$$

que es precisamente la proposición A_{r+1}. *b)* La proposición A_1 es evidente, ya que $1 = \frac{1 \cdot 2}{2}$. De donde, por el principio de inducción matemática, la proposición A_n es cierta para todo n, como se quería demostrar.

Corrientemente se suele probar escribiendo la suma $1 + 2 + 3 + \ldots + n$ de dos maneras:

$$S_n = 1 + 2 + \ldots + (n-1) + n$$

y

$$S_n = n + (n-1) + \ldots + 2 + 1.$$

Al sumar, se observa que cada par de números de la misma columna da como suma $n + 1$, y puesto que hay n columnas en total, se sigue

$$2S_n = n(n+1),$$

lo que prueba el resultado indicado.

De [1] se puede deducir de modo inmediato la fórmula de la suma de los $(n + 1)$ primeros términos de cualquier *progresión aritmética*,

$$P_n = a + (a+d) + (a+2d) + \ldots + (a+nd) = \frac{(n+1)(2a+nd)}{2} \qquad [2]$$

Puesto que

$$P_n = (n+1)a + (1 + 2 + \ldots + n)d = (n+1)a + \frac{n(n+1)d}{2} =$$

$$= \frac{2(n+1)a + n(n+1)d}{2} = \frac{(n+1)(2a+nd)}{2}$$

Para el caso $a = 0$ y $d = 1$, esta fórmula es la misma [1]:

3. Progresiones geométricas.—Se pueden estudiar las progresiones geométricas de modo análogo al precedente. Probaremos que para todo valor de n se tiene

$$G_n = a + aq + aq^2 + \ldots + aq^n = a\frac{1 - q^{n+1}}{1 - q} \qquad [3]$$

(Suponemos $q \neq 1$, ya que de otro modo el último miembro de [3] no tendría significado.)

La proposición es cierta para $n = 1$, ya que entonces

$$G_1 = a + aq = \frac{a(1 - q^2)}{1 - q} = \frac{a(1+q)(1-q)}{(1-q)} = a(1+q).$$

Y *si* suponemos que se tiene

$$G_r = a + aq + \ldots + aq^r = a\frac{1 - q^{r+1}}{1 - q},$$

resulta como consecuencia

$$G_{r+1} = (a + aq + \ldots + aq^r) + aq^{r+1} = G_r + aq^{r+1} = a\frac{1 - q^{r+1}}{1 - q} + aq^{r+1} =$$

$$= a\frac{(1 - q^{r+1}) + q^{r+1}(1 - q)}{1 - q} = a\frac{1 - q^{r+1} + q^{r+1} - q^{r+2}}{1 - q} = a\frac{1 - q^{r+2}}{1 - q}$$

Pero esto es precisamente la proposición [3] para el caso $n = r + 1$, lo que completa la demostración.

En los libros elementales, la prueba habitual procede como sigue. Pongamos

$$G_n = a + aq + \ldots + aq^n,$$

y multipliquemos los dos miembros por q. Se obtiene

$$qG_n = aq + aq^2 + \ldots + aq^{n+1}.$$

Restando entonces los miembros correspondientes de las dos igualdades, resulta

$$G_n - qG_n = a - aq^{n+1},$$
$$(1 - q)G_n = a(1 - q^{n+1}),$$
$$G_n = a\frac{1 - q^{n+1}}{1 - q}$$

4. Suma de los «n» primeros cuadrados.—Otra interesante aplicación del principio de inducción se refiere a la suma de los n primeros cuadrados. Mediante ensayos directos se encuentra, al menos para valores pequeños de n,

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n + 1)(2n + 1)}{6} \qquad [4]$$

y se *supone* que esta fórmula pueda ser válida para *todos los enteros n*. Para *probarlo*, haremos uso de nuevo del principio de inducción. Comenzaremos por observar que si la proposición A_n, que en este caso coincide con la ecuación [4], es válida para el caso $n = r$, de modo que se tenga

$$1^2 + 2^2 + 3^2 + \ldots + r^2 = \frac{r(r + 1)(2r + 1)}{6},$$

sumando $(r + 1)^2$ a los dos miembros de esta igualdad se obtiene

$$1^2 + 2^2 + 3^2 + \ldots + r^2 + (r + 1)^2 = \frac{r(r + 1)(2r + 1)}{6} + (r + 1)^2 =$$

$$= \frac{r(r + 1)(2r + 1) + 6(r + 1)^2}{6} = \frac{(r + 1)[r(2r + 1) + 6(r + 1)]}{6} =$$

$$= \frac{(r + 1)(2r^2 + 7r + 6)}{6} = \frac{(r + 1)(r + 2)(2r + 3)}{6},$$

que es precisamente la proposición A_{r+1} en el caso presente, puesto que se obtiene sustituyendo n por $r + 1$ en [4]. Para completar la demostración necesitamos únicamente observar que la proposición A_1, en este caso la ecuación

$$1^2 = \frac{1(1 + 1)(2 + 1)}{6},$$

es evidente. Por tanto, la ecuación [4] es válida para cualquier n.

Fórmulas análogas pueden hallarse para potencias superiores de los enteros, $1^k + 2^k + 3^k + \ldots + n^k$, donde k es un entero positivo cualquiera. Como ejercicio, el lector puede probar, mediante inducción matemática, que

$$1^3 + 2^3 + 3^3 + \ldots + n^3 = \left[\frac{n(n + 1)}{2}\right]^2 \qquad [5]$$

Debe observarse que, si bien el principio de inducción matemática es suficiente para *probar* la fórmula [5] una vez que se conoce ésta, la demostración no da indicación alguna sobre el modo en que dicha fórmula puede encontrarse; es decir, sobre el por qué debe suponerse la expresión $[n(n + 1)/2]^2$ como resultado para la suma de los n primeros cubos, en vez de la $[n(n + 1)/3]^2$ o $(19n^2 - 41n + 24)/2$ o cualquiera de las infinitas expresiones análogas que pudieran ser consideradas. El hecho de que la demostración de un teorema consista en la aplicación de ciertas reglas sencillas de lógica no disminuye el valor del elemento creador en matemáticas, el cual desempeña su papel en la elección de las posibilidades que deban ser tenidas en cuenta. La cuestión del origen de la *hipótesis* [5] pertenece a un dominio en el cual no pueden ser dadas reglas generales; ensayos, analogías e intuición constructiva tienen en esto papel importante. Pero una vez formulada correctamente la hipótesis, el principio de inducción matemática es con frecuencia suficiente para dar la demostración. En tanto que una demostración no proporcione una indicación para el acto del descubrimiento, debe llamarse más propiamente una *comprobación*.

***5. Una desigualdad importante.**—En otro capítulo haremos uso de la desigualdad

$$(1 + p)^n \geq 1 + np, \qquad [6]$$

que es válida para todo número $p > -1$ y todo entero positivo n. (Con objeto de dar mayor generalidad, anticipamos aquí el uso de números negativos no enteros admitiendo como p cualquier número mayor que -1. La demostración en el caso general es exactamente

la misma que en el caso en que p es un entero positivo.) Usaremos de nuevo la inducción matemática.

a) Si es cierto que $(1 + p)^r \geqslant 1 + rp$, multiplicando los dos miembros de esta desigualdad por el número positivo $1 + p$ se obtiene

$$(1 + p)^{r+1} > 1 + rp + p + rp^2.$$

Si prescindimos del término positivo rp^2, se tendrá

$$(1 + p)^{r+1} > 1 + (r + 1)p,$$

lo que prueba que la desigualdad [6] vale también para el entero siguiente $r + 1$. *b)* Evidentemente se tiene: $(1 + p)^1 \geqslant 1 + p$. Esto completa la demostración de que [6] es cierta para todo n. La restricción de que $p > -1$ es esencial. Si es $p < -1$, $1 + p$ es negativo y el razonamiento empleado para *a)* deja de ser válido, puesto que si se multiplican los dos miembros de una desigualdad por un número negativo, la desigualdad cambia de sentido. (P. ej., si multiplicamos los dos miembros de la desigualdad $3 > 2$ por -1, obtendríamos $-3 < -2$.)

***6. El binomio de Newton.**—Con frecuencia es importante tener una expresión explícita para la n-ésima potencia de un binomio, $(a + b)^n$. Mediante cálculo explícito se obtiene:

$$\text{para } n = 1, \quad (a + b)^1 = a + b.$$

$$\text{para } n = 2, \quad (a + b)^2 = (a + b)(a + b) = a(a + b) + b(a + b) = a^2 + 2ab + b^2,$$

$$\text{para } n = 3, \quad (a + b)^3 = (a + b)(a + b)^2 = a(a^2 + 2ab + b^2) + b(a^2 + 2ab + b^2) = a^3 + 3a^2b + 3ab^2 + b^3,$$

y así sucesivamente. ¿Cuál es la ley general de formación, implícita en la frase «y así sucesivamente»? Examinemos el proceso mediante el cual calculamos $(a + b)^2$. Puesto que es $(a + b)^2 = (a + b)(a + b)$, se obtiene el desarrollo de $(a + b)^2$ multiplicando primero cada término de la expresión $a + b$ por a y luego por b y sumando después. El mismo procedimiento se utilizó para calcular $(a + b)^3 = (a + b)(a + b)^2$. Se puede continuar del mismo modo para calcular $(a + b)^4$, $(a + b)^5$, y así indefinidamente. La expresión de $(a + b)^n$ se obtendría multiplicando cada uno de los términos de la expresión de $(a + b)^{n-1}$, calculada previamente, primero por a y luego por b, y sumando. Esto conduce al diagrama

$$\begin{aligned}
a + b &= a + b \\
(a+b)^2 &= a^2 + 2ab + b^2 \\
(a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\
(a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots ,
\end{aligned}$$

el cual da inmediatamente la regla general para formar los coeficientes del desarrollo de $(a + b)^n$. Para ello construímos un esquema triangular de números, partiendo de los coeficientes 1, 1 de $a + b$, y de tal modo que cada número del triángulo es la suma de los dos números inmediatos a él en la fila precedente. Esta disposición de los números es conocida con el nombre de *triángulo de Pascal.*

$$\begin{array}{ccccccccccccccc}
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots
\end{array}$$

La fila n*-ésima de este esquema da los coeficientes del desarrollo de* $(a + b)^n$ *según las potencias decrecientes de* a *y las crecientes de* b; así

$$(a + b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7.$$

Usando una notación concisa, mediante índices y subíndices, se pueden representar los números de la n-ésima fila del triángulo de Pascal por

$$C_0^n = 1,\ C_1^n,\ C_2^n,\ C_3^n,\ \ldots,\ C_{n-1}^n,\ C_n^n = 1.$$

Entonces, la fórmula general para $(a + b)^n$ puede escribirse

$$(a + b)^n = a^n + C_1^n a^{n-1}b + C_2^n a^{n-2}b^2 + \ldots + C_{n-1}^n ab^{n-1} + b^n. \qquad [7]$$

De acuerdo con la ley de formación del triángulo de Pascal, se tiene

$$C_i^n = C_{i-1}^{n-1} + C_i^{n-1}. \qquad [8]$$

Como ejercició, el lector puede utilizar esta relación, junto con el

hecho de que $C_0^1 = C_1^1 = 1$, para probar, mediante inducción matemática, que

$$C_i^n = \frac{n(n-1)(n-2)\dots(n-i+1)}{1 \cdot 2 \cdot 3 \dots i} = \frac{n!}{i!(n-i)!} \qquad [9]$$

(Para todo entero positivo n, el símbolo $n!$—léase «factorial de n»—designa el producto de los n primeros enteros: $n! = 1 \cdot 2 \cdot 3 \cdots n$. Y resulta conveniente definir también $0!$ por la igualdad $0! = 1$; de este modo [9] es válido para $i = 0$ e $i = n$.) Esta fórmula explícita para los coeficientes del desarrollo binómico se conoce con el nombre de *teorema del binomio*. (Véase también Cap. VIII, Suplemento, III, 1.)

