Capítulo 1

Funciones aritméticas

1.1. Introducción

Ejemplos básicos. Series de Dirichlet. Inversión de Möbius.

Hablar de *funciones aritméticas* en general, no es decir demasiado ya que se conoce bajo esta denominacióna cualquier función cuyo dominio son los naturales de toda la vida, $f : \mathbb{N} \longrightarrow \mathbb{C}$. La mayor parte de las veces la imagen también estará dentro de $\mathbb{N}$, o al menos de $\mathbb{R}$.

Entre las funciones aritméticas tienen especial interés las que dependen de la factorización en primos.

Definición: Se dice que una función aritmética f es *multiplicativa* si $f(nm) = f(n)f(m)$ siempre que n y m sean coprimos, y se dice que es *completamente multiplicativa* si $f(nm) = f(n)f(m)$ es cierto en general.

Si n se descompone en primos como $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ entonces cualquier f multiplicativa verifica

$$f(n) = \prod_{i=1}^{r} f(p_i^{\alpha_i}). \tag{1.1}$$

Por convenio se toma $f(1) = 1$.

Uno de los ejemplos más relevantes de funciones multiplicativas es la *función de Möbius* que tiene la extraña definición $\mu(1) = 1$, $\mu(p) = -1$ y $\mu(p^\alpha) = 0$ para $\alpha > 1$. Esto es

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ (-1)^r & \text{si } n = p_1 p_2 \cdots p_r \ \text{ (primos distintos)} \\ 0 & \text{en otro caso} \end{cases}$$

Por tanto μ^2 es la función característica de los *libres de cuadrados* (los no divisibles por ningún cuadrado mayor que 1).

Hay otras funciones sencillas multiplicativas (aunque comprobar que lo son por ahora no sea obvio) que tienen un puesto de palco en los textos de teoría de números:

- $d(n) = \#\{d \,:\, d|n\}$ (función divisor)
- $\sigma(n) = \sum_{d|n} d$
- $\phi(n) = \{1 \le m \le n \,:\, (m,n) = 1\}$ (función ϕ de Euler)
- $\lambda(n) = (-1)^{\alpha_1+\alpha_2+\cdots+\alpha_r}$ si $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ (función de Liouville).

De todas estas funciones, sólo la de Liouville es completamente multiplicativa.

Entre las funciones aritméticas no multiplicativas tiene especial interés la llamada *función de von Mangoldt*

$$\Lambda(n) = \begin{cases} \log p & \text{si } n = p^\alpha \\ 0 & \text{en otro caso} \end{cases}$$

También cabe mencionar dos funciones relacionadas con la distribución de los primos:

$$\pi(n) = \#\{p \le n\} = \sum_{p \le n} 1, \qquad \psi(n) = \sum_{m \le n} \Lambda(m).$$

En realidad estas funciones se suelen extender a funciones no aritméticas $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$ escribiendo $\pi(x) = \pi([x])$ y $\psi(x) = \psi([x])$, con $[\,\cdot\,]$ denotando la parte entera. Seguiremos la misma política en otras funciones que expresan sumas o promedios.

A cada función aritmética f se le puede asociar una *serie de Dirichlet* formal[1]

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

La serie de Dirichlet asociada a la función idénticamente uno es la famosa función ζ *de Riemann*

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

(en rigor esto sólo define el trozo de la función ζ correspondiente a $\Re(s) > 1$).

De acuerdo con lo anterior, una función es multiplicativa si y sólo si se tiene la igualdad formal (ejercicio)

$$F_f(s) = \prod_p D_{f,p}(s) \qquad \text{con} \qquad D_{f,p}(s) = 1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \frac{f(p^3)}{p^{3s}} + \ldots \tag{1.2}$$

Se dice que esta expresión es un *producto de Euler*. En el caso de la función ζ, para $\Re(s) > 1$ se puede sumar la progresión geométrica $D_{f,p}(s)$ y llegar a la igualdad de series convergentes

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1} \tag{1.3}$$

[1]Aquí "formal" significa que nos despreocupamos de la convergencia. Para algunas funciones puede que no converja para ningún valor.

que es el producto de Euler por antonomasia. Usando un poco de análisis[2] se tiene $\zeta(2) = \pi^2/6$, en particular

$$\frac{6}{\pi^2} = \prod_p (1 - p^{-2}).$$

Impresionante, a un lado la razón entre longitud de la circunferencia y diámetro, y al otro los primos. Además empleando que $6/\pi^2$ es irracional ([Sp] cap. 16) se concluye que ¡hay infinitos primos! Euclides y los pitagóricos temblarían de terror.

Dadas dos funciones aritméticas f y g se define su *convolución* como

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d).$$

Nótese que es conmutativa (ejercicio).

Es sencillo comprobar que

$$D_f(s)D_g(s) = D_{f*g}(s) \quad \text{y que} \quad D_{f,p}(s)D_{g,p}(s) = D_{f*g,p}(s).$$

Por tanto, empleando (1.2), hemos probado:

Proposición 1.1.1 *Si f y g son multiplicativas entonces $f * g$ también lo es y su serie de Dirichlet es $D_f(s)D_g(s)$.*

Antes de poner ejemplos, concedámonos una pausa para digerir las definición y dar un vistazo a la brújula de la intuición. Las funciones generatrices son un instrumento natural para tratar problemas aditivos. Así, cuando uno considera $f(x) = \sum_{a\in A} x^a$ (la función generatriz de A) su cuadrado tiene como coeficientes el número de representaciones como $a + a'$, y si multiplicamos f por $\sum_{b\in B} x^b$ se obtiene el número de representaciones de la forma $a + b$. En problemas multiplicativos estas funciones generatrices no sirven, porque al multiplicar se suman los exponentes, no se multiplican. La idea es intercambiar el papel que desempeñan a y x, si el exponente es el mismo, las bases se multiplicarán y eso es justamente una serie de Dirichlet. Lo de escribir $-s$ en lugar de x, es inofensivo.

Ejemplo. La función divisor es multiplicativa por ser convolución de dos funciones idénticamente uno, $d(n) = \sum_{d|n} 1$, y σ es multiplicativa por ser convolución con la identidad. En general $\sigma_k(n) = \sum_{d|n} d^k$ es multiplicativa (convolución de $f(n) = n^k$ y $g = 1$). Por ejemplo, la suma de los cuadrados de los divisores de un millón es

$$\sigma_2(10^6) = \sigma_2(2^6)\sigma_2(5^6) = (1^2 + 2^2 + \cdots + 2^{12})(1^2 + 5^2 + \cdots + 5^{12}) = (2^{13} - 1)(5^{13} - 1)/4.$$

Cuando (1.2) se aplica a la función μ se tiene $D_\mu(s) = \prod(1 - p^{-s})$ que combinado con (1.3) se puede leer como $\zeta(s)D_\mu(s) = 1$. Esta humilde relación permite invertir la convolución con 1:

$$\zeta(s)D_f(s) = D_{1*f}(s) \quad \Rightarrow \quad D_f(s) = D_\mu(s)D_{1*f}(s).$$

Escrito sin tantos símbolos:

[2]Por ejemplo la fórmula producto del análisis complejo $\operatorname{sen}(\pi x) = \pi x \prod_{n=1}^{\infty}(1 - x^2/n^2)$, o la identidad de Parseval aplicada al desarrollo de Fourier de la parte fraccionaria (véase [Co] para una prueba elemental).

Proposición 1.1.2 (Inversión de Möbius) *Dada una función aritmética f, sea la función $F(n) = \sum_{d|n} f(n)$, entonces $f(n) = \sum_{d|n} \mu(d)F(n/d)$.*

Observación: Desde el punto de vista histórico la inversión de Möbius original era la fórmula $F(x) = \sum_{n\leq x} f(x/n) \Rightarrow f(x) = \sum_{n\leq x} \mu(n)F(x/n)$ (véase [El] Th. 1.9).

Ejemplo. Se cumple $n = \sum_{d|n} \#\{1 \leq k \leq n \ : \ (n,k) = d\}$ (porque todo número tiene algún máximo común divisor con n). Entonces

$$n = \sum_{d|n} \#\{1 \leq l \leq n/d \ : \ (n/d, l) = 1\} = (1 * \phi)(n)$$

de donde $\phi(n) = (\mu * \mathrm{Id})(n) = \sum_{d|n} \mu(d)n/d$ y se concluye que ϕ es multiplicativa. De ello y (1.1) se sigue la fórmula bien conocida:

$$\phi(p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}) = \prod_j \big(p_j^{\alpha_j - 1}(p_j - 1)\big).$$

La importancia de la función Λ de von Mangoldt proviene de que da los coeficientes de la serie de Dirichlet de $-\zeta'/\zeta$, que es crucial en la demostración del teorema de los números primos. Tomando logaritmos en (1.3), derivando y usando la suma de una progresión geométrica (todo esto tiene sentido en $\Re(s) > 1$)

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

Multiplicando por $\zeta(s)$ en ambos miembros, se deduce

$$\log n = \sum_{d|n} \Lambda(n)$$

y por inversión de Möbius,

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}.$$

Simbólicamente $\log = 1 * \Lambda$ y $\Lambda = \mu * \log$. La primera igualdad conecta una función de las de cálculo de toda la vida (el logaritmo) con algo que depende de los primos (la función Λ). Sólo jugando adecuadamente con esta relación ya es posible avanzar un poco en el estudio de la distribución de los primos.

1.2. Promedios de funciones aritméticas

Sumación por partes. Fórmula de sumación de Abel. Teorema de Wirsing.

Los valores que toman algunas de las funciones aritméticas más habituales son bastantes caóticas pero podemos "domesticarlas" tomando promedios, lo que nos puede

ayudar a formarnos una idea global más acertada. Si la función viene dada por una suma sobre los divisores es pertinente emplear la identidad elemental:

$$\sum_{n\le x}\sum_{d|n} f(d) = \sum_{n\le x} f(n)\big[\frac{x}{n}\big] \tag{1.4}$$

cuya prueba se reduce a invertir el orden de sumación.

Raramente sabremos evaluar los promedios, y consideraremos suficiente una aproximación con un buen término de error, es ahí donde entra la notación de Landau.

Aplicando (1.4) a $f(n) = \Lambda(n)$ y teniendo en cuenta que $\log = 1 * \Lambda$, se obtiene un ejemplo espectacular:

$$\sum_{n\le x} \log n = \sum_{n\le x} \Lambda(n)\big[\frac{x}{n}\big].$$

La primera suma se puede aproximar por la integral [Sp] cap. 22 (más adelante veremos algo más preciso) que es $x\log x + O(x)$. Empleando que $[t/n] - 2[t/2n] = 1$ para $t/2 < n \le t$ se deduce $\sum_{t/2<n\le t}\Lambda(n) = O(t)$ y de aquí $\sum_{n\le x}\Lambda(n) = O(x)$. Esto permite quitar sin gran peligro la parte entera en el segundo miembro y obtener la llamativa *fórmula de Mertens*

$$\sum_{p\le x}\frac{\log p}{p} = \log x + O(1) \tag{1.5}$$

donde se ha acumulado la contribución de p^k con $k > 1$ en la constante.

Ejemplo. La función ϕ, como función multiplicativa que es, toma valores ligados a la factorización, lo que causa variaciones tan drásticas como $\phi(210) = 48$, $\phi(211) = 210$. Introduciendo la fórmula $\phi(n) = \sum_{d|n}\mu(d)n/d$ en (1.4) se obtiene

$$\sum_{n\le x}\frac{\phi(n)}{n} = \sum_{d\le x}\frac{\mu(d)}{d}\big[\frac{x}{d}\big] = \sum_{d\le x}\frac{\mu(d)}{d}\big(\frac{x}{d} + O(1)\big) = x\sum_{d\le x}\frac{\mu(d)}{d^2} + O(\log x).$$

Recordando que $D_\mu(s) = 1/\zeta(s)$, $\zeta(2) = \pi^2/6$, y empleando $\sum_{d>x} d^{-2} = O(x^{-1})$ se llega a

$$\sum_{n\le x}\frac{\phi(n)}{n} = \frac{6}{\pi^2}x + O(\log x).$$

Esta fórmula no está en el limbo de las aproximaciones asintóticas que sólo aproximan para infinitos muy infinitos. Con un pequeño programa se tiene que $\sum_{n\le 100}\phi(n)/n = 60'83268\ldots$ que no está lejos, en términos relativos de $600/\pi^2 = 60'7927\ldots$

Se podría juzgar, con razón, que un promedio de ϕ más natural es $x^{-1}\sum_{n\le x}\phi(n)$, en vez de la expresión del ejemplo anterior. Como $\phi(n)$ es "en media" $6n/\pi^2$ entonces $x^{-1}\sum_{n\le x}\phi(n)$ debe ser (?!) como $x^{-1}\sum_{n\le x} 6n/\pi^2 \sim 3x/\pi^2$. Esto es cierto y también demostrable. El artificio para poner o quitar coeficientes no oscilatorios se conoce bajo el nombre genérico de *sumación por partes*. Merece la pena detenerse en dos versiones: una discreta y otra continua (hasta cierto punto).

Lema 1.2.1 (Sumación por partes) *Se cumple la identidad*

$$\sum_{n=1}^{N} a_n b_n = a_N S_N + \sum_{n=1}^{N-1} (a_n - a_{n+1}) S_n$$

donde $S_n = \sum_{k=1}^{n} b_k$.

Demostración: Basta escribir $b_n = S_n - S_{n-1}$ y agrupar convenientemente los términos. □

Un uso habitual de la sumación por partes es deshacerse de coeficientes monótonos.

Corolario 1.2.2 *Si* $(a_n)_{n=1}^{N}$ *es una sucesión real monótona no creciente y positiva, entonces*

$$\big|\sum_{n=1}^{N} a_n b_n\big| \leq a_1 \sup_{1\leq n\leq N} |S_n|.$$

En la variante "continua" aparece una integral.

