NÚMEROS EN NÚM3ROS

Capi Corrales (*)

1. INTRODUCCIÓN

En el capítulo 5 de la serie de televisión *Numb3rs*, la hija de un matemático es raptada por unos secuestradores que creen que su padre ha demostrado la Hipótesis de Riemann (uno de los llamados *problemas del milenio*, siete problemas matemáticos por cada uno de los cuales el Clay Institute of Mathematics ofrece un millón de dólares a quien logre demostrarlo). Según explica Charlie, hermano matemático del agente del FBI encargado de resolver el caso, los secuestradores quieren utilizar la demostración del padre para romper los códigos de cifrado que se utilizan en las transacciones bancarias e Internet, y acceder así a cuentas bancarias. La seguridad de los cifradores en cuestión, explica, se basan en lo difícil que es factorizar números grandes en producto de números primos, y la Hipótesis de Riemann puede proporcionar una manera rápida de llevar a cabo esa factorización. Concretamente, nos dice que "la Hipótesis de Riemann es una herramienta para encontrar números primos grandes. Con su trabajo se pueden encontrar números primos grandes usando la criba de Eratóstenes".

Más adelante, el mismo Charlie es capaz de detectar a golpe de vista un error en los cálculos del matemático padre de la niña raptada y, tras hacerle admitir que se había equivocado y que no tenía una demostración válida de la Hipótesis en cuestión, le convence para inventar juntos un algoritmo falso con el que montan una trampa para capturar a los secuestradores.

Es cierto que la seguridad de los criptosistemas utilizados en Internet y bancos, lo mismo que muchos de los algoritmos más comunes para la construcción de cifradores se basa en la teoría de los números primos (específicamente, en la dificultad de factorizar números grandes en producto de primos). También es verdad que la Hipótesis de Riemann es un problema abierto en matemáticas que está relacionado con la teoría de los números primos (concretamente con la probabilidad de un número dado de ser primo), y por cuya resolución el Instituto Clay ofrece un millón de dólares. Sin embargo, la conexión entre ambos hechos está mucho menos clara de lo que los guionistas de la película nos quieren hacer creer.

2. CÓDIGOS: EL ALGORITMO RSA

Empecemos por lo más fácil de explicar, la parte de los códigos de cifrado. La característica del nuevo tipo de cifradores, llamados técnicamente *algoritmos criptográficos de clave pública*, al que pertenece el algoritmo RSA mencionado en *Numb3rs*, es que podemos dar públicamente toda la información necesaria para que quien quiera nos escriba mensajes codificados, sin por ello desvelar la información de cómo leerlos. Antes de seguir adelante, reflexionemos brevemente sobre esto.

Con un tipo de algoritmo más antigüo, si yo le diese a alguien la información necesaria para que me escribiese un mensaje, le estaría dando también la posibilidad de leer todos los mensajes que yo reciba codificados con él, pues la información necesaria para codificar es suficiente para descodificar. Esto no ocurre con los algoritmos de clave pública, y de ahí su enorme utilidad. Una compañia como Amazon.com, por ejemplo, que vende libros a través de la red a compradores de todo el planeta, quiere poder recibir mensajes codificados de mucha gente con sus números de tarjetas de crédito, y, a la vez, asegurarse de que nadie fuera de la compañía pueda leer esos mensajes.

(*) Profesora del Dpto. de Álgebra de la Universidad Complutense de Madrid.

En el algoritmo RSA cualquier persona que conozca un cierto número r (se tratará de un número realmente grande, un entero con miles de dígitos) puede codificar un mensaje m sin más que escribir m como un número que elevará a cierta potencia, obteniendo a continuación su resto al dividir por r. Ese resto será su mensaje codificado. Sin embargo,quien quiera descodificar el mensaje y hallar el m original, necesitará conocer los factores primos del número r y, si r es suficientemente grande, encontrarlos podría requerir décadas e incluso siglos. Ahí es donde la dificultad para factorizar números grandes en producto de sus factores primos entra en el juego.

Veamos con un ejemplo sencillo cómo funciona el algoritmo RSA.

Paso 1: Si quiero poder recibir un mensaje, necesito elegir tres números: dos primos p y q y un tercer número r que no tenga factores comunes con $(p - 1)\,(q - 1)$. Puedo elegir, por ejemplo, $p = 11$, $q = 3$ y $r = 3$, ya que 3 y $20 = (11 - 1)\,(3 - 1)$ no tienen factores comunes. Estos números son, claramente, demasiado pequeños como para que el algoritmo sea seguro (cualquiera puede factorizar sin problemas números pequeños) pero nos sirven como ilustración de cómo funciona el RSA.

Paso 2: A continuación debo encontrar un número $d < (p - 1)\,(q - 1)$ con la propiedad de que al dividir rd por $(p - 1)(q - 1)$ se obtiene 1 como resto. En nuestro ejemplo es $d = 7$, pues $3 \times 7 = 21$ produce un resto 1 al ser dividido por 20.