Ejercicios: Demuéstrese por inducción matemática:

1. $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \dots + \dfrac{1}{n(n+1)} = \dfrac{n}{n+1}$

2. $\dfrac{1}{2} + \dfrac{2}{2^2} + \dfrac{3}{2^3} + \dots + \dfrac{n}{2^n} = 2 - \dfrac{n+2}{2^n}$

*3. $1 + 2q + 3q^2 + \dots + nq^{n-1} = \dfrac{1 - (n+1)q^n + nq^{n+1}}{(1-q)^2}$

*4. $(1+q)(1+q^2)(1+q^4)\dots(1+q^{2^n}) = \dfrac{1 - q^{2^{n+1}}}{1-q}$

Hállese la suma de las siguientes progresiones geométricas:

5. $\dfrac{1}{1+x^2} + \dfrac{1}{(1+x^2)^2} + \dots + \dfrac{1}{(1+x^2)^n}$

6. $1 + \dfrac{x}{1+x^2} + \dfrac{x^2}{(1+x^2)^2} + \dots + \dfrac{x^n}{(1+x^2)^n}$

7. $\dfrac{x^2-y^2}{x^2+y^2} + \left(\dfrac{x^2-y^2}{x^2+y^2}\right)^2 + \dots + \left(\dfrac{x^2-y^2}{x^2+y^2}\right)^n$

Utilizando las fórmulas [4] y [5], pruébese:

*8. $1^2 + 3^2 + \dots + (2n+1)^2 = \dfrac{(n+1)(2n+1)(2n+3)}{3}$

*9. $1^3 + 3^3 + \dots + (2n+1)^3 = (n+1)^2(2n^2+4n+1)$.

10. Demuéstrense los mismos resultados directamente por inducción matemática.

***7. Algunas observaciones a propósito de la inducción matemática.**—El principio de inducción matemática puede modificarse un poco y darle la siguiente forma:

«Si una sucesión de proposiciones A_s, A_{s+1}, A_{s+2}, ..., donde s es un entero positivo, es tal que:

a) para todo $r \geqslant s$, de la verdad de A_r se sigue la de A_{r+1}, y

b) se sabe que A_s es cierta,

entonces todas las proposiciones A_s, A_{s+1}, A_{s+2}, ... son ciertas; es decir, A_n es cierta para todo $n \geqslant s$.» Se aplica aquí el mismo razonamiento utilizado para establecer la validez del principio ordinario de inducción matemática, con la única variante de sustituir la sucesión 1, 2, 3, ... por la sucesión análoga, s, $s + 1$, $s + 2$, ... Usando el principio en esta forma, se puede precisar más la desigualdad de la página 22, eliminando la posibilidad del signo «=». Se puede probar que: *Para todo* $p \neq 0$ y > -1 *y para todo entero* $n \geqslant 2$,

$$(1 + p)^n > 1 + np. \qquad [10]$$

Dejamos la demostración al cuidado del lector.

Íntimamente ligado con el principio de inducción matemática está el llamado «principio del menor entero», el cual establece *que todo conjunto* C, *no vacío, de números enteros positivos contiene un entero menor que todos los demás.* Un conjunto vacío es aquel que no contiene elementos; p. ej., el conjunto de las circunferencias rectilíneas o el conjunto de los enteros n tales que $n > n$. Por razones obvias excluímos tales conjuntos de nuestro principio. El conjunto C puede ser finito, tal como el conjunto 1, 2, 3, 4, 5, ó infinito, como el conjunto de todos los números pares 2, 4, 6, 8, 10, ... Cualquier conjunto C, no vacío, debe contener cuando menos un entero, p. ej., el n, y entonces el más pequeño de los enteros 1, 2, 3, ..., n, que pertenezca a C será el menor de los enteros contenidos en C.

Para comprender bien el significado de este principio se debe observar que deja de ser cierto si se aplica a cualquier conjunto C de números que no sean enteros; p. ej., el conjunto de fracciones positivas 1, $^1/_2$, $^1/_3$, ... no contiene una fracción menor que todas las demás.

Desde el punto de vista de la lógica es interesante observar que el principio del menor entero puede ser utilizado para *demostrar* el principio de inducción matemática como un teorema. Con tal objeto, consideremos cualquier sucesión de proposiciones A_1, A_2, A_3, ... tales que:

a) Para todo entero positivo r, la verdad de A_{r+1} se sigue de la de A_r.

b) Se sabe que A_1 es cierta.

Probaremos que la hipótesis de que cualquier A sea falsa es absurda. Pues si alguna de las A fuese falsa, el conjunto C de todos los enteros positivos n para los cuales fuera falsa A_n sería un conjunto no vacío. Por el principio del menor entero, C contendría un entero p más pequeño que todos los demás, y p debería ser > 1 a causa de *b)*. Por tanto, se tendría que A_p sería falsa mientras que A_{p-1} sería cierta; pero esto contradice *a)*.

Una vez más insistimos en que el principio de inducción matemática es completamente diferente de la inducción empírica de las ciencias naturales. La confirmación de una ley general en cualquier número finito de casos, por grande que sea dicho número, no suministra una demostración de la ley en sentido matemático riguroso, y esto aunque no sea conocida excepción alguna. Tal ley quedará únicamente como una *hipótesis* plausible, sujeta siempre a modificaciones por los

resultados de ulteriores experiencias. En matemáticas, una ley o un teorema quedan probados únicamente cuando se demuestra que son una consecuencia lógicamente necesaria de ciertos supuestos admitidos como válidos. Existen varios ejemplos de proposiciones matemáticas que han sido comprobadas en todos los numerosos casos particulares considerados, pero que hasta la fecha no han sido demostradas en general (para un ejemplo, véase pág. 38). Se puede *sospechar* que un teorema es cierto en general si resulta cierto en un número de ejemplos; entonces cabe intentar *probarlo* mediante la inducción matemática. Si el intento tiene éxito, el teorema queda demostrado; si se fracasa, el teorema puede ser cierto o falso y algún día podrá, posiblemente por otros métodos, ser probado o rechazado.

Al usar el principio de inducción matemática se debe estar seguro de que las condiciones *a)* y *b)* están efectivamente satisfechas. Descuidar esta precaución puede conducir a absurdos como el que sigue (invitamos al lector para que descubra el error en el razonamiento). «Probaremos» que *dos enteros positivos cualesquiera son iguales;* p. ej., que $5 = 10$.

Comenzamos con una definición: Si a y b son dos enteros positivos desiguales, definimos como máx. (a, b) aquel de los dos que sea mayor: si es $a = b$ pondremos máx. $(a, b) = a = b$. Así máx. $(3, 5) =$ máx. $(5, 3) = 5$, mientras que máx. $(4, 4) = 4$. Sea ahora A_n la proposición: «Si a y b son dos enteros positivos cualesquiera, tales que máx. $(a, b) = n$, se tiene $a = b$.»

a) Supongamos que A_r es cierta. Sean a y b dos enteros positivos cualesquiera, tales que máx. $(a, b) = r + 1$. Consideremos los dos enteros

$$\alpha = a - 1$$

$$\beta = b - 1;$$

se tendrá entonces: máx. $(\alpha, \beta) = r$. De aquí resulta $\alpha = \beta$, puesto que suponíamos que A_r era cierta. De donde se sigue $a = b$; en consecuencia, A_{r+1} es cierta.

b) A_1 es evidentemente cierta, pues si es máx. $(a, b) = 1$, se tendrá, ya que suponemos a y b enteros positivos, que estos dos números deben ser iguales a 1. Por consiguiente, en virtud de la inducción matemática, A_n es cierta para todo n.

Ahora, si a y b son dos enteros positivos cualesquiera, designemos máx. (a, b) por r. Puesto que hemos probado que A_n es cierta para todo n, se tendrá en particular que A_r es cierta. En consecuencia $a = b$.

SUPLEMENTO AL CAPÍTULO PRIMERO

TEORÍA DE NÚMEROS

Introducción.—Los enteros fueron poco a poco perdiendo su relación con supersticiones y misticismos, pero su interés para los matemáticos no disminuyó nunca. Euclides (hacia el año 300 a. de J.C.), cuya fama está ligada a la parte de sus *Elementos* que forma la base de la geometría elemental, parece haber obtenido resultados originales en la teoría de números, mientras que su geometría es, en su mayor parte, una compilación de resultados anteriores. Diofanto de Alejandría (hacia el año 275 d. de J.C.), uno de los primeros algebristas, dejó huella importante en la teoría de números. Pierre de Fermat (1601-1665), jurista en Toulouse y uno de los más grandes matemáticos de su tiempo, inició las investigaciones modernas en este campo. Euler (1707-1783), el más prolífico de los matemáticos, obtuvo gran cantidad de resultados en la teoría de números entre sus investigaciones matemáticas. Otros nombres preeminentes en los anales de la matemática—Legendre, Dirichlet, Riemann—deben añadirse a la lista anterior. Gauss (1777-1855), el más célebre matemático de los tiempos modernos, dedicado asimismo a distintas ramas de las matemáticas, resumió su opinión sobre la teoría de números en la frase: «La matemática es la reina de las ciencias, y la teoría de números es la reina de las matemáticas.»

I. LOS NÚMEROS PRIMOS

1. Hechos fundamentales.—La mayor parte de las proposiciones de la teoría de números, como en general de la matemática, no se refieren a un objeto particular—el número 5 ó el número 32—, sino a una clase completa de objetos que tienen alguna propiedad común; p. ej., la clase de todos los números pares:

$$2, 4, 6, 8, \ldots,$$

o la clase de los enteros divisibles por 3:

$$3, 6, 9, 12, \ldots,$$

o bien la clase de los cuadrados de los números enteros:

$$1, 4, 9, 16, \ldots,$$

y así sucesivamente.

De fundamental importancia en teoría de números es la clase de los números *primos*. Casi todos los enteros se pueden descomponer en producto de factores más pequeños: $10 = 2 \cdot 5$, $111 = 3 \cdot 37$, $144 = 3 \cdot 3 \cdot 2 \cdot 2 \cdot 2 \cdot 2$, etc. Los números que no pueden descomponerse de este modo se llaman números primos o simplemente primos. Con más precisión, *un número primo es un entero* p, *mayor que uno, que no admite más factores que él mismo y la unidad.* (Se dice que un *entero a* es *factor* o *divisor* de otro *entero b* si existe otro *entero c* tal que $b = ac$.) Los números 2, 3, 5, 7, 11, 13, 17, ... son primos, mientras que 12, p. ej., no es primo, ya que se tiene $12 = 3 \cdot 4$. La importancia de la clase de los números primos se debe al hecho de que cualquier entero puede expresarse como *producto de números primos*: si un número no es primo, se puede descomponer sucesivamente hasta que todos sus factores sean primos; p. ej., $360 = 3 \cdot 120 = 3 \cdot 30 \cdot 4 = 3 \cdot 3 \cdot 10 \cdot 2 \cdot 2 = 3 \cdot 3 \cdot 5 \cdot 2 \cdot 2 \cdot 2 = 2^3 \cdot 3^2 \cdot 5$. Un entero (distinto de 0 y 1) que no sea primo se llama *compuesto*.

Una de las primeras cuestiones que se presentan en la teoría de los números primos es la de saber si hay solamente un número finito de estos números o si, por el contrario, la clase de los números primos tiene infinitos elementos distintos, como ocurre con la clase de los números enteros de la cual forma parte. La respuesta es: *Existen infinitos números primos.*

La prueba de la infinitud de la clase de los números primos dada por Euclides ha quedado como modelo de razonamiento matemático. Procede por «reducción al absurdo», y parte de la hipótesis de que el teorema es falso. Esto significa que existiría únicamente un número finito de números primos, quizá muchos (p. ej., un millón) o, dicho de un modo general, n. Usando subíndices, podríamos indicar dichos números primos por $p_1, p_2, \ldots, p_n$. Cualquier otro entero sería compuesto y debería ser divisible por uno al menos de los $p_1, p_2, \ldots, p_n$. Vamos a ver que dicha hipótesis nos lleva a una contradicción; para ello, construyamos un número A que será distinto de los primos $p_1, p_2, \ldots, p_n$ por ser mayor que ellos y que, sin embargo, no será divisible por ninguno de los p. Tal número es el

$$A = p_1 p_2 \ldots p_n + 1,$$

es decir, se obtiene añadiendo 1 al producto de todos los números primos que suponíamos existentes. A es distinto de los p y, por tanto, debe ser compuesto. Pero dividiendo A por p_1, o por p_2, etc., se obtiene siempre de resto 1; en consecuencia, A no admite ningún p como divisor. Puesto que nuestra hipótesis inicial de que había sólo

un número finito de números primos nos conduce a una contradicción, dicha hipótesis es absurda y, por tanto, la contraria debe ser cierta; esto prueba el teorema.