Lema 1.2.3 (Lema de Abel) *Sea* $(c_n)_{n=1}^{\infty}$ *una sucesión arbitraria de números complejos y sea* $C(t) = \sum_{n\leq t} c_n$. *Dado* $x \geq 1$, *para cualquier* $g : [1,\infty) \longrightarrow \mathbb{C}$, $g \in C^1$, *se verifica*

$$\sum_{n\leq x} c_n g(n) = C(x)g(x) - \int_1^x C(t)g'(t)\, dt.$$

Demostración: Por continuidad podemos suponer que x no es entero, entonces el primer miembro es $\int_{1/2}^{x} h(t)g(t)\, dt$ con $g(t) = \sum c_n \delta(t-n)$ y δ la delta de Dirac. Como (?!) $C'(t) = h(t)$, el lema de Abel se reduce a integrar por partes. □

Para aplicar la primera versión al promedio de ϕ, tomemos $a_n = n$ y $b_n = \phi(n)/n$, entonces se sigue

$$\begin{aligned}\sum_{n=1}^{\infty} \phi(n) &= N\big(\frac{6N}{\pi^2} + O(\log N)\big) - \sum_{n=1}^{N-1} \big(\frac{6n}{\pi^2} + O(\log n)\big) \\ &= \frac{6}{\pi^2}\big(N^2 - \frac{N(N-1)}{2}\big) + O(N\log N) = \frac{3N^2}{\pi^2} + O(N\log N)\end{aligned}$$

que coincide con nuestras expectativas.

Por otra parte, el Lema de Abel con $c_n = \phi(n)/n$ y $g(x) = x$ lleva a un cálculo similar:

$$\begin{aligned}\sum_{n\leq x} \phi(n) &= x\big(\frac{6x}{\pi^2} + O(\log x)\big) - \int_1^x \big(\frac{6t}{\pi^2} + O(\log t)\big)\, dt \\ &= \frac{6}{\pi^2}\big(x^2 - \frac{x^2}{2}\big) + O(x\log x) = \frac{3x^2}{\pi^2} + O(x\log x).\end{aligned}$$

La sumación por partes es útil pero no la panacea. Por ejemplo, no podemos deducir el comportamiento de $\sum_{n\le x} t_n$ a partir de $\sum_{n\le x} t_n/n = 1+o(1)$, o lo que es lo mismo de $\sum_{n=1}^{\infty} t_n/n = 1$, porque hay muchas formas de hacer que la suma de la serie infinita sea uno. Por ejemplo $t_1 = 1$, $t_n = 0$ si $n > 1$ que cumple $\sum_{n\le x} t_n = 1$ o $t_n = (-1)^{n+1}/\log 2$ para los que $\sum_{n\le x} t_n$ tiene un comportamiento oscilante. Estos ejemplos prueban que ni siquiera $\sum_{n\le x} t_n/n = 1+O(x^{-1})$ puede llevarnos a deducir el comportamiento asintótico de $\sum_{n\le x} t_n$. Veamos qué parte de la maquinaria salta cuando aplicamos el lema de Abel con $c_n = t_n/n$ y $g(x) = x$:

$$\sum_{n\le x} t_n = x\big(1+O(x^{-1})\big) - \int_1^x \big(1+O(t^{-1})\big)\,dt = O(\log x).$$

Los términos principales se volatilizan y lo que resta es absorbido por el término de error. Nos chocamos contra un muro similar cuando intentamos deducir la asintótica de $\pi(x)$ a partir de la fórmula de Mertens (1.5), es analíticamente imposible, necesitaríamos más información.

Las técnicas de sumación por partes no son coto privado de la funciones aritméticas, uno puede emplearlas como un instrumento natural del análisis.

Ejemplo. Aplicando el lema de Abel con $c_n = 1$ y $g(t) = 1/t$ se deduce

$$\sum_{n\le x} \frac{1}{n} = \frac{[x]}{x} + \int_1^x \frac{[t]}{t^2}\,dt = \log x + \frac{[x]}{x} + \int_1^x \frac{[t]-t}{t^2}\,dt.$$

Escribiendo la integral como $\int_1^x = \int_1^\infty - \int_x^\infty = \text{cte} + O(1/x)$, se concluye el resultado clásico $\sum_{n\le x} n^{-1} = \log x + \gamma + O(1/x)$, donde $\gamma = 0'577\ldots$ es una constante llamada *constante de Euler*, [Sp] p. 566.

Ejemplo. Eligiendo $c_n = 1$ y $g(n) = \log n$, se tiene (tómese $x = N \in \mathbb{N}$)

$$\log N! = N\log N - \int_1^N [t]t^{-1}\,dt = N\log N - N + O(\log N).$$

Incluso se puede ir un poco más lejos. Manipulando la integral mediante integración por partes se deduce

$$\log N! = N\log N - N + \frac{1}{2}\log N + K - \int_N^\infty h(t)t^{-2}\,dt$$

donde $h(t)$ es la función periódica $1/6+\int_0^t (x-[x]-1/2)\,dx$. No es fácil identificar la constante K que se prueba que es $\frac{1}{2}\log(2\pi)$. Esto es suficiente para concluir la *aproximación de Stirling*:

$$N! \sim \sqrt{2\pi N}N^N e^{-N}.$$

Integrando por partes más veces se obtienen aproximaciones en las que el error relativo tiende a cero tan rápido como una potencia negativa arbitraria fijada (aunque el error absoluto no está acotado). El desarrollo "completo" de la aproximación depende de los números de Bernoulli [Sp] p. 713.

A medio camino entre el análisis y la artimética hay algunas aplicaciones de la sumación por partes relacionadas con la función ζ que son particularmente útiles. En primer lugar, tomando $c_n = 1$ y $g(x) = x^{-s}$ en el Lema de Abel (véase [Da] p. 32 si se quiere usar el Lema 1.2.1) se obtiene $\sum_{n=1}^{\infty} n^{-s} = s\int_1^{\infty}[x]x^{-s-1}\,dx$, que maquillado un poco se puede escribir como

$$\zeta(s) = \frac{s}{s-1} - s\int_1^{\infty} \frac{\mathrm{Frac}(x)}{x^{s+1}}\,dx. \tag{1.6}$$

Nótese que aunque originalmente $\zeta(s)$ sólo estaba definida para $\Re(s) > 1$, el segundo miembro permite extender la definición a $\Re(s) > 0$ y así se tiene que ζ se puede considerar como una función meromorfa[3] en $\Re(s) > 0$ con un único polo en $s = 1$. Con un poco de ingenio se puede integrar por partes (sacando la basura que pueda ensuciar la convergencia) y conseguir una extensión meromorfa a todo $\mathbb{C}$ de manera que $\zeta(s) - 1/(s-1)$ sea entera. Normalmente se llama función ζ de Riemann al fruto de esta extensión.

El mismo argumento aplicado a $-\zeta'(s)/\zeta(s)$ conduce a la fórmula:

$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{s}{s-1} + s\int_1^{\infty} (\psi(x) - x)x^{-s-1}dx. \tag{1.7}$$

A pesar de que (1.1) nos dice que una función multiplicativa f está determinada por sus valores en los primos y sus potencias, no parece nada claro que haya una relación entre promediar f sobre los primos y sus potencias y promediar la función sobre todos los números. Hay sin embargo algunas desigualdades obvias que se pueden establecer. Por ejemplo, si suponemos que f es completamente multiplicativa y $|f(p)| < \mathrm{cte} < 1$

$$\sum_{n\le x} f(n) \le \prod_{p\le x} \big(1 + f(p) + f(p^2) + \dots\big) = \prod_{p\le x} \big(1 - f(p)\big)^{-1}.$$

Tomando logaritmos se sigue $\log\sum_{n\le x} f(n) \le \sum_{p\le x} f(p) + \dots$ donde los puntos suspensivos son controlables si $f(p)$ tiende a cero. Todavía más, bajo buenas condiciones no debería haber gran diferencia entre $\log\sum_{p\le x} f(n)$ y $\sum_{p\le x} f(p)$. ¿Qué significa lo de "buenas condiciones"? En primer lugar que las potencias de los primos no molesten mucho, y en segundo lugar que se tenga un comportamiento adecuado de $\sum_{p\le x} f(p)$ (que $f(p)$ tienda correctamente a cero). E. Wirsing [Wi] logró materializar estas ideas en un teorema que, bajo ciertas hipótesis, traslada la asintótica de $\sum_{p\le x} f(p)$ a la de $\sum_{n\le x} f(n)$.

La versión que veremos aquí es una adaptación de la incluida en [Iw-Ko], ésta admite hipótesis más débiles que la formulación original. Además resuelve de forma elegante el tratamiento conjunto de los primos y sus potencias considerando la función de von Mangoldt generalizada $\Lambda_f(n)$ que se define formalmente como:

$$\sum_{n=1}^{\infty} \frac{\Lambda_f(n)}{n^s} = -\frac{D_f'(s)}{D_f(s)}.$$

[3]Si uno quiere ser pedante, la integral es holomorfa por el teorema de Morera

Como contrapartida algunas manipulaciones previas en algunos ejemplos concretos (para el teorema en su forma más conocida y su demostración véase [Po] o el original [Wi]).

Teorema 1.2.4 (Wirsing) *Sea f una función aritmética multiplicativa tal que*

$$\sum_{n\le x}\Lambda_f(n)=\kappa\log x+C+o(1)\qquad y\qquad \sum_{n\le x}|f(n)|\ll(\log x)^{|\kappa|}$$

para algún $\kappa\in\mathbb{R}$. Entonces

$$\sum_{n\le x}f(n)=K(\log x-C)^{\kappa}+o((\log x)^{|\kappa|-1})$$

donde K es la constante

$$K=\frac{1}{\Gamma(\kappa+1)}\prod_p\big((1-p^{-1})^{\kappa}(1+f(p)+f(p^2)+f(p^3)+\dots)\big).$$

Además el resultado también se cumple si se sustituye en la hipótesis y la conclusión "o" por "O" (en ese caso la C es superflua).

Observación: Nótese que el resultado sólo da una fórmula asintótica para $\kappa\ge -1/2$.

Aquí $\Gamma(z)$ es la *función Gamma*, $\int_0^\infty t^{z-1}e^{-t}\,dt$ para $\Re z>0$, que generaliza el factorial, $\Gamma(N+1)=N!$, y puede extenderse a una función homomorfa y sin ceros en $\mathbb{C}-\{0,-1,-2,\dots\}$ por medio de $\Gamma(z+1)=z\Gamma(z)$.

Demostración: Toda el truco está en emplear $f\log=f*\Lambda_f$ para relacionar $\sum\Lambda_f(n)$ con $\sum f(n)$:

$$\begin{aligned}\sum_{n\le x}f(n)\log n &= \sum_{n\le x}f(n)\sum_{m\le x/n}\Lambda_f(m)\\ &= (\kappa\log x+C)\sum_{n\le x}f(n)-\kappa\sum_{n\le x}f(n)\log n+o\big((\log x)^{|\kappa|}\big).\end{aligned}$$

Escribamos para abreviar $g(x)=\sum_{n\le x}f(n)$. Gracias al Lema de Abel se puede relacionar $\sum_{n\le x}f(n)\log n$ con $g(x)$, simplemente es igual a $g(x)\log x-\int_1^x g(t)t^{-1}\,dt$. Sustituyendo en la igualdad anterior se tiene que

$$\Delta(x)=(\log x-C)g(x)-(\kappa+1)\int_{x_0}^x g(t)t^{-1}\,dt$$

es muy pequeño, $o((\log x)^{|\kappa|})$. Ahora vamos a despejar g. Para ello lo mejor es soñar que todo es derivable (?!) y reducir el problema a una ecuación diferencial ordinaria en la que la incógnita es $g=g(x)$

$$\Delta'(x)=(\log x-C)g'(x)-\kappa g(x)x^{-1}.$$

Multiplicando por el factor integrante $(\log x-C)^{-\kappa-1}$ se tiene

$$\Delta'(x)(\log x-C)^{-\kappa-1}=\big((\log x-C)^{-\kappa}g(x)\big)'$$

que integrando conduce a

$$g(x) = \Delta(x)(\log x - C)^{-1} + (\kappa+1)(\log x - C)^{\kappa} \int_{x_0}^{x} \Delta(t)(\log t - C)^{-\kappa-2} t^{-1}\, dt.$$

Supongamos $\kappa \geq -1/2$, que es el único caso de interés. Si se completa la integral añadiendo la porción $\int_x^\infty$ se obtiene una constante y se pierde un factor $o((\log x - C)^{|\kappa|-\kappa-1})$. Lo cual prueba

$$g(x) = K(\log x - C)^{\kappa} + o((\log x)^{|\kappa|-1}).$$

El cálculo de la constante se lleva a cabo indirectamente. Sumando por partes se tiene

$$D_f(s) = s\int_1^\infty g(t)t^{-s-1}\, dt = Ks\int_1^\infty (\log t)^{\kappa}\big(1 + O((\log t)^{-1/4})\big)t^{-s-1}\, dt$$

que usando la fórmula $\Gamma(\kappa+1) = s^{\kappa+1}\int_1^\infty (\log t)^{\kappa} t^{-s-1}\, dt$ produce

$$K\Gamma(\kappa+1) \sim s^{\kappa} D_f(s) \sim (\zeta(s+1))^{-\kappa} D_f(s) = \prod_p \big((1-p^{-1})^{\kappa}(1+f(p)+f(p^2)+f(p^3)+\dots)\big)$$

cuando $s \to 0^+$. tomando límites se termina la evaluación de la constante. □

Veamos primero un par de ejemplos sencillos sólo para tantear por dónde respira el teorema.