Paso 3: Esta es la parte más interesante del algoritmo. Anuncio públicamente dos números: el producto $p \times q$ = n y r. Sin embargo, mantengo secretos los factores primos p y q de n, y el número d. Así pues, anuncio públicamente que $p \times q = 33$ y $r = 3$ y oculto que $p = 11$, $q = 3$ y $d = 7$.

Paso 4: Aquí interviene quien me envía un mensaje. Supongamos que eres tú quien me quiere enviar un mensaje. Para empezar, tu mensaje tiene que ser también un número (todo en los ordenadores se escribe mediante números). La única restricción es que tu mensaje ha de ser un número m menor que $n = p \times q$. En las situaciones reales esta restricción no plantea ningún problema pues p y q se eligen enormes. Supongamos, en nuestro ejemplo, que el mensaje que deseas enviarme es el número $m = 18$. Para codificar tu mensaje primero lo elevas a la potencia $r = 3$ y luego divides el resultado por $n = 33$.

El resto que te queda al dividir $18^3 = 5.832$ por 33, esto es, 24, es el número que tú me mandas [(1)]. Quieres que yo reciba el número 18, pero me lo envías codificado, disfrazado, como el 24. De esta manera, cualquier otra persona que conozca los dos números que he desvelado públicamente (el $n = 33$ y el 3), aunque sepa que tu mensaje codificado es el 24 no podrá descodificarlo si no conoce también los números que yo he mantenido en secreto (los factores de 33 y el número d; recordemos que aunque en este ejemplo 33 se puede factorizar sin problemas, y el número d calcular por la cuenta de la vieja, en situaciones reales este no será el caso.

Paso 5: Finalmente, para descodificar el mensaje y devolverlo a su forma original, necesito elevar el número que me has enviado a la potencia d y obtener el resto cuando el resultado se divide por $p \times q$. En nuestro ejemplo esto implica dividir 24^7 por 33 y quedarnos con el resto que es, precisamente 18.

Aunque nuestro ejemplo es un poco absurdo, pues si anuncio públicamente el número 33, no es ningún secreto que sus factores son 11 y 3, en los casos reales los número involucrados son tan enormes que se necesitarían cientos de años para romper el código y extraer la información codificada, y para entonces ésta ya no serviría de nada.

3. CONTANDO PRIMOS

En su libro *Elementos* (300 a.n.e.), Euclides demostró que hay una infinidad de números primos. Concretamente (recordemos que en la matemática griega se evitaba el infinito actual),

demostró que ninguna lista finita de primos {2, 3, 5, 7, . . . , p} los contiene todos. Basta considerar el número

$$N = 2 \times 3 \times 5 \times 7 \times 11 \times \ldots \times p + 1$$

que, o bien será él mismo primo, o bien tendrá un factor primo q distinto de 2, 3, 5, 7, . . . , p.

En principio, el argumento de Euclides nos dice cómo encontrar números primos nuevos: multipliquemos los que conocemos, sumemos uno al producto y factoricemos el número resultante. Por ejemplo, $N = 2 \times 3 + 1$ produce el nuevo primo 7, y $N = 2 \times 3 \times 5 + 1$ produce el nuevo primo 31. Desafortunadamente, pasados más o menos los doscientos primeros primos, el número N es tan grande que factorizarlo es imposible. (Curiosidad: el primo más grande que se conoce es $P = 2^{2596495} - 1$ con 7.816.230 dígitos, que se demostró primo en febrero 2005. Metiendo este número en *google* se puede leer su historia).

Saber que hay una infinidad de primos es sólo el primer paso. El siguiente es *cuantificar* esta infinidad, intentando saber, por ejemplo, el número de primos que hay menores que cualquier número N dado. Definimos, pues,

$$\pi(x) = \text{número de primos} \leq x$$

La herramienta más útil que se conoce para contar primos es la *Criba de Eratóstenes*, descubierta por Eratóstenes de Cirene (276-194 a.n.e.). La Criba de Eratóstenes es un *algoritmo* para eliminar los números compuestos entre 1 y x, y se basa en la observación de que si N es menor o igual a x y no es divisible por ningún primo menor o igual que $\sqrt{x}$, entonces N es primo, Comenzamos por hacer un listado de todos los enteros entre 1 y x, y eliminar de la lista todos los múltiplos de 2. A continuación borramos los múltiplos de 3, después los múltiplos de 5, etc., hasta que todos los múltiplos de los primos menores o iguales que $\sqrt{x}$ hayan desparecido de la lista. Los números que hayan sobrevivido a la criba serán todos primos.

Si pasamos por la Criba de Eratóstenes los números entre 1 y 1.000.000.000.000.000 y analizamos la tabla de primos resultante, contamos 1.177.209.242.304 parejas de primos gemelos (i.e., la distancia entre ellos es 2) menores que 1.000.000.000.000.000. ¿Hay una infinidad de parejas de primos gemelos? Respuesta: No se sabe.