Aunque la demostración es indirecta, puede modificarse fácilmente para dar un método constructivo, al menos teóricamente, de una sucesión infinita de números primos. Partiendo de un número primo cualquiera, p. ej., $p_1 = 2$, supongamos que hemos encontrado n primos $p_1, p_2, p_3, \ldots, p_n$; observemos entonces que el número $p_1p_2 \ldots p_n + 1$, o bien es primo o contiene factores primos que han de ser distintos de los n hallados previamente. Puesto que estos factores pueden hallarse por ensayos directos, estamos seguros de que, en todo caso, hay al menos un nuevo factor primo p_{n+1}; procediendo de este modo se ve que la sucesión de los números primos construíbles no tiene fin.

Ejercicio: Llévese a efecto dicha construcción a partir de $p_1 = 2$, $p_2 = 3$ y obténganse 5 números primos más.

Si un número ha sido expresado como producto de números primos, podemos disponer dichos factores primos en un orden cualquiera. La experiencia demostraría que, salvo la arbitrariedad en la ordenación, la descomposición de un número N en factores primos es única: *Todo entero* N, *mayor que* 1, *puede descomponerse en producto de números primos, y solamente de una forma.* Esta proposición parece a simple vista tan evidente que un profano podría inclinarse a admitirla sin prueba. Sin embargo, no es una trivialidad y la demostración, aunque elemental, requiere algunos razonamientos sutiles. La demostración clásica, dada por Euclides, de este «teorema fundamental de la aritmética» está basada en un método o «algoritmo» para el cálculo del máximo común divisor de dos números. Este método será considerado en la página 51. En vez de dicha demostración, daremos aquí otra de cosecha más reciente; más breve, pero quizá más artificiosa que la de Euclides. Será un ejemplo típico de demostración indirecta. Supondremos la existencia de un entero susceptible de dos descomposiciones esencialmente diferentes, y de esta hipótesis resultará una contradicción. Esta contradicción demostrará que la hipótesis de que existe un entero con dos descomposiciones esencialmente diferentes en factores primos es absurda, y, como consecuencia, resultará que la descomposición en factores primos de un entero cualquiera es única.

*Si existe un entero positivo capaz de descomponerse en dos productos esencialmente diferentes de primos, habrá uno *menor* que todos los demás para el que se verifique tal propiedad (véase pág. 26),

$$m = p_1p_2 \ldots p_r = q_1q_2 \ldots q_s, \qquad [1]$$

donde los p y los q son primos. Ordenando de nuevo, si es preciso, los p y los q, podemos suponer que se tiene

$$p_1 < p_2 < \cdots < p_r, \qquad q_1 < q_2 < \cdots < q_s.$$

Ahora bien: p_1 no puede ser igual a q_1, ya que en caso contrario, dividiendo los dos últimos miembros de [1] por $p_1 = q_1$, se obtendrían dos descomposiciones esencialmente diferentes para un entero menor que m, en contradicción con la elección hecha de m como el entero *más pequeño* para el cual ese hecho es posible. Por tanto, o bien es $p_1 < q_1$ o se tiene $q_1 < p_1$. Supongamos $p_1 < q_1$. (Si fuera $q_1 < p_1$, bastaría cambiar las letras p y q en lo que sigue.) Formemos el entero

$$m' = m - (p_1 q_2 q_3 \cdots q_s). \qquad [2]$$

Sustituyendo m por las dos expresiones dadas por [1], podríamos escribir el entero m' en una de las dos formas

$$m' = (p_1 p_2 \cdots p_r) - (p_1 q_2 \cdots q_s) = p_1(p_2 p_3 \cdots p_r - q_2 q_3 \cdots q_s) \qquad [3]$$

$$m' = (q_1 q_2 \cdots q_s) - (p_1 q_2 \cdots q_s) = (q_1 - p_1)(q_2 q_3 \cdots q_s) \qquad [4]$$

Puesto que es $p_1 < q_1$, de [4] se sigue que m' es un entero positivo, mientras que de [2] resulta que m' es menor que m. De donde se deduce que, salvo el orden de los factores, la descomposición de m' debe ser *única*. Pero de [3] resulta que p_1 es un factor de m'; por tanto, de [4] se concluye que p_1 debe aparecer como factor o del $(q_1 - p_1)$ o de $(q_2 q_3 \ldots q_s)$. (Esto resulta de suponer la descomposición única para m'; véase el razonamiento del párrafo siguiente.) La última hipótesis es imposible, porque p_1 es menor que todas las q. En consecuencia, p_1 debe ser un factor de $q_1 - p_1$, de modo que existirá un h tal que

$$q_1 - p_1 = p_1 \cdot h \qquad \text{o} \qquad q_1 = p_1(h + 1).$$

Pero esto expresa que p_1 es un factor de q_1, contrariamente al hecho de ser q_1 número primo. Esta contradicción prueba que nuestra hipótesis inicial era absurda y, por tanto, completa la demostración del teorema fundamental de la aritmética.

Un corolario importante de este teorema fundamental es el siguiente: *Si un número primo* p *es un divisor del producto* ab, p *debe ser factor de* a *o de* b. Pues si p no fuera divisor de a ni de b, el producto de factores primos que da la descomposición de *ab no contendría p*. Por otra parte, puesto que se supone que p es un divisor de *ab*, existirá un entero t tal que

$$ab = pt.$$

En consecuencia, el producto de p por una descomposición en factores primos de t daría una descomposición de *ab* en factores primos, que *contendría p*, lo que estaría en contradicción con el hecho de que la descomposición en factores primos es única.

Ejemplos.—Si se ha comprobado de una parte que 13 es un divisor de 2652, y de otra que es $2652 = 6 \cdot 442$, se puede concluir que 13 es un divisor de 442. Por el contrario, 6 es un factor de 240, y $240 = 15 \cdot 16$; sin embargo, 6 no es factor de 15 ni de 16. Esto último prueba que la hipótesis de que p es *primo* es esencial para el corolario.

Ejercicio: Para hallar todos los divisores de un número cualquiera a es suficiente descomponer a en un producto

$$a = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

donde las p son los distintos factores primos, cada uno de ellos elevado a una cierta potencia. *Todos* los divisores de a son los números

$$b = p_1^{\beta_1} \cdot p_2^{\beta_2} \cdots p_r^{\beta_r},$$

donde las β son enteros cualesquiera que satisfacen las desigualdades

$$0 \leqslant \beta_1 \leqslant \alpha_1, \quad 0 \leqslant \beta_2 \leqslant \alpha_2, \ldots, \quad 0 \leqslant \beta_r \leqslant \alpha_r.$$

Demuéstrese esta proposición. Como consecuencia, pruébese que el número de divisores distintos de a (incluídos los divisores a y 1) viene dado por el producto

$$(\alpha_1 + 1)(\alpha_2 + 1) \cdots (\alpha_r + 1).$$

Por ejemplo,

$$144 = 2^4 \cdot 3^2$$

tiene $5 \cdot 3$ divisores, que son: 1, 2, 4, 8, 16, 3, 6, 12, 24, 48, 9, 18, 36, 72 y 144.

2. Distribución de los números primos.—Puede construirse una lista de todos los números primos menores que un entero N dado escribiendo primero ordenadamente todos los enteros menores que N, tachando después todos los que sean múltiplos de 2; luego, todos los restantes que sean múltiplos de 3, y así sucesivamente, hasta que hayan sido eliminados todos los números compuestos. Este proceso, conocido como la «criba de Eratóstenes», conserva en el tamiz los números primos menores que N. Tablas completas de números primos hasta 10 000 000 han sido calculadas mediante perfeccionamientos del método anterior; gracias a ellas se dispone de una masa imponente de datos empíricos referentes a la distribución y otras propiedades de los números primos. Sobre la base de estas tablas se han adelantado conjeturas plausibles (del mismo modo que si la teoría de números fuese una ciencia experimental), muchas de las cuales resultan de demostración muy difícil.

a) *Fórmulas que dan números primos.*—Se han hecho muchos intentos para obtener fórmulas aritméticas simples que dieran única-

mente números primos, aunque no se obtuvieran todos. Fermat hizo la famosa conjetura (no en forma de afirmación definitiva) de que todos los números de la forma

$$F(n) = 2^{2^n} + 1$$

son primos. En efecto, para $n = 1, 2, 3, 4$ se obtiene

$$\begin{aligned} F(1) &= 2^2 + 1 = 5, \\ F(2) &= 2^{2^2} + 1 = 2^4 + 1 = 17, \\ F(3) &= 2^{2^3} + 1 = 2^8 + 1 = 257, \\ F(4) &= 2^{2^4} + 1 = 2^{16} + 1 = 65\,537, \end{aligned}$$

todos los cuales son números primos. Pero, en 1732, Euler obtuvo la descomposición $2^{2^5} + 1 = 641 \cdot 6\,700\,417$; de donde resulta que $F(5)$ no es primo. Más tarde se comprobó que otros de los «números de Fermat» eran compuestos, siendo necesarios para dichas comprobaciones métodos difíciles de la teoría de números, a causa de la insuperable dificultad de los ensayos directos. Hasta la fecha, no ha sido probado que sea primo ninguno de los números $F(n)$ para $n > 4$.

Otra sencilla y notable expresión que da varios números primos es

$$f(n) = n^2 - n + 41.$$

Para $n = 1, 2, \ldots, 40$ $f(n)$ es primo; pero para $n = 41$, se tiene $f(n) = 41^2$, que evidentemente no es primo.

La expresión

$$n^2 - 79n + 1601$$

da primos para n menor que 80, pero falla para $n = 80$. En resumen, puede decirse que la tarea de encontrar expresiones de tipo sencillo que den números primos ha resultado estéril. Ni que decir tiene que la posibilidad de obtener una fórmula algébrica simple que diera *todos* los números primos aparece como irrealizable.

b) *Números primos en una progresión aritmética.*—Mientras que es muy sencillo demostrar que existen infinitos números primos en la sucesión de enteros 1, 2, 3, ..., la extensión de un resultado análogo para sucesiones tales como la 1, 4, 7, 10, 13, ... o la 3, 7, 11, 15, 19, ... o, más en general, para toda progresión aritmética, a, $a + d$, $a + 2d$, ... $a + nd$, ..., donde a y d no tienen factores comunes, es bastante más complicada. Todas las observaciones parecen indicar el hecho de que *en toda progresión de dicho tipo existen infinitos números primos*, como ocurre para la progresión más simple: 1, 2, 3, ... El probar este teorema general requirió un gran esfuerzo. Lejeune-Dirichlet (1805-1859),

uno de los matemáticos más notables del siglo pasado, obtuvo ese brillante resultado mediante la aplicación de los métodos más avanzados del análisis matemático de su tiempo. Sus trabajos originales sobre este tema se cuentan aún hoy entre los resultados más destacados y, pasado ya un siglo, la demostración no ha sido lo suficientemente simplificada como para ser accesible a quienes no posean una buena preparación en la técnica del cálculo y de la teoría de funciones.