Ejemplo. Si $f(n) = |\mu(n)|/n$ entonces tomando derivadas logarítmicas en $D_f(s) = \prod(1+p^{-s})$ se sigue $\Lambda_f(p^k) = -(-p)^{-k}\log p$ y es cero en otro caso. Sabemos que

$$\sum_{n\leq x} \Lambda_f(n) = \log x + O(1)$$

y trivialmente $\sum_{n\leq x} f(n) \ll \log x$. La conclusión del Teorema de Wirsing es

$$\sum_{n\leq x} \frac{|\mu(n)|}{n} \sim \prod \big((1-p^{-1})(1+p^{-1})\big)\log x = \frac{6}{\pi^2}\log x.$$

En realidad la aplicación del Teorema de Wirsing es ridícula porque un argumento elemental (ejercicio) ya conduce a $\sum_{n\leq x}|\mu(n)| \sim 6x/\pi^2$ y basta sumar por partes.

Ejemplo. La fórmula de Mertens (1.5) implica, usando el teorema de Wirsing con $f(n) = 1/n$ que $\sum_{n\leq x} n^{-1} = \log x + O(1)$ (aquí $K = \kappa = 1$). Esto no es nada espectacular y ya lo habíamos obtenido en una versión más refinada $\sum_{n\leq x} n^{-1} = \log x + \gamma + O(x^{-1})$. Podemos ahora dar la vuelta al teorema con $C = -\gamma$ y concluir que

$$\lim_{x\to\infty}\left(\log x - \sum_{n\leq x}\frac{\Lambda(n)}{n}\right) = \gamma$$

si el límite existe. De hecho es fácil asegurar la existencia del límite a partir del teorema de los números primos (que aunque no se ha enunciado, seguro que el lector conoce). Aparentemente esta igualdad tan limpia no aparece reflejada en la literatura (por ejemplo [In] parece evitarla calculando otra constante menos natural).

Veamos ahora una aplicación más complicada[4], a cambio debemos creernos algún resultado.

Ejemplo. Intentemos decidir la asintótica de de los libres de cuadrados representables como suma de dos cuadrados, esto es, de

$$N(x) = \{n \leq x \ : \ n = a^2 + b^2 \text{ es libre de cuadrados}\}.$$

Como recordaremos en la siguiente sección, un número n libre de cuadrados es representable como suma de dos cuadrados si y sólo si $p|n \ \Rightarrow \ p = 2$ o $p \equiv 1 \ (4)$. Escojamos por tanto $f(n)$ como la función multiplicativa con $f(p^k) = c(p^k)p^{-k}$ donde $c(n)$ es la función característica de $\{p \equiv 1 \ (4)\} \cup \{2\}$. Procediendo como antes, $\Lambda_f(p^k) = -(-p)^{-k} \log p$ para este conjunto de primos (y cero en el resto). Como son la mitad de todos los primos (??) se debería tener

$$\sum_{n \leq x} \Lambda_f(n) = \frac{1}{2} \log x + O(1).$$

Dentro de dos capítulos tendremos las fuerzas necesarias para zarandear un poco al $O(1)$ hasta hacerle soltar una constante indeterminada:

$$\sum_{n \leq x} \Lambda_f(n) = \frac{1}{2} \log x + C + o(1).$$

Lo que moviendo el manubrio del teorema produce

$$\sum_{n \leq x} f(n) = K(\log x - C)^{1/2} + o\big((\log x)^{-1/2}\big)$$

y sumando por partes

$$N(x) = \sum_{n \leq x} nf(n) = Kx(\log x - C)^{1/2} - K \int_{x_0}^{x} (\log t - C)^{1/2} \, dt + o\big(x(\log x)^{-1/2}\big).$$

Integrando por partes dos veces, se tiene que la integral es $x(\log x - C)^{1/2} - \frac{1}{2}x(\log x - C)^{-1/2}$ más términos de orden inferior, por consiguiente

$$N(x) \sim \frac{Kx}{2\sqrt{\log x}}.$$

Con trabajo se puede limpiar un poco el aspecto de la constante.

1.3. Algunas técnicas algebraicas y analíticas

Teoría de anillos. Teoría de Galois. La delta de Dirac. Sumas trigonométricas.

Esta larga sección pretende ser algo así como un "repaso de todo" o al menos un repaso de todo lo que un estudiante de teoría de números deberái haber conservado

[4]Ésta es el ejercicio 4 de [Iw-Ko] p. 28 salvo una pequeña variante. Aparentemente no es suficiente con utilizar la versión del Teorema de Wirsing allí establecida.

de la licenciatura si se lo hubierna explicado. Por si no se sospechara por el forzado retruécano, la visión trata de ser original, en comparación con los cursos originarios, y está posiblemente muy sesgada.

Respetando el orden alfabético vayamos primero con el álgebra.

Comencemos con algunas mentiras piadosas que, al ser de repaso, el lector sabrá tolerar (para las definiciones reales véase [Cl], [Do-He], [Ja]).

Un *anillo* es un conjunto en el que podemos sumar, restar y multiplicar con las propiedades habituales. En álgebra se consideran anillos con multiplicación no conmutativa o sin elemento neutro pero aquí huiremos de ellos. En un anillo quizá se pueda dividir entre algunos elementos, con una notación un poco confusa se dice que son las *unidades*, en pocas palabras una unidad es un elemento con inverso multiplicativo. Si se puede dividir entre todos los elementos distintos de cero, se dice que tenemos un cuerpo.

Por ejemplo $\mathbb{Z}$ es un anillo con unidades 1 y -1 pero no es un cuerpo, $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \ : \ a, b \in \mathbb{Q}\}$ es un cuerpo (las unidades son $\mathbb{Q}(\sqrt{2}) - \{0\}$ porque siempre se puede racionalizar) y $\mathbb{Z}[i]$ no es un cuerpo pero sí un anillo y las unidades son 1, -1, i y $-i$.

En teoría de números los cuerpos más importantes son los *cuerpos de números* consistentes en tomar unos cuantos números algebraicos sobre $\mathbb{Q}$, $r_1, r_2, \ldots, r_k$ y hacer todas las posibles sumas, restas, multiplicaciones y divisiones, se escribe $\mathbb{Q}(r_1, r_2, \ldots, r_k)$. Y los anillos más importantes son los *anillos de enteros* que consisten en los elementos de un cuerpo (de números) que son raíces de polinomios mónicos de $\mathbb{Z}[x]$. Sus elementos se llaman *enteros algebraicos*. Esta definición tan rara viene motivada porque gozan de propiedades parecidas a las de los enteros, por ejemplo forman un anillo (aunque no sea trivial probarlo [St-Ta]) y (como veremos) se puede hacer aritmética con ellos. Por ahora notemos que el anillo de enteros de $\mathbb{Q}$ es $\mathbb{Z}$, el anillo de enteros de $\mathbb{Q}(i)$ es $\mathbb{Z}[i]$ y el de $\mathbb{Q}(\sqrt{-5})$ es $\mathbb{Z}[\sqrt{-5}]$, sin embargo el de $\mathbb{Q}(\sqrt{-7})$ es $\{a + b(1 + \sqrt{-7})/2 \ : \ a, b \in \mathbb{Z}\}$ (ejercicio: investigar esto calculando el anillo de enteros de $\mathbb{Q}(\sqrt{N})$ en general).

En un anillo (de enteros) A se los *ideales* son las combinaciones lineales de ciertos elementos dados, llamados *generadores*,. La notación para los ideales en función de sus generadores es:

$$\langle a_1, a_2, \ldots, a_n \rangle = \{\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n \ : \ \lambda_j \in A\}.$$

En $\mathbb{Z}$ los ideales son muy aburridos porque gracias al algoritmo de Euclides (o la identidad de Bezout) se cumple

$$\langle a, b \rangle = \langle (a, b) \rangle. \tag{1.8}$$

Un ideal consistente en los múltiplos de un solo número se dice que es *principal*. Todos lo son en $\mathbb{Z}$, según lo anterior. Los anillos en los que ocurre esto (y $a, b \neq 0 \Rightarrow ab \neq 0$) se dice que son *dominios de ideales principales*. Cuando éste es el caso números e ideales están en correspondencia biyectiva salvo multiplicar por unidades porque $\langle a \rangle = \langle b \rangle \Leftrightarrow a = ub$. Esto no ocurre siempre, por ejemplo $\langle 2, 1 + \sqrt{-5} \rangle$ no es principal en $\mathbb{Z}[\sqrt{-5}]$.

Los ideales se emplean como sustituto de los enteros algebraicos cuando éstos se portan mal. Por ello conviene saber cómo operarlos, sobre todo cómo multiplicarlos porque

la suma de ideales es una tontería, simplemente se unen los conjuntos de generadores. Para hacer el producto de dos ideales se multiplican los generadores de todas las formas posibles. En el caso de ideales principales el producto coincide con el de números $\langle a\rangle \cdot \langle b\rangle = \langle ab\rangle$ y en cualquier caso podemos definir los *ideales primos* del anillo A (se excluyen $\langle 0\rangle$ y $A = \langle 1\rangle$) como los que no se pueden factorizar.

Lo visto hasta ahora no deja de ser una sarta de definiciones, además imprecisas. La pregunta que subyace es ¿para qué quiero los ideales o los enteros algebraicos? ¿harán mi vida más sencilla? Remontándose a la historia, los ideales fueron introducidos por E. Kummer tratando de probar el último teorema de Fermat. Se puede encontrar una descripción fiel de la historia en [Ri] (o leyendo directamente [Sm]), lo que aquí veremos será un cuentecillo para aprender rápidamente la utilidad de las ideas.

Resulta que al resolver algunas ecuaciones diofánticas a uno le gustaría poder factorizar expresiones que no se pueden factorizar. Por ejemplo, compárense las siguientes ecuaciones:

$$a)\quad xy = 15^n, \qquad b)\quad x^2 + y^2 = 15^n.$$

La ecuación $a)$ es trivial, usando la factorización $15 = 3 \cdot 5$ se tiene que dado n todas las soluciones son $x = \pm 3^\alpha 5^\beta$, $y = \pm 3^{n-\alpha} 5^{n-\beta}$ con $0 \le \alpha, \beta \le n$, en total hay $2(n+1)^2$. Para la ecuación $b)$ no es inmediato cómo calcular ni siquiera el número de soluciones, en cambio si se sustituyera el "+" por "−" todo volvería a ser fácil porque $x^2 - y^2$ factoriza como $(x-y)(x+y)$ y todo se reduciría a despejar en $x-y = \pm 3^\alpha 5^\beta$, $x+y = \pm 3^{n-\alpha} 5^{n-\beta}$, de nuevo $2(n+1)^2$ soluciones ¿Podemos hacer trampas y factorizar $x^2+y^2 = (x+iy)(x-iy)$ y seguir adelante? Esto significa trabajar en el anillo de enteros algebraicos $\mathbb{Z}[i]$ y nos arriesgamos a que en él la factorización de 15 no sea $3\cdot 5$, de hecho es $15 = 3\cdot(2+i)(2-i)$ en el sentido de que 3, $2+i$ y $2-i$ no se pueden escribir como producto de más cosas si excluimos los "cuatro signos" (las unidades) de $\mathbb{Z}[i]$ que son $1, -1, i, -i$. Así pues tenemos

$$(x + iy)(x - iy) = 3^n (2 + i)^n (2 - i)^n.$$

La propiedad de que $x + iy$ y $x - iy$ sean conjugados restringe mucho las posibilidades, de hecho para n impar no hay ninguna solución, y para n par nos vemos obligados a escoger $3^{n/2}(2+i)^\alpha(2-i)^{n-\alpha}$ para un factor y el resto para el otro, salvo multiplicar por unidades. En resumen si n es par las soluciones son:

$$x = 3^{n/2}\Re\big(\epsilon(2+i)^\alpha(2-i)^{n-\alpha}\big), \qquad x = 3^{n/2}\Im\big(\epsilon(2+i)^\alpha(2-i)^{n-\alpha}\big),$$

con $\epsilon \in \{1, -1, i, -i\}$ y $0 \le \alpha \le n$, en total $4(n+1)$ soluciones.

Este invento de utilizar la aritmética de $\mathbb{Z}[i]$ para trabajar en $\mathbb{Z}$ se debe a Gauss, de ahí que los elementos de $\mathbb{Z}[i]$ reciban el nombre de *enteros gaussianos*. La justificación de que todo esto es lícito pasa por hurgar en la demostración del teorema fundamental de la aritmética y adaptarlo a $\mathbb{Z}[i]$. En pocas palabras todo lo que se necesita es un algoritmo de Euclides y para ello una división inexacta con resto menor que el divisor. En $\mathbb{Z}[i]$ se escoge como cociente de $a + bi$ entre $c + di$ al elemento de $\mathbb{Z}[i]$ más cercano a $(a + bi)/(c + di)$, y como resto lo que sobra.

Probemos ahora con otro ejemplo similar:

$$x^2 + 5y^2 = 21^n.$$

En $\mathbb{Z}[\sqrt{-5}]$ se tendría

$$(x+y\sqrt{-5})(x-y\sqrt{-5}) = (1+2\sqrt{-5})^n(1-2\sqrt{-5})^n$$

y se puede probar que $1+2\sqrt{-5}$, $1-2\sqrt{-5}$ son "primos": no se pueden factorizar salvo multiplicar por unidades, que en este caso son 1 y -1. Deberíamos por tanto tener que las soluciones son:

$$x = \pm\Re\big((1+2\sqrt{-5})^\alpha(1-2\sqrt{-5})^{n-\alpha}\big), \qquad x = \pm\frac{1}{\sqrt{5}}\Im\big((1+2\sqrt{-5})^\alpha(1-2\sqrt{-5})^{n-\alpha}\big),$$

con $0 \le \alpha \le n$, en total $2(n+1)$ soluciones. Los datos numéricos nos dan ahora un buen bofetón: resulta que por ejemplo para $n=2$ tenemos las soluciones

x	6	11	14	19	-6	-11	-14	-19	6	11	14	19	21	-6	-11	-14	-19	-21
y	9	8	7	4	9	8	7	4	-9	-8	-7	-4	0	-9	-8	-7	-4	0

esto hacen ¡18 soluciones! lejos de $2(2+1)$.