Muchas otras preguntas surgen de analizar esta tabla. ¿Hay una infinidad de parejas de primos cuya diferencia es 4? Respuesta: No se sabe. Todo número par menor o igual que 1.000.000.000.000.000 es suma de dos primos. ¿Es todo número par mayor que 2 la suma de dos primos? Respuesta: No se sabe. ¿Hay una infinidad de primos de la forma un cuadrado perfecto más 1? Respuesta: No se sabe. Dado un número N, ¿existe una fórmula que produzca el menor primo mayor que N? Respuesta: No se sabe.

La Criba de Eratóstenes nos permite evaluar la función $\pi(x)$ para valores pequeños de x: se trata de una función escalonada que da un salto 1 cada vez que aparece un primo nuevo. Sin embargo, para valores grandes de x, es imposible calcular el valor de $\pi(x)$ con exactitud, pero podemos intentar hacer una estimación de su valor basándonos en el patrón de su comportamiento para valores conocidos.

Aunque en principio no sabemos cuándo un número va a ser primo, y la distribución de los números primos en la recta real es bastante errática, lo cierto es que la función $\pi(x)$, pese a tener pequeñas oscilaciones, crece con bastante regularidad, y esta regularidad es mucho más extrema cuanto mayores son los valores de la variable x. Veamos una tabla de valores de $\pi(x)$.

x	$\pi(x)$	$x/\pi(x)$
10	4	2,5
100	25	4,0
1.000	168	6,0
10.000	1.229	8,1
100.000	9.592	10.4
1.000.000	78.498	12,7
10.000.000	664.579	15
100.000.000	5.761.455	17,4
1.000.000.000	50.847.534	19,7
10.000.000.000	455.052.512	22,0

Observando la tabla vemos que la razón de x a $\pi(x)$ aumenta aproximadamente en 2,3 cuando pasamos de una potencia de 10 a la siguiente: esto es, el el logaritmo de 10 en base e. Esto nos lleva a conjeturar que

$$\pi(x) \approx \frac{x}{\log x} \qquad (1)$$

donde $\approx$ significa que $\pi(x)$ y $\frac{x}{\log x}$ como funciones sobre **R** están cerca la una a la otra, en el sentido de que

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1$$

4. LA CONTRIBUCIÓN DE GAUSS Y UN PRIMER ENUNCIADO DE LA HIPÓTESIS DE RIEMANN

El primer estudio serio de la función $\pi(x)$ lo llevó a cabo Gauss (1777-1855) entre 1792 y 93, y lo describió en una carta a Encke escrita en 1849. Gauss estudió la densidad de los primos entre 1 y 3.000.000, y su distribución en intervalos de longitud 1.000, y conjeturó la fórmula (1), conocida como *el teorema de los números primos* y demostrada independientemente por Hadamard y de la Vallée Poussin en 1896.

Gauss calculó $\pi(x)$ y $\frac{x}{\log x}$ para $x = 3.000.000$ y obtuvo

$$\pi\,(3.000.000) = 216.745 \qquad \frac{3.000.000}{\log 3.000.000} = 216,\ 971$$

por lo que concluyó que $\frac{x}{\log x}$ aproxima $\pi(x)$ para $x = 3.000.000$ con un error de sólo 226 primos. De hecho el error es 161, aún menor, pues Gauss se equivocó en las cuentas y

$\pi(3.000.000) = 216.816$. En cualquier caso, Gauss supo ver que las funciones $\pi(x)$ y $\dfrac{x}{\log x}$ están muy cerca la una de la otra.

Sin embargo, aunque las gráficas de $\pi(x)$ y de $\dfrac{x}{\log x}$ son bastante parecidas, no se acercan lo suficiente como para explicar la regularidad de $\pi(x)$, por lo que resulta lógico buscar mejores aproximaciones. El propio Gauss encontró una estudiando en las tablas los valores de $\pi(x)$[(2)]. Gauss observó que la frecuencia de primos cerca de un número x grande es casi exactamente $\dfrac{x}{\log x}$, y por lo tanto la probabilidad de que un número grande x elegido al azar sea primo parece ser proporcional a

$$\frac{1}{\log_{10} x} \approx \frac{1}{\text{número de dígitos de } x},$$

observación que de hecho es la idea básica de la teoría de los números primos. Por ejemplo, la probabilidad de que un número de 100 dígitos sea primo es 1/230, mientras que la probabilidad de que un número de 1.000 dígitos sea primo es 1/2.302, etc. Así pues, concluyó Gauss, si estimamos $\pi(x)$ por

$$\pi(x) = \sum_{2 \le n \le x} \textit{Prob (n primo)} + \text{término de error}$$

tendremos la suma logarítmica

$$\pi(x) = \sum_{2 \le n \le x} \frac{1}{\log n} \textit{ Prob (n primo)} + E_1(x)$$

o, lo que es esencialmente lo mismo,

$$\pi(x) = Li(x) + E_2(x),$$

con $Li(x) = \displaystyle\int_2^x \frac{dt}{\log t}$ es la función llamada *logaritmo integral*. $E_1(x)$ y $E_2(x)$ son errores muy parecidos, de hecho,

$$\left| \sum_{2 \le n \le x} \frac{1}{\log n} - \int_2^x \frac{dt}{\log t} \right| \le 2$$

por lo que $\pi(x)$ puede ser aproximado por una suma o por una integral.