Aunque no vamos a intentar probar el teorema de Dirichlet en toda su generalidad, es fácil generalizar la demostración de Euclides sobre la infinitud de la sucesión de números primos, para extenderla a algunas progresiones aritméticas *especiales*, tales como $4n + 3$ y $6n + 5$. Para tratar la primera, observemos que todo número primo mayor que 2 es impar (ya que en otro caso sería divisible por 2) y, por tanto, será de la forma $4n + 1$ ó $4n + 3$, para ciertos enteros n. Por otra parte, el producto de dos números de la forma $4n + 1$ es también de esta forma, ya que

$$(4a + 1)(4b + 1) = 16ab + 4a + 4b + 1 = 4(4ab + a + b) + 1.$$

Supongamos ahora que no existiera más que un número finito de números primos $p_1, p_2, p_3, \ldots, p_n$, de la forma $4n + 3$, y consideremos el número

$$N = 4(p_1 p_2 \ldots p_n) - 1 = 4(p_1 \ldots p_n - 1) + 3.$$

O bien N es primo o puede ser descompuesto en producto de números primos, ninguno de los cuales puede ser $p_1, p_2, \ldots, p_n$, puesto que dividiendo N por cualquiera de éstos, da de resto -1. Además, no todos los factores de N pueden ser de la forma $4n + 1$, ya que N no es de esta forma y, como acabamos de ver, el producto de números de la forma $4n + 1$ es también de esta forma. Por tanto, al menos uno de dichos factores debe ser de la forma $4n + 3$, lo cual es imposible, puesto que hemos visto que ninguno puede ser de los p y éstos formaban, según nuestra hipótesis, *todos* los números primos de la forma $4n + 3$. En consecuencia, la hipótesis de que no hay más que un número finito de números primos de la forma $4n + 3$ nos ha llevado a una contradicción, lo que demuestra que dicho número debe ser infinito.

Ejercicio: Pruébese el correspondiente teorema para la progresión $6n + 5$.

c) *El teorema de los números primos.*—En el proceso de la investigación en busca de una ley que gobernase la distribución de los números primos, se dió un paso decisivo cuando los matemáticos, dejando de lado los fútiles intentos para hallar una fórmula simple que diera

todos los números primos o el número exacto de éstos comprendidos entre los n primeros enteros, buscaron en su lugar información relativa a la distribución *media* de los números primos dentro del conjunto de los enteros.

Para todo entero n, designemos con A_n el número de primos comprendidos entre los enteros 1, 2, 3, ..., n. Si subrayamos los números primos de la sucesión formada por los primeros enteros 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, ... podremos calcular los valores iniciales de A_n:

$$A_1 = 0,\ A_2 = 1,\ A_3 = A_4 = 2,\ A_5 = A_6 = 3,\ A_7 = A_8 = A_9 = A_{10} = 4,$$
$$A_{11} = A_{12} = 5,\ A_{13} = A_{14} = A_{15} = A_{16} = 6,\ A_{17} = A_{18} = 7,\ A_{19} = 8,\ \text{etc.}$$

Si tomamos ahora una sucesión de valores para n que crezca ilimitadamente; p. ej.,

$$n = 10,\ 10^2,\ 10^3,\ 10^4,\ \ldots,$$

la correspondiente sucesión de valores de A_n,

$$A_{10},\ A_{10^2},\ A_{10^3},\ A_{10^4},\ \ldots,$$

crecerá también sin límite (aunque menos rápidamente). Puesto que sabemos que existen infinitos números primos, los valores de A_n excederán más pronto o más tarde a cualquier número finito. La *densidad* de los números primos entre los n primeros enteros vendrá dada por el cociente A_n/n, y a partir de la tabla de números primos, los valores de A_n/n pueden ser calculados empíricamente para un gran número de valores de n.

n	A_n/n
10^2	0,168
10^4	0,078 498
10^9	0,050 847 478
...	.....

El último número decimal de esta tabla puede considerarse como la probabilidad para que un entero elegido al azar entre los 10^9 primeros enteros sea primo, ya que el número de casos posibles es 10^9, de los cuales A_{10^9} son primos.

La distribución de números primos particulares entre los enteros es muy irregular; pero esta irregularidad *en pequeño* desaparece si fijamos nuestra atención en la distribución media de los números primos dada por la razón A_n/n. La ley sencilla que rige el comportamiento de este cociente es uno de los más notables descubrimientos

de toda la matemática. Para establecer el *teorema de los números primos*, debemos previamente definir el «logaritmo natural» de un entero n. Con este objeto, tomemos en el plano dos ejes perpendiculares y consideremos el conjunto de los puntos del plano para los cuales el producto de sus distancias x, y a estos dos ejes es igual a la unidad. En función de las coordenadas x, y, este lugar, una hipérbola equilátera, queda definido por la ecuación $xy = 1$. El log n vendrá definido entonces como el área limitada, en la figura 5, por la hipérbola, el eje x y las dos verticales $x = 1$ y $x = n$. (Una exposición más detallada del logaritmo se dará en el capítulo VIII.) Del estudio empírico de la tabla de números primos, Gauss dedujo que el cociente A_n/n es aproximadamente igual a $1/\log n$, y que la aproximación mejora al crecer n. La bondad de la aproximación viene medida por el cociente $\dfrac{A_n/n}{1/\log n}$, cuyos valores para $n = 1000$, $1\,000\,000$, $1\,000\,000\,000$ se dan en la tabla siguiente.

n	A_n/n	$1/\log n$	$\dfrac{A_n/n}{1/\log n}$
10^3	0,168	0,145	1,159
10^6	0,078 498	0,072 382	1,084
10^9	0,050 847 478	0,048 254 942	1,053
...	.....	.....	....

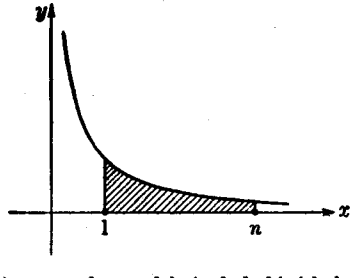

FIG. 5.—El área rayada por debajo de la hipérbola define log n.

Sobre la base de esta evidencia empírica, Gauss hizo la conjetura de que la razón A_n/n es «asintóticamente igual» a $1/\log n$. Lo que quiere decir que si tomamos una sucesión creciente de valores de n; p. ej., si tomamos para n

$$10,\ 10^2,\ 10^3,\ 10^4,\ \ldots$$

como antes hicimos, la razón de A_n/n a $1/\log n$,

$$\frac{A_n/n}{1/\log n},$$

calculada para estos valores sucesivos de n, se aproximará cada vez más a 1, y que la diferencia entre dicho cociente y 1 será tan pequeña como queramos para valores suficientemente grandes de n. Esta afirmación se expresa simbólicamente, mediante el signo $\sim$, en la forma:

$$\frac{A_n}{n} \sim \frac{1}{\log n}, \text{ lo que significa que } \frac{A_n/n}{1/\log n} \text{ tiende a 1 al crecer } n.$$

Que $\sim$ no puede ser reemplazado por el signo ordinario de igualdad $=$, resulta evidente si se observa que mientras A_n es siempre entero, $n/\log n$ no lo es.

El hecho de que el comportamiento medio de la distribución de los números primos pueda ser descrito mediante la función logarítmica es un descubrimiento notable, ya que resulta sorprendente que dos conceptos matemáticos que parecen tan inconexos estén en realidad tan íntimamente ligados.

Aunque el enunciado de la conjetura de Gauss es fácil de comprender, una demostración rigurosa de dicha conjetura era inaccesible a los métodos de la ciencia matemática de su época. Para probar el teorema, que se refiere únicamente a conceptos elementales, es necesario emplear los métodos más potentes de la matemática moderna. Fueron precisos casi cien años antes que el análisis se desarrollara lo suficiente para que Hadamard (1896), en París, y de la Vallée Poussin (1896), en Lovaina, pudieran dar una demostración completa del teorema de los números primos. Se han hecho simplificaciones y modificaciones importantes por v. Mangoldt y Landau. Mucho antes del resultado de Hadamard, Riemann (1826-1866) había contribuído al problema con una decisiva labor de exploración en un célebre trabajo, en el cual aparecen definidas las direcciones estratégicas del posible ataque. Recientemente, el matemático norteamericano Norbert Wiener ha sido capaz de modificar la demostración de modo que evita el uso de números complejos en un paso importante del razonamiento. En todo caso, la demostración del teorema de los números primos cotinúa siendo una cuestión difícil, aun para estudiantes avanzados. Volveremos a ocuparnos del tema en el capítulo VIII, suplemento IV. [1]

[1] En 1949 un joven matemático noruego, Selberg, y el matemático húngaro Erdös, en dos trabajos en cierto modo complementarios, dieron una demostración *elemental* (¡aunque no sencilla!) del teorema de los números primos. La demostración, que no depende de ideas analíticas ajenas al problema, constituye un descubrimiento de primordial importancia para la estructura lógica de la teoría de la distribución de los números primos. *(N. del T.)*

d) *Dos problemas no resueltos referentes a los números primos.*— Mientras que el problema de la distribución media de los números primos ha sido resuelto satisfactoriamente, quedan otras varias conjeturas que, sostenidas por la evidencia empírica, no han podido hasta ahora ser probadas como ciertas.

Una de ellas es la famosa *conjetura de Goldbach.* Goldbach (1690-1764) aparece en la historia de las matemáticas únicamente por esa conjetura que propuso como problema en una carta a Euler en 1742. Comprobó que en todos los casos observados, todo número par (excepto el 2, que es primo) puede ser representado como suma de dos números primos; p. ej.: $4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 5 + 7$, $14 = 7 + 7$, $16 = 13 + 3$, $18 = 11 + 7$, $20 = 13 + 7$, ..., $48 = 29 + 19$, ..., $100 = 97 + 3$, etcétera.

Goldbach preguntaba a Euler si era capaz de demostrar que esa propiedad es cierta para *todo* número par, o si podría encontrar un contraejemplo. Euler no pudo dar una respuesta, ni nadie ha podido darla hasta ahora. La evidencia empírica a favor de la proposición de que todo número par puede ser representado de ese modo es bastante convincente, como puede verificar cualquiera ensayando en algunos ejemplos. El origen de la dificultad reside en el hecho de que los números primos se definen mediante la *multiplicación*, mientras que el problema se refiere a la *adición*. En términos generales, resulta difícil establecer conexiones entre las propiedades aditivas y las multiplicativas de los enteros.

Hasta hace poco tiempo, una demostración de la conjetura de Goldbach parecía completamente inaccesible. Hoy, en cambio, no parece estar muy lejos la solución de este problema. Un éxito importante, completamente inesperado y que sorprendió a los expertos, fué alcanzado en 1931 por un joven matemático ruso desconocido hasta entonces, Schnirelmann (1905-1938), quien probó que *todo entero positivo puede ser representado como suma de a lo sumo 300 000 números primos.* Aunque este resultado pueda parecer ridículo en comparación con el problema inicial de probar la conjetura de Goldbach, fué, sin embargo, el primer paso en la dirección justa. La prueba es directa y constructiva, aunque no da un método práctico para encontrar la descomposición en suma de números primos para un entero arbitrario. Posteriormente, el matemático ruso Vinogradoff, usando métodos debidos a Hardy, Littlewood y a su gran colaborador indio Ramanujan, ha conseguido reducir el número de 300 000 a 4, lo que está mucho más cerca de la solución del problema de Goldbach. Sin embargo, hay una diferencia notable entre los resultados de Schnirelmann y Vinogradoff,

más significativa aún que la diferencia entre 300 000 y 4. El teorema de Vinogradoff ha sido demostrado únicamente para enteros «suficientemente grandes»; con más precisión: Vinogradoff ha probado que *existe* un entero N tal que todo entero $n > N$ puede ser representado como suma de, a lo más, cuatro números primos. La demostración de Vinogradoff no permite determinar N; en contraste con el teorema de Schnirelmann, el de Vinogradoff es esencialmente indirecto y no constructivo. Lo que realmente prueba Vinogradoff es que la hipótesis de que existen infinitos enteros que no pueden descomponerse en suma de 4 números primos es absurda. Aquí se tiene un buen ejemplo de la profunda diferencia existente entre los dos tipos de demostración, directa e indirecta. (Véase la discusión general en la página 95.)

El problema siguiente, más notable aún que el de Goldbach, está todavía muy lejos de ser resuelto. Se ha observado que los números primos se suceden frecuentemente en pares de la forma p y $p + 2$. Así, p. ej., 3 y 5, 11 y 13, 29 y 31, etc. Se cree que la proposición que afirma la existencia de infinitos pares de ese tipo es cierta; sin embargo, hasta el presente no se ha podido dar el menor paso aprovechable en el camino de demostrar tal proposición.

II. CONGRUENCIAS

1. Conceptos generales.—Siempre que se presenta la cuestión de la divisibilidad de enteros por un entero fijo d, el concepto y la notación de *congruencia* (debidos a Gauss) sirven para aclarar y simplificar el razonamiento.