¿Por qué Gauss puede invertarse nuevos enteros y nosotros no? Si intentamos definir la división inexacta en $\mathbb{Z}[\sqrt{-5}]$ como antes, no se consigue que el resto sea menor que el divisor. Peor todavía, no existe ni factorización única ni el concepto de máximo común divisor, así se cumple

$$3\cdot 7 = (1+2\sqrt{-5})(1-2\sqrt{-5})$$

y sin embargo 3 y $1\pm 2\sqrt{-5}$ no tienen divisores comunes no triviales en $\mathbb{Z}[\sqrt{-5}]$, ni 7 y $1\pm 2\sqrt{-5}$. De hecho todos estos números son "primos" en el sentido de que no se pueden descomponer más. Si nos pidieran un deseo, partiendo de $3\cdot 7 = (1+2\sqrt{-5})(1-2\sqrt{-5})$, nos gustaría que existiesen los divisores comunes antes indicados, digamos $\mathfrak{a}_\pm = \mathrm{mcd}(3, 1\pm 2\sqrt{-5})$, $\mathfrak{b}_\pm = \mathrm{mcd}(7, 1\pm 2\sqrt{-5})$, de forma que

$$(1.9)\quad 3 = \mathfrak{a}_+\cdot\mathfrak{a}_-, \qquad 7 = \mathfrak{b}_+\cdot\mathfrak{b}_-, \qquad 1+2\sqrt{-5} = \mathfrak{a}_+\cdot\mathfrak{b}_+, \qquad 1-2\sqrt{-5} = \mathfrak{a}_-\cdot\mathfrak{b}_-.$$

Como hemos mencionado, tales $\mathfrak{a}_\pm, \mathfrak{b}_\pm$ no existen. Pero según (1.8), al menos en $\mathbb{Z}$, un ideal con dos generadores es un sustituto para el máximo común divisor. Y así resulta que (1.9) pasa a ser cierto reemplazando 3, 7 y $1\pm 2\sqrt{-5}$ por los ideales que generan, $\mathfrak{a}_\pm$ por $\langle 3, 1\pm 2\sqrt{-5}\rangle$ y $\mathfrak{b}_\pm$ por $\langle 7, 1\pm 2\sqrt{-5}\rangle$.

Las cantidades $\mathfrak{a}_\pm$, $\mathfrak{b}_\pm$ son literalmente "ideales" en (1.9), no existen, y en general sólo corresponderían a cantidades "reales" cuando los ideales fueran principales (esta cantidad real sería el generador). Vemos también que el concepto de primalidad de números (como aquellos que no se pueden descomponer) sólo nos lleva a problemas y es mejor desterrarlo, a cambio se puede reestablecer el orden considerando ideales primos. Con ellos la factorización vuelve a se única.

Nuestro ejemplo para $n=2$ se transforma con el lenguaje de los ideales en

$$\langle x+y\sqrt{-5}\rangle\langle x-y\sqrt{-5}\rangle = \mathfrak{a}_+^2\mathfrak{a}_-^2\mathfrak{b}_+^2\mathfrak{b}_-^2$$

y las soluciones son $\langle x+y\sqrt{-5}\rangle = \mathfrak{a}_+^\alpha\mathfrak{a}_-^{2-\alpha}\mathfrak{b}_+^\beta\mathfrak{b}_-^{2-\beta}$, $0\le\alpha,\beta\le 2$ siempre que el segundo miembro sea principal, en este caso es así y $x+y\sqrt{-5}$ queda determinando salvo unidades, de ahí las $2\cdot 3\cdot 3 = 18$ soluciones.

Todo esto nos lleva a la conclusión de que la pregunta de cuántos ideales no principales hay y cómo detectarlos está íntimamente relacionada con el estudio del fallo en la factorización única y es muy importante en el estudio de ecuaciones diofánticas. Por ejemplo, Kummer sólo tuvo un éxito parcial al atacar el último teorema de Fermat porque es difícil saber en general si los ideales que aparecen en ciertas factorizaciones son principales[5], y por tanto si posibles soluciones "ideales" de la ecuación de Fermat son "irreales".

Para descontar los ideales principales uno puede considerar la relación de equivalencia entre ideales no nulos

$$\mathfrak{a} \sim \mathfrak{b} \iff (r)\mathfrak{a} \sim (s)\mathfrak{b} \qquad \text{con} \quad r, s \neq 0.$$

Al cociente

$$\mathcal{H} = \text{Ideales no nulos}/\sim$$

se le llama *grupo de clases* y tiene estructura de grupo (multiplicativo) porque para todo ideal $\mathfrak{a}$ siempre existe $\mathfrak{b}$ tal que $\mathfrak{ab}$ es principal. Los elementos de $\mathcal{H}$ se pueden identificar con "cosas" del tipo $\mathfrak{a}/(s)$ (se dice que son ideales fraccionarios) de la misma forma que $\mathbb{Q}$ es como $\mathbb{Z} \times \mathbb{Z}$ estableciendo la relación $(a, b) \sim (c, d) \Leftrightarrow ad = bc$.

Por ejemplo, el grupo de clases de $\mathbb{Z}[\sqrt{-5}]$ se puede probar que es isomorfo a $\mathbb{Z}_2$, como $\mathfrak{a}_+$, $\mathfrak{a}_-$, $\mathfrak{b}_+$, $\mathfrak{b}_-$ no son principales deben corresponder a $\overline{1}$ en $\mathbb{Z}_2$ y el producto $\mathfrak{a}_+^{\alpha}\mathfrak{a}_-^{2-\alpha}\mathfrak{b}_+^{\beta}\mathfrak{b}_-^{2-\beta}$ corresponde a la clase de $\alpha+(n-\alpha)+\beta+(n-\beta)$, que es siempre cero en $\mathbb{Z}_2$ y por tanto el ideal producto es principal, lo que prueba que el número de soluciones de $x^2 + 5y^2 = 21^n$ es $2(n + 1)^2$.

Cuando en primero nos dieron los números primos enseguida hicimos cocientes de $\mathbb{Z}$ por ellos para obtener $\mathbb{Z}_p$, que cuando fuimos mayores y supimos que era un cuerpo finito, llamamos $\mathbb{F}_p$. Además el resto de los $\mathbb{Z}_n$ eran peores, no se podía invertir siempre y al multiplicar dos elementos podía dar cero. ¿Serán igualmente buenos los cocientes de anillos de enteros por ideales primos? La respuesta es un sí sin matices: Si A es un anillo de enteros entonces $A/\mathfrak{a}$ es un cuerpo $\Leftrightarrow$ $\mathfrak{a}$ es un ideal primo. En realidad no hace falta que sea un anillo de enteros, se puede ampliar el resultado siempre que los ideales primos sean *maximales*, esto es, que no haya otros ideales propios que los contengan, esto ocurre por ejemplo en todos los anillos de polinomios de una variable con coeficientes en

[5]La ecuación de Fermat $x^n + y^n = z^n$, digamos con $n = p > 3$ por razones técnicas, se puede factorizar como

$$(x + \zeta y)(x + \zeta^2 y) \cdots (x + \zeta^n y) = z \cdot z^{\overset{n \text{ veces}}{\cdots}} z$$

con $\zeta = e^{2\pi i/n}$. Esto conduce a estudiar cuándo dos productos coinciden en el anillo $\mathbb{Z}[\zeta]$. En $\mathbb{Z}$ es evidente que si tenemos unos cuantos números que son coprimos dos a dos unos con otros y su producto es una n-potencia, cada uno de ellos lo es. Se puede probar que si $(x, y) = 1$ entonces los ideales generados por $x+\zeta^j y$ son "coprimos" así que una solución no trivial conduce a que cada uno de estos ideales es una n-potencia. Si los ideales son principales se puede escribir esto como una igualdad entre elementos de $\mathbb{Z}[\zeta]$ y se sabe llegar a una contradicción esencialmente usando congruencias [Bo-Sh], [St-Ta]. Con esto, Kummer sabía probar el último teorema de Fermat cuando se puede asegurar que la potencia n-ésima de cualquier ideal es principal, en el lenguaje que se introducirá a continuación, cuando el grupo de clases tiene orden no divisible por n.

un cuerpo. Este edificio no son alharacas lingüísticas. Por ejemplo, para $p \in \mathbb{N}$ primo de los de siempre:

$$\begin{aligned} \langle p\rangle \text{ primo en } \mathbb{Z}[i] \quad &\Leftrightarrow \quad \mathbb{Z}[i]/\langle p\rangle \text{ es un cuerpo} \quad \Leftrightarrow \ \mathbb{Z}_p[i] \text{ es un cuerpo} \\ &\Leftrightarrow \quad \mathbb{Z}_p[x]/\langle x^2+1\rangle \text{ es un cuerpo} \quad \Leftrightarrow \ x^2+1 \text{ es irreducible en } \mathbb{Z}_p \end{aligned}$$

Pero que $\langle p\rangle$ sea primo en $\mathbb{Z}[i]$ es lo mismo que decir que $p = (a+bi)(a-bi)$ (tomando normas se ve que no se puede descomponer más) y por uno de los casos especiales de la ley de reciprocidad cuadrática[6] (o más fácil) sabemos que -1 es un cuadrado en $\mathbb{Z}_p$ si y sólo si $p \equiv 3\ (4)$. En definitiva

$$p \text{ es suma de dos cuadrados} \quad \Leftrightarrow \ p \equiv 3\ (4)$$

y como subproducto se pueden caracterizar todos los ideales primos en $\mathbb{Z}[i]$.

Lo mismo que la factorización única en $\mathbb{Z}$ lleva asociada la función ζ con su producto de Euler, también es posible asociar funciones ζ (llamadas de Dedekind) a los ideales. La relación de los ideales con los números naturales de toda la vida se hace a través de la *norma*. La norma de un ideal $\mathfrak{a}$ en un anillo A se define como $N\mathfrak{a} = \#A/\mathfrak{a}$. Nos podemos sentir complacidos con la notación observando que si $z \in \mathbb{Z}[i]$, $N\langle z\rangle = |z|^2$.

Ejemplo. Sea

$$r(n) = \#\{(a,b) \in \mathbb{Z}^2 \ : \ a^2+b^2 = n\}.$$

Esta definición se puede reescribir como $\#\{z \in \mathbb{Z}[i] \ : \ |z|^2 = n\}$. El anillo $\mathbb{Z}[i]$ es un dominio de factorización única, concretamente de ideales principales, esto es, los números están en correspondiente biyectiva con los ideales salvo multiplicar por alguna de las cuatro unidades $\{1,-1,i,-i\}$. Entonces

$$r(n) = 4\#\{\mathfrak{a} \text{ ideal } \subset \mathbb{Z}[i] \ : \ N\mathfrak{a} = n\} \ \Rightarrow \ \sum_{n=1}^{\infty} \frac{r(n)}{n^s} = 4\sum_{n=1}^{\infty} \frac{1}{(N\mathfrak{a})^s}$$

como en el caso de números naturales (1.2), aquí también se puede descomponer en "factores locales"

$$\sum_{n=1}^{\infty} \frac{r(n)}{n^s} = 4\prod_{\wp} \left(1 + \frac{1}{(N\wp)^s} + \frac{1}{(N\wp)^{2s}} + \dots\right)$$

donde $\wp$ recorre los ideales primos de $\mathbb{Z}[i]$. Por el análisis anterior, los *primos racionales* p (los normales y corrientes) factorizan en $\mathbb{Z}[i]$ de la siguiente forma:

$$\langle p\rangle = \begin{cases} \langle a+bi\rangle\langle a-bi\rangle & \text{si } p \equiv 1\ (4) \\ \langle p\rangle & \text{si } p \equiv 3\ (4) \\ \langle 1+i\rangle^2 & \text{si } p = 2 \end{cases}$$

[6] Ésta afirma que para $p, q > 2$ primos distintos se cumple

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4} \qquad \text{además} \qquad \left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}, \qquad \left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

donde (a/r), con r primo $r \not| a$, es el símbolo de Legendre que vale uno si $x^2 \equiv a\ (r)$ tiene solución y -1 si no la tiene. Se escribe $(a/r) = 0$ si $r|a$.

Estos factores son por tanto todos los ideales primos en $\mathbb{Z}[i]$ y se sigue

$$\sum_{n=1}^{\infty}\frac{r(n)}{n^s}=4(1+2^{-s}+2^{-2s}+\dots)\prod_{p\equiv 1\ (4)}(1+p^{-s}+p^{-2s}+\dots)^2\prod_{p\equiv 3\ (4)}(1+p^{-2s}+p^{-4s}+\dots)$$

que se puede escribir como (ejercicio)

$$\sum_{n=1}^{\infty}\frac{r(n)}{n^s}=4\zeta(s)\prod_{p\equiv 1\ (4)}(1-p^{-s})^{-1}\prod_{p\equiv 3\ (4)}(1+p^{-s})^{-1}=\zeta(s)D_\chi(s)$$

donde $\chi=\chi(n)$ es la función multiplicativa que vale 1 si $n\equiv 1\ (4)$, -1 si $n\equiv 3\ (4)$ y 0 en el resto de los casos. Se deduce que $r(n)/4$ es multiplicativa y la fórmula

$$\frac{r(n)}{4}=\sum_{d|n}\chi(d)=\#\{d|n\ :\ d\equiv 1\ (4)\}-\#\{d|n\ :\ d\equiv 3\ (4)\}.$$

Por ejemplo, $r(10^6)=4(r(2^6)/4)(r(5^6)/4)=4\cdot 1\cdot 7=28$. En particular un número n es representable como suma de dos cuadrados si y sólo si los únicos posibles factores primos de n de la forma $p\equiv 3\ (4)$ aparecen con exponente par.

No abandonaremos los placeres del álgebra sin visitar brevemente a la teoría de Galois, una de las asignaturas más bellas de la licenciatura y una de las pocas que se muestra autocontenida con su planteamiento, nudo y desenlace.