Gauss conjeturó que la funciones $Li(x)$ y $\pi(x)$ están muy cerca la una de la otra, y que la probabilidad de que un número grande y arbitrario x sea primo está cerca de $\dfrac{1}{\log x}$.

El paso siguiente es saber exactamente cómo de cerca están estas funciones. Intentar responder a esta pregunta requiere, para empezar, entender la pregunta. *¿Cómo de cerca es cerca?* ¿Qué significa *cerca* en matemáticas? ¿Qué se considera una buena aproximación en teoría de números?

En teoría de números se considera *una buena aproximación* a una aproximación *de orden raíz cuadrada*. Si $a = \pm 10.000$ y lo aproximamos por un número b tal que $|a - b| \le 100$, decimos que tenemos un error de orden raíz cuadrada. Si $a = \pm 1.000.000$ y lo aproximamos por un número b tal que $|a - b| \le 1.000$, decimos que tenemos un error de orden raíz cuadrada. Si a tiene D dígitos y lo aproximamos por un número b con $D/2$ dígitos, decimos que

tenemos un error de orden raíz cuadrada. Si para n suficientemente grande, $|f(n)| \leq n^{0,5} = \sqrt{n}$, decimos que $f(n)$ tiene un tamaño de orden raíz cuadrada. Si para n suficientemente grande, $|f(n) - g(n)| \leq n^{0,5} = \sqrt{n}$, decimos que $f(n)$ y $g(n)$ están a una distancia de orden raíz cuadrada. Siendo algo menos estrictos, podemos definir una noción de *estar a distancia raíz cuadrada* con la enorme ventaja de ser una relación de equivalencia, y seguir siendo una aproximación fina [(3)]:

"Si para cualquier exponente d algo mayor que 0,5 (por ejemplo, 0,501 o 0,5000001, etc.), esto es, 0, 5 + ϵ, se verifica que

$$|\pi(n) - \pi'(n)| \leq n^{0,5+\epsilon},$$

para n suficientemente grande, decimos que $\pi(n)$ y $\pi'(n)$ están a distancia de orden raíz cuadrada."

Ciertamente, cuando Gauss aproximó $\pi(3.000.000)$ por 216.745, con un error de 226 (según pensaba él), se mantuvo muy por debajo del margen raíz cuadrada de error.

Ya podemos formular nuestra pregunta con precisión: ¿Está $Li(x)$ a distancia raíz cuadrada de $\pi(x)$? La respuesta nos la da una de las varias (y equivalentes) maneras de enunciar la *Hipótesis de Riemann*:

HR-1: *La función Li(x) de Gauss está a distancia raíz cuadrada de $\pi(x)$.*

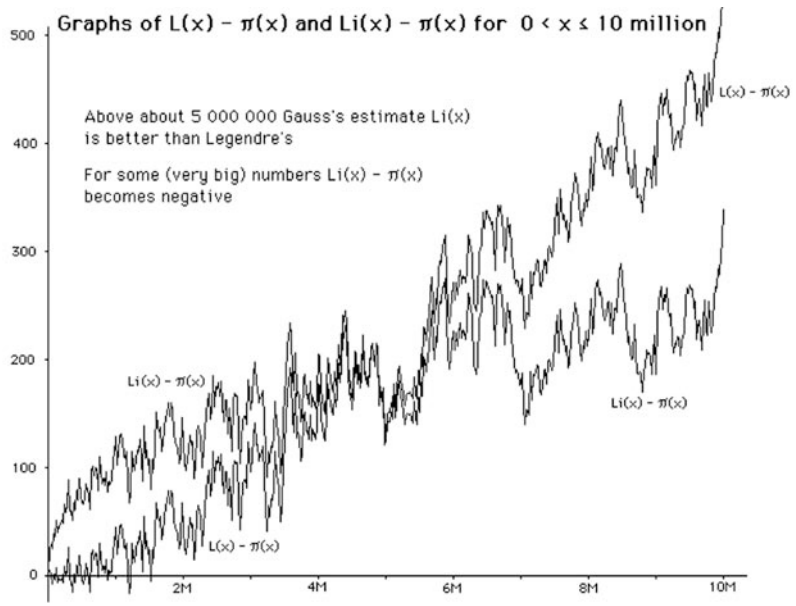
5. LA CONTRIBUCIÓN DE RIEMANN Y UN SEGUNDO ENUNCIADO DE LA HIPÓTESIS DE RIEMANN

Para poder entender la contribución que Riemann (1826-1866) hizo en 1859 al estudio de $\pi(x)$, necesitamos volver sobre la sugerencia de Gauss