Para introducir este concepto, consideremos los restos que dan los sucesivos enteros al dividirlos por 5; se tiene

$0 = 0 \cdot 5 + 0$	$7 = 1 \cdot 5 + 2$	$-1 = -1 \cdot 5 + 4$
$1 = 0 \cdot 5 + 1$	$8 = 1 \cdot 5 + 3$	$-2 = -1 \cdot 5 + 3$
$2 = 0 \cdot 5 + 2$	$9 = 1 \cdot 5 + 4$	$-3 = -1 \cdot 5 + 2$
$3 = 0 \cdot 5 + 3$	$10 = 2 \cdot 5 + 0$	$-4 = -1 \cdot 5 + 1$
$4 = 0 \cdot 5 + 4$	$11 = 2 \cdot 5 + 1$	$-5 = -1 \cdot 5 + 0$
$5 = 1 \cdot 5 + 0$	$12 = 2 \cdot 5 + 2$	$-6 = -2 \cdot 5 + 4$
$6 = 1 \cdot 5 + 1$	etc.	etc.

Observamos que el resto de cualquier entero es uno de los números 0, 1, 2, 3, 4. Diremos que dos enteros a y b son «congruentes módulo 5» si ambos dan el *mismo resto* al ser divididos por 5. Así, 2, 7, 12 17, 22, ..., −3, −8, −13, −18, ..., son todos congruentes módulo 5, puesto que todos dan de resto 2. En general, diremos que dos enteros a y b son *congruentes módulo d*, siendo d un entero dado, si a y b dan

el mismo resto al dividirlos por d, lo que equivale a decir que hay un entero n tal que $a - b = nd$; p. ej., 27 y 15 son congruentes módulo 4, puesto que

$$27 = 6 \cdot 4 + 3, \qquad 15 = 3 \cdot 4 + 3.$$

El concepto de congruencia es tan útil, que resulta conveniente disponer de una notación sencilla para él. Escribiremos

$$a \equiv b \pmod{d}$$

para expresar que a y b son congruentes módulo d. En los casos en que no haya duda acerca de cuál sea el módulo, el «mód d» de la fórmula puede ser omitido. [Si a no es congruente con b módulo d, escribiremos $a \not\equiv b \pmod{d}$.]

Las congruencias aparecen frecuentemente en la vida diaria; p. ej., las manecillas de un reloj indican la hora módulo 12, y el cuentakilómetros de un coche da el total de kilómetros recorridos módulo 100 000.

Antes de proceder a una discusión detallada de las congruencias, el lector debe observar las siguientes proposiciones, que son todas ellas equivalentes:

1. a es congruente con b módulo d.
2. $a = b + nd$ para un cierto entero n.
3. d divide a $a - b$.

La utilidad de la notación de Gauss para las congruencias se basa en el hecho de que la congruencia respecto de un módulo fijo tiene varias de las propiedades formales de la igualdad ordinaria. Las propiedades formales más importantes de la relación $a = b$ son las siguientes:

1) Se tiene siempre $a = a$.
2) De $a = b$ se sigue $b = a$.
3) De $a = b$ y $b = c$ se sigue $a = c$.

Además, de $a = a'$ y $b = b'$ se sigue

4) $a + b = a' + b'$
5) $a - b = a' - b'$.
6) $ab = a'b'$.

Estas propiedades continúan siendo válidas si se reemplaza la relación $a = b$ por la de congruencia $a \equiv b \pmod{d}$. Así:

1') Se tiene siempre $a \equiv a \pmod{d}$.
2') De $a \equiv b \pmod{d}$ se sigue $b \equiv a \pmod{d}$.
3') De $a \equiv b \pmod{d}$ y $b \equiv c \pmod{d}$ se sigue $a \equiv c \pmod{d}$.

Dejamos al cuidado del lector el comprobar estos hechos.

Además, de $a \equiv a'$ (mód d) y $b \equiv b'$ (mód d) se sigue

4') $a + b \equiv a' + b'$ (mód d).
5') $a - b \equiv a' - b'$ (mód d).
6') $ab \equiv a'b'$ (mód d).

Resulta, en consecuencia, que varias *congruencias respecto del mismo módulo pueden sumarse, restarse* y *multiplicarse*. Para probar estas tres proposiciones es suficiente observar que si

$$a = a' + rd, \qquad b = b' + sd,$$

se tiene

$$a + b = a' + b' + (r + s)d,$$
$$a - b = a' - b' + (r - s)d,$$
$$ab = a'b' + (a's + b'r + rsd)d,$$

de donde resultan las conclusiones deseadas.

El concepto de congruencia tiene una interpretación geométrica muy intuitiva. De ordinario, cuando se quieren representar los enteros geométricamente, se elige un segmento de longitud unidad y se

FIG. 6.—Representación geométrica de los números enteros.

llevan éste y sus múltiplos en los dos sentidos. De este modo se hace corresponder a cada entero un punto sobre la recta, como en la figura 6. Si se trata de los enteros módulo d, dos números congruentes se consideran como uno mismo en lo que a la división por d se refiere, puesto que los dos dan el mismo resto. Para ver esto geométricamente, utilicemos una circunferencia dividida en d partes iguales. Cualquier entero dividido por d da de resto uno de los d números 0, 1, ..., $d - 1$, que están situados a intervalos iguales sobre la circunferencia. Cualquier entero es congruente módulo d con uno de estos d números, y, por tanto, puede ser representado geométricamente por uno de esos puntos; dos números son congruentes si están representados por el mismo punto. La figura 7 está dibujada para el caso $d = 6$. La esfera de un reloj es otro ejemplo tomado de la vida corriente.

Como aplicación de la propiedad multiplicativa 6') de las congruencias, determinaremos los restos de las sucesivas potencias de 10 respecto de un número dado; p. ej.,

$$10 \equiv -1 \text{ (mód. 11)},$$

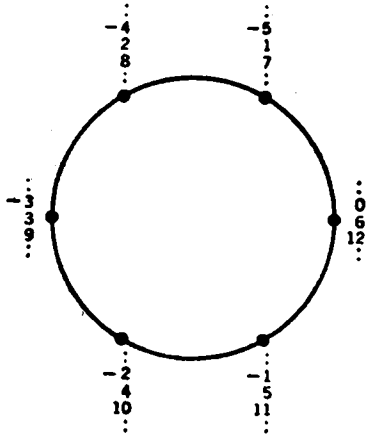

FIG. 7.—Representación geométrica de los números enteros respecto al módulo 6.

ya que $10 = -1 + 11$. Por multiplicaciones sucesivas de esta congruencia por sí misma se obtiene

$$\begin{aligned} 10^2 &\equiv (-1)(-1) = 1 \pmod{11}, \\ 10^3 &\equiv -1 \quad \text{"} \quad \text{"} \\ 10^4 &\equiv 1 \quad \text{"} \quad \text{"}, \text{ etc.} \end{aligned}$$

Como consecuencia, se puede probar que cualquier entero

$$z = a_0 + a_1 \cdot 10 + a_2 \cdot 10^2 + \ldots + a_n \cdot 10^n,$$

expresado en el sistema decimal, da al dividirlo por 11 el mismo resto que la suma de sus cifras, tomadas alternativamente con los signos más y menos,

$$t = a_0 - a_1 + a_2 - a_3 + \ldots.$$

En efecto, se puede escribir

$$z - t = a_1 \cdot 11 + a_2(10^2 - 1) + a_3(10^3 + 1) + a_4(10^4 - 1) + \cdots$$

Y como todos los números 11, $10^2 - 1$, $10^3 + 1$, ... son congruentes con 0 módulo 11, también lo será $z - t$; por consiguiente, z da el mismo resto que t al dividirlo por 11. Resulta en particular que un número es divisible por 11 (es decir, da de resto 0) cuando, y únicamente entonces, la suma alternada de sus cifras es divisible por 11. Por ejemplo, puesto que $3 - 1 + 6 - 2 + 8 - 1 + 9 = 22$, el número $z = 3\,162\,819$ es divisible por 11. Encontrar una regla de divisibilidad

por 3 ó por 9 es más sencillo aún, puesto que $10 \equiv 1$ (mód 3 ó 9), y, por consiguiente, $10^n \equiv 1$ (mód 3 ó 9) para todo n. Resulta así que para que un número z sea divisible por 3 ó por 9 es necesario y suficiente que la suma de sus cifras

$$s = a_0 + a_1 + a_2 + \ldots + a_n$$

sea también divisible por 3 ó por 9, respectivamente.

Para las congruencias módulo 7 se tiene

$$10 \equiv 3, \quad 10^2 \equiv 2, \quad 10^3 \equiv -1, \quad 10^4 \equiv -3, \quad 10^5 \equiv -2, \quad 10^6 \equiv 1.$$

Los restos sucesivos se repiten, de donde resulta que para que z sea divisible por 7 es necesario y suficiente que la expresión

$$r = a_0 + 3a_1 + 2a_2 - a_3 - 3a_4 - 2a_5 + a_6 + 3a_7 + \ldots$$

sea divisible por 7.

Ejercicio: Hállese una regla análoga para la divisibilidad por 13.

Sumando o multiplicando congruencias respecto de un módulo fijo, p. ej., $d = 5$, podemos evitar los números muy grandes al reemplazar cualquier número a por el del conjunto

$$0, \quad 1, \quad 2, \quad 3, \quad 4$$

congruente con él. Así, para calcular sumas y productos módulo 5 necesitamos únicamente las siguientes tablas de adición y multiplicación:

$a + b$

$a \equiv$ \ $b \equiv$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$a \cdot b$

$a \equiv$ \ $b \equiv$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

De la segunda de esas tablas resulta que un producto ab es congruente con 0 (mód 5) únicamente si a o b son $\equiv 0$ (mód 5). Esto sugiere la ley general

7) $ab \equiv 0$ (mód d) únicamente si $a \equiv 0$ ó $b \equiv 0$ (mód d),

la cual resultaría como una extensión de la ley ordinaria para enteros que dice que se tiene $ab = 0$ únicamente si es $a = 0$ ó $b = 0$. Ahora bien: *la ley* 7) *es válida solamente cuando el módulo* d *es un número primo*. En efecto, la congruencia

$$ab \equiv 0 \qquad (\text{mód } d)$$

significa que d divide a ab, y hemos visto antes que un número primo divide a un producto ab únicamente si divide a a o a b; esto es, únicamente si

$$a \equiv 0 \pmod{d} \quad \text{ó} \quad b \equiv 0 \pmod{d}.$$

Si d no es primo, esta regla no vale en general; puesto que si es $d = r \cdot s$, siendo r y s menores que d, se tiene

$$r \not\equiv 0 \pmod{d}, \qquad s \not\equiv 0 \pmod{d},$$

y, en cambio, es

$$rs = d \equiv 0 \pmod{d}.$$

Por ejemplo, $2 \not\equiv 0$ (mód 6) y $3 \not\equiv 0$ (mód 6); sin embargo, $2 \cdot 3 = 6 \equiv 0$ (mód 6).

Ejercicio: Pruébese que la siguiente *regla de simplificación* es válida para congruencias respecto de un módulo primo:

Si es $ab \equiv ac$ y $a \not\equiv 0$, se tiene $b \equiv c$.

Ejercicios:

1. ¿Con qué número entre 0 y 6 inclusive es congruente módulo 7 el producto $11 \cdot 18 \cdot 2322 \cdot 13 \cdot 19$?
2. ¿Con qué número comprendido entre 0 y 12 inclusive es congruente módulo 13 el producto $3 \cdot 7 \cdot 11 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 113$?
3. ¿Con qué número entre 0 y 4 inclusive es congruente módulo 5 la suma $1 + 2 + 2^2 + \ldots + 2^{19}$?

2. Teorema de Fermat.—En el siglo XVII, Fermat, el fundador de la moderna teoría de números, descubrió un teorema muy importante: *Si* p *es un número primo que no divide al entero* a, *se tiene*:

$$a^{p-1} \equiv 1 \pmod{p},$$

lo que significa que la potencia $(p - 1)$-ésima de a da resto 1 al dividirla por p.