Su propósito y gran logro es establecer una relación exacta entre la estructura de ciertos cuerpos y sus simetrías. Pensemos por ejemplo en cualquier identidad en $\mathbb{Q}(i)$, digamos $(3+2i)/(1+i)=5/2-i/2$, entonces al cambiar i por $-i$ la igualdad se conserva. Simetrías como éstas son lo que se llaman $\mathbb{Q}$-*automorfismos*: funciones biyectivas que respetan las sumas, restas, multiplicaciones y divisiones y que dejan a cada racional invariante. En general, dados dos cuerpos $L\supset K$ (se suele escribir L/K y se dice que es una *extensión* de cuerpos), se llama *grupo de Galois* de L/K al grupo $\mathcal{G}(L/K)$ formado por los automorfismos que respetan las operaciones y dejan fijo a cada elemento de K. Es muy fácil ver que si $\alpha\in L$ es raíz de un polinomio $P\in K[x]$, entonces $\sigma(\alpha)$ también lo es para todo $\sigma\in\mathcal{G}(L/K)$. Esto es, los elementos del grupo de Galois lo único que hacen es permutar raíces de polinomios. Una permutación de raíces (incluidas en L) puede no corresponder a ningún elemento de $\mathcal{G}(L/K)$ porque haya relaciones entre ellas. Por ejemplo, un $\mathbb{Q}$-automorfismo σ no puede mandar $\sqrt{2}$ a $\sqrt{3}$ aunque ambas sean raíces de x^4-5x^2+6 porque $\sigma(\sqrt{2})=\sqrt{3}\Rightarrow\sigma(\sqrt{2}\cdot\sqrt{2})=\sqrt{3}\cdot\sqrt{3}\Rightarrow 2=3$, que es muy feo. Sin embargo sí es cierto que si tenemos un polinomio irreducible $P\in K[x]$ con raíces en L siempre hay un elemento de $\mathcal{G}(L/K)$ que aplica una de las raíces en cualquier otra, ambas escogidas arbitrariamente.

Ejemplo. En $\mathbb{Q}(\sqrt{n})$ con $|n|\in\mathbb{N}$ libre de cuadrados, todo lo que se puede hacer es enviar $\sqrt{n}$ a $-\sqrt{n}$ o dejarlo fijo pues éstas son las dos raíces de x^2-n; lo que implica que $\mathcal{G}(\mathbb{Q}(\sqrt{n})/\mathbb{Q})=\{\mathrm{Id},\mathrm{conj.}\}$

Ejemplo. Si p es primo, el polinomio ciclotómico $x^{p-1}+\cdots+x+1=(x^p-1)/(x-1)$ es irreducible [St] y tiene como raíces a $\zeta,\zeta^2,\ldots,\zeta^{p-1}$ con $\zeta=e(1/p)$. Un elemento de $\mathcal{G}(\mathbb{Q}(\zeta)/\mathbb{Q})$ queda determinado por la imagen de ζ, ya que el resto de las raíces dependen de ésta. Así pues los únicos $\mathbb{Q}$-automorfismos son los dados por $\sigma_j(\zeta)=\zeta^j$, $j=1,2,\ldots,p-1$. En símbolos

$$\mathcal{G}(\mathbb{Q}(\zeta)/\mathbb{Q})=\{\mathrm{Id}=\sigma_1,\sigma_2,\ldots,\sigma_{p-1}\}\cong\mathbb{Z}_{p-1}.$$

El último isomorfismo no es evidente pero el lector debería saber al menos sutituir $\mathbb{Z}_{p-1}$ por $\mathbb{Z}_p^*$ (los elementos $\{\overline{1},\overline{2},\ldots,\overline{p-1}\}$ con el producto) y leer en algún sitio qué son las raíces primitivas [Ro].

El llamado *teorema fundamental de la teoría de Galois* afirma que suponiendo $L=K(r_1,r_2,\ldots,r_n)\subset\mathbb{C}$ donde r_j son todas las raíces de cierto polinomio de $K[x]$ (se dice que L/K es normal), entonces los subgrupos de $\mathcal{G}(L/K)$ están en correspondencia biyectiva con los subcuerpos de L conteniendo a K por medio de

$$\text{subcuerpo } M \quad\longrightarrow\quad \text{subgrupo } \mathcal{G}(L/M)\subset\mathcal{G}(L/K)$$

Además la dimensión de M como espacio vectorial sobre K, llamada *grado*, coincide con el índice de $\mathcal{G}(L/M)$ en $\mathcal{G}(L/K)$ (el cociente de sus órdenes). El teorema fundamental de la teoría de Galois tiene realmente una formulación más general y otro "además" importante [St]. Cualquier matemático debería conocer su enunciado completo.

Una utilidad de esta relación entre subcuerpos y simetrías es la facultad de poder "descomponer" una extensión de cuerpos L/K con sus elementos feísimos llenos de radicales, decimales y otras guarrerías indescriptibles (véase el teorema de Abel en [St]); gracias al estudio de los subgrupos de un grupo finito, algo que a primera vista tiene un aspecto tan limpio y discreto como los números naturales.

Ejemplo. Sea la raíz de la unidad $\zeta=e(1/17)$. Sabemos que $\mathcal{G}(\mathbb{Q}(\zeta)/\mathbb{Q})\cong\mathbb{Z}_{16}$, que es cíclico de orden 16 y por tanto la extensión $\mathbb{Q}(\zeta)/\mathbb{Q}$ tendrá sólo tres subcuerpos propios (conteniendo a $\mathbb{Q}$): los correspondientes a los subgrupos $\langle\overline{2}\rangle$, $\langle\overline{4}\rangle$ y $\langle\overline{8}\rangle$

$$\mathbb{Q}(\zeta)=M_0\supset M_1\supset M_2\supset M_3\supset M_4=\mathbb{Q}.$$

Cada extensión M_{i-1}/M_i tiene grado 2, entonces cada elemento de M_{i-1} es de la forma $a\cdot 1+b\cdot u$ con $a,b\in M_i$, $u\in M_{i-1}$. En particular u^2 es de esta forma y por tanto es raíz de un polinomio cuadrático. Por consiguiente M_{i-1} es lo mismo que M_i añadiendo cierta raíz cuadrada de algún elemento. Se deduce que el polígono de 17 lados es construible con regla y compás, porque $\zeta=\cos(2\pi/17)+i\sin(2\pi/17)$ y siempre podemos extraer raíces cuadradas con útiles de dibujo [Cl]. Este resultado se debe a Gauss y es una delicia leer de primera mano [Ga] cómo hacía teoría de Galois cuando todavía Galois no tenía ni bozo.

Pasemos ahora al análisis.

Al igual que la dinámica tiene como personaje canónico a la partícula (cero dimensional) de masa unidad, se puede conocer una buena porción del reparto de la película del análisis a través de su "función de densidad": la delta de Dirac[7].

Si llamamos δ a esta presunta función de densidad de una partícula unidad en el origen de la recta real, rápidamente topamos con un ramillete de contradicciones frente al concepto más básico de función, pues debería cumplir

$$\delta(x) = 0 \quad \forall x \neq 0, \qquad \delta(0) = \infty, \qquad \int_{-\infty}^{\infty} \delta(x)\, dx = 1,$$

y los rudimentos de la teoría de la integral (de Lebesgue) nos aseguran que esto no tiene sentido porque un cambio de una función en un conjunto de medida cero no puede influir en el valor de la integral, así que en sana puridad matemática debería cumplirse la igualdad $\int_{-\infty}^{\infty} \delta = \int_{-\infty}^{\infty} 0 = 0$ y no se podría ver la definición $\delta(0) = \infty$ más que como una exótica aberración. No obstante nuestra intuición no se ve fabulosamente perturbada al emplear $masa = \int densidad$ para todas las masas, incluso partículas unidad, todavía más, si en $\mathbb{R}$ no se emplea la medida usual dx sino $f(x)dx$ (con f una función decente), debería cumplirse

$$\int_{-\infty}^{\infty} f(x)\delta(x)\, dx = f(0)$$

que generaliza la relación anterior.

Puestos a hacer barbaridades, consideremos la δ de Dirac 1-periódica, esto es, $\delta_p(x) = \sum_{n\in\mathbb{Z}} \delta(x-n)$ que integrando contra f da lugar a

$$\sum_{n=-\infty}^{\infty} f(n) = \int_{-\infty}^{\infty} f\delta_p.$$

Esto es interesante porque permite escribir sumas como integrales (como en la "demostración" del Lema de Abel). Si calculamos formalmente los coeficientes de Fourier de δ_p son todos unos y se infiere la igualdad

$$\delta_p(x) = \sum_{n=-\infty}^{\infty} e(nx) \tag{1.10}$$

que no tiene sentido aparente pero sustituida en la igualdad anterior conduce a

[7]El artífice de este regalo, P.A.M. Dirac, es bien conocido por sus importantísimas contribuciones a la Física. En `http://www-groups.dcs.st-andrews.ac.uk/` podemos leer que su vocación eran las Matemáticas: "*Although mathematics was his favourite subject he chose to study an engineering course at university since he thought that the only possible career for a mathematician was school teaching and he certainly wanted to avoid that profession. He obtained his degree in engineering in 1921 but following this, after an undistinguished summer job in an engineering works, he did not find a permanent position. By this time he was developing a real passion for mathematics but his attempts to study at Cambridge failed for rather strange reasons*".

Teorema 1.3.1 (fórmula de sumación de Poisson) *Para* $f \in C_0^\infty$

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \widehat{f}(n).$$

Una fórmula utilísima[8] sin deltas fantasmagóricas que permite pasar una suma sobre enteros en otra, confiando en que la segunda sea más sencilla.

Ejemplo. Sea $\theta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}$ con $t > 0$. Aplicando la fórmula de sumación de Poisson a $f(x) = e^{-\pi t x^2}$, se obtiene $\theta(t) = t^{-1/2}\theta(1/t)$. Para apreciar los beneficios, nótese que por ejemplo $\theta(0'01) = \sum e^{-\pi n^2 0'01}$ es muy costoso de aproximar con precisión empleando una calculadora de bolsillo debido a la lenta convergencia inicial, mientras que la relación anterior permite asegurar $\theta(0'01) \approx 10$ con una precisión de más de 100 decimales.

Al igual que los sólidos se pueden estudiar en mecánica integrando partículas de masa pequeña, toda función 1-periódica se puede ver como una "combinación lineal continua" de deltas de Dirac:

$$f(x) = \int_{-1/2}^{1/2} f(t)\delta_p(x-t)\, dt,$$

lo que combinado con (1.10) implica

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} a_n e(nx)} \qquad \text{con} \qquad a_n = \int_0^1 f(t)e(-nt)\, dt.$$

Es decir, a partir de (1.10) se puede "deducir" el desarrollo de Fourier genérico de una función.

De la misma forma, a partir de $\widehat{\delta}(\xi) = 1$ si uno confía ciegamente en la fórmula de inversión de la transformada de Fourier, $\delta = \widehat{\widehat{\delta}}^{\,\vee}$, se debería cumplir

$$\delta(x) = \int_{-\infty}^{\infty} e(x\xi)\, d\xi.$$

De este caso de la fórmula de inversión, cambiando x por $x-t$ e integrando (en t) contra $f(t)$, se deduce

$$\boxed{f = \widehat{\widehat{f}}^{\,\vee}}$$

que es la fórmula de inversión en general.

Lo bueno de todo esto no es sólo que podemos ver y deducir intuitivamente fórmulas cruciales teniendo una fe vana en una función inexistente, sino que en muchos casos nos anticipa la demostración de verdad, lo que nos lleva a la definición matemática de δ:

[8] De acuerdo con la introducción de [Iw-Ko]: "Poisson summation for number theory is what a car is for people in modern communities –it transports things to other places and it takes you back home when applied next time– one cannot live without it". Lo de "takes you back home when applied next time" se deduce de que la transformada de Fourier es casi involutiva: $\widehat{\widehat{f}}(x) = f(-x)$.

Si $\eta \in C_0^\infty(\mathbb{R})$, digamos $\eta \geq 0$ con $\eta(0) = 1$, y suponemos $\int \eta = 1$, entonces la sucesión $\eta_n(x) = n\eta(nx)$, de algún modo tiende a δ porque $\lim_{n\to\infty} \int f(x)\eta_n(x)\, dx = f(0)$ para funciones "buenas" f. Para no pillarnos los dedos, podemos definir δ como la propia sucesión de funciones η_n (se dice que es una *aproximación de la identidad*) y entonces resulta que las "demostraciones" anteriores se vuelven demostraciones sin comillas a base de asegurar las condiciones de convergencia [Dy-Mc], [Fo]. Incluso en (1.10) se puede introducir la regularización $d_r(x) = \sum |r|^n e(nx)$ (núcleo de Poisson) y probar que realmente $\lim_{r\to 1^-} \int f d_r = f(0)$. Todas las pruebas requieren la dualidad, δ no aparece nunca como singularidad desnuda, sino bien arropada integrada contra una función buena para que no nos asustemos con el infinito. Se puede llevar esta idea al extremo definiendo herméticamente δ como el sencillo funcional que aplica una función en su valor en cero.

La gran maravilla del análisis de Fourier es que, de acuerdo con las fórmulas recuadradas, permite expresar una función como una combinación de tonos puros, a fin de cuentas sencillos senos y cosenos. Hay varias ideas básicas que permiten avanzar con soltura al razonar con este análisis. Una de ellas está basada en el comportamiento de la transformada de Fourier bajo cambios de escala:

$$g(x) = f(Tx) \quad \Rightarrow \quad \widehat{g}(\xi) = T^{-1}\widehat{f}(\xi/T).$$

Si pensamos que f y $\widehat{f}$ son funciones decentes que más o menos tienen casi toda su masa en $[-1, 1]$, lo que dice esta relación es que al achuchar f para confinarla al intervalo $[-1/T, 1/T]$, la transformada se extenderá a $[-T, T]$. Esto es una forma débil del *principio de incertidumbre*, que de una manera gráfica se puede enunciar de la manera siguiente:

Principio de incertidumbre. Una oscilación de frecuencia ω se salta los objetos de tamaño ω^{-1}.