$$\pi(x) \approx Li(x),$$

esto es,

$$\pi(x) = Li(x) + E(x).$$



Riemann, alumno de Gauss, consideró la posibilidad de describir con precisión el término de error $E(x)$. El estudio que para ello llevó a cabo de la distribución de los números primos, le llevó a sugerir una nueva función para aproximar $\pi(x)$, al observar que la probabilidad de que un número grande x elegido al azar sea primo es aún más cercana a $1/\log x$ si consideramos no sólo los primos, sino también las potencias de los primos, contando el cuadrado de un primo como medio primo, la potencia cúbica como un tercio de primo, etc., esto es,

$$\pi(x) + \frac{1}{2}\ Li(x^{\frac{1}{2}}) + \frac{1}{3}\ Li(x^{\frac{1}{3}}) + \frac{1}{5}\ Li(x^{\frac{1}{5}}) - \frac{1}{6}\ Li(x^{\frac{1}{6}}) - \frac{1}{7}\ Li(x^{\frac{1}{7}}) \ldots = Li(x)$$

o, equivalentemente,

$$\pi(x) \approx R(x),$$

con

$$R(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\ Li(x^{\frac{1}{n}})$$

donde $\mu(n)$ es la función de Möbius, que toma el valor 0 si n es divisible por el cuadrado de algún primo, 1 si n es el producto de un número par de primos distintos y -1 si n es un producto de una cantidad impar de primos.

La función $R(x)$, denotada así en honor a Riemann, representa una aproximación sorprendentemente buena a $\pi(x)$,

x	$\pi(x)$	$R(x)$
100.000.000	5.761.455	5.761.552
200.000.000	11.078.937	11.079.090
300.000.000	16.252.325	16.252.355
400.000.000	21.336.326	21.336.185
500.000.000	26.355.867	26.355.517
600.000.000	31.324.703	31.324.622
700.000.000	36.252.931	36.252.719
800.000.000	41.146.179	41.146.248
900.000.000	46.009.215	46.009.949
1.000.000.000	50.847.534	50.847.455

Dado que las funciones $Li(x)$ de Gauss y $R(x)$ de Riemann están a distancia de orden raíz cuadrada una de la otra, y que, como ya se ha mencionado, *estar a distancia de orden raíz cuadrada* es una relación de equivalencia, podemos reformular nuestro primer enunciado de la Hipótesis de Riemann:

HR-2: ***Las funciones $Li(x)$ de Gauss y $R(x)$ de Riemann están a una distancia de orden raíz cuadrada de $\pi(x)$.***

6. ANALIZANDO EL TÉRMINO DE ERROR $\pi(x) - R(x)$: LAS FUNCIONES ζ DE EULER Y RIEMANN

Aunque no logró demostrar el teorema de los números primos (ya se ha comentado que lo consiguieron en 1896 Hadamard y de la Vallée Poussin), en su artículo de 1859 ([Riemann, 1859]) Riemann hizo mucho más que dar una aproximación $R(x)$ a $\pi(x)$ más precisa que la $Li(x)$ de Gauss: consiguió dar una fórmula exacta para $\pi(x)$. Concretamente, Riemann estudió con precisión el error

$$R(x) - \pi(x)$$

y logró construir una serie infinita de términos correctores $C_1(x)$, $C_2(x)$, . . . a $R(x)$ de tal manera que las consiguientes *correcciones*

$$R_k(x) = R(x) + C_1(x) + C_2(x) + \ldots + C_k(x)$$

verifican

$$\lim_{k\to\infty} R_k(x) = \pi(x).$$

Riemann utilizó la *función zeta de Euler*, definida por

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

para construir los términos correctores $C_k(x)$. En general, las funciones *zeta* aparecen en teoría de números como series y/o productos infinitos que nos permiten organizar colecciones de datos numéricos de una manera única y compacta. Las propiedades analíticas de la función ayudan a entender la distribución de los datos como un todo, y, a veces, favorecen la emergencia de orden en un comportamiento aparentemente caótico. Un excelente estudio introductorio de estas funciones puede encontrarse en [Goldstein, 1999][(3)].

Euler fue el primero que introdujo este tipo de funciones, y lo hizo para estudiar los números primos y, entre otras cosas, dar una demostración distinta de la de Euclides a la existencia de una infinidad de primos. Veámosla.