Algunos de nuestros cálculos anteriores confirman este teorema; p. ej., encontrábamos que $10^6 \equiv 1$ (mód 7), $10^2 \equiv 1$ (mód 3), y $10^{10} \equiv 1$ (mód 11). Del mismo modo se vería que es $2^{12} \equiv 1$ (mód 13) y $5^{10} \equiv 1$ (mód 11). Para comprobar las últimas congruencias no es necesario calcular efectivamente las potencias indicadas, ya que pueden utilizarse ventajosamente las propiedades multiplicativas de aquéllas:

$2^4 = 16 \equiv 3$ (mód 13)
$2^8 \equiv 9 \equiv -4$ »
$2^{12} \equiv -4 \cdot 3 = -12 \equiv 1$ »

$5^2 \equiv 3$ (mód 11)
$5^4 \equiv 9 \equiv -2$ »
$5^8 \equiv 4$ »
$5^{10} \equiv 3 \cdot 4 = 12 \equiv 1$ »

Para demostrar el teorema de Fermat, consideremos los múltiplos de a:

$$m_1 = a, \qquad m_2 = 2a, \qquad m_3 = 3a, \ldots, \; m_{p-1} = (p-1)a$$

Ningún par de estos enteros puede estar formado por números congruentes módulo p; de lo contrario, p sería un factor de $m_r - m_s = (r-s)a$ para algún par r, s, siendo $1 \leqslant s < r \leqslant (p-1)$. Pero la regla 7) indica que esto no puede ocurrir; puesto que $r - s$ es menor que p, p no puede ser factor de $(r-s)$, y habíamos supuesto que p no dividía a a. Asimismo, ninguno de aquellos múltiplos puede ser congruente con 0. Por consiguiente, los números m_1, m_2, m_3, ..., m_{p-1} deben ser congruentes con los números 1, 2, 3, ..., $p-1$ tomados en orden conveniente. Resulta como consecuencia

$$m_1 m_2 \cdots m_{p-1} = 1 \cdot 2 \cdot 3 \cdots (p-1)a^{p-1} \equiv 1 \cdot 2 \cdot 3 \cdots (p-1) \pmod{p},$$

o, si escribimos por brevedad K en vez de $1 \cdot 2 \cdot 3 \ldots (p-1)$,

$$K(a^{p-1} - 1) \equiv 0 \pmod{p}.$$

Pero K no es divisible por p, ya que no lo es ninguno de sus factores; por tanto, en virtud de la regla 7), $(a^{p-1} - 1)$ debe ser divisible por p; es decir,

$$a^{p-1} - 1 \equiv 0 \pmod{p},$$

que es el teorema de Fermat.

Para comprobar el teorema una vez más, tomemos $p = 23$ y $a = 5$. Se tiene entonces (mód 23) $5^2 \equiv 2$, $5^4 \equiv 4$, $5^8 \equiv 16 \equiv -7$, $5^{16} \equiv 49 \equiv 3$, $5^{20} \equiv 12$, $5^{22} \equiv 24 \equiv 1$. Con $a = 4$, en vez de 5, se obtendría, también (mód 23), $4^2 \equiv -7$, $4^3 \equiv -28 \equiv -5$, $4^4 \equiv -20 \equiv 3$, $4^8 \equiv 9$, $4^{11} \equiv -45 \equiv 1$, $4^{22} \equiv 1$.

En el ejemplo anterior con $a = 4$, $p = 23$, y en otros, se observa que no solamente la potencia $(p-1)$-ésima de a, sino también otras potencias menores pueden ser congruentes con 1. Pero ocurre siempre que la menor de dichas potencias, en dicho caso 11, es un divisor de $p - 1$. (Véase el ejercicio 3 siguiente.)

Ejercicios:

1. Compruébese mediante cálculos análogos a los anteriores que $2^8 \equiv 1$ (mód 17); $3^8 \equiv -1$ (mód 17); $3^{14} \equiv -1$ (mód 29); $2^{14} \equiv -1$ (mód 29); $4^{14} \equiv 1$ (mód 29); $5^{14} \equiv 1$ (mód 29).

2. Compruébese el teorema de Fermat para $p = 5, 7, 11, 17$ y 23, con diferentes valores de a.

3. Demuéstrese el teorema general siguiente: El menor entero positivo e para

el que se tenga $a^e \equiv 1 \pmod{p}$, siendo p primo, debe ser un divisor de $p - 1$ [*Indicación*: Dividiendo $p - 1$ por e, se obtendría

$$p - 1 = ke + r,$$

siendo $0 < r < e$; utilícese el hecho de que $a^{p-1} \equiv a^e \equiv 1 \pmod{p}$.]

3. Restos cuadráticos.—Por los ejemplos considerados en relación con el teorema de Fermat, se ha visto que no solamente es siempre $a^{p-1} \equiv 1 \pmod{p}$, sino que (si p es un número primo distinto de 2, y, por consiguiente, impar y de la forma $p = 2p' + 1$) para algunos valores de a, $a^{p'} = a^{(p-1)/2} \equiv 1 \pmod{p}$. Este hecho sugiere otras varias cuestiones interesantes. Podemos escribir el teorema en la siguiente forma:

$$a^{p-1} - 1 = a^{2p'} - 1 = (a^{p'} - 1)(a^{p'} + 1) \equiv 0 \pmod{p}.$$

Puesto que un producto es divisible por p cuando uno al menos de los factores lo es, resulta inmediatamente que, bien $a^{p'} - 1$ o $a^{p'} + 1$ debe ser divisible por p, de modo que para cualquier primo $p > 2$ y cualquier número a no divisible por p, se tiene

$$a^{(p-1)/2} \equiv 1 \quad \text{o} \quad a^{(p-1)/2} \equiv -1 \pmod{p}.$$

Desde el comienzo de la moderna teoría de números los matemáticos se han interesado por el problema de determinar qué números a pertenecen a la primera clase y cuáles a la segunda. Supongamos que a es congruente módulo p con el cuadrado de otro entero x,

$$a \equiv x^2 \pmod{p}.$$

Entonces se tendrá: $a^{(p-1)/2} \equiv x^{p-1}$ y, de acuerdo con el teorema de Fermat, congruente con 1 módulo p. Un número a, no múltiplo de p, que sea congruente módulo p con el cuadrado de otro entero se llama *resto cuadrático de p*, mientras que un número b, no múltiplo de p, que no sea congruente con ningún cuadrado, se llama *no-resto cuadrático de p*. Acabamos de ver que todo resto cuadrático a de p satisface la congruencia $a^{(p-1)/2} \equiv 1 \pmod{p}$. Sin grandes dificultades se puede probar que para cualquier no-resto b se tiene la congruencia $b^{(p-1)/2} \equiv -1 \pmod{p}$. Por otra parte, vamos a demostrar que entre los números 1, 2, 3, ..., $p - 1$ hay precisamente $(p - 1)/2$ restos cuadráticos y $(p - 1)/2$ no-restos.

Aun siendo posible obtener muchos datos empíricos por cálculo directo, al comienzo no fué fácil descubrir leyes generales que rigieran la distribución de los restos y no-restos cuadráticos. La primera propiedad importante de los restos fué observada por Legendre (1752-

1833), y designada más tarde por Gauss con el nombre de *ley de reciprocidad cuadrática*. Esta ley se refiere al comportamiento de dos números primos distintos p y q, y dice que q es un resto cuadrático de p cuando, y sólo entonces, p es resto cuadrático de q, si el producto $[(p-1)/2]\ [(q-1)/2]$ es *par*. En el caso en que dicho producto sea *impar*, la situación es la inversa, de modo que p es resto de q si q es *no-resto* de p. Uno de los descubrimientos juveniles de Gauss fué la primera demostración rigurosa de este teorema, que durante mucho tiempo había desafiado al mundo matemático. La primera demostración de Gauss no era sencilla, y aún hoy la ley de reciprocidad no es fácil de establecer, a pesar de que se han dado de ella numerosas demostraciones. Su verdadera significación aparece claramente cuando se establece su conexión con los descubrimientos modernos en la teoría de los números algebraicos.

Como ejemplo para ilustrar la distribución de los restos cuadráticos, tomemos $p = 7$. Entonces, puesto que se tiene

$$0^2 \equiv 0, \quad 1^2 \equiv 1, \quad 2^2 \equiv 4, \quad 3^2 \equiv 2, \quad 4^2 \equiv 2, \quad 5^2 \equiv 4, \quad 6^2 \equiv 1,$$

todas módulo 7, y ya que los cuadrados restantes repiten esta sucesión, los restos cuadráticos de 7 son los números congruentes con 1, 2 ó 4, mientras que los no-restos son congruentes con 3, 5 ó 6. En el caso general, los restos cuadráticos de p son los números congruentes con $1^2, 2^2, \ldots, (p-1)^2$. Pero éstos son congruentes a pares, pues se tiene

$$x^2 \equiv (p-x)^2 \pmod{p} \qquad [\text{p. ej., } 2^2 \equiv 5^2 \pmod{7}],$$

por ser $(p-x)^2 = p^2 - 2px + x^2 \equiv x^2 \pmod{p}$. Resulta así que la mitad de los números $1, 2, \ldots, p-1$ son restos cuadráticos de p y la otra mitad son no-restos cuadráticos.

Como ejemplo de la ley de reciprocidad tomemos $p = 5$, $q = 11$. Por ser $11 \equiv 1^2 \pmod 5$, 11 es resto cuadrático (mód 5); y como el producto $[(5-1)/2]\,[(11-1)/2]$ es par, la ley de reciprocidad dice que 5 es resto cuadrático (mód 11). En confirmación de esto, observemos que se tiene $5 \equiv 4^2 \pmod{11}$. Por otra parte, si $p = 7$, $q = 11$, el producto $[(7-1)/2]\,[(11-1)/2]$ es impar, y en efecto 11 es un resto cuadrático (mód 7) [ya que $11 \equiv 2^2 \pmod 7$], mientras que 7 es no-resto cuadrático (mód 11).

Ejercicios:

1. $6^2 = 36 \equiv 13 \pmod{23}$. ¿Es 23 resto cuadrático (mód 13)?

2. Hemos visto que $x^2 \equiv (p-x)^2 \pmod p$. Demuéstrese que ésas son las *únicas* congruencias módulo p entre los números $1^2, 2^2, 3^2, \ldots, (p-1)^2$.

III. LOS NÚMEROS PITAGÓRICOS Y EL ÚLTIMO TEOREMA DE FERMAT

Una interesante cuestión de la teoría de números se halla relacionada con el teorema de Pitágoras. Los griegos sabían que un triángulo con lados 3, 4, 5 es un triángulo rectángulo. Se presentaba entonces la cuestión general: «¿Qué otros triángulos rectángulos tienen sus lados iguales a múltiplos enteros de la unidad de longitud?» El teorema de Pitágoras se expresa algébricamente por la ecuación

$$a^2 + b^2 = c^2, \qquad [1]$$

donde a y b son las longitudes de los catetos, y c, la de la hipotenusa. El problema de hallar *todos* los triángulos rectángulos con lados de longitud entera es equivalente al de hallar las soluciones enteras de la ecuación [1]. Una terna de dichos números se llama *terna de números pitagóricos.*

El problema de hallar todas las ternas de números pitagóricos puede resolverse fácilmente. Si a, b y c forman una terna pitagórica, de modo que $a^2 + b^2 = c^2$, pongamos para simplificar $a/c = x$, $b/c = y$, donde x e y son números racionales para los cuales se verifica $x^2 + y^2 = 1$. Entonces se tiene $y^2 = (1 - x)(1 + x)$, o lo que es lo mismo $y/(1 + x) = (1 - x)/y$. El valor común de los dos miembros de esta ecuación es un número t que puede expresarse como cociente de dos números enteros u/v. Podemos escribir entonces: $y = t(1 + x)$ y $(1 - x) = ty$; es decir,

$$tx - y = -t, \qquad x + ty = 1.$$

De este sistema de ecuaciones se deduce inmediatamente

$$x = \frac{1 - t^2}{1 + t^2}, \qquad y = \frac{2t}{1 + t^2}$$

y sustituyendo x, y, t por sus valores, se tiene

$$\frac{a}{c} = \frac{v^2 - u^2}{u^2 + v^2}, \qquad \frac{b}{c} = \frac{2uv}{u^2 + v^2}$$

Por consiguiente,

$$\begin{aligned} a &= (v^2 - u^2)r, \\ b &= (2uv)r, \\ c &= (u^2 + v^2)r, \end{aligned} \qquad [2]$$

con un factor arbitrario de proporcionalidad r. Esto prueba que si (a, b, c) es una terna pitagórica, a, b, c son proporcionales a $v^2 - u^2$,

$2uv$, $u^2 + v^2$, respectivamente. Recíprocamente, es fácil ver que toda terna (a, b, c) definida mediante [2] es pitagórica, ya que de [2] se sigue

$$a^2 = (u^4 - 2u^2v^2 + v^4)r^2,$$
$$b^2 = (4u^2v^2)r^2,$$
$$c^2 = (u^4 + 2u^2v^2 + v^4)r^2,$$

de forma que $a^2 + b^2 = c^2$.