Explicado en tiempo, en lugar de en espacio, si jugando al escondite inglés miro un instante dos veces por minuto y alguien se mueve sólo una vez y durante mucho menos de medio minuto, es difícil que lo vea moverse (ahí está la gracia del juego).

En términos matemáticos, las transformadas de Fourier (que dan el contenido en frecuencia) de dos funciones que sólo difieren a escala δ (pequeña) sólo empezarán a distinguirse sustancialmente para $|\xi| \gg \delta^{-1}$. Esto se aplica sobre todo al tratar regularizaciones. Por ejemplo, si

$$f(x) = \begin{cases} 1 & \text{si } x \in [-1, 1] \\ 0 & \text{si } x \notin [-1, 1] \end{cases} \qquad \text{y} \qquad g(x) = \begin{cases} 1 & \text{si } x \in [-1+\delta/2, 1-\delta/2] \\ 0 & \text{si } x \notin [-1-\delta/2, 1+\delta/2] \\ \text{redondeado } C^\infty \text{ en el resto} \end{cases}$$

entonces $\widehat{f}(\xi) = \text{sen}(2\pi\xi)/\pi\xi$, en particular decae como $|\xi|^{-1}$ cuando $\xi \to \infty$, mientras que $\widehat{g}(\xi)$, por ser $g \in C^\infty$, decae más rápido que $|\xi|^{-N}$ para cualquier N (simplemente integrando por partes). Lo que nos dice el principio de incertidumbre es que el precio a pagar por esta regularización de f a g es no ver este decaimiento cuando $|\xi| \ll \delta^{-1}$.

Lo habitual en los ejemplos concretos es que haya que optimizar el valor de δ para contentar a dos tendencias en oposición: 1) la función regularizada debe parecerse a la original; 2) la transformada de Fourier de la función regularizadora debe empezar a decaer rápidamente. La primera requiere disminuir δ y la segunda aumentarlo.

Todo esto funciona de la misma forma en más dimensiones cambiando valores absolutos por normas. En general el análisis armónico (fórmulas recuadradas, sumación de Poisson,...) se extiende al caso multidimensional simplemente añadiendo más variables. La transformada de Fourier pasa a ser $\int f(\vec{x})e(-\vec{x}\cdot\vec{\xi})\ d\vec{x}$ y la serie de Fourier $\sum a_{\vec{n}}e(\vec{n}\cdot\vec{x})$.

Ejemplo. ¿Cuántos puntos de coordenadas enteras hay en la esfera de radio R (con R grande)? Aproximadamente el volumen, esto es $4\pi R^3/3$. Pero ¿de qué orden es el error $E(R) = \text{nº de puntos} - 4\pi R^3/3$? Esto no tiene respuesta fácil. La acotación trivial, contando puntos cerca de la cáscara de la esfera, es $E(R) = O(R^2)$ (ejercicio). Veamos cómo mejorar este resultado.

Formalmente, tomando como f la función característica de la esfera de radio R en la fórmula de sumación de Poisson en tres dimensiones

$$E(R) = \sum_{\vec{n}\in\mathbb{Z}^3} f(\vec{n}) - \frac{4\pi}{3}R^3 = \sum_{\vec{n}\in\mathbb{Z}^3-\{\vec{0}\}} \widehat{f}(\vec{n}).$$

¡Una fórmula exacta! Nuestro entusiasmo se desvanece cuando hacemos los cálculos y vemos que $\widehat{f}(\vec{n})$ es como $R\|\vec{n}\|^{-2}$ salvo un factor acotado oscilante (se puede hallar la fórmula explícita [Dy-Mc] §2.11), por tanto la serie no converge absolutamente[9]. Para remediar este desaguisado tomamos una función $g_+ \in C^\infty(\mathbb{R}^3)$ con $g_+(\vec{x}) = 1$ si $\|\vec{x}\| \le R$, $g_+(\vec{x}) = 0$ si $\|\vec{x}\| \ge R+\delta$ y redondeada positiva entre medias. Igualmente se define $g_- \in C^\infty(\mathbb{R}^3)$ con $g_-(\vec{x}) = 1$ si $\|\vec{x}\| \le R-\delta$, $g_-(\vec{x}) = 0$ si $\|\vec{x}\| \ge R$. Entonces

$$\sum_{\vec{n}\in\mathbb{Z}^3-\{\vec{0}\}} \widehat{g}_-(\vec{n}) + O(R^2\delta) \le E(R) \le \sum_{\vec{n}\in\mathbb{Z}^3-\{\vec{0}\}} \widehat{g}_+(\vec{n}) + O(R^2\delta)$$

(el $R^2\delta$ viene de $R^3 - (R\pm\delta)^3$). Por el principio de incertidumbre, $\widehat{g}_+(\vec{n})$ y $\widehat{g}_-(\vec{n})$ serán como $R\|\vec{n}\|^{-2}$ hasta $\|\vec{n}\| \ll \delta^{-1}$ pero después ya se ve la suavidad y todo decae estupendamente, por tanto

$$E(R) \ll R \sum_{\|\vec{n}\|\ll\delta^{-1}} \|\vec{n}\|^{-2} + R^2\delta \ll R\delta^{-1} + R^2\delta.$$

Con la elección óptima $\delta = R^{-1/2}$ ambos sumandos se igualan y se tiene $E(R) \ll R^{3/2}$, un resultado de Landau [La].

[9] La serie $\sum \|\vec{n}\|^{-2}$ es divergente, no como su análoga unidimensional. Concretamente se verifica $\sum_{\|\vec{n}\|\le x} \|\vec{n}\|^{-2} \gg x$ (ejercicio).

Si el análisis de Fourier permite reducir las funciones a montones de senos y cosenos, el problema de estimar promedios con precisión se traslada a estimar sumas trigonométricas. Terminando este larguísimo repaso con algo que no lo es, adentrémonos en este árido terreno (véase [Gr-Ko] para más información y [Hu] si a alguno le quedan ganas de más).

Digamos que tenemos

$$S = \sum \phi(n)\ e(f(n))$$

donde $\phi \in C_0^\infty$ es una función adaptada a un intervalo entero $[a, b]$ (una regularización de su función característica). Por la fórmula de sumación de Poisson

$$S = \sum_n \int \phi(x)\ e(f(x) - nx)\, dx.$$

Supongamos que f' es monótona en el soporte de ϕ, por ejemplo creciente con $N_1 < f' < N_2$, $N_1, N_2 \in \mathbb{Z}$. Si n está lejos del intervalo $[N_1, N_2]$ la derivada de $f(n) - nx$ es grande y por tanto la integral pequeña (mucha oscilación $\Rightarrow$ mucha cancelación). El caso límite es cuando f' apenas varía, digamos $|f'| < \epsilon$, entonces el término que más contribuye a S es $\int \phi\, e(f) = \int f' e(f)\ \phi/f'$ y sumando por partes para sacar ϕ/f', todo queda mayorado por una cota inferior para f'. Poniendo todo esto en limpio (??)

Proposición 1.3.2 *Sea $f \in C^1([a, b])$ con f' monótona y $\lambda_1 < |f'| < 1/2$, entonces*

$$\sum_{a<n\leq b} e(f(n)) \ll \lambda_1^{-1}.$$

Típicamente f' varía más y con la notación anterior la parte del león en S es

$$\sum_{N_1<n<N_2} \int \phi(x)\ e(f(x) - nx)\ dx. \tag{1.11}$$

Ahora la derivada de $f(x) - nx$ puede ser nula. Si llamamos x_n al punto crítico (no puede haber más que uno si f' es monótona), entonces en las cercanías de x_n se tiene por Taylor $f(x) - nx = f(x_n) - nx_n + f''(x_n)(x - x_n)^2/2 + \ldots$. Pero es bien conocido[10] que $\int_{-\infty}^{\infty} e(\lambda x^2)\ dx = C/\sqrt{\lambda}$, por tanto con un cambio de variable es creíble que $\int \phi(x)\ e(f(x) - nx)\ dx$ se parezca a (??) $C\phi(x_n)e(f(x_n) - nx_n)/\sqrt{f''(x_n)}$. Entonces cada término de (1.11) es del orden de $(f'')^{-1/2}$ y se tienen $N_2 - N_1 = \Delta f' + 1 \asymp (b-a)f'' + 1$ términos. Multiplicando, se sigue (??)

Proposición 1.3.3 *Sea $f \in C^2([a, b])$, $a, b \in \mathbb{Z}$ con $\lambda_2 \ll |f''| \ll \lambda_2$, entonces*

$$\sum_{a<n\leq b} e(f(n)) \ll (b-a)\lambda_2^{1/2} + \lambda_2^{-1/2}.$$

Este resultado se debe a van der Corput (véase en [Gr-Ko] una prueba elemental).

[10] En realidad basta probar que $\int_{-\infty}^{\infty} e(x^2)\ dx$ converge y hacer un cambio de variable.

1.4. La distribución de los primos

El teorema de los números primos. Ideas básicas de la demostración. La función ζ.

Ya sea con argumentos heurísticos, por ejemplo a partir de la fórmula de Mertens (1.5) o con pruebas experimentales[11] no es difícil sospechar que la "densidad" de los primos en los naturales se comporta como $1/\log x$. En términos más precisos:

Teorema 1.4.1 (teorema de los números primos) *Se cumple*

$$\pi(x) \sim Li(x) \qquad Li(x) = \int_2^x \frac{dt}{\log t}.$$

Antes de seguir es justo confesar que $Li(x) \sim x/\log x$ (basta aplicar la regla de L'Hôpital) y por tanto el teorema se podría haber escrito más claramente como $\pi(x) \sim x/\log x$ o $\lim_{x\to\infty} \pi(x)\log x/x = 1$. Para acallar de golpe cualquier murmullo linchador basta exhibir la siguiente tabla:

	$\pi(x)/Li(x)$	$\pi(x)/(x/\log x)$
$x = 10^4$	0'986	1'132
$x = 10^6$	0'9983	1'084
$x = 10^8$	0'99987	1'061
$x = 10^{10}$	0'9999932	1'048

Si nuestra conjetura favorita (la hipótesis de Riemann) se cumpliera, entonces $Li(x)$ sería una aproximación para $\pi(x)$ con más o menos mitad de sus cifras significativas correctas, como sugieren los resultados numéricos, sin embargo se sabe positivamente (sin interrogaciones) que $x/\log x$ es una aproximación tan mala que no muchas más de las primeras $\log\log x$ cifras significativas pueden ser correctas, esto es prácticamente nada desde el punto de vista numérico.

Por razones que aparecerán más adelante, para probar el teorema de los números primos es más fácil trabajar con $\psi(x)$ (véase §1.1) que directamente con $\pi(x)$. Intuitivamente $\psi(x)$ es como $\pi(x)$ salvo un logaritmo. Como prontuario ínfimo de diferentes equivalencias naturales del teorema de los números primos se enuncia el siguiente resultado:

Lema 1.4.2 *Las siguiente afirmaciones son equivalentes*

$$a)\ \pi(x) \sim Li(x), \qquad b)\ \pi(x) \sim x/\log x, \qquad c)\ p_n \sim n\log n, \qquad d)\ \psi(x) \sim x.$$

Demostración: Ya hemos mencionado que $a) \Leftrightarrow b)$. Claramente $\pi(p_n) = n$, así pues $b)$ implica $p_n/\log p_n \sim n$ y tomando logaritmos $\log p_n \sim n$. Multiplicando estas relaciones se obtiene $c)$. El recíproco se prueba en las mismas líneas: $p_n \le x < p_{n+1}$, $c) \Rightarrow p_n \sim x$.

[11] Éstas pueden ser triviales en el actual mundo del *bit*, en el que cualquier tonto puede pulsar una tecla y tener una lista inmensa de primos y pulsar otra y analizarla estadísticamente, pero en tiempos de Gauss y Legendre fue algo notable.

Es fácil ver que $\pi(x) = \sum_{n\leq x} \Lambda(n)/\log n + O(x^{1/2}\log x)$. De hecho con un poco de esfuerzo (ejercicio) se puede reducir el error a $O(x^{1/2})$. Sumando por partes se deduce por tanto $\pi(x) = \psi(x)/\log x + \int_2^x (\psi(t)-t)/(t(\log t)^2)\, dt + O(x^{1/2})$, o equivalentemente

$$\pi(x) = Li(x) + \frac{\psi(x)-x}{\log x} + \int_2^x \frac{\psi(t)-t}{t(\log t)^2}\, dt + O(x^{1/2}) \tag{1.12}$$

que inmediatamente prueba $d) \Rightarrow a)$. Si se parte de $\psi(x) = \sum_{n\leq x}(\pi(n)-\pi(n-1))\log n + O(x^{1/2}(\log x)^2)$, un argumento similar prueba $b) \Rightarrow d)$. □

Más importante que el propio enunciado es la relación (1.12) y para subrayarlo machaconamente se enuncia de nuevo débilmente.

Lema 1.4.3 *Si $\psi(x) = x + O(E(x))$ para cierta E creciente, entonces se cumple $\pi(x) = Li(x) + O(x^{1/2} + E(x)/\log x)$.*

Observación: Insistiendo en las repeticiones, de $Li(x) - x/\log x \sim x/(\log x)^2$ se deduce que en cuanto se pruebe una acotación ligeramente buena para $E(x)$, el error $\pi(x) - x/\log x$ está prácticamente dominado por $x/(\log x)^2$.

Una equivalencia más profunda y curiosa tiene que ver con el promedio de la función de Möbius, como destaca la siguiente (meta-)proposición; cuya demostración elemental (originariamente debida a Landau) está tomada de [Iw-Ko].

Proposición 1.4.4 *El teorema de los números primos es equivalente a que el promedio de μ tienda a cero, esto es,*

$$\pi(x) \sim Li(x) \quad \Longleftrightarrow \quad \lim_{x\to\infty} \frac{1}{x}\sum_{n\leq x} \mu(n) = 0.$$

Demostración: Nos olvidaremos en ambas implicaciones de $\pi(x) \sim Li(x)$ y emplearemos $\psi(x) \sim x$ que ya hemos visto que es equivalente.