Sabemos que la serie armónica

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

diverge, pues si

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$$

es

$$s_{2^m} > 1 + \frac{m}{2}.$$

Sea $p = 2, 3, 5, 7, \ldots$ cualquier primo,

$$0 < \frac{1}{p} < 1 \Rightarrow \frac{1}{1 - 1/p} = 1 + \frac{1}{p} + \frac{1}{p^2} + \ldots;$$

si p_n denota el primo n-ésimo, como consecuencia del teorema fundamental de la aritmética obtenemos

$$\left(\frac{1}{1 - \frac{1}{p_1}} \cdot \frac{1}{1 - \frac{1}{p_2}} \cdot \ldots \cdot \frac{1}{1 - \frac{1}{p_n}} \cdot \ldots \right) = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots$$

y, por lo tanto, hay una infinidad de primos.

Consideremos la situación más general en que $p = 2, 3, 5, 7, \ldots$ es cualquier primo, y $s \geq 1$:

$$0 < \frac{1}{p^s} < 1 \Rightarrow \frac{1}{1 - 1/p^s} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \ldots$$

multiplicando, si p_n denota el primo n-ésimo, obtenemos

$$\lim_{x \to \infty} \left(\frac{1}{1 - \frac{1}{p_1^s}} \cdot \frac{1}{1 - \frac{1}{p_2^s}} \cdot \ldots \cdot \frac{1}{1 - \frac{1}{p_n^s}} \right)$$

$$= 1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots + \frac{1}{n^s} + \ldots = \zeta(s).$$

Euler estudió la función $\zeta(s)$ como función real de variable real,

$$\zeta : \mathbf{R} \to \mathbf{R} \cup \{\infty\}$$

$$s \to \zeta(s),$$

que verifica que si $s > 1$, $\zeta(s) < \infty$, e intentó utilizar esta función para estudiar los números primos. Se trata de una función continua (de hecho, converge uniformemente, pero Euler no conocía el concepto de convergencia uniforme), y esto permitió a Euler utilizar los métodos del cálculo para estudiarla. Sin embargo, considerada como una función real de variable real se trata de un objeto unidimensional, por lo que no tiene suficiente estructura geométrica como para poder desvelar (o codificar) el patrón de distribución de los números primos.

El gran salto lo dió Riemann, al extender $\zeta(s)$ a valores complejos de la variable $s \neq 1$, $s = a + bi$ con $i = \sqrt{-1}$ y $a, b \in \mathbf{R}$. Riemann consideró

$$\zeta : \mathbf{C} \setminus \{1\} \to \mathbf{C}$$

$$s \to \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \qquad (2)$$

Puesto que $n^s = e^{s \log n} = e^{a \log n + ib \log n}$, el valor absoluto de n^s es igual a $e^{a \log n}$, por lo que la serie converge absolutamente para $a = Re(s) > 1$ y uniformemente en todo conjunto compacto de este semiplano, por ejemplo para todo $a \geq a_o > 1$. En particular, la función definida por (2) es holomorfa para $Re(s) > 1$, y para estos valores de s, se tiene el llamado *producto de Euler*:

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

donde el producto recorre todos los primos p.

La función ζ admite una continuación analítica a todo el plano complejo como una función meromorfa que tiene un único polo, situado en $s = 1$, simple y con residuo 1[(4)].

Bernhard Riemann (1826-1866)

7. LOS CEROS DE LA FUNCIÓN ζ DE RIEMANN, Y UN TERCER ENUNCIADO DE LA HIPÓTESIS DE RIEMANN

Con respecto a los ceros de la función ζ, la situación es mucho más complicada. De la relación entre la función ζ y los números de Bernouilli que ofrecemos en el apéndice 3, se deduce la existencia de una infinidad de ceros simples, llamados *los ceros triviales de la función* ζ: si $n \in \mathbf{N}$, $\zeta(-2n) = 0$.

Si $\zeta(s) > 1$, la existencia del producto de Euler garantiza que $\zeta(s) \neq 0$, pues el producto converge y toma valores no nulos. No vamos a entrar aquí en detalles, que pueden encontrarse, por ejemplo, en [Tennenbaum, 1990], pero de hecho ζ no tiene ceros en el semiplano cerrado $Re(s) \geq 1$, y tan solo ceros triviales en el semiplano cerrado $Re(s) \leq 0$. Así pues, la única zona del plano donde la función ζ puede tener ceros no triviales es la franja vertical definida por $\{s = a + bi : 0 < a < 1\}$.

Riemann sabía que en esta franja vertical, llamada *franja crítica*, hay una infinidad de ceros de la función $\zeta(s)$. De hecho, en 1889 Riemann había conseguido calcular la primera docena de los ceros de la función zeta en la franja crítica, y había descubierto que todos ellos tenían como parte real precisamente 1/2.

Ya tenemos todos los *ingredientes* para entender la extraordinaria *receta* de Riemann para construir sus términos correctores $C_k(x)$, receta que nos permitirá formular un tercer enunciado de la Hipótesis de Riemann que involucra los ceros de la función ζ.