El resultado anterior puede simplificarse aún: a partir de cualquier terna pitagórica (a, b, c) se pueden obtener otras ternas pitagóricas (sa, sb, sc) para todo valor entero de s; así, de (3, 4, 5) se obtienen las (6, 8, 10), (9, 12, 15), etc. Tales ternas no son esencialmente distintas, ya que corresponden a triángulos rectángulos semejantes. Definiremos como *primitiva* una terna pitagórica cuando los números a, b y c no tengan ningún factor común. Se puede probar que las *fórmulas*

$$a = v^2 - u^2,$$
$$b = 2uv,$$
$$c = u^2 + v^2,$$

dan para todo par de enteros positivos u, v, *con* u $>$ v, *y tales que* u *y* v *no tengan factores comunes ni sean los dos impares, todas las ternas primitivas de números pitagóricos.*

***Ejercicio:** Demuéstrese la última proposición.

Como ejemplos de ternas primitivas de números pitagóricos se tiene: $u = 2$, $v = 1$: (3, 4, 5); $u = 3$, $v = 2$: (5, 12, 13); $u = 4$, $v = 3$: (7, 24, 25); ...; $u = 10$, $v = 7$: (51, 140; 149); etc.

Los resultados alcanzados para los números pitagóricos inducen de modo natural a plantear la cuestión de obtener enteros a, b, c para los que se verifique $a^2 + b^2 = c^2$ o $a^4 + b^4 = c^4$, o, en general, para un exponente entero positivo $n > 2$ dado, a determinar las soluciones enteras y positivas de la ecuación

$$a^n + b^n = c^n. \qquad [3]$$

Fermat dió a esta cuestión una respuesta de modo un poco espectacular. Estudiando la obra de Diofanto, el que más había contribuído entre los antiguos a la teoría de números, Fermat tenía la costumbre de escribir comentarios en las márgenes del libro. De esta forma enunció sin demostración muchos teoremas. Todos han sido probados posteriormente, con excepción de uno de ellos de especial importancia. Comentando la teoría de los números pitagóricos, Fermat escribió que

la ecuación [3] *no admite soluciones enteras para cualquier* n > 2, pero que la elegante demostración que había encontrado era desgraciadamente demasiado larga para el margen de que disponía.

De esta proposición general no han podido ser demostradas ni su validez ni su falsedad a pesar de los esfuerzos de muchos de los más grandes matemáticos posteriores a Fermat. El teorema ha sido probado para varios valores de n; en particular para todo $n < 619$, pero no para todo n, sin haberse encontrado hasta el presente ningún contraejemplo. Aunque el teorema en sí no es de gran importancia matemática, las tentativas hechas para demostrarlo han dado lugar a importantes investigaciones en la teoría de números. Este problema ha despertado también gran interés en círculos no matemáticos, debido en parte a un premio de 100 000 marcos ofrecido a la primera persona que diera una solución aceptada por la Real Academia de Gotinga. Hasta que la inflación que siguió en Alemania a la primera guerra mundial quitó todo valor económico a dicho premio, gran número de «soluciones» incorrectas eran enviadas todos los años a la Academia. Incluso matemáticos serios han creído a veces haber encontrado e incluso han publicado soluciones «correctas»; sin embargo, hasta el presente en todas ellas ha sido descubierto algún error. Con la devaluación del marco, el interés general parece haber desaparecido, aunque de cuando en cuando la prensa anuncia que el problema ha sido resuelto por un genio hasta entonces desconocido.

IV. EL ALGORITMO DE EUCLIDES

1. Teoría general.—Es bien conocida la regla de división de un entero a por otro b, y el lector sabe que el proceso de división no termina hasta que se llega a un resto más pequeño que el divisor. Así, si $a = 648$ y $b = 7$ se obtiene un cociente $q = 92$ y un resto $r = 4$.

$$\begin{array}{r|l} 648 & 7 \\ \cline{2-2} 63 & 92 \\ \cline{1-1} 18 & \\ 14 & \\ \cline{1-1} 4 & \end{array} \qquad 648 = 7 \cdot 92 + 4.$$

Lo que se puede enunciar como un teorema general: *Si* a *es un entero cualquiera y* b *es un entero mayor que* 0, *se pueden encontrar siempre dos enteros* q *y* r *tales que*

$$a = b \cdot q + r, \qquad 0 \leqslant r < b. \qquad [1]$$

Para probar esta proposición sin utilizar la regla ordinaria de división de enteros, basta observar que cualquier entero a es o bien múltiplo de b,

$$a = bq,$$

o está comprendido entre dos múltiplos consecutivos de b,

$$bq < a < b(q + 1) = bq + b.$$

En el primer caso se tiene [1] con $r = 0$, y en el segundo resulta, de la primera desigualdad anterior,

$$a - bq = r > 0,$$

mientras que la segunda desigualdad da

$$a - bq = r < b,$$

de modo que resulta $0 < r < b$, como se dice en [1].

De este sencillo hecho se deduce una gran variedad de consecuencias importantes; la primera de ellas es un método para hallar el máximo común divisor de dos enteros.

Sean a y b dos enteros cualesquiera, ninguno de los cuales sea 0, y consideremos el conjunto de todos los enteros positivos que dividen simultáneamente a a y b. Este conjunto es evidentemente finito, puesto que si, p. ej., es $a \neq 0$, ningún entero mayor que a puede ser divisor de a. En consecuencia, no puede haber más que un número finito de divisores comunes a a y b, y entre ellos habrá uno mayor que todos, d. Este número entero d se llama *máximo común divisor* de a y b, y se escribe $d = (a, b)$. Así, para $a = 8$ y $b = 12$ se obtiene por ensayos directos $(8, 12) = 4$, mientras que para $a = 5$ y $b = 9$ resulta $(5, 9) = 1$. Si a y b son números grandes, p. ej., $a = 1804$ y $b = 328$, el método de ensayos sería muy laborioso e incierto. Un método breve y seguro lo da el llamado *algoritmo de Euclides*. (Un algoritmo es un método sistemático de cálculo.) Está basado en el hecho de que de toda relación de la forma

$$a = b \cdot q + r \qquad [2]$$

se sigue

$$(a, b) = (b, r), \qquad [3]$$

puesto que todo número u que divida a a y b

$$a = su, \qquad b = tu,$$

divide también a r, ya que es $r = a - bq = su - qtu = (s - qt)u$; y recíprocamente, todo número v que divide a b y r,

$$b = s'v, \qquad r = t'v.$$

divide también a a, ya que es $a = bq + r = s'vq + t'v = (s'q + t')v$. Por consiguiente, *todo* divisor común de a y b es también divisor común de b y r, y recíprocamente. Por tanto, siendo el conjunto de *todos* los divisores comunes a a y b idéntico con el conjunto de *todos* los divisores comunes a b y r, el *mayor* de los divisores comunes a a y b debe ser igual al mayor de los divisores comunes a b y r, lo que demuestra [3]. Veremos inmediatamente la utilidad de esta relación.

Volvamos a la cuestión de hallar el máximo común divisor de 1804 y 328. Por división entera obtenemos

$$\begin{array}{r|l} 1804 & 328 \\ \cline{2-2} 1640 & 5 \\ \cline{1-1} 164 & \end{array}$$

de donde resulta

$$1804 = 5 \cdot 328 + 164.$$

De aquí se sigue, en virtud de [3],

$$(1804,\ 328) = (328,\ 164).$$

Se observa que el problema de hallar (1804, 328) ha sido reducido a otro análogo que se refiere a números más pequeños. Continuemos el proceso; de

$$\begin{array}{r|l} 328 & 164 \\ \cline{2-2} 328 & 2 \\ \cline{1-1} 0 & \end{array}$$

resulta $328 = 2 \cdot 164 + 0$, de modo que $(328, 164) = (164, 0) = 164$. En consecuencia, $(1804, 328) = (328, 164) = (164, 0) = 164$, que es el resultado buscado.

Este proceso para hallar el máximo común divisor de dos números está expuesto en forma geométrica en los *Elementos* de Euclides. Para enteros arbitrarios a y b, no simultáneamente nulos, dicho proceso puede ser descrito aritméticamente en los términos siguientes.

Podemos suponer $b \neq 0$, puesto que $(a, 0) = a$. Entonces, por divisiones sucesivas escribiremos:

$$\begin{array}{ll} a = bq_1 + r_1 & (0 < r_1 < b) \\ b = r_1q_2 + r_2 & (0 < r_2 < r_1) \\ r_1 = r_2q_3 + r_3 & (0 < r_3 < r_2) \\ r_2 = r_3q_4 + r_4 & (0 < r_4 < r_3) \\ \cdots\cdots\cdots & \cdots\cdots\cdots \end{array} \qquad [4]$$

mientras los restos r_1, r_2, r_3, ... son distintos de 0. De las desigualdades que aparecen a la derecha en [4] resulta que los restos sucesivos forman una sucesión decreciente de números positivos:

$$b > r_1 > r_2 > r_3 > r_4 > \cdots > 0. \qquad [5]$$

Por tanto, a lo sumo al cabo de b divisiones (bastante menos, ya que la diferencia entre dos restos sucesivos es en general mayor que 1), se debe llegar al resto 0:

$$r_{n-2} = r_{n-1}q_n + r_n$$
$$r_{n-1} = r_n q_{n+1} + 0.$$

Cuando esto ocurre, sabemos que es

$$(a, b) = r_n ;$$

en otros términos, (a, b) es *igual al último resto mayor que* 0 *en la sucesión* [5]. Esto se deduce de la aplicación sucesiva de la igualdad [3] a las ecuaciones [4], puesto que de las igualdades de [4] resulta

$$(a, b) = (b, r_1), \quad (b, r_1) = (r_1, r_2), \quad (r_1, r_2) = (r_2, r_3)$$
$$(r_2, r_3) = (r_3, r_4), \ldots, (r_{n-1}, r_n) = (r_n, 0) = r_n.$$

Ejercicio: Aplíquese el algoritmo de Euclides para hallar el máximo común divisor de: *a)* 187, 77; *b)* 105, 385; *c)* 245, 193.

A partir de las ecuaciones [4] puede obtenerse una importante propiedad de (a, b). *Si es* d = (a, b), *existen dos enteros, positivos o negativos,* k *y* l *tales que*

$$d = ka + lb. \qquad [6]$$

Para probarlo, consideremos la sucesión [5] de restos. De la primera ecuación en [4] resulta

$$r_1 = a - q_1 b,$$

de modo que r_1 puede escribirse en la forma $k_1 a + l_1 b$ (en este caso $k_1 = 1$, $l_1 = -q_1$). De la ecuación siguiente en [4] se obtiene

$$r_2 = b - q_2 r_1 = b - q_2(k_1 a + l_1 b) = (-q_2 k_1)a + (1 - q_2 l_1)b = k_2 a + l_2 b.$$

Evidentemente, este proceso puede repetirse para los restos sucesivos r_3, r_4, ... hasta llegar a una relación

$$r_n = ka + lb,$$

como queríamos probar.