$\Rightarrow$) La relación $\log = 1 * \Lambda$ multiplicando por $\mu(n)$ y sumando, da lugar a

$$\sum_{n\leq x} \mu(n)\log n = \sum_{n\leq x} \psi(x/n).$$

Por otra parte la igualdad $f = \mu * 1$ con $f(1) = 1$ y $f(n) = 0$ si $n \neq 1$, implica después de sumar $[x] = \sum_{n\leq x} \mu(n)[x/n]$ (ejercicio). Restando estas igualdades y usando que por hipótesis $\psi(x/n) - [x/n] = o(x/n)$ se sigue $\sum_{n\leq x} \mu(n)\log n = o(x\log x)$ y sumando por partes se tiene el resultado deseado.

$\Leftarrow$) Se relaciona ψ con μ "despejando" Λ en $\log = 1 * \Lambda$, lo que implica $\Lambda = \mu * \log$ y por tanto

$$\psi(x) = \sum_{n\leq x}\sum_{d|n} \mu(d)\log\frac{n}{d} = \sum_{m\leq x} \log m \sum_{d\leq x/m} \mu(d).$$

Por otro lado, empleando de la misma forma $1 = \mu * d$ y $f = \mu * 1$ (como en la implicación directa), se deduce combinando las fórmulas resultantes

$$\psi(x) - x + C = \sum_{m \le x} (\log m - d(m) + C) \sum_{d \le x/m} \mu(d)$$

para cualquier constante C. Dado $M \in \mathbb{N}$, usando la hipótesis y sumación por partes en el rango $m \le M$ e intercambiando el orden de sumación en $M < m \le x$, se obtiene

$$\psi(x) - x = -C + o(x(\log M)^2) + \sum_{d \le x/m} \mu(d) \sum_{M < m \le x/d} (\log m - d(m) + C).$$

De un ejemplo anterior, sabíamos que $\sum_{n \le x} d(n) = \sum_{n \le x} [x/n]$ pero notando que a cada divisor $d|n$ con $d \le \sqrt{n}$ le corresponde un divisor $n/d \ge \sqrt{n}$ y viceversa, se tiene

$$\sum_{n \le x} d(n) = 2 \sum_{n \le x} \left[\frac{x}{n}\right] - [\sqrt{x}]^2$$

donde $[\sqrt{x}]^2$ descuenta los divisores $d = n/d$, contados dos veces. Aproximando $[x/n] = x/n + O(1)$ y usando la fórmula para las sumas parciales de la serie armónica, se deduce sumando por partes (ejercicio)

$$\sum_{M < m \le x/d} (\log m - d(m) + C) = (C - 2\gamma)(x/d - M) + O(x^{1/2} d^{-1/2} + M^{1/2}).$$

Escogiendo $C = 2\gamma$ y sustituyendo en la fórmula para $\psi(x) - x$, se sigue

$$\limsup |\psi(x)/x - 1| \ll M^{-1/2}.$$

Como M es arbitrario, el límite de $\psi(x)/x - 1$ existe y es nulo. □

Antes de pasar a esbozar demostraciones del teorema de los números primos, veamos una prueba falsa utilizando el comportamiento de $\zeta(s)$ cuando $s \to 1$. Todo el razonamiento es una línea:

$$\zeta(s) \sim (s-1)^{-1} \quad \overset{?}{\Rightarrow} \quad \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \sim s - 1 \quad \overset{?}{\Rightarrow} \quad \sum_{n \le x} \mu(n) = o(x) \quad \overset{?}{\Rightarrow} \quad \text{TNP}.$$

En favor de la primera implicación está la relación $\zeta(s) D_\mu(s) = 1$, en favor de la segunda, la relación $\sum_{n \le x} \mu(n) = \sum_{n \le x} \sum_{m \ge n} \mu(m)/m$ y en favor de la última que la acabamos de probar. Se deja al lector el pasatiempo de localizar el error.

Las demostraciones clásicas del teorema de los números primos tienen como motivación tratar de despejar de alguna forma los primos de la identidad de Euler (1.3). Como allí están metidos de una manera complicada que involucra un producto, se muestra más sencillo despejar $\psi(x)$ de la relación

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

Veamos un esquema de dos demostraciones[12], como son sólo esquemas el lector debe imaginar interrogaciones donde sean necesarias.

a) Con análisis de Fourier:

Restemos $\zeta(s)$ en ambos miembros para quitar la singularidad de $s=1$. Escribiendo $n^{-\sigma-it} = n^{-\sigma}e(-t(\log n)/2\pi)$, al integrar contra una función f en los dos miembros (supóngase que es buena y hágase la vista gorda cuando se elija, o viceversa, más adelante habrá una justificación), se tiene para $\sigma > 1$

$$\int_{-\infty}^{\infty} g_\sigma(t)f(t)\,dt = \sum_{n=1}^{\infty}(\Lambda(n)-1)\frac{\widehat{f}((\log n)/2\pi)}{n^\sigma} \qquad \text{con} \quad g_\sigma(t) = -\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)} - \zeta(\sigma+it).$$

Eligiendo una función f tal que

$$\widehat{f}(u) = \begin{cases} e^{2\pi\sigma u} & \text{si } u \in [0,(2\pi)^{-1}\log x] \\ 0 & \text{si } u \notin [0,(2\pi)^{-1}\log x] \end{cases}$$

se tiene una expresión exacta en términos de la función ζ para $\psi(x)-x$. Si nos molestamos en calcular la transformada inversa veremos que $f(t) = (x^{\sigma+it}-1)/(2\pi(\sigma+it))$. En un pequeño entorno del origen la función g_σ no difiere mucho de una constante, mientras que para x grande f se comporta como $x^{\sigma+it}$, entonces la contribución correspondiente es del orden de $x^\sigma/\log x$ y esto es malo si $\sigma > 1$, lo cual sugiere (¿obliga?) tomar $\sigma = 1$ entendiéndolo como una especie de límite. Hay dos problemas al respecto. El primero es el más serio y consiste en que no está claro que g_1 esté bien definida (un cero de ζ en $\Re s = 1$ causaría un desastre). Esto se resuelve con un estudio cuidadoso de la función ζ que prueba que g_1 crece como una potencia de logaritmo. El otro problema, más técnico, es que la convergencia de la integral no está asegurada, necesitábamos una función buena y hemos escogido una mala que sólo decae como t^{-1} y tiene transformada de Fourier discontinua. La solución pasa por regularizar un poco la f para ganar todas las potencias de logaritmo que deseemos aunque el principio de incertidumbre esté contra nosotros.

b) Con variable compleja:

Se puede despejar la suma de los coeficientes de una serie de Dirichlet gracias a la fórmula mágica

$$\lim_{T\to\infty} \frac{1}{2\pi i}\int_{L_T} \frac{y^s}{s}\,ds = \begin{cases} 0 & \text{si } 0 < y < 1 \\ 1 & \text{si } y > 1 \end{cases} \tag{1.13}$$

donde L_T es el segmento $\{s \;:\; \Re s = \sigma_0 > 0,\ |\Im s| < T\}$ orientado hacia arriba. La demostración se reduce al teorema de los residuos aplicado a la derecha o a la izquierda de L_T.

[12]Por si el lector está interesado, hay algunas pruebas muy simplificadas que siguen las líneas clásicas, por ejemplo las incluidas en [Ne] y en [Iw-Ko] p.40. También hay diferentes pruebas elementales (pero no sencillas) comenzando por las que dieron Erdős y Selberg (véase una en [El]). A pesar de su interés, tienen la desventaja de no revelar la verdadera naturaleza del término de error.

Si se tiene la convergencia adecuada (y si no se trunca la integral y se acota el error [Da]) entonces

$$\sum_{n\le x} f(n) = \frac{1}{2\pi i}\int_{L_\infty} D_f(s)\frac{x^s}{s}\,ds \qquad \text{para } x\in\mathbb{R}^+-\mathbb{N}.$$

En nuestro caso elegimos $\sigma_0>1$, $f(n)=\Lambda(n)$ y $D_f(s)=-\zeta(s)/\zeta(s)$. Sabemos que esta última función tiene una extensión meromorfa con al menos un polo simple de residuo 1 en $s=1$ (porque $\zeta(s)\sim(s-1)^{-1}$). Si $\mathcal{R}$ es una región que contiene al semiplano $\Re s\ge 1$ (y por tanto a L_∞) en la cual no hay ceros de ζ, entonces por el teorema de los residuos

$$\psi(x) = x - \frac{1}{2\pi i}\int_{\partial\mathcal{R}} \frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds \tag{1.14}$$

donde $\partial\mathcal{R}$ es el borde de $\mathcal{R}$ (orientado como L_∞). Si se tiene cierto control sobre ζ'/ζ y se asegura la existencia de tal región con $\partial\mathcal{R}$ contenido en $\Re s<1$, de (1.14) se deduce el teorema de los números primos. También queda el cabo suelto de la convergencia de las integrales, pero como ya hemos insinuado se resuelve sustituyendo (1.13) por una fórmula con T finito y término de error.

La prueba que acabamos de ver es muy ilustrativa. ¿Qué pasaría si se amplía la región $\mathcal{R}$ llevándola más y más a la izquierda sin importarnos que aparezcan ceros de ζ en su interior? Entonces x^s tendería a cero cuando $\Re s\to-\infty$ y habida cuenta que los polos del integrando en (1.14) son $s=0$ y $s=\rho$ donde ρ recorre los ceros de ζ, se tendría la llamada *fórmula explícita*

$$\psi(x) - x = -\frac{\zeta'(0)}{\zeta(0)} - \sum_\rho \frac{x^\rho}{\rho}$$

(como antes, para $x>0$ no natural). ¡Una fórmula explícita! A decir verdad, debido a la convergencia pobre es mucho más útil una fórmula no explícita sino truncada, como

$$\psi(x) = x - \sum_{|\rho|<T}\frac{x^\rho}{\rho} + O\big(\frac{x}{T}\log^2(xT)+\log x\big) \tag{1.15}$$

que es válida para cualquier $x>0$ porque la diferencia entre la contribución a $\psi(x)$ de un número entero y uno casi entero es absorbida por $O(\log x)$. Con esta fórmula tenemos que cuanto menor sea la parte real de los ceros, más pequeño será el error en el teorema de los números primos, recuérdese (1.12), y si un solo cero malvado tiene parte real casi uno, nos tendremos que aguantar con un error malo. ¿No es increíble que la distribución de los primos dependa tan estrechamente de una función de variable compleja?

Se sabe que hay infinitos ceros con $\Re\rho=1/2$ (esto se debe a Hardy, véase [El]) por tanto lo mejor que podemos esperar es lo que ya esperaba Riemann en 1860 (véase el apéndice de [Ed] con la traducción inglesa de la memoria original).

Hipótesis de Riemann: *Todos los ceros de ζ en el semiplano derecho tienen parte real* $1/2$.

Lo de restringirse al semiplano derecho se debe a que se conoce que los únicos ceros en el semiplano izquierdo son $-2, -4, -6, \ldots$, se dice que éstos son *ceros triviales*. Por otra parte, si hubiera ceros en $0 < \Re s < 1/2$, como veremos, también estarían sus simétricos por $\Re s = 1/2$ en $1/2 < \Re s < 1$.

Se sabe que hay $O(\log T)$ ceros de ζ con $T < |\rho| < T + 1$, $T > 2$, con lo cual de (1.15) y del Lema 1.4.3 se deduce

Corolario 1.4.5 *Si la hipótesis de Riemann es cierta, entonces*

$$\psi(x) = x + O(x^{1/2}(\log x)^2) \qquad y \qquad \pi(x) = Li(x) + O(x^{1/2} \log x).$$

La hipótesis de Riemann es un problema abierto desde hace casi 150 años (a pesar de los premios del paso a la posteridad y un millón de dólares) lo cual es suficiente para ahuyentar a casi todos los investigadores y atraer a no pocos diletantes. El avance actual es tan leve que siquiera se conoce ningún semiplano abierto conteniendo a $\Re s \geq 1$ en el que no haya ceros.

Para no sucumbir al desánimo ante un muro infranqueable, terminaremos con una breve lista de utensilios que sí podemos adquirir y que constituyen los puntos fundamentales que se han dejado en suspenso en los esquemas anteriores y los comentarios que los siguen.

1. Simetría de los ceros y ceros triviales.

2. Crecimiento del módulo de los ceros.

3. Región libre de ceros.

4. Acotaciones de ζ y ζ'/ζ.

El primer punto es la ecuación funcional por antonomasia y se debe al propio Riemann.

Proposición 1.4.6 *La función ζ tiene una extensión holomorfa a $\mathbb{C} - \{1\}$ y verifica la ecuación funcional*

$$\Gamma(s/2)\zeta(s) = \pi^{-(1-s)/2}\Gamma((1-s)/2)\zeta(1-s).$$

Observación: Recuérdese que $\Gamma(z)$ está definida por $\int_0^\infty t^{z-1}e^{-t}\, dt$ si $\Re z > 0$ y se extiende analíticamente a $\mathbb{C} - \mathbb{Z}^- - \{0\}$ mediante $\Gamma(z+1) = z\Gamma(z)$. Casi todas las propiedades de la función Γ se pueden deducir de la fórmula (véase [Ah]) $1/\Gamma(z) = se^{\gamma s}\prod(1 + z/n)e^{-z/n}$ donde n recorre $\mathbb{Z}^+$ y γ es la constante de Euler. La aproximación de Stirling es válida para $\Gamma(z+1)$ cambiando N por z, se cumple $|\Gamma(z)/\Gamma'(z)| = O(\log|z|)$ y $\Gamma(z) \neq 0$.