Cada uno de los términos $C_k(x)$ se corresponde con un cero $\rho_k = a_k + b_k\, i$ de la función $\zeta(s)$ en la franja crítica. Concretamente,

$$C_k(x) = -R(x^{\rho_k}) - R(x^{\overline{\rho}_k})$$

y el tamaño de $E_k(x)$ viene dado por la parte real a_k de ρ_k, en el sentido de que para valores grandes de x, $|C_k(x)| \leq x^{Re(\rho_k)}$. Por lo tanto, si un cero de la función zeta en la franja crítica tiene la forma $\rho_k = 1/2 + b_k i$, el correspondiente $C_k(x)$ tiene un tamaño de orden raíz cuadrada.

Esto es extraordinariamente importante desde el punto de vista de la estimación del error

$$E(x) = \pi(x) - R(x)$$

pues si los $C_k(x)$ tienen un tamaño de orden raíz cuadrada, esto implica que las resultantes correciones

$$R_t(x) = R(x) + \sum_{k=1}^{t} C_k(x) = R(x) - \sum_{k=1}^{t} R(x^{\rho_k}) + R(x^{\overline{\rho}_k})$$

son aproximaciones de $R(x)$ de orden raíz cuadrada, y, a su vez, $R(x)$ (y, consecuentemente, también $Li(x)$) es una aproximación de orden raíz cuadrada a $\pi(x)$.

Ya podemos reformular nuestro segundo enunciado de *la Hipótesis de Riemann*, que, recordamos, era:

HR-2: Las funciones Li(x) de Gauss y R(x) de Riemann están a una distancia de orden raíz cuadrada de $\pi(x)$.

La nueva formulación, equivalente a la anterior, es:

HR-3: Todos los ceros no triviales de la función zeta están sobre la recta vertical formada por los números complejos con parte real 1/2 , esto es, en mitad de la franja crítica.

8. CONCLUSIÓN

Hay dos problemas distintos, ambos muy importantes para la construcción de códigos, y ambos involucran números primos: factorizar un entero en producto de sus factores primos, y encontrar primos grandes.

La seguridad de los algoritmos criptográficos utilizados en Internet se basa en el primero de ellos, esto es, en la dificultad de factorizar números grandes en producto de primos.

La Hipótesis de Riemann, al conjeturar cuál es la probabilidad de que un número grande arbitrario sea primo, y decirnos hasta qué grado esta probabilidad acierta, está relacionada con el segundo de los problemas que hemos mencionado. Muchos algoritmos para construir primos grandes que funcionan de manera muy eficiente están construidos sobre la base de que la conjetura de Riemann es cierta, algo que se da por hecho en prácticamente toda la comunidad matemática.

Concluimos, pues, que el que la Hipótesis de Riemann sea verdadera o falsa no afecta la seguridad de los códigos utilizados en internet. Cabe la posibilidad de que un día se encuentre una demostración de dicha hipótesis que incluya un algoritmo para factorizar números grandes. Pero dado que, insistimos, factorizar números grandes y encontrar primos grandes son problemas bien distintos, esta posibilidad resulta poco creíble, incluso si viene avalada por los guionistas de Hollywood.

APÉNDICES

Apéndice 1

Utilizando un poco de teoría de funciones, se obtiene que la función de Riemann,

$$R(x) = Li(x) - 1/2\ Li(x^{1/2}) - Li(x^{1/3}) - Li(x^{1/5}) + Li(x^{1/6}) + \ldots$$

$$= \sum_{n=1}^{\infty} \frac{\mu(n)}{n} Li(x^{1/n})$$

es una función entera de log x, dada por una serie de potencias que converge rápidamente:

$$R(x) = 1 + \sum_{k=1}^{\infty} \frac{(\ln x)^k}{kk!\zeta\ (k+1)}.$$

Apéndice 2

Cortesía de Catherine Goldstein, [Goldstein, 1999], p. 60:

"Es sencillo dar una continuación analítica a la función ζ en el semiplano Re(s) > 0 (ver [Lang, 1970], p. 157): introducimos la función ζ *alternada*

$$\zeta_2(s) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^s},$$

que, como la función ζ, converge absolutamente para $Re(s) > 1$. Además, por el teorema de las series alternadas, ζ_2 converge también (aunque no absolutamente) en el semiplano abierto $Re(s) > 0$, en el que define una función holomorfa, y se tiene la relación

$$\zeta_2(s) = (2^{1-s} - 1)\ \zeta(s),$$

que inmediatamente permite deducir la continuación analítica de la función ζ a $Re(s) > 0$, con posibles polos si $2^{1-s} = 1$, esto es, para $s = 1 + \frac{2\pi in}{\log 2}$. Claramente, $s = 1$ es un polo, pues $\zeta_2(1) = -\log 2 \neq 0$. Para eliminar los otros polos potenciales, Lang sugiere el truco de considerar la función

$$\zeta_3(s) = -1 - 1/2^s + 2/3^s - 1/4^s - 1/5^s + 2/6^s - \ldots$$

para la que existe una relación del mismo tipo con la función ζ, pero que ofrece $s = 1 + \frac{2\pi in}{\log 3}$ como posibles polos. La única posibilidad común es $s = 1$".