Como ejemplo, consideremos el algoritmo de Euclides para hallar (61, 24); el máximo común divisor es 1, y la representación deseada para 1 puede calcularse a partir de las ecuaciones

$$61 = 2 \cdot 24 + 13; \quad 24 = 1 \cdot 13 + 11; \quad 13 = 1 \cdot 11 + 2;$$
$$11 = 5 \cdot 2 + 1; \quad 2 = 2 \cdot 1 + 0.$$

De la primera de estas ecuaciones obtenemos

$$13 = 61 - 2 \cdot 24;$$

de la segunda,

$$11 = 24 - 13 = 24 - (61 - 2 \cdot 24) = -61 + 3 \cdot 24;$$

de la tercera,

$$2 = 13 - 11 = (61 - 2 \cdot 24) - (-61 + 3 \cdot 24) = 2 \cdot 61 - 5 \cdot 24,$$

y de la cuarta,

$$1 = 11 - 5 \cdot 2 = (-61 + 3 \cdot 24) - 5(2 \cdot 61 - 5 \cdot 24) = -11 \cdot 61 + 28 \cdot 24.$$

2. Aplicación al teorema fundamental de la aritmética.—El hecho de que $d = (a, b)$ pueda escribirse siempre en la forma $d = ka + lb$ puede utilizarse para dar una demostración del teorema fundamental de la aritmética, independiente de la dada en la página 30. Primero probaremos, como lema, el corolario de la página 31, y luego, a partir de este lema, deduciremos el teorema fundamental, invirtiendo así el orden de la demostración anterior.

LEMA: *Si un número primo* p *divide a un producto* ab, *divide a uno de los factores* a, b.

Si p no divide a a, por ser primo p, se debe tener $(a, p) = 1$, ya que los únicos divisores de p son p y 1. Por consiguiente, se podrán encontrar dos enteros k y l tales que

$$1 = ka + lp.$$

Multiplicando los dos miembros de esta igualdad por b, se obtiene

$$b = kab + lpb.$$

Ahora bien: si p divide a ab, se puede escribir

$$ab = pr$$

de modo que

$$b = kpr + lpb = p(kr + lb),$$

de donde resulta evidente que p divide a b. Hemos demostrado así que si p divide a ab y no divide a a, debe dividir a b; por consiguiente, en todo caso p dividirá a a o a b si divide a ab.

La extensión de este resultado a productos de más de dos factores es inmediata; p. ej., si p divide a abc, aplicando dos veces el lema podemos demostrar que p debe dividir al menos a uno de los enteros a, b o c. Ya que si p no divide ni a a, ni a b, ni a c, no puede dividir a ab y, en consecuencia, tampoco puede dividir a $(ab)c = abc$.

Ejercicio: La extensión de este razonamiento a productos de cualquier número n de enteros requiere el uso tácito o expreso del principio de inducción matemática. Complétense los detalles de tal razonamiento.

Del resultado anterior se sigue de modo inmediato el teorema fundamental de la aritmética. Supongamos que se tienen dos descomposiciones de un entero N en producto de números primos:

$$N = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s .$$

Puesto que p_1 divide al segundo miembro de estas igualdades, debe dividir también al tercero, y, por tanto, por el ejercicio anterior, debe dividir a uno de los factores q_k. Pero q_k es primo; por consiguiente, p_1 debe ser igual a q_k. Suprimiendo este factor común en los dos últimos miembros, resultará que p_2 debe dividir a uno de los q_t restantes, y en consecuencia ser igual a él. Suprimiendo p_2 y q_t, procederíamos análogamente con p_3, ..., p_r. Al final de este proceso se habrán suprimido todas las p, dejando solamente la unidad en el segundo miembro. En el último miembro no podrá quedar ninguna q, puesto que todas las q son mayores que uno. Por consiguiente, las p y las q aparecen en parejas de números iguales, lo que prueba que, salvo quizá el orden de los factores, las dos descomposiciones son idénticas.

3. La función φ de Euler. De nuevo el teorema de Fermat.—Dos enteros a y b se llaman *primos relativos* cuando su máximo común divisor es 1:

$$(a, b) = 1.$$

Por ej., 24 y 35 son primos relativos, mientras que 12 y 18 no lo son. *Si* a *y* b *son primos relativos, existen dos enteros positivos o negativos* k *y* l *tales que*

$$ka + lb = 1.$$

Esto se sigue de la propiedad de (a, b) establecida en la página 53.

Ejercicio: Demuéstrese el teorema siguiente: *Si un entero* r *divide a un producto* ab *y es primo relativo con* a, r *divide a* b. *(Indicación: si* r *es* primo relativo con *a*, existen dos enteros *k* y *l* tales que

$$kr + la = 1.$$

Multiplíquense los dos miembros de esta igualdad por *b*.) Este teorema incluye como caso particular el lema de la página 54, ya que un número primo *p* es primo relativo con un entero *a* cuando, y solamente entonces, *p* no divide a *a*.

Para todo entero positivo *n*, designemos con $\varphi(n)$ *el número de enteros primos con n y menores que* n. Esta función, introducida por Euler, es una función de gran importancia en teoría de números. Los valores de $\varphi(n)$ para los primeros valores de *n* pueden calcularse fácilmente:

$\varphi(1) = 1$ puesto que 1 es primo relativo con 1,
$\varphi(2) = 1$ » 1 » » con 2,
$\varphi(3) = 2$ » 1 y 2 son primos relativos con 3,
$\varphi(4) = 2$ » 1 y 3 » » con 4,
$\varphi(5) = 4$ » 1, 2, 3, 4 » » con 5,
$\varphi(6) = 2$ » 1, 5 » » con 6,
$\varphi(7) = 6$ » 1, 2, 3, 4, 5, 6 son primos relativos con 7,
$\varphi(8) = 4$ » 1, 3, 5, 7 » » con 8,
$\varphi(9) = 6$ » 1, 2, 4, 5, 7, 8 » » con 9,
$\varphi(10) = 4$ » 1, 3, 7, 9 » » con 10.
etc.

Se observa que $\varphi(p) = p - 1$ si *p* es primo, pues un número primo *p* no tiene más divisores que él mismo y la unidad, y, por tanto, es primo relativo con todos los enteros 1, 2, 3, ..., $p - 1$. Si *n* es compuesto, con la descomposición en factores primos

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

donde las *p* representan números primos distintos, cada uno elevado a una cierta potencia, se tiene

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_r}\right)$$

Por ej., puesto que $12 = 2^2 \cdot 3$,

$$\varphi(12) = 12(1 - {}^1/_2)(1 - {}^1/_3) = 12({}^1/_2)({}^2/_3) = 4,$$

como debía resultar, previo el cómputo de primos relativos con 12 y menores que él. La demostración general es completamente elemental, pero no la daremos aquí.

*Ejercicio: Basándose en la función φ de Euler, generalícese el teorema de Fermat de la página 44. El teorema general se enuncia así: *Si* n *es un entero cualquiera, y* a *es primo relativo con* n, *se tiene*

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

4. Fracciones continuas. Ecuaciones diofánticas.—El algoritmo de Euclides para el cálculo del máximo común divisor de dos enteros conduce de modo natural a un importante método para representar el cociente de dos enteros mediante una fracción compuesta; p. ej., aplicando a los números 840 y 611 el algoritmo de Euclides resultan las igualdades

$$\begin{aligned} 840 &= 1 \cdot 611 + 229, & 611 &= 2 \cdot 229 + 153, \\ 229 &= 1 \cdot 153 + 76, & 153 &= 2 \cdot 76 + 1, \end{aligned}$$

las cuales demuestran, incidentalmente, que (840, 611) = 1. De estas ecuaciones se pueden deducir las siguientes expresiones:

$$\frac{840}{611} = 1 + \frac{229}{611} = 1 + \frac{1}{611/229},$$

$$\frac{611}{229} = 2 + \frac{153}{229} = 2 + \frac{1}{229/153},$$

$$\frac{229}{153} = 1 + \frac{76}{153} = 1 + \frac{1}{153/76},$$

$$\frac{153}{76} = 2 + \frac{1}{76}.$$

Por combinación de estas igualdades se obtiene el desarrollo del número racional $\dfrac{840}{611}$ en la forma

$$\frac{840}{611} = 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{76}}}}$$

Una expresión de la forma

$$a = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}}, \qquad [7]$$

donde las a representan enteros positivos, se llama *fracción continua*. El algoritmo de Euclides da un método para representar todo número fraccionario en esta forma.

Ejercicio: Hállense los desarrollos en fracción continua de

$$\frac{2}{5}, \frac{43}{30}, \frac{169}{70}$$

*Las fracciones continuas son de gran importancia en la rama de la aritmética superior conocida con el nombre de análisis diofántico. Una *ecuación diofántica* es una ecuación algebraica, con una o más incógnitas, de coeficientes *enteros* y de la que interesan únicamente las soluciones *enteras*. Tales ecuaciones pueden carecer de soluciones o tener un número finito o infinito de ellas. El caso más sencillo es el de la ecuación diofántica *lineal* con dos incógnitas,

$$ax + by = c, \qquad [8]$$

donde a, b y c son enteros dados, y se buscan soluciones x, y enteras. La solución completa de una ecuación de esta forma puede hallarse por el algoritmo de Euclides.

Para comenzar, hallemos $d = (a, b)$ por el algoritmo de Euclides; se tendrá entonces, para k y l convenientes,

$$ak + bl = d, \qquad [9]$$

de donde resulta que la ecuación [8] admitirá la solución particular $x = k$, $y = l$ en el caso en que sea $c = d$. En general, si c es un múltiplo de d,

$$c = d \cdot q,$$

se obtendría de [9]

$$a(kq) + b(lq) = dq = c,$$

de modo que [8] tiene la solución particular $x = x^* = kq$, $y = y^* = lq$. Recíprocamente: si [8] tiene una solución x, y para un c dado, c debe ser múltiplo de $d = (a, b)$, puesto que q, por dividir a a y b, debe dividir a c. Hemos probado así que, para que la ecuación [8] admita soluciones, es necesario y suficiente que c sea múltiplo de (a, b).

Para determinar otras posibles soluciones de [8] observemos que si $x = x'$, $y = y'$ es otra solución cualquiera, distinta de la $x = x^*$, $y = y^*$ obtenida por el algoritmo de Euclides, las diferencias $x = x' - x^*$, $y = y' - y^*$ forman una solución de la ecuación «homogénea»

$$ax + by = 0. \qquad [10]$$

Pues si es

$$ax' + by' = c \qquad \text{y} \qquad ax^* + by^* = c,$$

restando miembro a miembro estas dos ecuaciones se obtiene

$$a(x' - x^*) + b(y' - y^*) = 0.$$

Ahora bien: la solución general de la ecuación [10] es $x = rb/(a, b)$, $y = -ra/(a, b)$, donde r es un entero cualquiera. (Dejamos al lector la demostración como ejercicio.

Indicación: Divídase por (a, b) y utilícese el ejercicio de la pág. 56.) Resulta entonces de modo inmediato que

$$x = x^* + rb/(a, b), \qquad y = y^* - ra/(a, b).$$

En resumen: La ecuación diofántica lineal $ax + by = c$, donde a, b y c son enteros, tiene soluciones cuando, y únicamente entonces, c es múltiplo de (a, b). En este caso, se puede hallar una solución por el algoritmo de Euclides, y la solución general es de la forma

$$x = x^* + rb/(a, b), \qquad y = y^* - ra/(a, b),$$

donde r es un entero cualquiera.

Ejemplos: La ecuación $3x + 6y = 22$ no admite soluciones enteras, puesto que $(3, 6) = 3$ no divide a 22.

La ecuación $7x + 11y = 13$ tiene la solución particular $x = -39$, $y = 26$, obtenida en la forma siguiente:

$$11 = 1 \cdot 7 + 4, \qquad 7 = 1 \cdot 4 + 3, \qquad 4 = 1 \cdot 3 + 1, \qquad (7, 11) = 1.$$
$$1 = 4 - 3 = 4 - (7 - 4) = 2 \cdot 4 - 7 = 2(11 - 7) - 7 = 2 \cdot 11 - 3 \cdot 7.$$

De donde

$$7 \cdot (-3) + 11(2) = 1,$$
$$7 \cdot (-39) + 11(26) = 13.$$

Las demás soluciones vienen dadas por las fórmulas

$$x = -39 + 11r, \qquad y = 26 - 7r,$$

siendo r un entero arbitrario.

Ejercicio: Resuélvanse las ecuaciones diofánticas: *a)* $3x - 4y = 29$; *b)* $11x + 12y = 58$; *c)* $153x - 34y = 51$.