Demostración: Hay varias pruebas de la ecuación funcional [Ti], aquí seguiremos la de la memoria de Riemann, que parte de lasiguiente fórmula fruto de la definición de la función Γ en $s/2$, tras el cambio variable $t \mapsto \pi n^2 t$

$$\pi^{-s/2}\Gamma(s/2)\sum_{n=1}^{\infty} n^{-s} = \sum_{n=1}^{\infty}\int_0^{\infty} t^{s/2-1}e^{-\pi n^2 t}\,dt \qquad \text{para} \quad \Re s > 1,$$

o equivalentemente

$$\pi^{-s/2}\Gamma(s/2)\sum_{n=1}^{\infty} n^{-s} = \frac{1}{2}\int_0^{\infty} t^{s/2-1}(\theta(t)-1)\,dt \qquad \text{donde} \quad \theta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t}.$$

Lo que se gana es que se puede aplicar la fórmula de Poisson a $\theta(t)$ dentro de la integral. Con ello esencialmente t pasará a $1/t$ y por tanto la parte de la integral $\int_0^1$ se transformará en $\int_1^{\infty}$. Esto es interesante para llevar a cabo la extensión ya que la divergencia de $\int_0^1 t^{s/2-1}t^{-1/2}dt$ para $\Re s < 1$ es la responsable de que se pueda extender el segundo miembro. Con esta idea en mente separando el rango de integración y utilizando $\theta(t) = t^{-1/2}\theta(1/t)$, el segundo miembro es

$$\frac{1}{2}\int_0^1 + \frac{1}{2}\int_1^{\infty} = \frac{1}{2}\int_0^1 t^{s/2-1}(t^{-1/2}\theta(1/t)-1)\,dt + \frac{1}{2}\int_1^{\infty} t^{s/2-1}(\theta(t)-1)\,dt.$$

Con el cambio $t \mapsto 1/t$ en la primera integral se llega a que para $\Re s > 1$

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2}\int_1^{\infty}(t^{s/2-1} + t^{-1/2-s/2})(\theta(t)-1)\,dt \tag{1.16}$$

Ahora la integral tiene sentido para todo $s \in \mathbb{C}$ y como $\Gamma(s/2)$ no se anula, la función ζ así definida (que coincide con $\sum n^{-s}$ en $\Re s > 1$) es meromorfa en $\mathbb{C}$ y holomorfa en $\mathbb{C} - \{0, 1\}$. Usando que $\lim_{s\to 0} s\Gamma(s/2) = 2$ y $\Gamma(1/2) = \pi^{1/2}$ (véase [Ah]), se deduce que ζ es de hecho meromorfa en $\mathbb{C}$ con un único polo en $s = 1$ de residuo 1. Además la invariancia del segundo miembro de (1.16) al cambiar s por $1 - s$ termina la prueba. □

Corolario 1.4.7 *La función ζ tiene ceros simples en $s = -2, -4, -6$ (los llamados ceros triviales) y todos los ceros restantes están en la banda crítica $0 \le \Re s \le 1$. Además si ρ es un cero no trivial, también lo son $\overline{\rho}$, $1 - \rho$ y $1 - \overline{\rho}$.*

Demostración: De (1.3) se deduce que $\zeta(s) \neq 0$ en $\Re s > 1$ y por la ecuación funcional, $\Gamma(s/2)\zeta(s)$ no se anula en $\Re s < 0$. Como $\Gamma(s/2)$ tiene polos simples en $s = -2, -4, -6, \ldots$, es fácil deducir que éstos son ceros simples de ζ.

Finalmente, las simetrías de los ceros se deducen de la ecuación funcional y la relación $\zeta(s) = \overline{\zeta(\overline{s})}$ que es obvia en $\Re s > 1$ por $\zeta(s) = \sum n^{-s}$ y se extiende analíticamente. □

Para el resto de las propiedades se aplica la teoría de funciones de orden finito de de Hadamard [Ah] (quien la creó en relación con la función ζ) para obtener una bella fórmula.

Proposición 1.4.8 *Si s no es un polo de $\zeta'(s)/\zeta(s)$, se cumple*

$$\frac{\zeta'(s)}{\zeta(s)} = K - \frac{1}{s-1} - \frac{1}{2}\frac{\Gamma'(s/2+1)}{\Gamma(s/2+1)} + \sum_{\rho}\big(\frac{1}{s-\rho}+\frac{1}{\rho}\big)$$

donde ρ recorre los ceros no triviales de ζ y K es cierta constante.

Demostración: De acuerdo con (1.16) y la ecuación funcional, la función

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$$

es entera, cumple $\xi(s) = \xi(1-s)$ y $|\xi(s)| = O\big(e^{C|s|\log|s|}\big)$ para cierta constante $C > 0$ cuando $|s| \to \infty$. La teoría de funciones de orden finito asegura en este caso [Ah] que

$$\xi(s) = e^{A+Bs}\prod(1-s/\rho)e^{s/\rho}$$

donde el producto es sobre todos los ceros ρ no triviales de ζ, esto es $0 \leq \Re\rho \leq 1$. Tomando logaritmos y derivando se tiene la relación buscada. □

Corolario 1.4.9 *Para $T > 2$ el número de ceros con $T \leq |\Im\rho| \leq T+1$ es $O(\log T)$.*

Demostración: Si en la proposición anterior escribimos $s = 2+iT$ y tomamos partes reales, se tiene

$$1 \gg -\log T + \sum_{\rho}\mathrm{Re}\,\big(\frac{1}{2+iT-\rho}+\frac{1}{\rho}\big).$$

Después de calcular la parte real (recuérdese que $0 \leq \Re\rho \leq 1$) se sigue $\log T \gg \sum(1+(T-\Im\rho)^2)^{-1}$ y de aquí el resultado. □

Proposición 1.4.10 *Existe una constante C (efectiva) tal que la región*

$$\big\{\sigma+it \;:\; \sigma > 1 - \frac{C}{\log(|t|+2)}\big\}$$

está libre de ceros de la función ζ.

Demostración: El argumento parece mágico y tiene una base intutiva que no se debe ocultar: si existiera un cero no trivial σ_n+it_n muy cerca de $\Re s = 1$ entonces para $\sigma \to 1^+$ se tendría que $-\zeta'(\sigma+it_n)/\zeta(\sigma+it_n)$ tiene parte real muy grande y negativa. En ese caso, $\zeta'(s)/\zeta(s) = \sum\Lambda(n)n^{-s}$ sugiere que $\cos(t_n\log p)$ toma muchas veces valores negativos. Entonces, recíprocamente, $\cos(2t_n\log p)$ debe tomar muchas veces valores positivos y $-\zeta'(\sigma+2it_n)/\zeta(\sigma+2it_n)$ debe tener parte real grande y positiva. Controlando el tamaño de esta última cantidad controlaremos la cercanía del posible cero a la línea $\Re s = 1$. Lo más ingenioso, a la par que simple, es la manera de cuantificar los tamaños relativos al evaluar en $\sigma+it_n$ y en $\sigma+2it_n$. Se emplea para ello la sencilla desigualdad trigonométrica

$$3 + \cos(2\alpha) \geq -4\cos\alpha \qquad \forall\alpha \in \mathbb{R}$$

y con ello comienza la prueba.

Sustituyendo $\alpha = t_n \log p$ y sumando con coeficientes adecuados se tiene, para $\sigma > 1$,

$$-3\frac{\zeta'(\sigma)}{\zeta(\sigma)} - \Re\frac{\zeta'(\sigma+2it_n)}{\zeta(\sigma+2it_n)} \geq 4\Re\frac{\zeta'(\sigma+it_n)}{\zeta(\sigma+it_n)}$$

De la Proposición 1.4.8 cuando $\sigma > 1$ está suficientemente cercano a 1 se cumple $-\zeta'(\sigma)/\zeta(\sigma) < (\sigma-1)^{-1} + \text{cte}$, y además

$$-\Re\frac{\zeta'(\sigma+2it_n)}{\zeta(\sigma+2it_n)} < \text{cte } \log(|t_n|+2) \quad \text{y} \quad -\Re\,\frac{\zeta'(\sigma+it_n)}{\zeta(\sigma+it_n)} < \text{cte } \log(|t_n|+2) - \frac{1}{\sigma-\sigma_n}.$$

Para probar estas desigualdades utilícese, aparte de la Proposición 1.4.8, $\Gamma'(s)/\Gamma(s) = O(\log|s|)$ y que $\Re((s-\rho)^{-1}+\rho^{-1}) > 0$ para $\Re s > 1$. Sustituyendo se obtiene

$$3/(\sigma-1) + \text{cte}\log(|t_n|+2) \geq 4/(\sigma-\sigma_n).$$

Tomando $\sigma = 1 + \epsilon/\log(|t_n|+2)$ con ϵ pequeño se sigue que $\sigma_n \leq 1 - \text{cte}/\log(|t_n|+2)$ para cierta constante positiva, que equivale al resultado. □

Por último veamos acotaciones para ζ y ζ'/ζ en $\Re s = 1$ que permitirían completar la prueba con análisis de Fourier del teorema de los números primos. (En realidad con mucha menos precisión es suficiente [Dy-Mc] §3.10)

Proposición 1.4.11 *Para $|t| > 2$ se cumple*

$$\zeta(1+it) = O(\log t) \qquad y \qquad -\frac{\zeta'(1+it)}{\zeta(1+it)} = O\big((\log t)^2\big).$$

Demostración: La primera es más elemental, sólo requiere sumar por partes en $\sum_{n>t} n^{-s}$ (con $\Re s > 1$) para concluir como en (1.6) que la fórmula

$$\zeta(s) - \sum_{n\leq t} n^{-s} = O(1) - s\int_t^\infty \frac{u-[u]}{u^{s+1}}\,du$$

es válida para $\Re s \geq 1$. Eligiendo $s = 1+it$ y haciendo las acotaciones triviales, se obtiene el resultado.

Para la segunda, tomando $s = 1+iT$ en la Proposición 1.4.8 y restando lo obtenido al sustituir $s = 2+iT$, se sigue

$$\frac{\zeta'(1+iT)}{\zeta(1+iT)} \ll \log T + \sum_\rho \left|\frac{1}{1+iT-\rho} - \frac{1}{2+iT-\rho}\right|.$$

Los $O(\log T)$ sumandos correspondientes a $|T - \Im\rho| \leq 1$ contribuyen $O(\log^2 T)$ en total por la condición de la distancia. La contribución de los correspondientes a $|T - \Im\,\rho| > 1$ es menor sin más que emplear la acotación para $\sum(1+(T-\Im\rho)^2)^{-1}$ (véase el final de la prueba del Corolario 1.4.9). □

Bibliografía

[Ah] L.V. Ahlfors. Análisis de variable compleja: Introducción a la teoría de funciones analíticas de una variable compleja. Aguilar, Madrid 1971.

[Bo-Sh] A.I. Borevich, I.R. Shafarevich. Number theory. Pure and Applied Mathematics, Vol. 20 Academic Press, New York-London 1966.

[Cl] A. Clark. Elementos de álgebra abstracta. Alhambra, 1987.

[Co] A. Córdoba. Disquisitio numerorum. Gac. R. Soc. Mat. Esp. 4 (2001), no. 1, 249–260.

[Da] H. Davenport. Multiplicative number theory (2nd ed.). Graduate texts in Mathematics 74. Springer-Verlag, New York-Berlin, 1980.

[Do-He] J.R. Dorronsoro, E. Hernández. Números, grupos y anillos. Addison-Wesley Iberoamericana–UAM, 1996.

[Dy-Mc] H. Dym, H.P. McKean. Fourier series and integrals. Probability and Mathematical Statistics 14. Academic Press, New York-London, 1972.

[Ed] H.M. Edwards. Riemann's zeta function. Pure and Applied Mathematics, Vol. 58. Academic Press, New York-London, 1974.

[El] W.J. Ellison. Les nombres premiers. En collaboration avec Michel Mendès France. Publications de l'Institut de Mathématique de l'Université de Nancago, No. IX. Actualités Scientifiques et Industrielles, No. 1366. Hermann, Paris, 1975.

[Fo] G.B. Folland. Fourier analysis and its applications. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992.

[Ga] C.F. Gauss. Disquisitiones arithmeticae. Springer-Verlag, New York, 1986.

[Gr-Ko] S.W. Graham, G. Kolesnik. Van der Corput's method of exponential sums. London Mathematical Society lecture note series 126. Cambridge University Press, 1991.

[In] A.E. Ingham. The distribution of prime numbers. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1990.

[Iw-Ko] H. Iwaniec, E. Kowalski. Analytic number theory. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[Hu] M.N. Huxley. Area, lattice points, and exponential sums. London Mathematical Society Monographs 13. The Clarendon Press, Oxford University Press, New York, 1996.

[Ja] N. Jacobson. Lectures in Abstract Algebra. Vol.1,2,3. Van Nostrand, 1964.

[La] E. Landau. Über die Anzahl der Gitterpunkte in Gewissen Bereichen. Göttinger Nachr. (1912) 687–771.

[Ne] D.J. Newman. Analytic number theory. Graduate Texts in Mathematics, 177. Springer-Verlag, New York, 1998.

[Po] A.G. Postnikov. Introduction to analytic number theory. Translations of Mathematical Monographs, 68. American Mathematical Society, Providence, RI, 1988.

[Ri] P. Ribenboim. 13 Lectures on Fermat's Last Theorem. Springer-Verlag, 1979.

[Ro] H.E. Rose. A course in number theory. Second edition. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1994.

[Sm] D.E. Smith. A Source Book in Mathematics. Dover Publications Inc., 1959.

[Sp] M. Spivak. Calculus. Vol. 1 y 2. Reverté, Barcelona, 1984.

[St] I. Stewart. Galois Theory. Chapman & Hall/CRC Mathematics, Boca Raton, FL, 2004.

[St-Ta] I. Stewart, D. Tall. Algebraic number theory. Chapman and Hall Mathematics Series. Chapman & Hall, London, 1987.

[Ti] E.C. Titchmarsh. The theory of the Riemann zeta-function. Second edition. The Clarendon Press, Oxford University Press, New York, 1986.

[Wi] E. Wirsing. Das asymptotische Verhalten von Summen über multiplikative Funktionen. Math. Ann. 143 (1961) 75–102.