Apéndice 3

Los números de Bernouilli son números racionales definidos por la serie

$$\frac{z}{e^z - 1} = \sum_{m=0}^{\infty} B_m \frac{z^m}{m!},$$

con $B_0 = 0$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$, $B_6 = 1/42$, etc.

Para $r = 1, 2, 3, \ldots$, se tiene:

$$B_{2r+1} = 0$$

$$\zeta(-r) = -\frac{B_{r+1}}{r+1}$$

$$\zeta(2r) = (-1)_{r-1}\, B_2 r 2^{2r-1} \frac{\pi^{2r}}{(2r)!}$$

BIBLIOGRAFÍA:

[E-M] **W. J. Ellison,** 1975: M. Mend`es France, Les Nombres Premiers. Hermann. Paris.

[G] **Catherine Goldstein,** 1999: *Introduction to ζ and L -functions*, en "Cuatrocientos años de matemáticas en torno al ´Ultimo Teorema de Fermat", pp. 59-80. C. Andradas y C. Corrales Rodrigáñez, eds. Editorial Complutense.

[La-1] **Serge Lang,** 1970: *Algeraic Numbre Theory*. Addison-Wesley.

[La-2] **Serge Lang,** 1992: *¿Qué hace un matemático y por qué? Los números primos*, texto de una conferencia dada el 16 de mayo de 1981 en el Palais de la Découverte (Museo de Ciencias de París), publicado en castellano en *El placer estético de las matemáticas* pp. 15-50, A.U. 737.

[Ma] **Barry Mazur,** *Are there still unsolved problems about the numbers 1,2,3,4,...?*, notas de la conferencia dada el 3 de mayo de 2005 en MIT organizada por el Clay Mathematics Institute (ver www.claymath.org).

[Od] **Andrew Odlyzko,** http://www.dtc.umn.edu/ odlyzko/zeta.tables/ (Los 1000.000 ceros de la función zeta con una precisión de $3 \cdot 10^{-9}$).

[Ri] **Bernhard Riemann,** 2000: *Sobre el número de primos menores que una cantidad dada* (1859), trad. al castellano de José Ferreirós en *Riemanniana Selecta*, pp. 79-85. Ediciones del CSIC. Madrid.

[Te] **G. Tenenbaum,** 1990: *Introduction à la théorie analytique et probabiliste des nombres*. Publ. Inst. Elie Cartan, Nancy.

[Za] **Don Zagier,** 1977: *The first 50 million prime numbers*, The mathematical Intelligentzer, 0 7-19.

NOTAS

(1) Con números grandes es imposible llevar a cabo de una vez la exponenciación m^r y a continuación tomar el resto de dividir el número obtenido por n, pues salen cantidades enormes e imposibles de manejar. El cálculo se hace por etapas, elevando al cuadrado sucesivamente y calculando el resto módulo n cada vez.

(2) En 1808 Legendre encontró otra aproximación a $\pi(x)$ especialmente buena,

$$\pi(x) = \frac{x}{\log x - 1{,}08366}$$

(3) Un margen de error de magnitud exactamente raíz cuadrada es, de hecho muy fino, y no es frecuente encontrarlo en las ciencias empíricas ni en las estadísticas de grandes poblaciones. Por ejemplo, en el censo de EEUU que aparece en noviembre de 2006 en *google* (que es del año 2000), se lee que en noviembre del 2000 hubo 2.957.000 desempleados entre la *fuerza laboral civil* de los EEUU con un error estandar de 91.000. Un margen de error de orden raíz cuadrada supondría un error de 2.000 personas como mucho.

(3) En el apéndice 1 a este texto puede encontrarse una primera relación entre las funciones ζ y $R(x)$.

(4) La existencia de esta continuación analítica, así como su construcción, puede encontrarse en bibliografía especializada, por ejemplo [Ellison y Mend`es France, 1975], ó [Tenenbaum, 1990]. Conocer la función ζ para números s con $Re(s) > 0$ es esencialmente suficiente, pues existe una ecuación funcional que relaciona los valores $\zeta(s)$ y $\zeta(1 - s)$. En el apéndice 2 ofrecemos, para lector interesado, una continuación analítica sencilla de la función ζ al semiplano $Re(s) > 0$.

METHODUS
INVENIENDI
LINEAS CURVAS
Maximi Minimive proprietate gaudentes,
SIVE
SOLUTIO
PROBLEMATIS ISOPERIMETRICI
LATISSIMO SENSU ACCEPTI
AUCTORE
LEONHARDO EULERO,
Professore Regio, & Academiæ Imperialis Scientiarum Petropolitanæ Socio.
LAUSANNÆ & GENEVÆ,
Apud MARCUM-MICHAELEM BOUSQUET & Socios.
MDCCXLIV.

Methodus inveniendi lineas curvas. Leonhard Euler (1744)